A Multimodal Approach to Persuasion in Oral Presentations

The case of conference presentations, research dissemination talks and product pitches

PhD dissertation presented by Julia Valeiras Jurado

Supervised by:

Inmaculada Fortanet Gómez

Noelia Ruiz Madrid

Geert Jacobs

25 April 2017

ACKNOWLEDGEMENTS

This PhD would not have been possible without the fruitful cooperation established between Universitat Jaume I and Ghent University. The venture of a joint PhD has clearly translated into added value in this research project, facilitating access to data and providing an unusual and enriching variety of points of view.

I also want to express my gratitude to the participants in this study, the speakers in the presentations, who have kindly given consent to use their presentations as object of my analysis and have patiently made themselves available for follow-up interviews.

Likewise, I remain indebted to my three supervisors, not only for their wise guidance in this project, but also for being an example of good team work and for turning this long journey into an enjoyable experience, worthwhile at the academic and the personal level.

Last but not least, I want to thank my husband, for being both my technical and my emotional support throughout.

To all of you, thank you.

RESUMEN

Esta tesis presenta un estudio multimodal y etnográfico del uso de estrategias persuasivas en tres géneros orales. Los géneros en cuestión son presentaciones en conferencias, charlas de divulgación científica, y presentaciones de productos. Estos géneros comparten un importante componente persuasivo: los tres se dirigen a una audiencia tratando de convencerles del valor de un producto, servicio, o investigación. Sin embargo, se usan en dos contextos profesionales diferentes: el académico y el económico, por lo que cabe esperar que consigan su propósito comunicativo de forma diferente. Por otra parte, recientes estudios muestran como distintos discursos, entre el que se incluye el discurso académico, tienden a adoptar cada vez más rasgos promocionales (lo que se ha denominado promocionalización o marquetinización del discurso). En vista de ello, es factible establecer como hipótesis que los tres géneros están relacionados intertextual e interdiscursivamente. En mi opinión un estudio multimodal y etnográfico del uso de la persuasión en dichos géneros puede ayudar a clarificar las relaciones existentes entre ellos, así como sus diferencias.

La motivación de este estudio es doble. Por un lado, intenta elucidar cómo estos géneros consiguen sus propósito comunicativo (persuadir a una audiencia). Con este estudio intento averiguar lo que hace que estas presentaciones sean efectivas y convincentes. Por otro lado, intenta reconocer el importante papel que juegan en esta tarea los rasgos kinésicos y paralingüísticos del discurso oral. A menudo dichos rasgos son subestimados, ya que tanto los ponentes como los investigadores tienden a centrar su atención en la parte verbal-lingüística (el “texto”) y en los elementos visuales, que habitualmente toman la forma de presentaciones en formato electrónico (p. ej. PPT, Prezi, entre otras). Por ejemplo, los ponentes suelen centrarse más en el *qué* van a decir (el contenido de su mensaje) que en *cómo* lo van a decir (es decir, la forma en que se transmite el mensaje). Cuando prestan atención al *cómo*, se centran en diseñar diapositivas visualmente atractivas y un buen guión, pero no dedican la misma atención a otros aspectos kinésicos y paralingüísticos (gestos, movimientos de cabeza, entonación) que también juegan un papel central en este *cómo*. Esta es una tendencia que estudios anteriores han detectado, y que los resultados de mi análisis confirman. La importancia de estos aspectos kinésicos y paralingüísticos se ha tenido en cuenta principalmente en contextos de promoción comercial o en el ámbito de la comunicación empresarial, y mayormente en forma de guías que ofrecen consejos prácticos. Esto demuestra que los profesionales que utilizan estos géneros en su día a día parecen ser conscientes de que estos aspectos afectan profundamente el resultado final de una presentación. Con esta tesis

quiero investigar cuál es exactamente este efecto, y cómo puede ayudar a diseñar una presentación persuasiva. Dicho de otro modo, la gran mayoría de nosotros prefiere a un ponente dinámico, que usa una amplia gama de gestos y tonos, a un ponente monótono que da la impresión de estar leyendo en alto. La cuestión es *¿por qué?* Este estudio es mi contribución a la línea de investigación en el área específica del análisis multimodal del discurso que intenta dar respuesta a esta pregunta. Hasta ahora, esta línea se ha centrado mayormente en discurso académico, y en particular en las posibles aplicaciones pedagógicas: con esta tesis es mi intención expandir el ámbito a otros géneros diferentes, aunque potencialmente relacionados.

Esta tesis tiene tres objetivos. En primer lugar, propone una metodología que combina multimodalidad y etnografía para estudiar el uso de estrategias persuasivas de forma sistemática, sin priorizar ningún recurso semiótico y permitiendo la triangulación de resultados. En segundo lugar, en este estudio pongo a prueba esta metodología en un corpus de 15 presentaciones, lo que ha dado como resultado una taxonomía que recoge 7 estrategias persuasivas llevadas a cabo de forma multimodal (es decir, mediante un conjunto de recursos semióticos coherentemente orquestado). En tercer lugar, aplicando este marco metodológico analizo el uso de estrategias persuasivas y sus realizaciones semióticas en los tres géneros objeto de este estudio, centrándome en las palabras, la entonación, los gestos y los movimientos de cabeza como recursos semióticos.

El estudio intenta dar respuesta a tres preguntas principales:

1. ¿Qué estrategias persuasivas se pueden identificar desde un enfoque multimodal?

2. ¿Cómo contribuye el conjunto de recursos semióticos incluidos en el estudio (palabras, entonación, gestos y movimientos de cabeza) a la persuasión en los géneros analizados?

3. ¿Hasta qué punto es el uso de estrategias persuasivas específico para cada género y hasta qué punto hay similitudes entre ellos?

La tesis está estructurada en 8 capítulos. Los primeros tres capítulos establecen el marco teórico para mi estudio. El capítulo 4 sirve de transición entre teoría y práctica, detallando mi marco metodológico. Los cuatro capítulos restantes constituyen la parte empírica: el capítulo 5 se centra en el estudio (recogida de datos y análisis), el capítulo 6 presenta los resultados y el capítulo 7 ofrece una discusión de dichos resultados. Finalmente, en el capítulo 8 presento mis conclusiones, sugiero aplicaciones pedagógicas e identifico áreas que pueden ser objeto de futuras investigaciones.

En respuesta a la primera cuestión, la combinación de multimodalidad y etnografía como marco metodológico hace posible estudiar el uso de estrategias persuasivas de forma sistemática, en este caso identificando siete estrategias persuasivas que se realizan de forma multimodal en el corpus analizado.

En relación a la segunda cuestión, los resultados del estudio muestran una correlación entre densidad modal (el número de recursos semióticos utilizados), coherencia modal (la forma en que estos recursos semióticos se combinan formando un conjunto coherente) y la persuasión. Es decir, los ponentes son más persuasivos cuando son conscientes de los recursos semióticos que tienen disponibles y cuando son capaces de usarlos de forma coherente.

En lo que respecta a la tercera cuestión, el estudio etnográfico y multimodal del uso de estrategias persuasivas en el corpus analizado ha demostrado que la relación establecida entre ponente y audiencia es un rasgo distintivo de cada uno de los tres géneros a la vez que un requisito para persuadir. Los ponentes necesitan establecer la relación correcta con la audiencia para poder convencerles, y para ello utilizan distintas estrategias persuasivas en diferentes frecuencias, así como distintos tipos de gestos. Otras diferencias en el uso de recursos semióticos son atribuibles a diferencias individuales o a los medios existentes (p. ej. la distribución de la sala, o la tecnología disponible), y no constituyen una característica distintiva del género. El análisis también sugiere relaciones intertextuales e interdiscursivas entre los tres géneros: en el corpus analizado las presentaciones en conferencias asimilan rasgos de las charlas de divulgación científica, y éstas a su vez asumen rasgos de las presentaciones de productos para conseguir la relación correcta con la audiencia.

Estos resultados pueden ser de utilidad para futuros estudios de género y contribuir a las definiciones existentes de los géneros en cuestión, lo cual a su vez puede ayudar a mejorar el material didáctico para la enseñanza de idiomas con fines específicos.

RESUMEN

ABSTRACT

This thesis is a multimodal and ethnographic study of the use of persuasive strategies in three oral genres used in two different professional settings: business and academia. In particular, the genres concerned are conference presentations, research dissemination talks and product pitches. These presentations share a strong persuasive component in their communicative purpose: the three of them address an audience to convince them of the value of a product, a service or a piece of research. However, they are used in business and academia by different discourse communities in different contexts, and therefore they can be expected to achieve their communicative goals in different ways. In addition, research suggests that there is a trend towards promotionalization (also called marketization) of different discourses, among which academic discourse is included. In view of this, I hypothesize that these three genres are intertextually and interdiscursively related, and that a multimodal and ethnographic study of the use of persuasion in them can help to shed some light on these relationships and differences.

The motivation of this study is twofold. First, it is intended to gain insight into how these three genres fulfil their main communicative purpose: persuade an audience. With this study I want to probe into what makes these presentations effective and convincing. Second, it tries to acknowledge the important role that kinesic and paralinguistic features play in this task. Often, these aspects are underestimated in oral presentations as both speakers and researchers tend to pay more attention to the verbal, linguistic component of the presentation and the visual support, usually in the form of electronic slide decks. For example, speakers tend to devote more attention to *what* they say (meaning the content of their message), than to *how* they say it (meaning the way this message is conveyed). In addition, when they care about the *how* they tend to focus on attractive slide design and on crafting an effective script, without devoting the same attention to other kinesic and paralinguistic semiotic modes (e.g. gestures, head movements and intonation), which are nonetheless integral parts of this *how*. This is a tendency already spotted by previous research that the results of this study have corroborated. The effect of kinesic and paralinguistic aspects such as body language or voice qualities in persuasive oral discourse has been often acknowledged, in particular in business contexts, but mostly in the form of guides offering practical tips. This shows that professionals using these oral genres in their everyday work seem to be aware of the fact that kinesic and paralinguistic aspects have an effect on their performance. With this thesis, I want to probe into the exact contribution of these aspects and how they can help design a persuasive presentation. In other words, most people prefer a dynamic speaker who uses varied gestures and intonation to a monotonous speaker who seems to be reading out loud. The question is, *why is this so?* This study is my contribution to the line of research in multimodal discourse analysis that tries to answer this question. This line of research has so far focused mainly on academic discourse, with special interest in effective lecturing and the influence of different semiotic modes on the learning process: with this thesis I try to expand the scope to other different, but potentially related, discourses and genres.

This thesis has three main objectives. First, it suggests a methodology that combines multimodality and ethnography to approach the use of persuasive strategies in a systematic way, without prioritizing any semiotic mode in particular and allowing for the triangulation of results. Second, in this study I test this methodology on a corpus of 15 presentations, which has yielded a taxonomy of seven persuasive strategies that are realized as multimodal ensembles. Third, applying this methodological framework I analyse the use of persuasive strategies and their modal realisations (focusing on words, intonation, gestures and head movements) in the three genres which are object of the study.

The following three research questions guide my study:

RQ I. What type of persuasive strategies can be identified using an MDA approach?

RQ II. What is the contribution of the multimodal ensemble orchestrated through words, intonation, head movements and gestures to the creation of persuasion in the genres included in the analysis?

RQ III. To what extent is the use of persuasive strategies genre-specific and to what extent is there overlap between the three genres?

The thesis is structured in 8 chapters. The first three chapters provide the theoretical framework for my study. Chapter 4 makes the transition from theory to practice detailing my methodological framework. The remaining four chapters describe the empirical analysis: Chapter 5 focuses on the study, the results are presented in Chapter 6, a discussion of these results follows in Chapter 7 and finally Chapter 8 provides some conclusions, suggests some pedagogical applications and identifies areas for further research.

In answer to my first research question, the combined multimodal and ethnographic methodology has made it possible to approach the use of persuasive strategies systematically,

identifying seven persuasive strategies that are realized multimodally in my corpus of presentations.

Regarding my second research question, the results of the study reveal a correlation between modal density (i.e. the number of modes used), modal coherence (i.e. the way these modes are orchestrated into a coherent multimodal ensemble) and persuasion. In other words, speakers in oral presentations become more persuasive when they are aware of the modes they can use, and when they are able to use them coherently.

Concerning my third research question, the multimodal and ethnographic analysis of the use of persuasive strategies in my corpus of presentations has shown that the type of relationship that is established with the audience stands out as a defining trait of each genre, and as a crucial element to pave the way for persuasion, because speakers need to target at the right interpersonal relationship in order to convince their audiences. Furthermore, this generic difference is revealed in the frequency with which speakers use persuasive strategies and in the type of gestures that they use, while other differences in the use of modes are attributable to individual differences or the media available, and cannot be considered a distinctive trait of the genres. The study has also suggested intertextual and interdiscursive connections among the three genres: the conference presentations in the corpus draw on the research dissemination talks and these in turn draw on the product pitches to achieve the right interpersonal relationship with the audience.

These results can pave the ground for further genre studies and contribute to existing descriptions of the three genres. This knowledge can then be fed into Language for Specific Purposes (LSP) teaching materials, ultimately helping students and trainees design more persuasive presentations.

OVERZICHT

Deze thesis is een multimodale, etnografische studie naar het gebruik van persuasieve strategieën in drie mondelinge genres, in twee professionele omgevingen: de bedrijfswereld en de academische wereld. Dit onderzoek focust zich meer specifiek op drie genres: congrespresentaties, populariserende wetenschappelijke presentaties en productpresentaties. Al deze presentaties hebben een sterk persuasief component in hun communicatief doel: ze dienen steeds een publiek te overtuigen van de waarde van een product, service of academisch onderzoek. Per genre zijn er echter verschillende discursieve gemeenschappen in verschillende contexten aan het woord, dus zijn er verschillen te verwachten in hoe deze gemeenschappen hun communicatief doel effectief bereiken. Daarnaast suggereert de bestaande literatuur dat er een trend is naar commercialisering (of vermarkting) van de verschillende discours, het academische discours incluis. In het licht van deze bevindingen stel ik voorop dat deze drie genres intertekstuele en interdiscursieve verbanden zullen vertonen, en dat een multimodaal en etnografisch onderzoek naar persuasieve strategieën deze onderlinge relaties kan blootleggen, evenals de verschillen.

De motivering voor dit onderzoek is tweeledig. Vooreerst wil het inzicht verschaffen in hoe de drie genres slagen in hun voornaamste communicatieve doel, een publiek overtuigen, en welke elementen hiertoe bijdragen. Ten tweede wil ik de belangrijke rol van lichaamstaal en paralinguïstische elementen in dit proces belichten. Deze aspecten worden namelijk vaak onderschat in mondelinge presentaties; zowel sprekers uit de bedrijfswereld en onderzoekers besteden gewoonlijk meer aandacht aan de verbale en linguïstische componenten van de presentatie, en aan de visuele ondersteuning, meestal in de vorm van slides. Wat ze zeggen, oftewel de inhoud van de boodschap primeert boven hoe ze het zeggen, of de manier waarop de boodschap wordt overgebracht. Daarnaast blijkt dat, als ze aandacht besteden aan hoe ze de boodschap overbrengen, dit meestal gefocust is op het maken van aantrekkelijke slides en aan het opstellen van een effectief scenario, en niet op lichaamstaal en paralinguïstische semiotische modi (bijvoorbeeld handgebaren, bewegingen van het hoofd, en intonatie), hoewel deze modi integraal deel uitmaken van hoe men een boodschap overbrengt. Deze tendens is reeds beschreven in voorgaand onderzoek; het effect van lichaamstaal en paralinguïstische elementen zoals lichaamstaal en stem in persuasief mondeling discours is erkend, meer specifiek in de context van de zakenwereld, meestal in de vorm van handleidingen met praktische tips. Dit toont dat professionals die deze mondelinge genres in hun dagelijkse praktijk beoefenen zich bewust lijken te zijn van het effect van lichaamstaal en paralinguïstische elementen op hun prestaties. Met deze thesis wil ik peilen naar de exacte rol van deze elementen en hoe ze kunnen bijdragen tot het opstellen van een overtuigende presentatie. Het publiek verkiest gewoonlijk een spreker die varieert in handgebaren en in intonatie, boven een monotone spreker die slechts hardop voorleest. De vraag is echter waarom dit zo is, en deze vraag uit het onderzoek naar multimodale discoursanalyse wil deze studie beantwoorden. Dit veld heeft tot nu toe voornamelijk gefocust op academisch discours, met een speciale interesses in effectief lesgeven en de invloed van verschillende semiotische modi op het leerproces. Met deze studie wil ik deze focus verbreden en uitbreiden naar andere potentieel vergelijkbare discours en genres.

Dit onderzoek heeft drie doelen; ten eerste stelt het een methodologie voor die multimodaal methodes en etnografie combineert om de persuasieve strategieën systematisch te analyseren, zonder een specifieke semiotische modus voorop te stellen, en die toelaat op resultaten te trianguleren. Ten tweede wordt deze methodologie getest op een corpus van 15 presentaties, wat geleid heeft tot een taxonomie van zeven persuasieve strategieën, die gerealiseerd worden door het combineren van verschillende multimodale elementen. Tot slot wordt dit methodologisch kader gebruikt om persuasieve strategieën en hun modale realisaties (met een focus op woorden, intonaties, handgevaren en hoofdbewegingen) te analyseren in de drie reeds benoemde genres.

De volgende drie onderzoeksvragen staan voorop in deze studie:

OV1: Welke verschillende persuasieve strategieën kunnen we identificeren met behulp van multimodale discoursanalyse?

OV2: Wat is de bijdrage van het multimodale ensemble van woorden, intonatie, hoofdbewegingen en handgebaren in het creëren van een persuasieve presentatie in de drie bovengenoemde genres?

OV3: In welke mate zijn het gebruik van persuasieve strategieën genre-specifiek en in welke mate overlappen de drie verschillende genres elkaar?

De thesis is opgedeeld in acht hoofdstukken. De eerste drie hoofdstukken omvatten het theoretisch kader, en hoofdstuk vier legt de link tussen theorie en onderzoekspraktijk met verdere uitwerking van het theoretisch kader. De overige vier hoofdstukken omvatten het empirische luik: hoofdstuk 5 focust op de datacollectie en -analyse; hoofdstuk 6 is gewijd aan de empirische resultaten; die verder uitgewerkt worden in de discussie in hoofdstuk 7.

Hoofdstuk 8 bevat de conclusies, evenals enkele pedagogische toepassingen, en pistes voor verder onderzoek.

Wat betreft de eerste onderzoeksvraag heeft de gecombineerde multimodale en etnografische methodologie het mogelijk gemaakt om persuasieve strategieën systematisch te analyseren, door 7 persuasieve strategieën te onderscheiden die op multimodale wijze gerealiseerd worden in het corpus van presentaties.

Beantwoordend aan de tweede onderzoeksvraag brengen de resultaten van mijn studie een correlatie aan het licht tussen modale dichtheid (het aantal gebruikte modi), modale coherentie (de manier deze modi samengebracht worden in een coherent multimodaal ensemble) en persuasieve strategieën. Met andere woorden, sprekers in mondelinge presentaties worden persuasiever als ze zich bewust zijn van de verschillende modi die ze kunnen gebruiken, en als ze in staat zijn om deze op coherente wijze te combineren.

In verband met de derde onderzoeksvraag heeft de multimodale en etnografische analyse omtrent persuasieve strategieën in het presentatiecorpus aangetoond dat het type relatie dat de spreker heeft met het publiek een bepalende factor is voor elk genre. Het is een cruciaal element om overtuigend over te komen, aangezien sprekers zich moeten richten op de relevante interpersoonlijke relatie die ze met hun publiek hebben om hen te overhalen. Daarnaast is er een algemeen verschil in de persuasieve strategieën en in types handgebaren die sprekers hanteren, terwijl andere verschillen veeleer te wijten zijn aan individuele verschillen en de beschikbare media, en daarom niet kunnen gezien worden als gemarkeerde kenmerken van de specifieke genres. De studie heeft ook enkele intertekstuele en interdiscursieve connecties tussen de drie genres aan het licht gebracht; de conferentiepresentatie in het corpus gebruiken elementen van populariserende presentaties uit de academische wereld, terwijl deze populariserende presentaties op hun beurt gebruik maken van elementen uit productpresentaties, om de juiste interpersoonlijke relatie met het publiek te verkrijgen.

Deze resultaten kunnen gebruikt worden in verder onderzoek naar genre, en om de drie relevante genres beter te omschrijven. Daarnaast kunnen ze verwerkt worden in lesmateriaal voor taalonderwijs voor specifieke doeleinden, om studenten en leerlingen te ondersteunen in het opstellen van meer persuasieve presentaties.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

CPs: Conference presentations

DA: Discourse Analysis

DI: Discourse Intonation

MDA: Multimodal Discourse Analysis

OHP: Open Hand Prone

OHS: Open Hand Supine

SFL: Systemic Functional Linguistics

USP: Unique Selling Proposition or Unique Selling Point

INTRODUCTION

This thesis is a multimodal and ethnographic study of the use of persuasive strategies in three oral genres used in two different professional settings: business and academia. In particular, the genres concerned are conference presentations, research dissemination talks and product pitches. These presentations share a strong persuasive component in their communicative purpose: the three of them address an audience to convince them of the value of a product, a service or a piece of research. However, they are used in business and academia by different discourse communities in different contexts, and therefore they can be expected to achieve their communicative goals in different ways. In addition, as it will be discussed in Chapter 1, research suggests that there is a trend towards promotionalization of different discourses, among which academic discourse is included. In view of this, I hypothesize that these three genres are intertextually and interdiscursively related, and that a multimodal and ethnographic study of the use of persuasion in them can help to shed some light on these relationships.

The motivation of this study is twofold. First, it is intended to gain insight into how these three genres fulfil their main communicative purpose: persuade an audience. With this study I want to probe into what makes these presentations effective and convincing. Second, it tries to acknowledge the important role that kinesic and paralinguistic features play in this task. Often, these aspects are underestimated in oral presentations as both speakers and researchers tend to pay more attention to the verbal, linguistic component of the presentation and the visual support, usually in the form of PowerPoint slide decks. For example, speakers tend to devote more attention to *what* they say (meaning the content of their message), than to *how* they say it (meaning the way this message is conveyed). In addition, when they care about the *how* they tend to focus on attractive slide design and on crafting an effective script, without devoting the same attention to other kinesic and paralinguistic semiotic modes (e.g. gestures, head movements and intonation), which are nonetheless integral parts of this *how*. This is a tendency already spotted by previous research that the results of this study have corroborated. The effect of kinesic and paralinguistic aspects such as body language or voice qualities in persuasive oral discourse has been often acknowledged, in particular in business contexts, but mostly in the form of guides offering practical tips. This shows that professionals using these oral genres in their everyday work seem to be aware of the fact that kinesic and paralinguistic aspects have an effect on their performance. With this thesis, I want to probe into the exact contribution of these aspects and how they can help design a persuasive presentation. In other words, most people prefer a dynamic speaker who uses varied gestures and intonation to a monotonous speaker who seems to be reading out loud. The question is, *why is this so?* This study is my contribution to the line of multimodal research that tries to answer this question. This line of research has so far focused mainly on academic discourse, with special interest in effective lecturing and the influence of different semiotic modes on the learning process: with this thesis I try to expand the scope to other different, but related, discourses and genres.

This thesis has three main objectives. First, it suggests a methodology that combines multimodality and ethnography to approach persuasive strategies in a systematic way, without prioritizing any semiotic mode in particular and allowing for the triangulation of results. Second, in this study I test this methodology on a corpus of 15 presentations, which has yielded a taxonomy of seven persuasive strategies that are realized as multimodal ensembles. Third, applying this methodological framework I analyse the use of persuasive strategies and their modal realisations (focusing on words, intonation, gestures and head movements) in a corpus comprised by examples of the three genres which are object of the study. The results can pave the ground for further genre studies and contribute to existing descriptions of the genres, which in turn can help improve Language for Specific Purposes (LSP) teaching materials.

The following three research questions guide my study:

RQ I. What type of persuasive strategies can be identified using an MDA approach?

RQ II. What is the contribution of the multimodal ensemble orchestrated through words, intonation, head movements and gestures to the creation of persuasion in the genres included in the analysis?

RQ III. To what extent is the use of persuasive strategies genre-specific and to what extent is there overlap between the three genres?

The thesis is structured in 8 chapters. The first three chapters provide the theoretical framework for my study. Chapter 4 makes the transition from theory to practice detailing my methodological framework. The remaining four chapters describe the empirical analysis: Chapter 5 focuses on the study, the results are presented in Chapter 6, a discussion of these results follows in Chapter 7 and finally Chapter 8 provides some conclusions, suggests some pedagogical applications and identifies areas for further research.

Chapter 1 focuses on discourse and genre analysis, the two theoretical pillars of my approach to oral presentations. It is structured in three sections. In the first section I review different approaches to discourse, highlighting a trend towards descriptive stances that avoid looking at language as an abstract system independent from its users, and showing an increasing interest in how people use language in real situations. Another trend in Discourse Analysis has been to broaden the understanding of communication, going beyond the purely linguistic and verbal towards a multimodal take on language. I then zoom in on three disciplines that I consider of particular interest to the present research because of their potential to shed light on the speakers' intentionality while taking into account the variety of semiotic resources that characterises oral discourse: Pragmatics, Systemic Functional Linguistics and Multimodal Discourse Analysis. I also consider in this section the contribution of ethnography to multimodal studies. I close the section narrowing down the scope to the focus of this thesis: oral academic and business discourse, paying particular attention to aspects that can potentially influence persuasion. Section 2 runs parallel to Section 1, but focusing on the concept of genre. The review of previous research that I provide in this section highlights the relevance of aspects such as the dynamism that characterizes genres, the need to adapt to social situations and discourse communities, and the importance of the production and interpretation processes. These are aspects that I argue make a multimodal approach to genre particularly appropriate, in particular in the case of oral genres. Finally, the last section in the chapter narrows down to the three genres which are object of this study: conference presentations, research dissemination talks and product pitches. In this section I review existing definitions and descriptions of these genres, to which I try to contribute with the results of this study.

Chapter 2 is devoted to persuasion, and it is also organised in three sections. In the first section I provide a review of previous literature on persuasion that sets the ground for the persuasive strategies that I later suggest in my taxonomy. In this review I also draw particular attention to the fact that persuasion has non-verbal realisations, which justifies a multimodal approach. The next section of the chapter shifts attention to genre, and presents persuasion as a powerful drive of generic change. Crucial concepts for this study, such as *intertextuality* or *interdiscursivity* are discussed in detail. The final section of this chapter focuses on the use of persuasion in the context of conference presentations, research dissemination talks and product pitches, and argues that persuasion is a defining characteristic of these three genres.

Chapter 3 focuses on the paralinguistic and kinesic semiotic modes that are the object of this study. The first section focuses on multimodality in academic and business discourse. Here, I review previous literature that reveals the variety of semiotic modes that play a role in academic and business genres. In the following section, I discuss previous studies dealing with the paralinguistic and kinesic modes included in my analysis: intonation, gestures and head movements. I pay special attention to several approaches that I find particularly useful to account for communicative intent, mainly Brazil's Discourse Intonation and Kendon's classification of gestures. Throughout this section I also put particular emphasis on the interaction between modes. I close the chapter zooming in on the use of these modes within the specific context of the three genres included in the analysis. Therefore, in the final section I review previous studies on oral academic and business discourse, focusing on the extent to which the role of intonation, gestures and head movements has been considered.

In Chapter 4 I explain how I use the theories and concepts described in the three previous chapters to develop my methodological framework for a multimodal discourse analysis of persuasive strategies in oral presentations. First I state and explain in detail the objectives and research questions that guide the study, which I have outlined earlier in this introduction. I then justify my decision to use a combined methodology: video-based multimodal research and ethnography (interviews and observation sheets). In the following section I illustrate how the theories discussed in the previous chapters inform and feed my analysis of gestures, intonation and head movements. Finally, the chapter concludes with a proposal for an analytical framework that can facilitate the multimodal analysis of persuasive strategies in oral genres. The framework is conceived as a data-driven and cyclical process that draws from three sources: i) input from previous literature, ii) multimodal data (results from the multimodal analysis), and iii) ethnographic data (results from interviews and observation sheets).

Chapter 5 provides a detailed account of my methodology. In this chapter I first describe the data gathering process and the corpus of presentations that I have compiled. I then describe the different steps of the ethnographic and multimodal analysis: i) observation sheets and initial interviews, ii) macroanalysis and selection of a subcorpus in each genre dataset for multimodal analysis iii) microanalysis of this subcorpora, including computer-assisted multimodal analysis, iv) triangulation of results through interviews, v) building a taxonomy of persuasive strategies and vi) comparative analysis between subcorpora. In this chapter I also describe the specialised software used to assist in the multimodal analysis (step iii). In addition, I pay particular attention to the selection of subcorpora for later fine-grained analysis (step ii), which draws from the ethnographic concept of *rich point*.

Chapter 6 presents the results of the analysis. The chapter is organised in three sections that correspond to my three research questions. As a first outcome of my analysis, and regarding the question of how to study persuasive strategies from a multimodal perspective, in the first section I focus on the taxonomy of seven multimodal persuasive strategies resulting from applying the methodological framework devised in Chapter 4 to the corpus described in Chapter 5. In the following section I present the results of the microanalysis, that is, the way these strategies are used in each of the three subcorpora. The analysis focuses on identifying the persuasive strategies used and their modal realisations. This provides an answer to my second research question concerning the contribution of multimodal ensembles to persuasion. These results are also the basis of the comparative analysis presented in the last section of this chapter, which answers my third research question related to the degree of specificity or overlap in persuasive strategies across the three genres.

In Chapter 7 I provide my interpretations of the results I have obtained, relating the results of the multimodal study with those of the ethnographic study and with insights gained from previous literature. The structure of the chapter mirrors that of Chapter 6, and has three sections that address my three research questions. Therefore, the first section focuses on the methodology applied for a multimodal approach to persuasive strategies, the second discusses the results of the microanalysis for each genre, and the third and last section centres on the results of the comparative analysis.

Chapter 8 turns to the potential interest of my results for discourse and genre analysis. Here I reflect on what we can learn about a genre by looking at the use of persuasive strategies in it from a multimodal perspective. I also discuss the pedagogical applications and implications of the study, focusing on the importance of raising awareness of modes and promoting a skilful use in seamless integration. The final section acknowledges the limitations of the study and suggests areas for further research. A list of references and appendixes is included after Chapter 8.

INTRODUCTION

CHAPTER 1. THE STUDY OF ORAL DISCOURSE AND GENRES

This thesis is a multimodal and ethnographic study of the use of persuasive strategies in oral presentations. In this first chapter I will present the theoretical foundations of my methodological approach. The chapter is divided in three sections. Section 1 deals with different approaches to discourse on which I draw to build my methodological framework. I close this section zooming in on academic and business discourse, which are most relevant for my study. Section 2 runs parallel to Section 1, but focusing on the concept of genre. Finally, the last section in the chapter narrows down to the three genres which are object of this study: conference presentations, research dissemination talks and product pitches.

1.1 Approaches to discourse

In this section I provide an overview of how different approaches to discourse have gradually progressed to a multimodal take on language, in which the diversity and complexity of meaning-making is fully acknowledged. The approaches to discourse that I discuss in this chapter share an interest in how speakers use language in real communicative situations and a broad understanding of communication that goes beyond the purely linguistic and verbal. I also consider in this section the contribution of ethnography, which, although not fully rooted in linguistics, is a fundamental pillar of my methodological framework as I will describe in Chapter 4. I close the section narrowing down the scope to the focus of this thesis: academic and business discourse.

1.1.1 Discourse analysis

Celce-Murcia and Olshtain (2000: 4) define discourse as

> an instance of spoken or written language that has describable internal relationships of form and meaning (e.g. words, structures, cohesion) that relate coherently to an external communicative function or purpose and a given audience /interlocutor. Furthermore, the external function or purpose can only be properly determined if one takes into account the context and participants (i.e., all the relevant situational, social and cultural factors) in which the piece of discourse occurs.

This definition highlights interesting aspects about the concept of discourse to be taken into account for the purposes of this work. First it foregrounds that discourse needs to be coherently related to an audience and a communicative function, and second it states that in order to determine the purpose, the context needs to be considered. According to Swales (1990), the purpose is usually determined by the discourse community and its social practice, and this communicative purpose allows us to differentiate between different types of discourse. We can associate a discourse community with a professional community, which makes it possible to speak of professional discourses such as business discourse, legal discourse, etc. Each type of discourse is, in turn, manifested in different genres. Different approaches to discourse have been adopted. In this section I will provide an overview of the most relevant for the present research.

Discourse analysis (often abbreviated as DA) is an approach to the study of discourse that focuses on how language is actually used (Fasold 1990). According to Slembrouck (n.d.) DA has three main concerns: the first one is language use beyond the boundaries of a sentence or utterance, the second one is the interrelationships between language and society, and the third one is the interactive properties of everyday communication.

The first of these concerns falls within what Schiffrin (1994) calls the formal approach to discourse. It focuses on how language is organised beyond the sentence level, since traditional grammar is able to account for sentence structure, but not so easily for longer stretches of discourse. The functional approach, on the other hand, is concerned with language use. It involves the study of text and context and regards text-external factors as relevant. The concern of DA for text external factors is also foregrounded by Alba-Juez (2009), who points out that discourse analysts pay attention to aspects that are not purely linguistic, such as the social, political and cultural context in which language is produced, since they are considered to have a strong influence on the way people use language:

> The current aim now in DA is to describe language where it was originally found, i.e. in the context of human interaction. In this respect, it is important to point out that this interaction often involves other media besides language. (…) it is also the discourse analyst’s job to explain the connection between these systems and language. (…) Discourse is multimodal because it uses more than one semiotic system and performs several functions at the same time (2009: 12, 13)

On the other hand, Brown and Yule (1983) highlight an interesting aspect of DA: the need to take into account the purpose of linguistic forms instead of studying them in isolation. According to them, DA studies ordinary language data, and its approach is not prescriptive but descriptive, i.e. it does not look for rules, but for regularities. In addition, it does not see language as a static product but rather as a dynamic process. When discourse analysts look at the way a speaker said something, they do not consider it in terms of correctness. Rather, they are interested in why a particular option is chosen and what this reveals about communicative intentions. Therefore, emphasis is not on language as a result of rules, but on language as a process of meaning-making and interpretation. This is a view that I share and I adopt throughout this study: throughout my analysis I look at the choices that speakers make when using different semiotic modes, and I try to elucidate what their communicative intention was.

What these different approaches to DA have in common is that they avoid looking at language as an abstract system independent from its users. Instead, DA focuses on how people use language to show their attitude, to express feelings, to exchange information, or to socialize, among others. Due to this ambitious scope, Discourse Analysis is a broad area of study which embraces different disciplines. I consider three of them are of particular interest for the present research: Pragmatics, Systemic Functional Linguistics and Multimodal Discourse Analysis. An explanation of them will follow in the next sections. In my approach to genre I share DA's basic tenet that how and why something is said is as important as what is said. Using the aforementioned disciplines as a framework I intend to gain an insight into how participants in a communicative process dynamically cooperate to construe meaning and achieve communicative goals by means of resources that are not always linguistic or verbal. For this reason throughout this thesis I will use the term "text" in a broad sense referring not only to written documents, but also oral messages, and including their non-verbal and non-linguistic features.

1.1.2 Pragmatics

Morris defines pragmatics as the "relation of signs to interpreters" (1938: 6). It is concerned with the study of how language reveals our attitudes, beliefs and intentions, and therefore pays special attention to the intended interpretation of a communicative event and the expected reaction from the receiver (Grice 1957; Green 1989). I consider these as crucial aspects for the objective of this study, i.e. to understand how persuasion gets across in oral genres, since effective persuasive communication depends to a great extent on the ability to foresee and control the listener's reaction (O'Keefe 2002; Perloff 2003). There are four concepts in particular that are essential in the tradition of pragmatics and are very relevant for the study of communicative intentions and communicative effect: i) speech acts, ii) implicatures, iii) presuppositions, and iv) politeness. An explanation of each follows in the next pages.

Speech Act Theory was first put forward by Austin (1962) and Searle (1969, 1971, 1975, 1976). The basic tenet of this theory is that language is a form of (inter)acting. In other words, speakers do things through language. For example they express their feelings and regulate social relations (e.g. apologise, make compliments, give condolences, etc.), express wills and intentions (e.g. make a promise, swear, make a vow), or provide information they consider to be true. All expressions of language can be considered acts, hence the term speech act.

According to Austin (1962), there are three acts implicit in any utterance: a *locution*, or the act of physically pronouncing an utterance; an *illocution* or the act committed by producing an utterance (e.g. make a promise, give an order, complain, congratulate, etc.) and the *perlocution*, or the effect provoked as the result of the processing of the locution and the illocution.

He suggests the following taxonomy of speech acts:

i) Verdictives: delivery of a finding;

ii) Exercitives: giving a decision in favour or against a certain course of action;

iii) Commissives: commit the speaker to a certain course of action;

iv) Expositives: expounding of views, conducting of arguments;

v) Behavitives: reaction to other people's behaviour and attitude to someone else's conduct.

Austin's theory highlights the vital importance of purpose in communication. He reveals that we always speak with a purpose in mind and we act through language. He further claims that it is possible to differentiate utterances according to the act we perform as we say them. This is a view of language that will influence genre studies. Authors such as Bazerman (1994) apply speech act theory to genre (see Section 1.2.1).

Nonetheless, Searle (1976) argues that Austin's taxonomy does not really apply to speech acts but to illocutionary verbs (e.g. protest, apologise, claim, etc.). He further argues that Austin's taxonomy does not have a consistent principle of classification, and because of this there is overlap among the categories and heterogeneity within them. He therefore revisits Austin's taxonomy and suggests his own, based on three criteria: 1) communicative purpose, 2) direction of fit (whether the words are supposed to match reality as in a description, or reality is supposed to match the words as in a set of instructions to build a piece of furniture), and 3) psychological state expressed, yielding the following categories:

i) Representatives or assertives state what the speaker believes to be true. By using a representative, the speaker is committed to the truth of the proposition.

ii) Directives: the speaker expresses his will, and usually tries to make the addressee do something. They include commands, orders, requests, and suggestions.

iii) Commissives: speakers express their intentions, committing themselves to a future course of action. Commissives include promises, threats, refusals, pledges.

iv) Expressives: the speaker expresses an attitude about a state of affairs.

v) Declarations: the speaker alters the status quo by making the utterance, as when a priest pronounces a couple husband and wife.

Austin (1962) and Searle (1969, 1976) use the term illocutionary force to refer to that specific aspect of meaning transmitted by the utterance which conveys the speaker's intentions and expected reaction from the addressee. I consider that this illocutionary force is precisely what my study of persuasive strategies tries to elucidate by means of a multimodal and ethnographic analysis of the speakers' performance during their presentations. In other words, I try to determine the illocutionary force of the speakers' persuasive multimodal behaviour, which I then confirm through ethnographic methods.

The illocutionary force of an utterance is sometimes made explicit by certain markers, which are called IFIDS, or Illocutionary Force Indicating Devices (Searle 1969). Some IFIDS are performative verbs (command, request, etc.), word order (e.g. to mark a question), mode of the verb (imperative, indicative), intonation or accent. As it can be noted, IFIDS are not always linguistic aspects of the speech. Paralinguistic features such as intonation and accent also convey illocutionary force. As I will discuss in Chapter 3, kinesic features such as gestures can also serve the same purpose.

In some cases the syntactic structure of the sentence is in itself an IFID, as when an imperative construction indicates a command. However, this is not always the case. An utterance that contains IFIDs for one kind of illocutionary force can sometimes perform, in addition, another type of illocutionary act. In other words, when the form of an utterance does not match its illocutionary force, we find what Searle (1975) calls indirect speech acts. These speech acts add an additional burden to the interpretation of the illocutionary force. In fact, a major concern of pragmatics has been the study of what makes an illocution successful, especially in the case of indirect speech acts. This is only possible when the listener perceives the utterance as valid. According to Habermas (1981) for an utterance to be valid it has to fulfil three conditions: 1) *truth*, meaning that it has to provide truthful information about the reality to which it refers, i.e. it has to be true; 2) *sincerity*, meaning that the intention of the speaker has to be perceived as sincere (i.e. to perceive an utterance as a question we need to be sure that the speaker really lacks the requested information) and finally 3) *legitimacy*, meaning that the intended effect that we want to provoke on our listener has to be legitimate from the point of view of capacity (I cannot ask you to do something you cannot possibly do) and social acceptability (I cannot ask you to do something you are not supposed to do).

Despite validity conditions and IFIDS, the process of interpreting an illocution correctly is not an easy one, especially with indirect speech acts. In fact, much of what happens in the communication process is that one interlocutor guides the other towards the intended interpretation. This is particularly salient in the persuasive oral genres which are the object of this study. It is my contention in this thesis that speakers in persuasive presentations use a variety of semiotic resources to achieve this goal and obtain the intended interpretation, sometimes in particularly subtle ways. As I discuss in more detail in Chapter 2, persuasion tends to be more effective when it is subtle, and indirect speech acts are an effective way of subtly guiding the audience towards a particular intended interpretation.

If there is something that studies in pragmatics reveal, it is that communication is not a simple process, and that the possibilities of misunderstandings are actually great. A question that has long raised the interest of researches in pragmatics is how listeners are able to understand much more than what is explicitly said. Grice (1975) suggested that this is made possible because speakers agree to abide by certain rules when communicating, which he grouped under what he called the *cooperative principle*: “Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged” (1975: 45). According to this author, this cooperative principle can be divided into four maxims:

i) Maxim of quantity: give the appropriate amount of information;

ii) Maxim of quality: do not say something which you believe to be false;

iii) Maxim of relation: be relevant;

iv) Maxim of manner: be brief, orderly and unambiguous.

People engaging in communication assume that their interlocutors are following these maxims. When one of the maxims is flouted they try to find a reason for it, and they infer meaning form the context to explain it. This is the way in which utterances can convey more meaning than what they explicitly say, or, in other words, *implicatures* are produced. For example, by saying the opposite of what you really mean (i.e. flouting the maxim of quality), an utterance can be interpreted as ironic.

It is reasonable to expect persuasive discourse to be rich in implicatures, since implying more than what you actually say has a powerful communicative effect, and can be used manipulatively. For example, a negative evaluation of a competitor product can be implied by a comparison like:

"Our product is safer, and faster and easier to use than waxing kits"

This example, by the way, is taken from one of the product pitches in my corpus, which will be described in detail in Chapter 5.

Sometimes, when producing an utterance, speakers take certain information for granted without explicitly stating it. In pragmatics this is called *presupposition*. For example in uttering: "How many people would be happier if they had a job that is customised to them?" the speaker is presupposing that regular jobs are making people unhappy. This example is also part of my corpus of research dissemination talks.

Renkema (1993) points out how presuppositions are sometimes prompted by intonation. In the example "Other products may be cheaper" (my own), a change of intonation in "cheaper" presupposes the interpretation "but in other aspects they are worse".

Presuppositions, therefore, can make a powerful persuasive tool with potential to be realised multimodally, because it is easier to cajole someone into believing something by means of a false presupposition than with a direct lie, and it is possible to do this with semiotic modes other than words. In the abovementioned example, the presuppositions may not be actually true, but the addressee will probably be more willing to accept it as such than if the untrue claim was stated directly (Hornby 1974).

The fourth pragmatic concept that I consider relevant for my study of persuasive strategies is *politeness.* The term *politeness* in pragmatics is used to refer to the strategies used for the regulation of interaction. Every communicative activity involves interacting with other people. People reaffirm or readjust their interpersonal relationship (e.g. social distance, dominance) as they speak to each other and they do so resorting to politeness. Lakoff (1973) explains the use of politeness in communication by means of three rules to which speakers adhere: 1) do not impose, 2) give options and 3) make your receiver feel good.

These rules account for many conventionalised expressions that recur in oral speech, such as “I’m sorry to trouble you but…” Green (1989) further develops this approach to politeness and explains that the three rules apply to different levels of formality. Rule 1 applies to formal situations, in which there is a difference of power and status. It implies being impersonal and avoiding emotional language and taboos. Rule 2 applies to more informal contexts, in which interactants have a similar status. It implies expressing opinions in a way that they can be ignored without being openly rejected. Rule 3 applies to informal situations in which interactants are socially close. It implies showing interest in the other person, and showing trust by talking openly about yourself.

Brown and Levinson (1987) offer a different approach to politeness, based on the psychological concept of *face*. Face is the image you project to society, or the way you want people to see you. It has a negative and a positive side. In its negative side it refers to the need to preserve one's space. In its positive side it refers to the need to be appreciated. According to them, the role of politeness is to prevent or soften potentially face threatening acts that may occur in interaction. The greater the risk, the more polite speakers need to be. Risk to face is measured according to rate of imposition (which refers to the way the face threatening act is perceived in a specific culture), social distance and power. Depending on these parameters speakers will resort to one of the following politeness strategies:

i) No redressive action: the face threatening act is performed blatantly;

ii) Positive redressive action: appeal to interactant's need to be appreciated;

iii) Negative redressive action: appeal to interactant's need to protect personal space;

iv) The face threatening act is performed implicitly, so that it is hardly recognisable as such;

v) The face threatening act is supressed.

Politeness is interesting for a study of the expression of an inherently persuasive message, as is the case in oral presentations. It can be very revealing about the relationship established between interactants, and this is a factor that highly influences the success of a persuasive message. It is also worth noticing that different kinesic and paralinguistic features (for example a use of a specific tone) can make a difference in terms of politeness, and a misuse of them in a presentation can achieve the wrong effect on the audience.

In conclusion, the study of speech acts, implicatures, presuppositions and politeness, contribute to a better understanding of the communicative intentions of speakers and the communicative effects they achieve. Every choice in language is meaningful. This idea is the basic tenet of Systemic Functional Linguistics, which is the object of the next section.

1.1.3 Systemic functional linguistics

The systemic functional approach was developed by Halliday (1978, 1985). It perceives language as a semiotic system in which speakers make meaningful choices at different levels as they speak. These choices reveal the speaker's assessment of the communicative situation and to some extent determine the social relations established between communicators. In Systemic Functional Linguistics (hereafter SFL) *the way* speakers say something is as important as *what* they say (Thompson and Muntigl 2008). This is a position that I share throughout this study.

According to SFL the language system is made up of three main strata: semantic, lexico-grammatical and phonological. In each stratum speakers have a set of options, and each stratum is related to the next through realization: meanings are realised through lexis and wordings are realized in sounds. Language is, therefore, a system of interrelated sets of choices with which speakers construe meaning (Halliday 1985; Halliday and Greaves 2008).

In addition, the available choices of the language system convey different types of meaning. This allows us to differentiate between three metafunctions of language (Halliday 1978, 1985; Halliday and Greaves 2008):

i) *Ideational* function: it construes human experience. We use language to understand and refer to our experience. This includes the expression of experiential meaning (i.e. perceptions of the real world) and logical meaning (e.g. logical relations between stretches of discourse). The ideational function can therefore be subdivided into experiential function and logical function.
ii) *Interpersonal* function: it enacts human relationships as we interact using language.
iii) *Textual* function: it makes it possible to construct texts and engender human discourse.

In his study of scientific discourse, Lemke (2005) renames these functions as *presentational*, *orientational* and *organisational* respectively, and he notes how they can be realised by semiotic resources that go beyond language as understood by SFL (e.g. graphic representations, images and maths).

According to Matthiessen (2007) a semiotic system can be paralleled to a puzzle: if it is seen from above, i.e., from the culture in which they operate, what is perceived is the whole puzzle as a unified whole. This would correspond to a representation of reality (ideational function). However, when seen from below, what can be perceived are the different pieces that make up the puzzle, or the semiotic resources that interact to create meaning. These pieces need to be put together in an organised way that allows for the interpretation of meaning (textual function) in order for the message to get across from one person to another, enabling interaction (interpersonal function). When speakers use the system of language they produce what Halliday calls an *instantiation* of language. In order to explain the system-instance relation he uses the following analogy: the system is like the climate, i.e. a generalisation of weather patterns across time, equivalent to the full set of options that a speaker has, while instantiations are equivalent to the specific weather in a particular day, or the specific choice that the speaker makes. As I will later argue in Chapter 3, the Discourse Intonation (DI) approach that I adopt for the study of intonation is very consistent with this view of language as a system of meaningful choices.

As mentioned above, the way experience is presented has a strong influence on persuasion. According to Halliday (1985, 1994) the grammatical means to express experience is the clause. In a clause we conceptualise our experience as a situation type that consists mainly of a process involving certain participants. The process is usually expressed by the verb. Halliday differentiates 5 types of processes:

i) material: processes of doing and causing;
ii) mental: processes of perception cognition and affection;
iii) relational: processes that characterised or identify or situate a participant;
iv) verbal: processes of communicating;
v) existential: processes of existing or happening.

According to Halliday the experiential content of a clause can be perceived from two perspectives: how the process affects participant(s), which he calls transitivity, and whether the process is understood and caused by another participant or presented as happening by itself, what he calls ergativity.

As mentioned above, the way experience is presented can make a huge difference in how a message is interpreted, as the following examples (my own) show:

(a) We produced a revolutionary device

(b) This device is revolutionary

Sentence a) assigns merit to the producers, while sentence b) provides the product with an inherent characteristic. Persuasive texts that promote a product tend to portray the addressee as an experiencer in a mental process (enjoying the experience of using the product), while the speaker is shown as an agent in a material process (producing the product). Likewise, presenting a situation as an ergative process that is not caused by any participant makes it more difficult for the addressee to challenge it (Thompson 1998).

In addition to the expression of ideational meaning, the expression of interpersonal meaning is particularly relevant in the study of persuasive genres, because, as it will be discussed in Chapter 2, persuasion is also achieved by appeals to emotion that are conveyed through interpersonal meaning.

As Thompson and Muntigl (2008) point out, in the interrelated set of choices that conforms language, several systems are particularly relevant for the expression of interpersonal meaning. At the stratum of lexico-grammar there are the systems of *mood, modality and appraisal,* while at the stratum of semantics there is the system of *speech function (*Halliday 1985, 1994; Martin 2000; Martin and Rose 2003; Martin and White 2005). By combining choices in these systems speakers are able to interact with addressees by making clear what they expect from them. These systems also allow speakers to express their attitude and establish relative power status (e.g. obtaining information without expressing dominance, asking the addressee to do something dominantly, asserting something without leaving room for discussion, receiving confirmation of something the speaker is not sure of, or evaluating something). A description of each system follows in the next paragraphs.

Firstly, within the system of mood speakers have the choice of imperative mood vs. indicative mood, and within indicative there is a further choice of declarative and interrogative. Finally, within interrogatives speakers can choose between a yes/no question and a wh-question. Figure 1.1 summarizes these set of options:

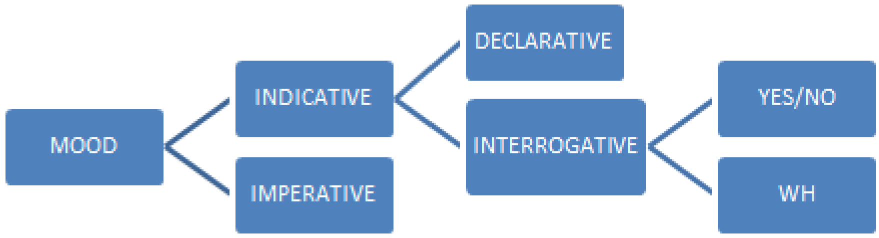


Figure 1. 1 System of mood (Halliday 1985, 1994)

Secondly, at the stratum of semantics we find the system of speech function *(*Halliday 1985, 1994). Within this system speakers can choose between 'give' and 'demand' either 'goods and services' or 'information', resulting in the following possibilities:

i) Give information = statement;
ii) Give goods and services = offer;
iii) Demand information = question;
iv) Demand goods and services = command.

Halliday (1994) calls utterances that deal with the exchange of information *propositions*, and those that deal with the exchange of goods and services *proposals*. Usually speakers associate a speech function to a particular mood, so that questions are usually realised by means of interrogatives, commands by means of imperatives, statements by means of declaratives, etc. These are the unmarked options. However, very frequently we find what in SFL is called *form-function mismatches*, and in pragmatics is referred to as *indirect speech acts*. This is what happens when a command is realised through a question, or a question as an imperative, as in the following examples (my own) respectively:

(a) Would you mind waiting a second?

(b) Tell me what you need.

Sentence a) is an interrogative that expresses a command, and b) is an imperative that can be interpreted as an offer or a question. As Thompson and Muntigl (2008) point out, these marked choices are especially meaningful. A command realised as a question adds politeness and soothes the interaction. On the other hand, a question realised as a command implies authority and dominance in conversation. Therefore, form-function mismatches are very revealing about the way speakers in presentations position themselves with respect to their audiences.

A third set of choices available to speakers in the expression of interpersonal meaning is the system of modality *(*Halliday 1985). By means of modality speakers can adjust the strength of a proposition in terms of probability and usuality (modalization) and proposals in terms of inclination and obligation (modulation). Figure 1.2 summarizes the set of options within the system of modality.

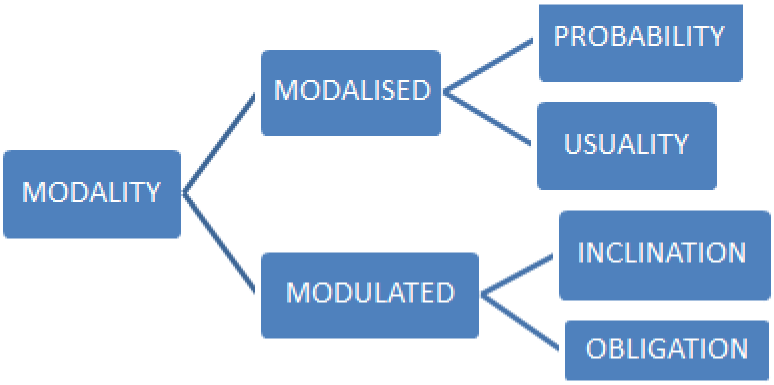


Figure 1. 2 System of modality (Adapted from Halliday 1994: 357)

Finally, a fourth system with which speakers can construe interpersonal meaning is the system of appraisal (Martin 2000, Martin and Rose 2003, Martin and White 2005, Thompson and Muntigl 2008). Appraisal deals with the expression of evaluation. In this system speakers can make choices regarding three aspects: *attitude*, or assessment itself; *engagement*, or negotiation of the intersubjective status of the assessment; and *graduation* (amplification or tone down of assessment).

Further choices available within each aspect are described in Figure 1.3.

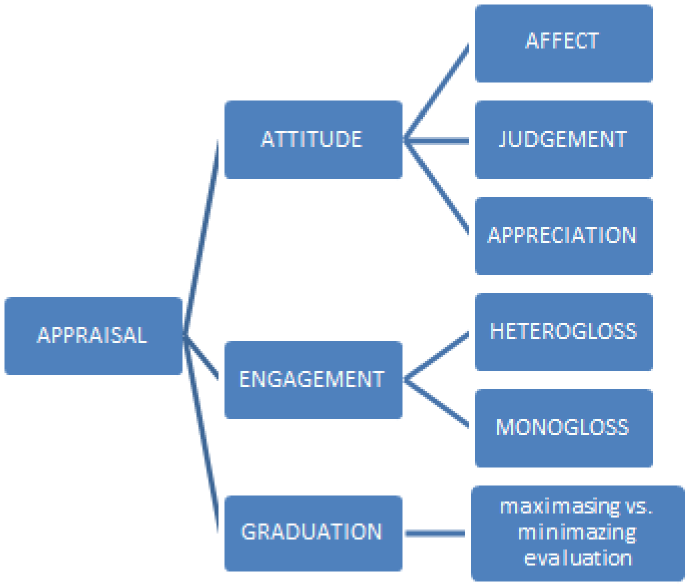


Figure 1. 3 System of appraisal (Adapted from Martin and Rose 2003: 54)

As Figure 1.3 shows, regarding attitude speakers first decide whether they present the assessment as a mental process they direct to the assessed entity (*affect*, e.g. I love it) or as an attribute of that entity (*judgement and appreciation*), and then they choose to assess in moral/ethical terms (*judgement*, e.g. It is not fair) or in aesthetic terms (*appreciation*, e.g. It is a powerful tool). The expression of attitude becomes more complex when it is indirect, as in "Our product is safer", which is not overtly negative, but implies a negative evaluation which the listener is expected to understand. This also proves that appraisal is more than just expression of speaker's personal feelings: it is intersubjective, and with it speakers try to get listeners to share their values so that the information provided is correctly understood. In a way, it is an appeal to solidarity (Thompson and Muntigl 2008). This turns the expression of evaluation into a persuasive tool, because as speakers evaluate, they expect their audiences to agree with it (Bamford 2007, Querol-Julián 2011).

A further choice with which speakers are presented within the appraisal system is engagement. They decide whether to acknowledge other alternative points of view (*heterogloss*) or not (*monogloss*). Monogloss utterances represent an attempt to avoid any negotiation process over something and present it as the only possible option, as in "If you want to heat stuff really fast you are not going to use the oven, you’re going to use the microwaves" (example from my research dissemination corpus). They are a recurrent device in persuasive discourse, since they can prevent potential counterclaims.

Finally, speakers have the option of maximizing or minimizing the appraisal. This can help foreground or background claims, and it can also help regulate interaction by matching (or intentionally mismatching) the strength of evaluations to the level speakers assume acceptable for their interlocutor (Thompson and Muntigl 2008).

In conclusion, SFL shows how every choice speakers make when communicating is meaningful. The above discussion has presented the different systems as operating at the strata of lexico-grammar and semantics, but as Halliday and Greaves (2008) point out, the way speakers realize their choices in the different systems is not always verbal. Intonation can realize meaningful choices in the modality system, for example, turning a declarative utterance into an interrogative.

Muntigl (2004) suggests that non-linguistic semiotic systems are organised in strata and metafunctions just like language is, and they interact with the system of language in the communicative process. According to Halliday (1978) these systems cannot be considered languages because they lack the stratum of lexico-grammar. He refers to them as *protolanguages*. Kress and Van Leeuwen (1996), however, questioned this suggesting that images indeed have a grammar, and that they could be analysed in terms of the three metafunctions of language. Martinec (2001) undertook a similar analysis of kinesic features showing how gestures, facial expression, head movements and proxemics can all contribute to the three SFL metafunctions. Muntigl (2004) agrees with this idea. He claims that some gestures represent an aspect of the real world, clearly expressing ideational meaning. Gestures can also signal the informational organization of the text (new vs. given), therefore expressing textual meaning. Finally, other gestures show the attitude of the speaker or regulate interaction, expressing interpersonal meaning (see Chapter 3 for a detailed discussion of gestures). This line of research questions McNeill's (1992) description of gesture as global and synthetic. According to McNeill, gestures are not divisible into meaningful parts because their parts acquire their meaning from the whole, and they are idiosyncratic. Muntigl (2004) argues that, although we need to take into account the whole gesture to grasp its meaning, its parts do contribute individually to the general meaning. For example the position of the hands or fingers may represent an object and the movement may indicate the action that is performed on that object. He further claims that although the relation between meaning and form is much looser than in speech, this does not imply that any form is valid for any meaning, and that certain regularities can be found. Therefore, it is not unreasonable to talk about a possible grammar of gestures. Kendon's (2004) gesture families certainly point in this direction, as I will discuss in Chapter 3.

The claim that paralinguistic features and gestures are more organised than we have usually acknowledged is therefore not unjustified, and the same could be reasonably expected of other semiotic systems that have been traditionally considered ancillary to language. In fact, Van Leeuwen (2004) claims that we should talk about *communicative acts* instead of *speech acts*, because all the signs present in these microevents (linguistic or not) combine to determine their communicative purpose. It is for this reason that Multimodal Discourse Analysis, the focus of the next section, foregrounds the role of different semiotic modes, and not necessarily linguistic, to transmit meaning. As Muntigl (2004: 47) puts it, "we should broaden our analytic lens beyond speech and writing to include other meaning-making systems".

1.1.4 Multimodal Discourse Analysis

The basic tenet of Multimodal Discourse Analysis (hereafter MDA) is that communicators resort to different ways of expressing their meaning, i.e. they use different *semiotic modes*, in MDA terminology, in the meaning-making process. Some of these modes are verbal, while others are not. This approach has two important implications. On the one hand the study of discourse must broaden its scope to include non-verbal aspects. On the other hand, in order to fully comprehend any communicative process, every mode employed should be considered.

Iedema (2003: 33) claims that the interest in semiotic modes other than language is a result of changes in our semiotic landscape, which are the consequence of social changes:

> Our semiotic landscape is becoming more and more populated with complex social and cultural discourse practices. Here, the influence of electronic communication, the globalisation of trade and commerce and the increasingly political-cultural mix of the countries in which we live mark important facets of this changing landscape.

He points out how these changes imply on the one hand a "decentralisation" of the role of language in communication and, on the other hand, a "blurring of the boundaries between roles" allocated to different semiotic modes. What has been traditionally expressed through written language is now expressed through other modes such as image, and it is no longer constrained to the logic of the linguistic system (Iedema 2003).

Iedema (2003) summarises the main concerns of MDA in four main areas. First and foremost is acknowledging other semiotic modes apart from language. The second one is studying how the different semiotic modes work together in communication. Third is analysing how social changes affect the use of semiotic modes (e.g. technology can provoke a displacement of one semiotic mode by another). Finally, MDA also investigates how the potential of the different semiotic modes can influence interaction.

Kress (2009), to a large extent in agreement with these four main concerns, states that there are four theoretical assumptions of multimodality:

i) Language is part of a multimodal ensemble;
ii) Each mode realises different communicative work: modes have been shaped through cultural, historical and social uses to realise social functions;
iii) Meaning is built selecting different modes and making them interact;
iv) The meaning we interpret from semiotic resources is social; it is determined by our social context.

Reflecting on the origins of MDA, Constantinou (2005) suggests that SFL, as advocated by Halliday, evolved into a new trend of Social Semiotics, which sees meaning-making as a social practice and studies the ways of meaning-making in specific social and cultural contexts. This in turn was extended into MDA by the work of Kress and Van Leeuwen *Reading Images* (1996). According to Jewitt (2014) there are three main approaches to multimodality. The first is social semiotics multimodality, and is concerned with how modes are used in certain social contexts. The interest is the "articulation and realisation of an individual's relation to an object or event, acting out of that social complex at a particular moment, in the context of an interaction with other constitutive factors of the situation which are considered as relevant by the individual" (Kress 1993: 174). The second is rooted in SFL, and focuses on the meaning potential of semiotic modes and the way they interact creating a system. In Kress' words,

> It takes a systematic focus to semiotic resources as sets of metafunctionally oriented systems (...) in addition to understanding and examining social semiotic resources as systems of meaning, the focus is on exploring the system in use and the social contexts that it is embedded in. Thus, the multimodal phenomenon, not the sign maker, is the focus of this perspective on multimodality (Kress 2009: 36).

The third is multimodal interactional analysis, and focuses on context and situated interaction, i.e., how participants in a communicative situation react to each other's discourse. As Kress explains,

> the focus is on (…) how a variety of modes are brought into and made constitutive of social interaction, identities and relations. The modal system is no longer a primary concern. Modes are understood as made and constituted through interaction and action, and thus there is no notion of modal system outside of interaction (Kress 2009: 36).

This study can be located within the SFL trend of multimodality, because it is aimed at investigating the meaning-making potential of particular modes and the way they coherently and systematically interact to express persuasion. The focus is, therefore, the multimodal phenomenon of persuasion, rather than the sign maker. In addition, I side with the view of language as a system of choices postulated by SFL. However, I also make use of some notions and methods that are common in Multimodal Interactional Analysis such as foregrounded higher level action, modal configuration, modal density or playback methodology (Norris 2004). I will come back to them later on in this section.

Two concepts are central to MDA: *medium and mode*. Different definitions of them have been suggested. Kress and Van Leeuwen (2001: 21) define modes as "semiotic resources which allow the simultaneous realisation of discourses and types of (inter)action (...). Modes can be realised in more than one production medium". They argue that modes are in constant evolution to suit changing socio-cultural interests, and also point out that modes are not specialized for specific types of meaning. Instead, the same semiotic principles work across modes. However, as discussed in more detail in section 1.2.4, they do acknowledge that modes have different affordances (the materiality in which a mode is brought into existence can impose some advantages and limitations, for example depending on whether the mode follows the logic of space or the logic of time), epistemological commitments (due to the way different materializations of modes are processed by addresses when they make meaning out of these modes) and social valuation (highly dependent on culture). Norris (2004) offers an additional perspective on the concept of mode, and claims that a mode is a tool for analysis that makes it possible to study how a text builds meaning. She points out that modes are not necessarily clearly identifiable discrete units, but rather theoretical and analytical abstractions. Later, Kress has defined mode as a "socially shaped and culturally given resource for making meaning" (Kress 2009: 60). To answer the question of what can be considered a mode, he suggests that "socially, a mode is what a community takes to be a mode and demonstrates that in its practices; it is a matter for a community and its representational needs" (Kress 2009: 63). From a formal perspective, he adds that a mode needs to be able to fulfil the three metafunctions postulated by SFL.

Norris (2004) and Kress and Van Leeuwen (2001) agree that modes are resources to produce meaning, and that it is something abstract, as opposed to media, which are "material resources used in the production of meaningful products or communicative events" (Kress and Van Leeuwen 2001: 22). Jewitt (2004) clarifies the difference between the concepts modes and media stating that modes enable *representation* of meaning while media allow for the *dissemination* of meaning. For example a book is a medium that transmits a message in which different modes may operate (e.g. writing and image, and even sound if the book includes a CD).

A question that is raised when adopting a multimodal approach is whether modes have an inherent hierarchical importance in a message. Norris (2004, 2004b) agrees with Iedema (2003) that language is not necessarily central in communicative interaction, as there are other semiotic modes that play an important role and are not ancillary to language. Among these semiotic modes she makes a difference between embodied modes (among which she includes paralanguage and kinesics) and disembodied modes (elements in the material world that play a role in human communicative interaction). In addition, she claims that the number of modes in itself is not significant: what is relevant is their intensity and the complexity of the interrelationships among them. She refers to this as *modal density*. Communicative events are complex, and participants engage in different (inter)actions at the same time with varying degrees of attention and awareness. These are called *higher level actions*, and are in turn composed by a chain of different *lower level actions*. An example of a higher level action can be the action of supporting a claim with an argument. In order to do this speakers may resort to different lower level actions, such as pointing to visual material or raising their voice, each lower level action using one or more semiotic modes. Modal density makes it possible to organise higher level actions in a continuum of foreground-midground-background according to the intensity and interplay of modes employed in their realisation. The *modal configuration* (Norris 2004) refers to the relationships among the modes of a higher level action that are at play at a given moment. It is a way of analysing a higher level action into the lower level actions that compose it and the modes that make them possible. These three concepts, i.e. modal density, higher level action, and modal configuration are key in my methodological framework, as I will explain in chapters 4 and 5, and they are the basis for my interpretation of the results in Chapter 7.

In relation to the relative importance of modes, Baldry and Thibault (2006: 27) point out that multimodal texts can be considered "composite products of the combined effects of all the resources used to create and interpret them". In this study I side with this idea and consider that the persuasive effect achieved by a presentation is the result of a carefully orchestrated *ensemble of modes* (Kress 2003) which are mutually interdependent and cannot be fully understood in isolation. The meaning created by this multimodal ensemble is, therefore, something more than the mere sum of each mode. Lemke (2005) refers to this phenomenon as the *multiplying effect*.

The present research argues that a multimodal approach to the study of persuasive oral genres will be able to shed some light on the role that different semiotic modes play in the expression of persuasion. These genres resort to a variety of modes to achieve their communicative goals, and since not all of them are linguistic, a multimodal approach is required to obtain a full picture of the communicative process. As I argue in the next section, ethnographic methods combined with multimodal analysis provide a useful resource for a multimodal approach to genre in general, and persuasion in oral genres in particular.

1.1.5 Ethnography and multimodality

Harklau (2005: 179) defines ethnography as "a range of diverse and ever-changing research approaches […] originating in anthropological and sociological research and characterized by first-hand, naturalistic, sustained observation and participation in a particular social setting". According to this author the purpose of ethnography is to "come to a deeper understanding of how individuals view and participate in their own social and cultural worlds" (2005: 179).

Street, Pahl and Rowsell (2014: 227) explain that multimodal approaches that apply ethnographic methods are characterised by taking "equal account of where, how, and by whom a text is made as it does of the physical features of a text as signifiers of contextual meaning". This dual focus on situated meaning-making and meaning as social practice has important implications. Researchers adopting ethnographic research methods analyse multimodal texts and genres taking into account not only production (as writers or speakers) and consumption (as addressees) practices, but also the motivations and values underlying these practices (O'Halloran 2016).

Among the different available ethnographic methods, two of them are particularly relevant for the present study, because they are part of my methodology: qualitative interviews and observation sheets. Qualitative interviews enjoy increasing popularity in recent multimodal studies. Norris (2004) was among the first researchers in multimodal studies to use interviews in order to contrast results with the participants, including visualising videos of their interactions to prompt their interpretations of the communicative activity. She calls this practice ‘playback methodology’. Querol-Julián and Fortanet-Gómez (2012) or Morell (2014) have later adopted similar methodologies, the latter using qualitative interviewing not only to include participants’ interpretations but also to obtain contextual information about the speakers’ use of the particular genre under study (e.g. previous knowledge and experience with the genre). This type of ethnographic interviews is an invaluable source of information that can help the researcher to probe into the reasons behind speakers' meaningful choices in their orchestration of modes. Nonetheless, they present some drawbacks. They are time consuming for both the researcher and the interviewee, and they require a good predisposition on the part of the informant (i.e. the quality of the data gathered depends to a great extent on the willingness of the informants to reflect about their answers). In addition, the reliability of the data gathered depends also on the quality of the interview design, the type of questions included and the interviewer’s skills (Rubin and Rubin 1995).

Observation sheets are also frequent in MDA studies in which the researcher is physically present in the event to gather data (see for example Querol-Julián 2011 or Querol-Julián and Fortanet-Gómez 2012). They are a helpful complement to video-data and greatly contribute to a scientifically solid interpretation of the communicative activity being recorded. According to Punch (2009) the type of observation most typically used in multimodal studies is naturalistic (i.e. the situation is not created specifically for research purposes) and non-interventionist (i.e. the researcher does not manipulate the behaviour of the participants, or at least not on purpose). These observations also present some limitations. For example, they require the researcher to be physically present, which can limit the number of options for data gathering. As it was the case with the ethnographic interviews discussed in the previous paragraph, they are also time consuming and require prior permission and good predisposition from the participants. Finally, it can be difficult to be fully non-interventionist and avoid at least some influence on the participants. Chapter 4 and 5 offer more details about the way I have applied ethnographic methods in the present study.

1.1.6 Academic and business discourse

This study is to a great extent motivated by the initial hypothesis that the increasing trend to promotionalization has originated interdiscoursive links between academic and business discourse. For this reason, the three genres included (i.e. conference presentations, research dissemination talks and product pitches) are likely to be intertextually related, and a multimodal approach can help elucidate these relationships. In this section I discuss aspects of academic and business discourse that are particularly relevant for the topic of this thesis. Therefore, I focus on oral discourse in particular, and on aspects that can affect persuasion. I begin by discussing academic discourse and close the section with a focus on business discourse.

Hyland (2009: 1) defines academic discourse as “ways of thinking and using language which exist in the academy". In the study of academic discourse, Hyland (2009) identifies the following main trends:

i) A textual trend, which focuses on language choices, meanings and patterns in texts. Within these trend, we could include corpus analysis and genre analysis;
ii) A contextual trend, which takes the wider situational aspects as a point of departure and looks at how language is socially enacted in specific situations;
iii) A critical trend, which tries to bring to the fore the ideologies and relations of power revealed through texts.

The present study fits into the textual trend, and is concerned with how speakers, as members of a *discourse community*, make choices in their meaning-making process as they design their texts. In the case of this thesis these *texts* are oral presentations, and the *choices* I pay attention to are not always within the system of language, but concern other non-linguistic semiotic modes. Swales (1990) describes a *discourse community* as composed of members with a suitable degree of relevant content and discoursal expertise, who share a set of common public goals and use one or more genres and specific lexis in the achievement of those goals (Swales 1990: 24). As Hüttner (2005: 32) says, discourse communities play a crucial role in genre analysis:

> The importance of the discourse community, which ‘owns’ particular genres, lies in its status as a source of ethnographic information regarding the prototypicality of particular texts as exemplars of a genre, and the position of individual genres within the professional field in which they occur. Obtaining such information is seen as a necessary factor in the analysis of particular genres.

Mastering the discourse of a professional community can be considered the key to succeed as a member, because it is the conventions of this particular discourse (in the form of genres) that determine what is accepted or not as true and valid. It is therefore a prerequisite to design persuasive messages within that community. As Hyland explains:

> In engaging with others we enter into a community of shared belief or value concerning what is interesting or worth discussing and extend what has been said before. The notion of community therefore seeks to offer a framework within which these actions occur and so characterise how speakers position themselves with and understand others. It is a means of accounting for how communication succeeds" (Hyland 2009: 47)

In the specific case of academic discourse, identifying a unified and homogenous discourse community is complicated due to the differences existing between *disciplines*. These discipline-related differences also influence the design of a persuasive text. According to Hyland, discipline-related differences reveal different ways of establishing truth and engaging with readers. “Scholarly discourse is not uniform and monolithic, differentiated merely by specialist topics and vocabularies. It is an outcome of a multitude of practices and strategies, where what counts as convincing argument and appropriate tone is carefully managed for a particular audience" (Hyland 2004: 3). A wide-spread distinction within disciplines is that of hard vs. soft sciences. Hyland (2005, 2006) has provided an extensive description of their differences. Hard sciences are empirical and objective, and show a rather linear progress as new problems are solved based on former ones. This usually means that validity can be experimentally proved and the contribution of each new study tends to be quite clear. On the other hand, soft sciences are more interpretative, and as a consequence its validity cannot always be empirically proved (i.e. studies cannot always be replicated) and new contributions are not so linearly derived from previous research. This has important consequences for the way authors project themselves into their messages in order to sound convincing: within the hard sciences, authors can rely on replicability and use an impersonal style to suggest that results would be the same irrespective of the individual doing the study; in soft sciences, authors rely on their own interpretations, and they need to invoke an intelligent, credible and engaging persona (Hyland 2000).

A feature of academic discourse that has been extensively discussed and also directly affects the persuasiveness of a message is *hedging*. Hyland (1995, 2006) claims that hedging in English academic writing has three important functions: 1) to ‘‘distinguish the actual from the potential and imply that a proposition is based on the writer’s plausible reasoning rather than certain knowledge’’; 2) to ‘‘help writers avoid personal responsibility for statements in order to protect their reputations and limit the damage which may result from categorical commitments"; and 3) to ‘‘contribute to the development of the writer-reader relationship, addressing the need for deference and cooperation in gaining reader ratification of claims’’ (Hyland 2006: 34-35). I believe hedging needs to be carefully crafted in order to achieve a persuasive message. In the case of oral presentations in particular, the level of hedging needs to be fine-tuned to the audience, in order to build an appropriate speaker-audience relationship.

Previous research also suggests that academic discourse can be, at least partly, culture-dependent, and in particular what concerns the use of hedging. As Uysal (2014: 182) claims in his study of conference proposals,

> existing literature suggests that use of an appropriate level of hedging and indirectness is crucial in scientific academic discourse, and the degree of indirectness in opinion-statements or claim-making in written argumentation varies across cultures. Thus, cross-cultural differences in indirectness/hedging are also likely to exist in conference proposals — another type of persuasive scientific genre.

Morell (2014, 2015) also finds cultural differences in her multimodal study of conference presentations, in particular in the field of social sciences, where her participants revealed that there are differences between NS and NNS in delivery style and in the use of humour to build rapport (also a key to persuasion, as I will discuss in Chapter 6).

A final aspect of academic discourse that is particularly relevant for this study is its so-called *promotionalization* or *marketization.* According to Fairclough (1993), Hyland (2004) and Bhatia (2005) social and economic changes have provoked a degree of promotional discourse leaking into academic discourse. Bhatia (2010, 2012) has called this phenomenon *interdiscursivity*. A sign of it is the increasing importance that academics attach to novelty, relevance, credibility and recognition. As Hyland claims, "while academic ideologies may claim that research is driven by the disinterested pursuit of truth, individual academics generally put peer approval and institutional recognition high on their list of motivating forces" (Hyland 2009: 14). This provides a sound rationale for my contrastive study of three genres (i.e. conference presentations, research dissemination talks and product pitches) that come originally from two different discourses (academic and business). Given this promotionalization of academic discourse, I hypothesize that conference presentations may borrow features from research dissemination talks and product pitches. In other words, I expect to find intertextual and interdiscursive relationships among them. I resume this idea in

Chapter 3, as I discuss persuasion as a drive for generic change and genre mixing in the form of intertextual and interdiscoursive relations across the three genres.

Turning now to business discourse, research on this area has apparently paid more attention to written genres than to oral ones. Within oral business discourse, interactional genres seem to have attracted more attention, especially meetings, and particularly in multicultural contexts. Boden (1994) performed an innovative study that focused on what business talk (i.e. who talks to whom, what is said, how it is said and when) can reveal about the way a company works. Firth (1995) also undertook a study of selling interactions among speakers of English as Lingua Franca in a variety of workplaces, applying a multidisciplinary approach to the communicative process of negotiations. Bargiela-Chiappini and Harris (1995, 1997) provides a cross-cultural study of business meetings in Italian and British contexts. Attention is paid to non-linguistic aspects such as power or politeness, but the focus is on their linguistic realisation. Poncini (2002) also approaches multicultural meetings, focusing on how language is used to construe business relationships. Crawford-Camiciottoli (2011) directs her attention to a different and more innovative oral genre: corporate earning calls. These are telephone-mediated meetings between company executives and investment analysts with the aim of providing financial information. She studies the rhetorical strategies used by participants, which are often aimed at guiding interpretation in the desired direction and promoting corporate image.

As for corpus-based studies, McCarthy and Handford (2004) undertook an analysis of spoken business English, using the Nottingham Corpus of Business English developed at the University of Nottingham, School of English Studies. This corpus covers a variety of genres such as meetings, office talk, sales presentations or telephone conversations. They set themselves the following broad research questions: to what extent is spoken business English unlike everyday informal casual conversation? And can it be compared to academic talk? They concluded that Spoken Business English is an institutional form of talk, with an associated community of practice, although it also shares features of casual conversation. For example, interpersonal meaning has the same importance as transactional meaning both in business English and in conversation.

What this review seems to bring to the fore is that effective oral business discourse is highly dependent on the interpersonal aspect of meaning, so much so that the overt communicative intention of a genre can be eclipsed by a more covert intention (e.g. an overtly informative genre can have the hidden agenda of persuading or selling). Consequently, the rhetorical strategies that speakers use have attracted considerable interest among discourse analysts, since they determine the effectiveness of communication. Interestingly, the same seems to be true to a great extent for academic discourse (see for example the notions of intertextuality, interdiscursivity and promotionalization discussed above). It is my intention with this thesis to probe into the role of persuasion in oral academic and business discourse, and in particular its multimodal realisation in the context of conference presentations, research dissemination talks and product pitches. I believe this can shed light on the way these genres achieve they communicative purpose and greatly improve the way we understand and use them.

1.2 Genre studies

In this section of the chapter I focus on the concept of genre gradually narrowing down my scope to a multimodal approach, which is the one that feeds my analysis. I will argue that a multimodal approach is particularly suitable for the study of persuasive oral genres, since oral genres in general, and persuasive oral genres in particular, are especially prone to resort to a variety of semiotic modes.

1.2.1 An overview of the concept of genre

Genre studies have a very long tradition. Ever since Aristotle's work (see for example *Poetics*), innumerable authors have dealt with the issue of genre and, especially in the last decades, genre theory has developed to become a major trend in the study of professional and academic discourses (Miller 1984; Swales 1990; Bhatia 1993; Freedman and Medway 1994, Paré and Smart 1994; Paltridge 1995; Bazerman 1994; Berkenkotter and Huckin 1995; Yates and Orlikowski 2002; Kwasnik and Crowston 2005, among others). Most early genre studies have focused on *written* texts, and especially on *literary* texts. More recent trends, however, consider non-literary oral texts as amenable to being studied in terms of genre, including texts that have long been regarded as showing little regularity, such as for example conversations (Bakhtin 1986; McCarthy 1998).

The concept of genre is a complex one, and a proof of this is the numerous definitions of the concept provided in the literature, as well as the different proposed taxonomies. In general terms, an evolution can be perceived from a focus on text-internal elements (Longacre 1976; Biber et al 1998) to a greater acknowledgment of the importance of text external elements such as the production and interpretation processes and the context in which the text is produced (Miller 1984; Swales 1990; Bhatia 1993; Paré and Smart 1994; Paltridge 1995;

Berkenkotter and Huckin 1995; NT&T 2011). In addition, some recent trends account for the flexibility of genres and the emergence of new genres (Kwasnik and Crowston 2005; Halmari and Virtanen 2005) and the complexity of the meaning-making process through the use of multiple semiotic modes (Kress and Van Leeuwen 2001; Kress 2003; Van Leeuwen 2004). In the following pages I will provide an overview of different approaches to genre that reveal this evolution. This review will set the grounds for a methodological proposal to approach persuasive presentations from a multimodal perspective.

Miller (1984) was among the first authors to consider in her definition of genre the restrictions that the specific situations imposed on users. Likewise, she argues that genres have a recognisable coherent pragmatic force, or, in other words, a conventionally acknowledged purpose.

Bazerman (1988, 1994, 2005), on the other hand, perceives genres as a standardised response to a recurring communicative situation. Therefore, in his definition both context and purpose play a crucial role. In fact, he equates instantiations of a genre with *macro-speech acts*, which are defined by communicative purpose and need to fulfil certain felicity conditions to be successful, in the line of Austin and Searle's theory of speech acts (see Section 1.1.2 for an explanation of Speech Act Theory). According to Bazerman (1994) genres are similar to moves or acts in conversations, which can be followed by a set of possible responses. To describe how genres interact with each other in specific settings he uses the concepts of *genre set* and *systems of genres*. A genre set is the full range of texts that professionals must produce in the course of their work, while a system of genres includes all the texts produced –usually in some typical sequence - in the communicative situations deriving from professional activity, including all the parties involved. As I will explain in section 1.3, these two concepts are frequently present in the definitions of the conference presentation and the research dissemination talk, although occasionally varying in terminology.

Probably one of the most widely accepted definitions of genre, especially in the context of LSP, is the one provided by Swales (1990: 34)

> A genre comprises a class of communicative events, the members of which share some set of communicative purposes. These purposes are recognised by the experts members of the parent community, and thereby constitute the rationale for the genre.

He perceives genres as a communicative event, i.e. language in action. According to this author a genre has a recognizable communicative purpose, which is not exclusively the purpose of the individual, but one acknowledged by the community of users. This purpose lies at the heart of his approach to the analysis of genres, which is based on the identification of communicative purpose and the corresponding *moves* (realised by means of rhetorical strategies) which are used to fulfil that purpose. Later, Askehave and Swales (2001) fine-tuned their notion of purpose as a defining feature of genre claiming that purpose is more complex than it might initially seem, and it cannot be used as an unequivocal concept to classify genres. Rather, the identification of the purpose should be an outcome of the analysis, and it is necessary to confirm initial hypothesis and *repurpose* the genre accordingly after analysis. In this study I side with this view and I consider that purpose is a defining trait of any genre. I also argue that a combined multimodal and ethnographic approach can be useful in this after-analysis repurposing of the genre suggested by Askehave and Swales.

Swales's definition has proved very influential. Many authors have developed and expanded on the concepts introduced by Swales, as this review will highlight. For example Yates and Orlikowski (1992) agree with Swales that genres are communicative actions, and that they have a particular communicative purpose. They further point out that genres entail expectations regarding content, participants, form, time and place. However they also acknowledge that changes sometimes occur in these discursive expectations as speakers adapt to the specific situation. When these changes are widely adopted they alter the genre, leading to what they call *corollary genres* and eventually to the emergence of new genres. An example of this generic change, they note, can be found in presentations. As new technologies found their way into presentations, the amount of visual aids increased considerably. The structure has also been affected, as software like PowerPoint promotes the use of bulleted lists, and even the media, as the presentations began to be projected directly from a computer. Finally, the communicative situation can also change as it is now possible to save and send the presentations by email, or even stream them, so that presenter and audience do not have to coincide in space and time. These variations will result in corollary genres.

Bhatia (1993) also foregrounds the flexibility of genres, claiming that skilled users are able to use conventions and rules creatively to provoke particular effects that suit their individual needs. He shares with Swales (1990) the idea that the members of a particular community use the knowledge gained through their experience as members of that community to shape the texts they produce, resulting in different genres, and considers the communicative purpose a determining factor in the definition of a genre. He proposes the following seven steps for a comprehensive analysis of unfamiliar genres (which draws partly from Swales' moves model):

1. Placing the given genre in a situational context;
2. Surveying existing literature;
3. Refining situational analysis (author, audience, community, other aspects of extra-textual reality);
4. Selecting corpus according to purpose;
5. Studying the institutional context (the rules and conventions that determine the way language is used in that context);
6. Levels of linguistic analysis:
    a. Lexico-grammatical features: features of language which are predominantly used;
    b. Text-patterning: this should answer the question "Why does the author use this feature in particular and not another? At this stage, the analyst will try to find the function of each predominant feature;
    c. Structural interpretation: At this stage Swales' model of rhetorical moves provide a deeper insight into the cognitive structure of the genre. Each move is realized by different rhetorical strategies, and shows a particular communicative intention, which in turn contributes to the overall communicative purpose of the genre;
7. Specialist information: check results with specialist members of the community in which the genre is used.

This methodology is particularly interesting due to its ability to present and explain textual features and rhetorical strategies as contributing to the communicative purpose of the genre. Bhatia (2002) also agrees with Bazerman on the concepts of genre set and system of genres to explain the complex relations between genres, and adds the notions of *disciplinary genres*, or group of genres belonging to a disciplinary domain (e.g. law or business), and *colonies of genres*, which are genres operating across disciplines but sharing a more or less common communicative purpose. He identifies, for example, a colony of promotional genres (Bhatia 2005), all sharing the communicative purpose of promoting a product. Central in this colony are the advertising genres, but equally promotional are other academic genres such as University websites or professional genres such as CVs or reference letters. As Bhatia (2002) remarks, they may differ in medium and linguistic and rhetorical devices, but they all aim to persuade the addressee of the excellence of a product (a university) or a person (in the case of the CV or reference letter). With time, these genres may even become more and more similar as they borrow from each other, generating genre change, hybridization and eventually emergence of new genres (Halmari and Virtanen 2005). Other examples of this tendency are video-CVs, that borrow the media form TV advertisements, and company profiles in social networks, that borrow the layout from personal profiles but are still pitches. Recently Bhatia has broadened his notion of relationships between genres to include not only appropriations of text-internal resources, but also contextual and text-external ones. He calls this phenomenon *interdiscursivity* (Bhatia 2010, 2012), and it entails not only the mixing and embedding of genres, but also professional practices and professional cultures.

Paré and Smart (1994) follow Bazerman (1994) in their definition of genre. They conceive genres as patterns of regularities, not only in textual features (i.e. structure, rhetorical moves and style), but also in the production and interpretation processes (when, where, how and why the text is produced and received) and social relations of interactants (what can or cannot be said and done, when, how, where and to whom). I share this view of genre, and in my analysis both the production and interpretation process and the social relations established among interactants play a crucial role.

Berkenkotter and Huckin (1995) also share Bazerman’s concept of genre as a response to a recurrent situation. Furthermore, they draw on Swales’ idea of the parent community to argue that genres are linked to particular disciplines and their methodologies. In fact they claim that a genre cannot be considered as such until it is accepted by a community. The practices of this parent community –and particularly the changes in them– will therefore deeply affect the way genres change and evolve through time to better suit the needs of this community. They suggest five parameters to describe genre:

i) *Dynamism*: genres change over time in response to the needs of their users.

ii) *Situatedness*: they are embedded in communicative activities, so genre knowledge is situated cognition and it is not learnt, but rather enculturated as the user becomes member of a community. This notion has been later taken up by other researchers (e.g. Rowley-Jolivet and Carter Thomas 2005; Morton 2009).

iii) *Form and content*: what to say and how to say it are also integral aspects of the knowledge about a genre, as well as the background knowledge assumed.

iv) *Duality of structure*: there is a reciprocal influence between genre and the situation in which it is used. Genres are the result of these situations and social practices, and at the same time reinforce them and shape them through recurrent use.

v) *Ownership by a community*: as mentioned before, conventions applied to genre are those of a particular discourse community, which reveal its epistemological assumptions.

Their description of genre is perhaps more flexible than other proposals. Instead of providing a taxonomy of genres they focus on identifying generic elements in different texts that illustrate a particular genre. These texts are related to each other because they share similar discursive and linguistic features. Likewise, within the same text it is possible to find generic elements characteristic of different genres. This relationship between texts is called *intertextuality*, and it highlights the fact that a correct interpretation of a particular text is dependent on knowledge of a previous text (De Beaugrande and Dressler 1981).

These relationships of intertextuality and interdiscursivity between genres mentioned in the previous paragraphs and also in section 1.1.6 are the rationale behind my decision to compare three genres that come from different disciplines and could be in principle considered to be totally unrelated. As I will argue in Chapter 2, the need to be persuasive is a common drive that can be hypothesized to prompt intertextual and interdiscoursive relationships between the three, as indeed Bhatia (2010, 2012) has found for other colonies of genres that are intertextually and interdiscursively related.

Paltridge's (1995) definition of genre is also based on the concept of intertextuality and *prototype*. A prototype is a model from which different genres inherit qualities, but genres can vary in their prototypicality. The different instances of the same genre are related to each other through intertextuality. He points out that we assign texts to a particular genre on the basis of pragmatic and perceptual aspects, rather than linguistic and text internal features. In fact, when a text does not resemble closely its prototype, we assign it to a particular genre based on a pragmatic basis. He coincides with Bazerman that felicity conditions as proposed in Speech Act theory can be applied in this process. These conditions emerge from the context and are non-linguistic. As examples of felicity condition he mentions a particular type of author and audience, a particular medium for dissemination, or a particular communicative purpose.

Finally, Bateman (2014) offers a different and more SFL-oriented idea of genre. He perceives genre as a multi-stratal phenomenon, in which he identifies regular choices across the options available in each strata. This patterned collection of choices across strata is what defines genre.

1.2.2 The case of oral genres

The flexibility of genres pointed out in the previous section of this chapter is even more salient in the case of oral genres. In fact, it is precisely their flexibility and dynamism, together with their potential as a tool for social interaction that is widely highlighted in previous research on oral genres. For example Hymes (1972) emphasizes the dynamism of oral genres and differentiates them from *speech events*, pointing out that the same genre can show variations in different speech events. In the same line, Coupland (1983) points out the variability in service encounters: buying a stamp is different from buying a holiday (the latter is less prone to be enacted as a ritual and allows for more interactional talk), and yet both are the same oral genre. Bargiela-Chiappini and Harris (1995) make the same point in comparing service encounters with business meetings. On the other hand, Ylänne-McEwen (1997) considers that in service encounters the transactional element (i.e. getting business done) and relational element (i.e. establishing social relations) intertwine and need to be equally regarded in a description of the genre. Likewise Thompson (1997, 2002) treats the building of an appropriate relationship between speaker and audience and the assumption of roles as integral parts of oral research presentations.

A particularly influential view of oral genres is provided by Bakhtin (1986). His idea of genre is based on the concept of *utterance*. An utterance is a unit of talk, with full communicative value and capable of prompting a response from the interlocutor. In this sense it differs from the sentence, which is a unit of grammar with no communicative value until it is actually used in a specific communicative situation by a specific speaker and addressed to a particular interlocutor. Utterances, therefore, reflect a communicative situation within an area of human activity. Each area of human activity in which communicative situations emerge develops stable types of utterances, which in turn constitute an oral genre. In this sense, he points out that

> each separate utterance is individual, of course, but each sphere in which language is used develops its own relatively stable types of utterances. These we may call speech genres. (…) A particular function (scientific, technical, commentarial, business, everyday) and the particular conditions of speech communication specific for each sphere give rise to particular genres, that

> is, certain relatively stable thematic, compositional and stylistic types of utterances. (Bakhtin 1986: 61, 64)

Bakhtin (1986) differentiates between *secondary* and *primary genres*. Secondary genres are the result of complex and highly developed communicative situations (e.g. a political speech) and in their construction they "absorb and digest various primary (simple) genres that have taken form in unmediated speech communion" (1986: 62). A conversation would be an example of a primary genre, which could be used in the construction of a more complex secondary genre such as a novel. Looking at genre from its social aspect, he distinguishes between generic forms that we use to shape our discourse and language forms that we use to construct our message. The latter are stable and compulsory, the former are much more flexible. This accounts for the diversity and flexibility of oral genres, which can differ according to the situation and relationship of the participants. Despite this flexibility, Bakhtin (1986) acknowledges regularities in oral genres. In fact, he claims that each utterance carries what he calls "echoes" of other utterances to which it is related due to a common generic form. On the other hand, looking at genre from a more individual perspective, Bakhtin points out how the genre-related choices speakers make in their utterances are partly determined by addressees and their potential response. In his own words: "The various typical forms this addressivity assumes and the various concepts of the addressee are constitutive, defining features of various speech genres" (1986: 99).

McCarthy (1998) agrees with Bakhtin that a genre is both the product of the individual mind and a social construct. According to him, oral genres are characterised by dynamism, fluidity, variability, mixing and negotiation. This is due to the fact that speakers use language with some practical goals in mind, and these goals are usually not set in advance, but emerge as discourse goes along. In this sense he describes genre as a

> useful concept that captures the recurrent, differing social compacts (i.e. cooperative sets of behaviour) that participants enter upon in the unfolding discourse process, whether writing or speaking. (…) Participants orient themselves towards necessary actions such as establishing roles and identities, protecting face, achieving goals, and so on. (…) It is the task of the genre analyst to construct the bridge between texts and socially constituted activity (1997: 32).

This view of genre claims for a focus on different types of generic activity rather than a static notion of genres. McCarthy (1998) suggests that the features identifying oral texts as genres can be classified along four dimensions. The first is *formulations*, which are recapitulations or comment on progress of on-going generic activity. The second is e*xpectations* regarding the kind of generic activity. The third is r*ecollections*, or participants' past experiences that echo in their discourse (in the sense given by Bakhtin). The fourth and final is *instantiations*, or examples of new aspects that participants agree to incorporate to the generic activity and which account for genre dynamism.

In view of the flexibility of oral genres, the task of identifying and classifying them on the basis of recurrent features of form, content and purpose becomes more complicated. In the description of oral genres, the role played by the interaction is crucial and probably more defining, as it explains the decisions made by participants regarding the way they decide to say things. In this sense, the contribution of paralinguistic and kinesic features is important, since they can reveal how speakers perceive the interaction as well as their communicative intentions. For example Bakhtin (1986) notes how intonation is employed in oral genres to carry the expressive tone. He further points out how the expressive tone carried by intonation can produce a reversal of the expectations of a genre (e.g. a comical epitaph can achieve its humorous effect through the way it is read aloud). In addition, intonation expresses the individuality of the speaker (his emotionally speech intent) in highly conventionalized and standardized oral genres such as greetings or congratulations. This is also the case in the presentations I include in my study. As I will illustrate when I discuss my results, speakers can use intonation or a kinesic feature such as a gesture or facial expression to express their communicative intent (i.e. if they expect the audience to concur with a particular evaluation) and also to project the way they perceive their interaction with the audience (e.g. if they expect the audience to assume something as already agreed or as something open to discussion).

1.2.3 Points of agreement and shortcomings of traditional approaches to genre

The most relevant aspects that this review of the concept of genre has revealed can be summarized as follows:

i) *Recursiveness*: a genre happens more than once, and is characterized by recurrent use.
ii) *Regularity in purpose, content and form*: a genre is recognizable by recurrent patterns of form and purpose, and according to some authors such as Paré and Smart (1994) and Yates and Orlikowski (1992) also regularities in content.
iii) *Social phenomenon*: genres are language in use in particular situations and allow social interaction.

iv) *Reciprocal influence society-genres*: Genres are socially constructed, but in turn they also influence and shape social situations and conventions through recurrent use. Paré and Smart (1994) refer to this phenomenon as *duality of structure*.

v) *Dynamism*: genres change according to users' needs and also respond to social changes. Rather than a product, they are dynamic responses to particular situations.

vi) *Acceptance by a community*: this is a requirement for a genre to be considered as such. Through this acceptance, a genre raises expectations in its community, and users can make predictions of what they are likely to encounter.

The concepts of genres as (inter)action, the existence of a community of use, the dynamism that characterizes genres to adapt to social situations and discourse communities, and the importance of the production and interpretation processes are all revealing aspects that suggest a need to move to a wider description of genre, capable of accounting for non-linguistic aspects which also have an influence on genres and on the way they are used and understood (Kress 2003). In the case of oral genres, I think that the role paralinguistic and kinesic aspects play in the meaning-making process becomes crucial. Likewise, social changes and technological advances multiply the ways in which meaning can be transmitted and construed. This leads me to revisit the concept of genre adopting a multimodal approach that accounts for a multiplicity of semiotic modes to make meaning, in order to achieve more accurate and comprehensive descriptions of genres in general and of conference presentations, dissemination talks and product pitches in particular. This is the aim of the following section.

1.2.4 Multimodal approaches to genre

As I have argued in the literature review presented in the previous sections of this chapter, as society and the way people communicate evolve, new texts that do not seem to fit into existing genre typologies emerge. Widely accepted definitions become unable to account for such emerging genres and therefore a revision of the concept of genre is needed. Along this line, Kress (2003: 35-36) emphasizes the need to move from a theory that accounts for language alone to "a theory that can account equally well for gesture, speech, image, writing, 3D object, colour, music and others". Bateman (2014: 237) also suggests that, as this extension in the concept of genre takes place, the questions arises whether "definitions and conceptions of genre derived in less multimodally-driven times continue to be adequate"

Likewise, in genre studies a shift can be perceived from a focus on the final product to a growing interest in the production and interpretation processes. As the review of previous literature on the concept of genre has revealed, there is a wide acceptance that genres deal with language in use, and therefore they are to be treated as dynamic processes and not as stative products. For example, according to authors such as Kress (2003) or Van Leeuwen (2004), genres are dynamic and the result of a creative process. This is particularly salient in the case of persuasive genres (Halmari and Virtanen 2005). Kress defines genre as the "shaping of text which reflects and is brought into existence as a result of the social relations of the participants in the making (writing/ speaking) and in the use (reading/ hearing/ interpretation) of a text" (Kress 2003: 103). He also foregrounds two main aspects of genres: 1) they are the result of the social relationships of the participants, i.e. they are a social product; and 2) they are conceived as a dynamic process of meaning- making and interpretation, i.e. they are not stative.

This means that the production and interpretation processes are relevant in the description of genres. Furthermore, these processes are *multimodal*, because users select among available semiotic modes to express their meaning (Kress and Van Leeuwen 2001; Kress 2003; Norris 2004; Kress and Van Leeuwen 2006). In other words, rather than relying exclusively on written language or verbal elements, in most cases users resort to a combination of different ways of expressing their meaning. Just to mention a few evident examples, nowadays written texts include an increasing amount of images, and oral texts are almost inevitably accompanied by hand gestures, head movements, gaze and facial expressions.

A mode, as mentioned in section 1.1.4, is a semiotic resource which allows the realisation of discourse. Kress and Van Leeuwen (2001) argue that modes are not strictly specialized in the type of meaning they convey, but rather share common semiotic principles instead. In addition, new technologies make it possible to work across modes almost seamlessly, something that is particularly salient in the case of oral presentations, that make use of visual aids and sound. *Modes* are not to be confused with *medium*, which are "material resources used in the production of semiotic products and events" (Kress and Van Leeuwen 2001: 22). Written language is a mode, a semiotic resource, but it can be realised by different production media, such as a book, a post-it, a computer, etc. (see Section 1.1.4 for an overview of the basis of multimodal theory).

Kress (2003, 2009) claims that the production process of a text is closely related to design. Users design the most appropriate way of communicating their meaning in a given situation, and this implies choosing the most suitable combination of modes, bearing in mind rhetorical and ontological issues. Each mode has different affordances and epistemological commitments, and they implicitly carry different *takes on the world* (Kress 2009: 71). That is, they allow users to do certain things, but at the same time they also impose certain restrictions. As an illustration of this idea we might take the example of a description of a room in writing vs. image. The differences between both modes are clear: representing through image forces the user to accurately locate each element in space and in relation to one another, while in the case of writing, a relation of recipient-content is established instead between the room and the elements inside it, and reference to location may be avoided, as shown in Figure 1.4:

Vincent's Room Ares 1888, Vincent Van Gogh Foundation, Amsterdam

"The walls are pale violet. The floor is of red tiles.

The wood of the bed and chairs is the yellow of fresh butter, the sheets and pillows very light greenish-citron.

The coverlet scarlet. The window green.

The toilet table orange, the basin blue.

The doors lilac.

And that is all--there is nothing in this room with its closed shutters.

The broad lines of the furniture again must express inviolable rest.

Portraits on the walls, and a mirror and a towel and some clothes.

The frame--as there is no white in the picture--will be white."

Extract from a letter of the painter to his brother Theo Van Gogh. [1]

Figure 1. 4 Comparison of writing and image as modes

As argued by Kress and Van Leeuwen (2006), it is also necessary to take into account that not all modes enjoy the same social valuation. In Western cultures language –especially written language- has enjoyed a privileged status in many contexts, while other modes such as image have been neglected or even ruled out from certain communicative situations. Most readers would not expect to find a picture in a legal document, for example, but they would

[1] Retrieved 3/2/2017 from http://www.vggallery.com/painting/p_0482.htm

find it perfectly acceptable in a children's book. It is also interesting to note that this valuation is entirely culturally influenced, and therefore varies from one culture to another. Equally remarkable is the fact that these valuations change according to social changes. It is clear that new technologies have had an influence on the way different modes combine to create coherent messages, and that social changes are making modes such as image or music gain importance in areas where they have long been neglected (e.g. text books) (Iedema 2003).

Apart from advantages and restrictions inherent to modes themselves, Kress and Van Leeuwen (2001) point out that the process of making meaning is influenced by three main aspects:

i) *Human physiology*: when we perceive a message this is a physiological process, which means that the physical senses used when perceiving a message also affect the way meaning is construed. Different modes with their different production media may offer different sensory appeals, which is why hearing or reading the name of a colour is not the same as seeing that colour.
ii) *Culture*: communication is a tool for humans to act as social actors. The modes themselves that we use in the construction of messages are actually "conventionalisations produced through cultural action over time" (Kress and Van Leeuwen 2001: 28). Likewise, it is undeniable that our interpretation of a message is also mediated by our culture.
iii) *Meaning potential*: the elements that make up the message –the material signifiers- can contribute to meaning through the potential associations that they may prompt. That is, they may acquire meaning derived from what we associate them with. For example, a paper napkin can be used as *medium* for the transmission of a text (written language as *mode*), and the paper itself will add to the meaning of the message because of the associations that it may prompt (i.e. what it is that we do and how we feel when we use this product).

From this discussion it follows that not all modes and media are equally suitable for all discourses in every situation, and this is where design comes into play. Kress (2009, 2009b) uses the concept of *modal fixing* to explain how the design of a specific multimodal ensemble enables a particular meaning to get across. Modal fixing is similar to developing a photograph and capturing the image on paper: a mode captures meaning. These multimodal ensembles are then used in genres, which he perceives as frames that capture the social relationships between the participants in the communicative situation. In the case of oral genres, their intrinsic flexibility and dynamism (as discussed in section 1.2.2) make this modal fixing more difficult to determine, as the modal configuration is constantly changing, producing different multimodal ensembles at each specific moment. This is a consequence of the low level of materiality of many of the modes that are salient in oral discourse, such as intonation or spoken words. In Pirini's (2016) terms, these modes have a *fleeting materiality*, because they do not persist throughout the interaction. This makes their analysis more challenging. Nonetheless, future research with larger-scale studies can reveal regularities regarding the use of multimodal ensembles in oral genres that can contribute to our understanding of these genres.

Bateman (2008, 2014) proposed the model "Genre and Multimodality" ("GeM") to study multimodal genres, in particular written documents with a strong reliance on visual elements. Drawing on Rhetorical Structure Theory (Mann and Thompson 1988), the focus of his approach to genre is the relationship between elements and the meaningfulness of their combination. He claims that in the same way that segments of a text contribute to coherence, different modes contribute to a meaningful whole. The GeM model pays particular attention to how layout, typography, images and language, as well as physical properties of the document (e.g. size or display technology) are all closely interconnected in the process of meaning-making. Bateman (2014) argues that a complete description of a genre needs to take into account not only the rhetorical strategies that fulfil the communicative goal, but also the semiotic modes that realise those strategies, and consider both as two sides of the same coin. This is the approach that I take to persuasion in oral genres, as I will describe in detail in chapters 4 and 5. These modes are, according to Bateman, made available by the medium that supports the genre. A medium is a "historically stabilised site for the deployment of some selection of semiotic modes for the achievement of varied communicative purposes" (Bateman 2014: 256). A medium can thus materialise a whole range of genres. Books, for example are a medium that materialise a wide variety of genres. He also warns that media have but an indirect relation with genre considerations, and that "information that is inherent in the medium should not be considered as offering an identifying feature for some presumed variety of genre in that medium" (Bateman 2014: 256).

In sum, multimodal approaches to genre studies reveal the importance of aspects that, to a great extent, have been neglected in the description of genres. These aspects, however, play a central role in the transmission of meaning and communicative intentions, which, in turn, are at the core of what makes a genre recognisable as such. This is what leads me to adopt a multimodal approach in the study of persuasive oral presentations. However, as Bateman (2014) very relevantly points out, multimodal approaches to genre do not always agree on which aspects or features can be considered defining of the genre and which are not and might even distort the definition of the genre. In addition, a multimodal approach to genre can challenge the validity of what have traditionally been considered central properties of genre. Bateman provides the example of moves, which are "sequenced staging of activities within any genre" (Bateman 2014: 243) dependent on a linearity in space (written) and time (spoken). The problem is that this linearity very easily disappears when we move away from the strictly verbal. One possible solution, he suggests, is to focus the definition of a genre in the figure of the recipient, because even if genre itself can no longer be considered a linear artefact, human attention remains time-based. However, this still leaves open the question of how strictly are interpretation paths intended to be understood. According to Bateman, genre implies restrictions and constraints in production, in the form of conventions or common practices, which at the same time prompt in the addressee some interpretive frames and expectations. In the present study, and as I will argue in Chapter 3, I consider that kinesic and paralinguistic features are characteristics of oral presentations which need to be taken into account in their analysis as genres. I also focus on the speakers' intended interpretation path, which I confirm with the speakers with the help of ethnographic interviews (see Chapter 5).

1.3 Oral genres in academic and business discourse: conference presentations, research dissemination talks and product pitches.

In this final section of the chapter I zoom in on the three genres which are the object of my study, i.e. conference presentations, dissemination talks and product pitches. I review previous literature on these genres to which I hope to contribute with the results of my study. It will be noted that the amount of interest and subsequent research on these three genres is quite unbalanced. So far, it has clearly favoured conference presentations, with only a recent interest in research dissemination and very little scientific research on product pitches to date, to my knowledge.

1.3.1 The conference presentations as a genre

The conference presentation can be classified as an academic genre since it is mainly used by discourse communities within academia. In order to better understand the nature of this genre and the way it is related to other genres within academic discourse, I draw on

Fortanet-Gómez's (2005) classification of spoken academic genres, which uses the purpose of the genre as the main criterion and distinguishes three main categories:

i) *Classroom genres*, within which we can include lectures, seminars, student's presentations and oral exams, among others;

ii) *Institutional genres*, which include academic year opening lectures, Honoris Causa speeches and Rector's addresses to the faculty, among others;

iii) *Research genres*, which can be further subdivided into a) *conference genres,* that include plenary lectures, poster presentations, workshops and the conference presentations that are central to this study and b) *other research genres* such as PhD thesis defences, Master's thesis presentations, etc.

Conference presentations have attracted the attention of many scholars who have tackled its study from different angles. Among these, we can mention the first publication on this topic by Dubois (1982) who presents a detailed account of biomedical conferences. Later Rowley-Jolivet (1999: 179) claimed what she termed the "pivotal role" of conference presentations in academic research, and identified three main functions of this genre: to present a scientific novelty, to give visibility to research and to reinforce social cohesion within the discourse community. This work was expanded in Rowley-Jolivet and Carter-Thomas (2003, 2005), in which they studied presentations from three different approaches: microscopic bottom-up analysis, macroscopic top-down analysis and multimodal approach. They also compared the rhetorical structure of conference introductions with research article introductions. In these studies, they highlight an important feature of conference presentations that is of particular relevance for this study: they are not only novelty-oriented and informative, but they are also inherently persuasive, something which will be discussed in more detail in Chapter 2.

Another issue of particular relevance raised in the literature on conference presentations is the immediacy in time and place in relation to an expert audience with whom a great amount of shared knowledge can be expected. This is discussed by Hood and Forey (2005) and Rowley-Jolivet and Carter-Thomas (2005), who maintain that the presence of an audience and the consequent need for real-time processing strongly influence the way a presentation is designed (e.g. less dense, simpler, with a considerable amount of signposting and inclusive use of pronouns) and the way different semiotic modes are used. In spite of the fact that they are semi-prepared or rehearsed, they require constant adaptation on the part of the speaker according to the reaction of the audience. For example Hood and Forey (2005) note how during introductions, speakers use evaluation to subtly pose their research in positive terms and also to seek alignment with the audience. They further notice how the use of inclusive gestures that bring the audience into the discussion can contribute to this. Rendle-Short (2006) follows up on this idea claiming that the presentation is designed for a particular audience. In conversation analysis (CA) this is called the principle of "*recipient design":*

> Recipient design of a turn at talk therefore refers to the speakers ability to take account of who the recipient is and what they know, as well as referring to the ability of the speaker to continually monitor the recipient and to modify their talk in response to the interactional needs of the recipient (Rendle-Short 2006: 12)

According to her CA approach to presentations (previously also supported by Brown and Yule 1983), a presentation is an interactive dialogue between a speaker and an audience. CA looks at how speakers interact with an audience that does not respond audibly but is none the less reacting. Atkinson (1984), for example, showed how in public speeches speakers interactively elicit applause from their audiences, and not only through verbal devices such as three part lists and contrastive pairs, but also through gaze shifts to the audience.

The relationship of conference presentations with other genres has also been the focus of several studies that see this genre as part of a macro-generic event embedded in the wider conference experience. With the term *semiotic spanning*, Ventola (1999, 2002) tries to explain the connection between the presentation itself and the discursive practices of the participants during the whole conference event. This implies that every presentation is necessarily influenced by the previous presentations in the conference event and the future presentations that will come after it, as well as by related genres such as the abstract or the research article. This relationship has also been explained through the concept of *genre chain,* which identifies a chronological sequence of genres that are produced during the unfolding of a communicative event, e.g.:

Call for papers→Abstract→Presentation→Research article (Räisänen 2002).

Shallom (2002) for example identifies the following macrogeneric event of the conference:

-pre-event: call, evaluation, programme

-conference

-post-event: papers, proceedings

The idea that a conference presentation is associated with (and influenced by) a written paper is also discussed by Rowley-Jolivet (1999), Rowley-Jolivet and Carter-Thomas (2005) and Hood and Forey (2005). Presentations are often a way of testing research before a definitive paper is published. They may also present research that is still ongoing. This dynamic process is usually diluted and idealized in the written paper (Thompson 2002; Hyland 2009). For this reason, we can argue that a conference presentation is a process genre (Swales 1990; Rowley-Jolivet and Carter-Thomas 2005), since it often takes place while research is still in progress.

Hyland (2009) summarizes in the following way the differences between a conference presentation and a written paper that previous literature has pointed out:

i) a greater use of active verbs in comparison to passive forms (Rowley-Jolivet 2002);

ii) more informal boundary markers, e.g. OK, right, now, etc. (Webber 2005);

iii) more use of self-mention;

iv) more imprecise description of results (Dubois 1987);

v) more humour and self-irony (Frobert-Adamo 2002);

vi) less use of extraposition;

vii) more use of existential there (Carter-Thomas and Rowley-Jolivet 2001)

Apart from the boundary markers mentioned in the list above, Webber (2005) also finds differences in the use of other interactive features between oral presentations and written papers (e.g. personal deictics or imprecise quantifiers) . She argues that these features are used to create rapport with the audience in a communicative event that is influenced by two important factors already mentioned in the previous paragraphs: the presentation of research in progress and the co-presence of an audience.

There are also some interesting studies that deal with specific sections of conference presentations. The discussion session or Q&A has been dealt with by Shalom (1993), Wulff, Swales and Keller (2009) and Querol-Julián (2011). Wulff, Swales and Keller (2009) consider that the discussion session is part of the genre of the conference presentation, but still they note some important differences. For example the presentation is more prepared and closer to written language, while the discussion session is more conversational and closer to spoken language. According to Querol-Julián (2011) and Räisänen (1999, 2002), the Q&A section can be considered a separate, but closely related, genre, because during this section a rather different communicative situation is built, entailing a different relationship between presenter and discussant. In this thesis I side with this latter view, and for that reason I focus my analysis on the monologic part of the presentation.

Other studies have dealt with the language of presentations (Ruiz-Garrido and Fortanet-Gómez 2008) or the narrative discourse used within them. Thompson (2002), for example, claims that narratives in past tense during presentations relate difficulties and failures during research process, which in the written paper become present tense results. This is a persuasive appeal and creates rapport: it shows the person behind research and provides credibility.

Focusing on the way speakers deliver a presentation, Hyland (2009) proposes three styles (based on Dudley-Evans and Johns 1981 and Dudley-Evans 1994). The first is a reading style, when the speaker is reading from written notes. The second is a conversational style, when the presentation is closer to a conversation with the audience. Finally, the third is a rhetorical style, when the speaker is fully oriented to his performance. He also points out that these styles need not appear in isolation and are frequently combined.

Also focusing on the way speakers deliver presentations, and in particular on the roles that speakers assume and project, Zareva (2013) finds that presenters prefer to project their scholarly selves in their presentations by staying close to the written academic genres, but they also show their personal and social selves in relation to the topic content. This is consistent with Thompson's (2002) account of narratives in oral presentations mentioned in the previous paragraph.

Finally, Hood and Forey (2005) and Hyland (2009) have identified several challenges when studying conference presentations as a genre. On the one hand, we need to consider their inherent *variety*. Presentations can range from one hour plenary to a short parallel session. They can present research at various levels of completion, for example Räisänen (2015) discusses what she calls *after-print conferences*, which are held after the paper is published. Also their audience can differ in size and homogeneity. On the other hand, we cannot ignore the difficulties of collecting and coding *multimodal data*, something which I will discuss in detail in Chapter 5.

1.3.2 Research dissemination talks as a genre

The present study is concerned with a particular type of research dissemination talks, i.e. TED talks (TED being an acronym for Technology, Entertainment and Design). Caliendo

(2014) and Caliendo and Compagnone (2014: 105) define this genre as "a series of short popularizing talks (of approximately twenty minutes), addressing a mass audience and delivered by top-level experts in a wide variety of domains." They explain that TED talks as a genre provide experts with the chance to disseminate knowledge outside their disciplinary communities to a lay audience. Since the talks are made available on-line, this is possible at two levels: the audience physically present and the web-users at home. Because of this, they consider TED as a hybrid web-mediated genre connected with other genres such as university lectures, conference presentations or TV documentaries. In their study they compare TED talks to university lectures. Although in both genres an expert conveys specialized content to an audience of lay, non-expert people using different semiotic modes in a monologic speech, when moving from one genre to the other both the speaker's communicative purpose and the recipient's expectations change considerably: "While in a classroom a lecturer's main objective is to train an audience of students, these being aware of the fact that they will be assessed at the end of the course, on the TED stage academics mostly present and 'promote' their research while the audience expects to listen to inspiring and ground-breaking ideas, as well as receive a 'smart' form of entertainment" (Caliendo and Compagnone 2014: 106). They conclude that TED talks are less informative than lectures, and are mainly used by experts to construe their image stressing their affiliations and promoting their group's research. In view of this, Compagnone (2015) argues that TED talks are an example of interdiscursivity (see Sections 1.1 and 1.2 in this chapter) where informational and promotional aims overlap, and that "the 'cult of personality'-intertwined with advertising purposes - has now become a common trend even in academia, as is also evidenced by other web-mediated genres (e.g. academia.edu, LinkedIn, Twitter, Facebook) used by academics to interact with their scientific communities as well as with the lay public" (Compagnone 2014: 66).

The different purpose of the speaker and expectations of the audience means that a different relationship is established between speaker and audience, which, I would argue, is a defining trait of the TED talk. The curator of TED, Chris Anderson, points out in his article published in Harvard Business Review (2013) that a key to a successful talk is to get the right connection with the audience. For example, assuming that the audience has more knowledge and interest than they actually do, using specialised language that excludes them or sounding as if you are reciting your talk from memory will make you fail as a speaker. The clue is to introduce your topic, explain why you care deeply about it and convince the audience that they should care too. Results of later studies (Caliendo 2014, Scotto di Carlo 2014, 2014b, 2014c, 2015, Caliendo and Compagnone 2014, Compagnone 2015) have corroborated these views, demonstrating how TED talks differentiate themselves from lectures or CPs in several ways: 1) they establish a closer relationship with the audience, 2) they are meant for the general public, so they are accessible, 3) they are more emotional than rational, 4) they are more persuasive than informative: they are a call to action. The relationship of TED talks with other genres has also been noted by Caliendo (2014). She describes TED talks as a genre lying at the intersection of a number of other genres, including university lectures, newspaper articles, conference presentations and TV science programmes. Interestingly, Caliendo (2014) and Compagnone (2015) also acknowledge its multimodal nature noting that it mixes different semiotic modes (i.e. spoken, written, video and audio).

The type of TED talks analysed in the present study, as described in detail in Chapter 5, are slightly different from the ones described above because they are part of a smaller scale event organised independently by Ghent University, and are not made available on-line. The motivations of the speakers in my corpus can also be considered to differ from the ones described by Caliendo and Compagnone (2014), since the event was actually a contest and speakers competed to gain access to bigger-scale TED events. However, the underlying objective of promoting research while entertaining an audience still holds.

1.3.3 Product pitches as a genre

Product pitches are probably one of the most important genres in the business field and one that professionals are likely to deal with in one way or another in the development of their professional activity. The effectiveness of these presentations is decisive for many business transactions. Product pitches vary depending on the type of product, the communicative situation in which they occur and the community of use. However, all product pitches share one main purpose: to persuade the addressee of the excellence of the product.

Literature about product pitches seems to have focused less on defining it as a genre and more on identifying the features that make them successful. Consequently, scientific literature about this genre does not abound, and studies undertaking a definition of the genre or dealing with its multimodal nature are virtually inexistent. In contrast, there are plenty of guides to pitching that provide advice for practitioners. One example is Burzinski's (2013) "How to pitch a product: Elements of a successful pitch", where he advices speakers giving pitches to show enthusiasm, find a personal connection with the audience (e.g. by telling stories), sell the benefit of the product and focus on what is new and innovative. Nancy Duarte (2015), in “Find the perfect mix of data and narrative” also takes on the idea of the narrative and recommends telling a compelling story of how people will use the product and why it will change their lives. Another example is Clemence Leper’s (2015) “How to pitch an idea to anyone in 7 steps”. In this guide she advices to adopt a problem-solution approach, and emphasize why your suggested solution is the best. She also recommends to use Cialdinis' (1984, 2001, 2006) principles of persuasion (discussed in Chapter 2) and to appeal to senses to convince the audience of the value of a product. The strong focus on making the pitch consumer-oriented (or audience-oriented) is also taken up by Rackham (1998), who argues that a possible cause for the failure of a pitch is to focus on the innovative features of the product instead of on the customer needs. The guide “Engineering innovative products" by Woods et al. (2014), includes a chapter entitled “The perfect pitch”. In it they suggest structuring the pitch in five sections: 1) pain and solution; 2) value proposition; 3) market and competition; 4) company traction and go-to-market strategy and 5) finance. They stress the important role that evaluation plays in describing market and competition, and the need to fine-tune it appropriately: without highlighting the importance of a competitor and without choosing questionable comparison criteria.

A very noticeable characteristic of product pitches is that they tend to be short and concise. There is a variety of pitch that is actually called *elevator pitch*. The term is credited to Ilene Rosenzweig and Michael Caruso (former editor of Vanity Fair), and reflects the idea that it should be possible to deliver the pitch in the time span of an elevator ride. A golden rule for this type of pitches is, therefore, “be concise and to the point”.

What these recommendations by practitioners with first-hand experience with the genre can be said to bring to the fore is that product pitches are inherently persuasive, and that presenting a product as a (better) solution to a problem that directly affects the audience lies at the core of a product pitch.

In this chapter I have provided a review of approaches to discourse and genre that have been relevant in the construction of my methodological framework. I have paid particular attention to oral discourse and genre, and in particular to the three genres which are object of this study: conference presentations, research dissemination talks and product pitches. I have argued that a multimodal approach is particularly appropriate for the study of persuasive oral genres. The following chapter will be devoted to persuasion as a characteristic trait of these three genres.

CHAPTER 2. PERSUASION: DIFFERENT REALISATIONS AND ITS RELATION TO ACADEMIC AND PROFESSIONAL DISCOURSE

In this chapter I deal with the concept of persuasion, which is the focus of my multimodal study. In particular, I am interested in the persuasive activity, or persuasive strategies that speakers use in oral presentations. In the first section of this chapter I draw attention to the fact that persuasion is also, and to a great extent, non-verbal. This justifies a multimodal approach to the study of persuasive strategies. In the second section I resume the concepts of intertextuality and interdiscursivity introduced in Chapter 1 to show how persuasion in one genre can draw from other related genres. The last subsection narrows down the scope to persuasion in the three oral genres which are the object of this study. As I mentioned in Chapter 1, I will use the term *text* in a broad sense referring not only to written documents, but including oral messages.

2.1 Verbal and non-verbal persuasion

The question of how a text becomes persuasive has an extremely long tradition since Aristotle's times (see for example *Ars Rhetorica*), and yet it is still unanswered to a great extent. Aristotle established three types of persuasive appeals to the audience: ethos, or appeal to the credibility and reliability of the communicator; pathos, or appeal to emotions; and logos, or appeal to rationality. They are still relevant and taken into account in modern studies on persuasion (Halmari and Virtanen 2005; Hyland 2005; Bamford 2007; Scotto di Carlo 2014). However, modern research points at the need for a more flexible and dynamic concept of persuasion. Instead of trying to find a magic recipe for a persuasive text, good speakers need to adapt their ideas to the situational context and in particular to the audience they are addressing. What is persuasive in one situation may not work in a different context or with a different audience (O'Keefe 2002; Perloff 2003; Hyland 2005). Because of this intrinsic dynamic aspect of persuasion, and because persuasive techniques may have different effects on different audiences, modern research on persuasion has focused on the intention of the communicator, so that a text could be considered persuasive if the speaker is trying to persuade (Jucker 1997; O'Keefe 2002). This is the trend I adopt in my study, and for this reason I focus on the the use of persuasive strategies by the speakers. The following pages will review previous research on persuasion to highlight some of the most salient aspects of persuasive texts.

Lakoff (1982) defines persuasion as an attempt to make the addressee change their mind or course of action by communicative means. He points out how these means can be either linguistic or non-linguistic (e.g. gestures, images, etc.). Perloff (2003) adds other features to the concept of persuasion: it necessarily constitutes a conscious attempt to influence and it occurs in an atmosphere of free choice (this is what differentiates persuasion from coercion). That is, it is ultimately the recipient and not the persuader who decides to accept the persuasive influence and alter their behaviour. From this it follows that the recipient plays a more active role in persuasion than it has generally being acknowledged. At the core of persuasion, therefore, there is a process of change. The object of this change is what most researchers agree to call *attitude*. Attitude is defined as a “learned global evaluation of an object (person, place or issue) that influences thought and action” (Perloff 2003: 39).

Therefore, one of the crucial questions persuasion research has tried to answer is how to craft a message that can trigger this attitude change. Early research suggested that attitude change was mediated through learning. Once the speaker has the attention of the recipient, the message will persuade if recipients understand it, accept it as true and make it part of their knowledge (Hovland, Janis and Kelley 1953). More recent approaches, however, argue in favour of a more active role of the recipient in the persuasive process. What is more important in persuasion is not understanding and recalling the message itself, but the way the recipient reacts to it, and the thoughts it stirs (Perloff and Brock 1980; Perloff 2003). One of the models that research on persuasion has produced to account for different ways in which addressees react to a persuasive message is the Elaboration Likelihood Model (hereafter ELM) (Petty and Cacioppo 1986; Petty and Wegener 1999; Perlof 2003). ELM claims that under certain circumstances, for example when the topic is important to us or when we are knowledgeable about it, we are more likely to process a message in detail, considering each argument, example, evidence, etc. This is called *elaboration based persuasion* or *central route*. In other situations, for example when we are not concerned or we do not have sufficient knowledge, we tend to rely on peripheral cues such as the speaker’s appearance. This is called *heuristic-based persuasion* or *peripheral route*. Consequently, under conditions that prompt high elaboration, powerful arguments are important, while in the case of heuristic-based persuasion other aspects become more relevant. Among these aspects are the credibility and likability of the speaker. A speaker gains credibility as he is perceived as expert and trustworthy (i.e. able and willing to say and share the truth). To this aim speakers often resort to evidence, reliable sources and credentials. A high speed of speech and a powerful speech, avoiding hesitations or hedges, can also enhance credibility (Perloff 2003). Humour, on the other hand, makes the speaker sound more trustworthy. Likability can also enhance heuristic-based persuasion (for this reason models or successful sportsmen usually appear in advertisements). It prompts positive feelings that can then be transferred to a product being promoted (Perloff 2003). In fact, emotions considerably affect persuasion, either by actual arousal or invocation (O'Keefe 2002). References to how good a product can make you feel are frequent in all genres that promote a product, and are sometimes more effective than factual objective information.

Petty et al. (2002) draw on ELM to develop their *self-validation* hypothesis. Their research focuses on how the way individuals reflect on the thoughts provoked by a message (and in particular their confidence in them) can affect persuasion. They conclude that if a message prompts positive thoughts, people who have confidence in these thoughts are more likely to be persuaded than people with low confidence. The opposite happens with negative thoughts: low confidence will result in persuasion while high confidence will not prompt persuasion. This effect is greater under high elaboration conditions. This means that when individuals under high elaboration conditions have doubts about the thoughts that the message has stirred in them, the quality of arguments loses efficacy.

Taking a more sociological approach, Cialdini (1984, 2001, 2006) has postulated that persuasion works by appealing to a limited set of *human drives and needs*. According to him persuasion is governed by six basic principles that can be summarized as follows:

i) Reciprocity: we feel an obligation to repay favours;
ii) Consistency and commitment: we feel a need for personal alignment, so we like to be consistent;
iii) Social proof: we look for support in what others do;
iv) Liking: we are more positive to the opinions and requests of people we like;
v) Authority: we tend to believe in the opinions of experts;
vi) Scarcity: something exclusive is more attractive.

Although I do not adopt a psychological or cognitive approach to persuasion, I do believe that these human drives and needs lie behind many rhetorical decisions that speakers make in their presentations.

Moving away from the figures of the persuader and persuadee, there are features of the message itself that have received considerable attention in the literature about persuasion.

Following ELM, they are expected to work particularly well under high elaboration conditions (i.e. when the addressee is carefully considering the message), although they occasionally contribute in a more heuristic way (i.e. promoting credibility, or making arguments more salient). Among these we can mention:

i) *Evaluation*: Evaluation can be very persuasive, because when speakers evaluate something they are implicitly inviting the listener to accept their point of view (Bamford 2007), therefore, evaluation is a frequent device in persuasive discourse. Bhatia (2005) also points out how positive evaluations are used in advertising genres to differentiate the product they advertise. Pomerantz (1986) notes how extreme case formulations are frequently used to legitimise claims when speakers expect possible counterclaims. Interestingly enough, these claims are commonly accompanied by head shakes, which seem to deny in advance a potential counter-argument (Kendon 2002). Extreme case formulations also provide credibility. For example something that happens "every time" cannot be easily disregarded, and something "everybody does" simply seems the right thing to do and is not easily questioned. On the other hand, evaluation can be especially effective when its source is hidden or attributed to people in general. The information is thus presented as common ground, and not as the particular opinion of an individual with whom the addressee may disagree. This positions both the addressee and the persuader within the same opinion group, creating consensus. In fact, evaluation can depend on shared background knowledge and become very implicit (Hunston 2000).

ii) *Framing or orientation* (Goffman 1979; Pomerantz 1986; Levin, Schneider and Gaerth 1998) is a powerful persuasive device, and it refers to the way information is presented so as to guide a particular interpretation of the content. This accounts for the different effects created by "30% mortality rate" vs. "70% survival rate" or "provide benefits" vs. "prevent losses". Perloff (2003) claims that positively framed messages are more effective when preventive actions are promoted (because they highlight the benefit of the preventive action), but negatively framed messages work better in other circumstances, particularly when the threat is perceived as real and likely and when the preventive action is perceived as feasible (because they highlight what addressees will lose if they do not change their mind).

iii) *Examples vs. statistics*: A similar difference can be spotted in the use of examples vs. statistical data. Lischinski (2008) points out that the use of examples can contribute to the credibility of the speaker, showing that some practical knowledge is shared with the addressee, and that the speaker is knowledgeable enough so as to apply theories to real situations. When examples are provided instead of raw data, the information is perceived as more likely, desirable or important (Zillman, Perkins and Sundar 1992). Even when examples contradict data, the addressee tends to rely on the example and disregard abstract data, because the latter are more complex to process and attract less attention. On the other hand, generalizing from an example is an easier process. Even if statistics represent a larger number of cases, examples representing a single case are more persuasive (Brosius and Bathelt 1994). In addition, it is perceived as a neutral and active process of opinion-building on the part of the addressee, who does not feel any type of imposition (Goffman 1979). Adding several examples together also achieves a sense of normality or typicality (Brosius and Bathelt 1994, Lischinski 2008), which may not in fact be supported by statistics. Due to their apparent reliability and neutral footing (Goffman 1979), examples are a very powerful persuasive device.

iv) *Language intensity*: the use of metaphors, strong, vivid language and emotionally charged words can also contribute to a persuasive message. They raise interest, and make arguments stand out (Sopory and Dillard 2002; Perloff 2003). As Perloff (2003) points out, intense language can work particularly well when the speaker enjoys credibility or when the audience does not have well developed attitudes towards the topic dealt with.

v) *Stylistic devices* such as three-part lists, symmetric structures, contrasting pairs or lexical creativity (Lakoff 1982, Bowring et. al 2013) can help make the text more memorable and in turn more persuasive.

vi) The use of vocabulary from the semantic field of *novelty* can also highlight the innovative character of what is presented in order to boost persuasion (Lakoff 1982, Bamford 2007, 2008).

vii) *Direct addresses* to the audience (Bamford 2007) can also become persuasive when they are used to prompt a particular response. This is usually the case with rhetorical questions and question tags, which anticipate and assume a preferred response from the audience. On the other hand, direct addresses to the audience contribute to rapport, which is also beneficial for persuasion.

As O’Keefe (2002) and Perloff (2003) note, one of the challenges in persuasion research is the fact that it is extremely complicated to isolate variables and assign a particular effect to them. It is not always easy to attribute the final effect on the addressee to a specific variable. For example a text may be persuasive due to its framing of information or its use of positive evaluation, but the addressee may also be influenced by the perceived credibility of the speaker, and it is difficult to tell which of the variables is more influential. On the other hand, one variable can have different functions. The same variable under certain circumstances can act as a peripheral cue, while in other circumstances it may influence high elaboration processing. This is the case for example of speaker's credibility or the use of evidence. Because of this, persuasion needs to remain flexible, and good persuaders need to learn to adapt to the communicative situation.

The ELM approach to persuasion outlined above, therefore, is far from being applicable as a magic recipe for a persuasive text. Nonetheless, it explains why certain persuasive techniques are usually effective in certain circumstances and with certain audiences. For example it can explain why persuasion becomes more effective the more unexpected and subtle it is. Forewarning is a persuasion killer, because it activates the recipients mind and stirs potential counterarguments in advance (Petty and Cacciopo 1977, Perloff 2003), or in other words, activates elaboration-based persuasion. For the same reason, distractions are sometimes used in persuasion as an inhibition of counterarguments: they prevent the recipient from reflecting on the message and prompt a heuristic-based persuasion. This idea is shared by Cook (2001: 229), who in his study of advertising discourse claims that "if an ad departs from expectation, it will attract attention and / or convince (…) On the other hand, ads do not need to explicitly identify themselves as such or reveal their persuasive intention, because this is normally inferred, mainly by their position within other discourse in which they appear ".

Halmari and Virtanen (2005) also argue that persuasion is more effective when it is implicit. This accounts for the tendency to innovation which many inherently persuasive genres show. Advertisements are constantly striving to surprise the audience, and oral presentations also tend to become increasingly innovative. New technologies have greatly contributed to this, opening a wide range of possibilities. This tendency to novelty can be realised through a flouting of Grice conversational maxims (see Section 1.1.2 in Chapter 1). Faced with an unusual message, the addressee's attention is deeply engaged as they are prompted to find a suitable interpretation, and this process leaves a longer lasting impression (Lakoff 1982).

The role of culture in persuasion has been considered by Fitch (2003). According to her, studies on persuasion can be divided into three main trends: *social sciences*, with special emphasis on the influence on attitude as a process; *rhetoric*, with focus on the audience and how to adapt messages to them; and finally *mass media*, with a particular interest on understanding the effects of messages on audiences. Neither of them, she claims, considers that persuasion is a culturally-determined phenomenon. It is generally assumed that whenever there is choice there is room for persuasion (when there is no choice, we are talking about coercion). However, culture narrows down the scope, delimiting what you can be persuaded of (i.e. persuadables) and how much persuasive effort is required. For example, you cannot be convinced to do something your culture generally approves of, or of something you are expected to do anyway (i.e. taken for granted), because these are non-persuadables. Likewise, your culture determines how much effort is required to convince you of other aspects that do constitute persuadables. The relevance of this for a persuasive presentation is that speakers need to take into account potential cultural differences with their audiences as they design their message.

Turning now to non-verbal realisations of persuasion, the fact that persuasive techniques can also be non-linguistic has long been acknowledged (Woodall and Burgeon 1981; Lakoff 1982; Morrison and Vogel 1988; Messaris 1997; Perloff 2003; Bamford 2007). Several studies focus on the effect of different modes on persuasion. Chaiken and Eagly (1976) studied the effect of mode on message comprehension and persuasive effect. They concluded that difficult messages are better understood and have greater persuasive effects when they are written. In the case of easy messages comprehension was not affected by mode, but persuasion was greater in videotaped messages. Chaiken and Eagly (1983) focus on the effect of mode on the perception of the speaker and final persuasive effect. They found out that likable speakers were more persuasive in audio and video tape, while unlikable communicators were more persuasive in writing, and concluded that the mode used can foreground the figure of the speaker making it more salient and influencing the final persuasive result. Sparks, Areni and Cox (1998) also studied the effect of mode on persuasion, concluding that writing prompts a central route to persuasion, while audio and audio-visual modes seem to encourage a more peripheral, heuristic-based processing. In addition, they found out that the influence of language style (powerful vs. powerless) on persuasion was affected by mode: in audio messages the effect of language style was greater. It follows from this discussion that speakers need to be particularly conscious of the different effects that audio-visual modes (e.g. intonation, gestures and head movements) can have on their audiences, because they are particularly salient in oral genres such as presentations.

Among these audio-visual modes, the role of visuals has received particular attention. Morrison and Vogel (1988) have shown how the use of visuals in presentations improves comprehension and retention, which would be helpful in central routes to persuasion, and at the same time improves the perception of the presenter, which in turn results in higher attention and yielding (peripheral route to persuasion). Messaris (1997) studies the role of images in advertising. He points out three main characteristics of persuasive images:

i) Iconicity: they can simulate the visual appearance of reality and elicit emotions and attitudes to real people, places and objects;

ii) Indexicality: they can be used as evidence supporting claims, especially photographs, which are generally considered trustworthy and reliable;

iii) Lack of propositional syntax: the link between images is implicit. There is no explicit and well-established way of expressing analogy, contrast, causality, etc. as there is with language. Because they are more implicit as a semiotic system, advertisers use them to convey messages that they are reluctant to overtly express with words.

Although in this study I do not focus on visuals as a mode, I acknowledge their importance and interconnectedness with other modes, especially in the case of pointing to projections of visual materials. This is a practice that can be used with a clear persuasive intent, as I will discuss in chapters 6 and 7.

The role of intonation and gestures in persuasion has also been acknowledged. Woodall and Burgeon (1981) showed that kinesic cues synchronised with the verbal message resulted in a higher recall and more effective persuasion, demonstrating how non-verbal synchrony affects processing and acceptance of a message. Burgoon, Birk and Pfau (1990) also examined the effect of non-verbal behaviour on credibility and persuasion. They found out that aspects such as facial expressiveness, kinesic and proxemic immediacy, relaxed posture, fluency and pitch variety contributed to the perception of the speaker as a competent and sociable person and resulted in increased persuasiveness. Chebat, Chebat and Vaninsky (1996) investigated the effect of intonation and intensity of voice in persuasion, concluding that both affect the credibility of the speaker, but in different ways. The effect of voice intensity is greater under conditions of low involvement (when the topic is not personally relevant for the addressee), while intonation has a stronger effect under conditions of high involvement.

Political discourse has provided a certain amount of research on non-verbal persuasion. For example Atkinson (1984) provides a study of body language in political communication, and shows how speakers elicit applause from their audiences not only through verbal devices such as three part lists and contrastive pairs, but also through gaze. Bull (1986) examines the use of hand gestures in political speeches and their relationship with the style and content of the speech, concluding that hand gesture is closely related to intonation, particularly to vocal stress and tone group boundaries.

Rosenberg and Hirschberg (2005) analysed the influence of prosody on the perceptions of charisma in political speeches. A charismatic person is one who is able to lead without the need of being dominant. They found that voices that were judged to be more charismatic had a higher and more variable pitch pattern; higher loudness and a faster pace. A charismatic speaker is not in itself a guarantee of persuasion, but it can be considered to pave the way for it making audiences more willing to listen and retain the message. Later, Niebuhr et al. (2016) applied a similar study to Steve Jobs' product presentations, obtaining congruent results and showing how Jobs also adjusted his level of "charisma" to his audience, depending on whether he was addressing customers or investors. A more detailed discussion of the results of this study is offered in Chapter 3.

Poggi and Pelachaud (2008) undertake the study of persuasive gestures in political discourse, concluding that gestures are not inherently persuasive, but some of them express meanings that are relevant to persuasion. Among these they find:

i) Importance, which enhances the perception of the product or idea. Gestures such as Kendon's finger bunch or beats can express this meaning (see Chapter 3).

ii) Certainty and exactitude, which enhance credibility of the speaker. Kendon's ring (see Chapter 3) would be an example.

iii) Positive evaluation also enhances the perception of the product or idea. For example a finger bunch gesture on the lips to express "delicious" or "beautiful".

iv) Benevolence transmits speaker's goodwill, for example putting a hand on your chest or showing your palms to mean "honesty".

v) Strong emotions, such as pride, satisfaction or impatience which can be expressed by rubbing hands for example, can also trigger emotions in the audience and contribute to persuasion.

The use of gestures that I have identified in this study is to a great extent congruent with the description offered by Poggi and Pelachaud (2008), since speakers in my corpus also tend to use gestures to emphasize important parts, convey a sense of accuracy, evaluate or stir emotions in the audience.

This study by Poggi and Pelachaud (2008) makes clear that gestures make a great contribution to persuasion, and need to be taken into account even if their interpretation is not always straightforward, especially when they are studied in isolation. In particular, gestures have been found to play a role in *rapport*. Tickle and Rosenthal (1990) define rapport as a relationship of sympathy and mutual understanding. Rapport plays a crucial role in persuasion, because it is an important means by which speakers adapt to their audience and achieve the required flexibility that is required for a persuasive message. As these authors point out, rapport can be expressed non-verbally, something which I have also found in my corpus. They identify three essential components of rapport: mutual attentiveness, positivity and coordination. Mutual attentiveness refers to mutual attention prompted by interest, which can be expressed through spatial configurations and bodily postures that signal accessibility. Positivity is positive behaviour prompted by feelings of friendliness and warmth, which can be materialised through smiling and nodding that signal liking and approval. The third component, coordination, refers to a coordinated interaction prompted by feelings of balance and harmony which can take the form of postural mirroring and interactional synchrony.

Concerning head nods, Buñol and Petty (2003) have shown how head shakes and nods affect our own confidence in our thoughts in response to a message (i.e. self-validation hypothesis). Under high elaboration conditions, nodding reinforces our attitude towards a message, while shaking downplays it. With a combination of strong arguments and positive thoughts persuasion is increased by nodding. On the other hand, with a combination of weak arguments and negative thoughts persuasion is increased by head shakes (because they make you doubt your own negative response). The reason why this tendency occurs under high elaboration conditions is because the conditions that promote high level of thinking also promote high level of meta-reflection about this thinking. They conclude that "manipulation of head movements affects confidence in the validity of one's thoughts, not confidence in the validity of the persuasive message per se" (Buñol and Petty 2003: 1135).

A particularly interesting study for the present research is the one by Kast (2008), which provides a rhetorical analysis of one example of a product pitch: the presentation of Apple's iPhone in 2007. Kast shows how Steve Jobs builds a persuasive message by appealing to ethos, pathos and logos. Regarding ethos, Kast points out how Jobs performs three roles, or projects three different identities: the descriptive informer, the innovator and the enthusiastic customer. In relation to pathos, Jobs addresses values that he assumes the audience to share (e.g. taste for diversity challenging environments or self-expression), and he uses materials like photos, movie and TV clips to prompt emotions in alignment with the content of the presentation (excitement, tension, curiosity, etc.). In addition, humour is used to create a relaxed atmosphere. At the linguistic level, appeals to emotions are achieved through the use of rhetorical figures of speech. Kast finds that the verbal message is supported by non-verbal and paraverbal communication, as well as by visual aids (the catchy projected slogans can be considered the media counterpart of the linguistic figures of speech). Finally, as regards logos, Kast notes how the presentation is structured in parts, each part containing several subunits which in turn have a short introduction and a summary. These are usually metadiscursively marked (e.g. “the first thing I’d like to do”). Argumentations take the form of causal schemes, example schemes, comparing schemes, and opposition schemes, but the most remarkable feature is that they are very often implicit or even replaced by multimedia aspects (e.g. an image in a slide can make explicit the conclusion the speaker wants to make, or prove a claim). What this study reveals is that persuasion in the presentation of the iPhone is achieved by a complex interplay of semiotic resources, which calls for a multimodal approach.

2.2 Persuasion across genres

Genre can also play a role in creating novelty and implicitness for persuasive effects. For example, persuasion can spring from a reversal of expectations regarding genre. Likewise, generic features can be used to mislead the audience into interpreting a text as belonging to a different genre and treating information differently as a consequence. For example, advertising can be presented as objective information (Halmari and Virtanen 2005). Persuasion, therefore, leads to a great flexibility, generic change and generic mixing. As Bamford (2007: 141) explains:

> Persuasion creates generic change and evolution since the speaker has to craft the message to make it effective as a persuasive tool, yet at the same time must try to conceal the act of persuasion from the audience. This persistent tension between the two discursive necessities leads to ever new ways of presenting the message.

Other authors have noted generic change derived from persuasion. Bhatia (2005, 2010, 2012) points out that there is a general trend to marketization of academic and professional genres that is building a colony of promotional genres. Some of these genres are borrowing rhetorical processes and generic values from advertising genres such as 'endorsement or testimonials', or 'offer incentives'. Halmari and Virtanen (2005) describe this generic change and its relation to persuasion in terms of *intertextual gap* (deviation from a generic prototype) and *interdiscursive gap* (distance between different genres). As the intertextual gap is increased, the text may acquire characteristics of other genres, and the interdiscursive gap is minimised. As a consequence persuasion becomes more implicit. On the other hand, this leads to the appearance of hybrid genres, and eventually to the appearance of new genres. Bhatia (2010) discusses three examples of interdiscursivity: the genre of the corporate annual report, that combines accounting discourse and public relations discourse; the influence of litigation discourse in arbitration practices; and the culture of marketing and advertising leaking into fundraising discursive practices. As I have argued in Chapter 1, interdiscursive and intertextual relations are the rationale for my contrastive analysis of the three oral genres included in my study. I hypothesize that the three genres can be interdiscursively related as they draw from each other to fulfil their communicative aims, and in particular to achieve the right interpersonal relationship between speaker and audience.

2.3 Persuasion in conference presentations, research dissemination talks and product pitches

The importance of persuasion in written academic discourse has been extensively studied by Hyland (1997, 2005, 2009). He stresses the importance of being aware of the audience. In order to get claims accepted, the audience needs to perceive them as 1) holding a plausible relation with reality (i.e. adequate) and 2) interpersonally acceptable, "with the writer adopting a professionally acceptable persona and a tenor consistent with the norms of the disciplinary community" (Hyland 1997: 440).

According to Hyland, authors need to find the right balance between tentativeness and assertiveness, and they have to be able to project a suitable relationship with their data, arguments and audience. It follows from this that, in order to achieve adequate and acceptable arguments, both textual and interpersonal metadiscourse is needed. Textual metadiscourse signals the writer's intention guiding the interpretation of propositional meanings. Interpersonal metadiscourse shows the author's perspective towards both the propositional content and the addressees. In short, academic writing requires constant anticipation of potential counterclaims, and metadiscourse is the writer's main tool to achieve this.

Later, Fuertes Oliveira et al. (2001) and Dafouz-Milne (2008) have applied this notion of textual and interpersonal metadiscourse to study persuasion in advertising and journalistic discourse, respectively. The former finds that textual metadiscourse guides an interpretation of slogans and headlines that is in accordance with the intended audience's epistemological understanding and genre expectations, while interpersonal metadiscourse makes it possible to achieve a balance between informing and persuading. The latter concludes that "the key to an effectively persuasive text is the artful combination of weakening expressions (i.e. hedges) and strengthening ones (i.e. certainty markers and/or attitudinal markers) with the final intention of producing a discourse that is neither too assertive nor too vague" (Dafouz-Milne 2008: 108). They also show the importance of well-structured texts and solid argumentations as key factors leading to persuasion in journalistic discourse.

Hyland (2005) has also identified specific features in written academic papers that are meant to boost persuasion: *citation*, as a way to contextualise research within a particular discipline; *interaction and engagement* of the audience through-audience oriented features (e.g. inclusive pronouns, questions and directives) and *self-mention* that presents the author as informed and reliable. Particularly interesting for its application to oral presentations is the idea of engaging with an audience. It implies recognition of the "need to adequately meet readers' expectations of inclusion and disciplinary solidarity" (Hyland 2006: 33), addressing readers as participants in an argument. It also implies "positioning the audience.(...) the writer pulls readers into the discourse at critical points, predicting possible objections and guiding them to particular interpretations with questions, directives (mainly obligation modals and imperatives) and explicit references to shared knowledge" (Hyland 2006: 33).

These studies deal with written academic discourse, but the principles of how persuasion is achieved and some of the persuasive features of written papers seem to hold for oral academic discourse as well, in particular because the presence of an audience that has been evidently taken into account in written academic discourse is much more immediate and conspicuous in oral presentations, as discussed in Chapter 1.

Rowley-Jolivet (1999) and Rowley-Jolivet and Carter-Thomas (2003, 2005) have claimed that conference presentations are inherently persuasive. This notion that conference presentations are persuasive is shared by numerous authors. Räisänen (1999, 2002), for example, suggests that the function of conference presentations is to publicize, critically discuss and ratify research. Therefore, because their primary aim is to convince an audience of the validity, originality and usefulness of a piece of research, conference presentations can be considered to be a persuasive genre. Swales and Burke (2003), Wullf et al. (2009) or Querol-Julián (2011) describe the genre of conference presentations as belonging to the *contingent* repertoire of academic discourse as opposed to the *empiricist* repertoire, which includes genres that are more objective and impersonal. As Ruiz-Garrido (2015) points out, this persuasive communicative goal has important implications for the genre, because it requires the development of an interpersonal relationship with the audience who is physically present. Fernandez-Polo (2014) summarizes these views when he says that

> [m]odern scientific communication is strongly argumentative and persuasive. In the highly competitive context of present-day science, stressing and persuading others of the value of our claims has become crucial, and consequently the building up in the text of favourable personal relationships with the audience (Fernandez-Polo 2014: 166).

As he points out, two of the interpersonal skills that speakers use in conferences to develop an appropriate interpersonal relationship are self-mention and humour, and in his study he finds differences between NSs and NNs in both aspects.

An example of this trend towards persuasion in oral academic discourse is also shown by Thompson (2002), who claims that speakers in conference presentations resort to what she calls "involvement strategies" (a term she borrows from Tannen 1989), such as *storytelling* and *constructed dialogues,* to persuade an audience to accept their ideas. She points out how storytelling has the effect of engaging the listener. She demonstrates how narratives in past tense during conference presentations relate difficulties and failures during the research process. Sometimes these narratives include self-deprecating humour and are spiced up with direct speech, colloquialisms and idiomatic expressions. In contrast, in the written paper these stories become plain present tense results. The use of narratives is a persuasive appeal that helps create rapport: it shows the person behind the study and provides credibility. Taken a step further, this vivid account of research practices can even take the form of enactments that are performed multimodally. They can raise interest, and, as mentioned before, also contribute to rapport and credibility. For example, Ochs et al. (1994) argue that dynamic grammar and gesturing help physicists create a visual representation of experiments so that they are perceived as credible. Morton (2009) also found similar attempts to increase vividness in student architecture presentations through gesture, language, visuals and physical objects.

The role of persuasion in research dissemination has been dealt with in recent literature, with particular emphasis on TED talks. Although the presentations I analyse in my study are slightly different form the web-based, mass-addressed TED talks that this literature refers to, they are indeed a subtype of them, and many of the features, in particular in what regards how persuasion is achieved, still apply to my corpus. For example, Scotto di Carlo (2014) has performed a case study that shows how persuasive appeals through ethos, pathos and logos are remarkably different in a TED talk and a CP dealing with the same topic and delivered by the same speaker. First, the TED talk appeals much more to pathos through the use of stories, videos, pictures, graphs and other visual media (note how these resources include a variety of semiotic modes). The speaker also uses evaluative adjectives and engages the audience using direct addresses and inclusive pronouns. In addition, her position on stage and delivery style create proximity with the audience and enhance the sharing of emotions. In contrast, the CP projects a greater distance with the audience. Second, the way the speaker appeals to ethos in the TED talk differs from the CP: In the TED talk *trustworthiness* (i.e. ability to be consistent with the message) and *similarity* (i.e. ability to make the audience identify with you) (Dlugan 2010) play a more important role than *authority* or *reputation*, the focus being on establishing *alignment* (Goffman 1981). In contrast, in the CP authority derived from belonging to a scientific community and reputation derived from expertise seem to play a more important role. Third, concerning logos, in the TED talk the speaker uses examples, metaphors and visuals to enhance the comprehension of the message, while in the CP the use of specialised terminology limits the type of audience that will be able to follow the presentation. She concludes that both presentations have very different aims: while the CP tries to be informative, the TED talk is also proactive, in that it not only raises awareness of a situation, but creates a need to solve it.

Scotto di Carlo (2014b) looks more closely into the relationship established between TED speakers and their audiences. The features of TED talks that have been discussed in the previous paragraph serve to bring the audience closer and overcome the expert vs. layman distinction. While in CP speakers tend to create proximity of membership by emphasizing their legitimate role as members of the scientific community, in TED talks speakers create proximity of commitment by personal involvement in the topic. Scotto di Carlo (2014c) zooms in on the use of ethos in TED talks, specifically in the role of credibility. She follows Gili's (2013) proposal of the three reasons why speakers are recognised as credible: *knowledge*, *affectivity* and *shared values*. She points out that in TED talks transmission of knowledge in itself is not sufficient. The speakers need to convey dynamism, immediacy, energy and emotional involvement. *Affectivity* becomes especially important, and TED speakers need to create a direct emotional connection with their audiences, because their aim is not only trying to inform, but prompt a change in society: they are a call to action. *Sharing values* also plays an important role, and, unlike in CPs these are not the values of a professional community, but general social values that the audience also shares. Thus, credibility in TED talks is built around shared values and affection rather than knowledge and expertise. Some credibility enhancing strategies that she identifies in TED talks are: personal stories or anecdotes that demonstrate deep personal involvement with the topic; audience engagement (familiar and accessible language, inclusive pronouns, direct addresses) and use of material (i.e. visuals, quotes, examples) that resonates with the audience. A particular way of engaging the audience is making them feel as members of a community and participants in the "TED experience". Speakers usually achieve this by explicitly referring to the event and positioning themselves as equals with the audience, as "one of them". Interestingly, a similar trend can be spotted in CPs, but with a focus shift: emphasis is not on co-participating but on being accepted as legitimate members.

Finally, Scotto di Carlo (2015) focuses on the use of *evaluation* in TED talks. An important difference between CPs and TED talks, she finds, is that while the former show a tendency to cautiousness, hedging and objectivity, the latter tend to emphasize originality and uniqueness of findings by boosting claims and showing subjective personal engagement. In a way, TED talks are designed to transmit enthusiasm and excitement. A particular way in which they do this is through *evaluation*, which enhances credibility and strengthens the speaker-audience connection. As Scotto di Carlo (2015) notes, this connection is to a great extent expressed through non-linguistic terms (e.g. intonation, facial expressions, visuals), but words, and in particular evaluative adjectives, greatly contribute to it. For example, adjectives describing objects are used to recreate the experience of using that object, making the audience feel the same emotions as the speakers. There is also a higher frequency of adjectives expressing personal opinions (usually attributing positive aesthetic or emotive properties) to convey the affective response of the researcher. As in the case of CPs, TED talks also try to show how the results presented will add to the scientific community, but they do so in a way that emphasizes the pragmatic consequences in everyday life, for ordinary people.

Turning now to product pitches, persuasion is a salient feature of many business genres, among which product pitches are included. This topic has generated some research, such as Swales and Rogers' (1995) study of mission statements, an organisational corporate genre, and how their rhetorical design makes them persuasive, despite differences in their specific communicative purposes. Hyland (1998) also provides a study of persuasion in CEOs letters included in annual reports, focusing on rhetorical strategies.

Regarding persuasion in business presentations in particular, Bamford (2007) performed a study of evaluation and persuasive discourse in business presentations which is particularly relevant to the present research. Drawing on Aristotle's three main aspects of persuasion ethos pathos and logos, she claims that persuasion is achieved in this genre by appeals to reason (complex argumentations), appeals to the credibility of the speaker (achieved linguistically by statements of belief, positive evaluations, emphatics and hyperboles) and emotional appeals (achieved linguistically by means of personal testimonies and direct addresses to the audience). Emotional appeals help create rapport with the audience, and in turn, a good interpersonal relationship is key to make the audience willing to change their attitudes. These persuasive techniques that Bamford points out have been also considered by other researchers and fit into the ELM (O'Keefe 2002; Perloff 2003). According to the ELM, complex argumentations would promote persuasion under conditions of high elaboration, and appeals to credibility and emotions work as peripheral cues. Interestingly, Bamford also acknowledges that aspects such as body language and intonation play a role in persuasion and that they can reveal useful information about the communicative intention of the speaker. However, she does not undertake a systematic study of these features.

Other interesting studies about persuasion in oral business presentations are provided by Bowker (2006, 2007). Bowker focuses on linguistic features, and notes how certain narrative techniques are employed to make oral presentations more persuasive. One of these techniques is the change from indirect to direct mode, or impersonating instead of referring to a third person. The information presented in this way is loaded with affected force, and perceived as particularly relevant and real. A second technique is the use of conceptual clarification questions, or rhetorical questions used to guide interpretation, e.g. elaborating on abstract concepts, rephrasing or exemplifying them in understandable ways. Third, extension questions can also be used to prompt the audience to relate ideas to their own experience, and check their relevance. They serve as a counterbalance to the unidirectional and monologic expert authoritative voice.

It will be noted that these narrative techniques aim basically at easing interpretation and presenting information as relevant for the addressee, which is expected to enhance persuasion under conditions of high elaboration.

Also worth mentioning is May's (2014) proposal for persuasive business presentations. He recommends organising business presentations according to a problem-solution pattern, according to the following scheme:

i) Opening of the presentation: gain attention, clarify purpose and provide an overview;
ii) Body: discuss the problem, present a solution discuss alternative solutions and stress the benefits of the suggested solution;
iii) Closing: summarize, call to action and close memorably.

This problem solution structure is actually a basic pattern that I have been able to identify in the product pitches included in my study.

As possible attention-getters for the openings he mentions rhetorical questions, anecdotes, startling statistics or quotes. For the closing, he recommends showing that the presentation is coming to a close, not only through words (i.e. through metadiscourse), but also changing the delivery style, so that the audience resumes attention. As possible closing techniques he suggests relevant quotations, linking back to the opening or offering a vision story, some of which I have actually found in the product pitches that I have studied, but interestingly, also in the research dissemination talks and in the conference presentations.

Regarding delivery, he suggests focusing on four aspects. The first one is keeping eye contact with the audience, being particularly careful to engage randomly with all members. The second is body language. He suggests moving hands naturally, adopting a comfortable posture and not hiding behind a podium. Next he recommends speakers to project their voice, using a conversational tone, pausing and avoiding fillers. The fourth aspect is enthusiasm, that is, showing that you are happy to be there, and that you are enthusiastic about the topic.

Finally, concerning preparation, he recommends starting with a drafted script and reading it aloud with its accompanying visual support. Then, he suggests repeating the same procedure but in front of an audience, and focusing on content and design. The next step is to reduce the script to key words and concepts. After that another rehearsal is recommended, this time making it as realistic as possible and focusing on delivery.

Although these recommendations are originally meant for business presentations, I have seen them being used (more or less successfully) in the other presentations in my corpus as well.

From the above discussion it can be concluded that persuasion is a defining characteristic of conference presentations, dissemination talks and product presentations. It determines its communicative efficiency and its appropriateness to the communicative situation and discursive community. The discussion provided in this section also shows that persuasion adopts a variety of forms, and therefore it can be better analysed from a multimodal perspective. The following chapter focuses on multimodality in academic and professional discourse, with particular attention to the modes that are the object of this study: words, gestures, intonation and head movements.

CHAPTER 3. MULTIMODALITY IN ACADEMIC AND BUSINESS DISCOURSE

The focus of this chapter is multimodality, and specifically, the use of non-verbal modes included in my analysis (intonation, gestures and head movements) in academic and business discourse. The first section of this chapter discusses the multimodal nature of oral academic and business discourse, with special focus on the three oral genres I have selected for analysis: conference presentations, research dissemination talks and product pitches. Section 2 focuses on words, intonation, gestures and head movements as semiotic modes. The third and last section zooms in on the use of intonation, gestures and head movements in the three genres under analysis.

3.1 Multimodality in academic and business discourse

The multimodal character of many genres in academic and business discourse has long been acknowledged in literature, especially in what concerns the genre of CPs, which seems to have been more extensively studied. In this section, I want to provide a literature review of multimodal studies that provide the foundations of my multimodal approach to persuasion in conference presentations, research dissemination talks and product pitches.

Dubois (1982) was among the first to draw attention to the multimodal nature of presentations, pointing out the increasing importance and abundance of visuals in medical conferences, as well as their role in helping to structure the presentation.

Later, Carter-Thomas and Rowley Jolivet (2003) have suggested that the study of conference presentations must be approached from three different perspectives:

i) A microscopic bottom-up analysis to identify recurrent linguistic features and relate these choices to the communicative context (but which fails to account for the rhetorical macrostructure of the text);
ii) A macroscopic top-down analysis (analysis of moves) to explore the rhetorical structure (but which fails to account for the variety of modes employed);
iii) A multimodal approach that allows them to identify three main modes in CPs: language, visual communication and gesture.

They claim that language is the main mode and the other two are used to reinforce language. However, they acknowledge that the role of visuals becomes more important in science conference presentations, because the affordances of this mode make it particularly suitable for easy processing during real time delivery.

In the process of applying these levels of analysis to the introductions of CP they found certain methodological problems. To begin with, in an oral genre the boundaries of the different sections are not as clear as in a written paper, so delimiting the introduction proved a difficult task. Then they found significant differences in the number and order of moves between the oral and the written genre. Finally, since there is no one-to-one correspondence between moves and their realizations, applying the categories coherently proved problematic. Analysing an oral text in terms of rhetorical moves is, therefore, more challenging than in the case of written texts because of the inherent flexibility of spoken genres. Speakers have practical goals in mind as they speak which emerge as discourse goes along, and as a consequence oral genres are more dynamic and flexible. However, a moves analysis can still provide a valuable insight into the communicative purpose of the text, which in turn justifies and explains the choices made at a more microscopic level.

The role of visuals pointed out by this study is also dealt with in Rowley-Jolivet (2002), where she claimed that visuals constitute a major resource for meaning-making in presentations. Moreover, because they provide immediate access to the data discussed, visuals serve to reinforce newness and immediacy of the presentation. Hyland (2009) further claimed that visuals fulfil the three metafunctions postulated in Systemic Functional Linguistics: they can provide information, help to structure the talk and appeal to the imagination of the audience. This line of research is followed up by Diani (2015), who, in a detailed study of the slides used in CPs as visual support, finds divergences with respect to the traditional moves structure of written research articles. This is explained as a consequence of the twofold purpose of slides in CPs: they serve as support for speakers (as a digital version of written notes), but also as a meaning-making resource for the audience.

From this review, we see that visuals seem to have enjoyed a privileged treatment in multimodal analyses of academic discourse and conference presentations in particular. However, some studies have also looked at the role of kinesic or paralinguistic features in academic discourse. With reference to classroom settings, English (1985) discussed the role of kinesics in academic lectures to facilitate understanding. Along the same line, Pozzer-Ardenghi and Roth (2005) pointed out how gestures in lectures can reduce the ambiguity of visuals, while Crawford-Camiciottoli (2007, 2015) has included nonverbal behaviour in her description of business and humanities lectures, revealing how stress, gaze or gestures were used to facilitate and reinforce comprehension. Zhang (2015) shows how facial expression, gaze and posture play significant roles in the expression of disagreement in plenary lectures and points out the pedagogical implication of this finding stressing the need to provide students with appropriate resources to express disagreement in academic contexts. Plenary addresses are also the focus of Ruiz-Madrid and Fortanet-Gómez (2015), where they offer a multimodal contrastive study of two plenary lectures: one delivered in Spanish and another delivered in English. Their study shows how kinesic and paralinguistic features help both speakers to engage interpersonally with their audiences during digressions, in particular using humour and modesty to evoke empathy. Interestingly, they find no particular differences attributable to native language in the use of non-linguistic resources.

Focusing on conference presentations, Fernandez Polo (2014) notes how speakers in CPs resort to gesture and facial expression (e.g. a smile) as a redressive action to soften the intrinsically face-threatening act of citing somebody who is physically present in the audience. A contrastive study of the use of intensifying adverbs in Spanish and English CPs is provided by Ruiz-Garrido (2015). He finds that in both languages, linguistic expressions of intensification were reinforced by kinesic and paralinguistic features such as stress, higher pitch, longer duration, hand and arm gestures, head movements and body movements (the latter more frequent among Spanish speakers). Morell (2014, 2015) finds that a skilful combination of modes makes presentations easier to understand through effective redundancy and can compensate for deficiencies in language proficiency (specially the visual modes). She also notes how more experienced presenters know how to keep eye contact and use gestures to interact with the audience without being distracting or blocking the view. She determines the effectiveness of a presentation from a SFL and multimodal perspective: a presentation is effective when the speaker expresses ideational, textual and interpersonal meaning using the four modes she identifies: verbal, written, non-verbal material and body language. The more aware presenters are of these modes, the more able they are to combine them effectively, foregrounding them at specific moments of their presentations and achieving effective redundancy. She also recommends using modes consecutively in case of complex propositions. She further argues that "it appears to be that having the ability to orchestrate semiotic resources or modes such as images, writing, layout, sound, gestures, speech and 3D objects (Kress 2010) may be more important than having a good command of the spoken language or verbal mode” (Morell 2014: 154). From a more pedagogical perspective, Busá (2010) used a multimodal approach to the teaching of English communication skills, and she found that when students were made aware of the complex interplay of semiotic resources, their oral presentation skills improved. She paid particular attention to prosody, body language, cultural awareness and genre awareness. She concludes that “skills such as eye contact, body language, style, understanding the audience, adapting to the audience, active and reflective listening, using formulas, conventions and discourse strategies as is appropriate in different linguistic situations and social interactions are as important to communication as language itself and should be integrated in classes focusing on spoken language” (Busá 2010: 63). The role of intonation and body language in oral presentations is a topic that will be resumed in section 3.3 in this chapter. They are not, however, the only modes along with visuals that play a role in the communicative event of the CP. Räisänen (2015) draws attention to the fact that material aspects such as the organization of space, light and technical props (i.e. microphones, pointers, etc.) are also integral parts of the communicative process and deeply affect meaning-making.

Among the studies dealing with business presentations, Bamford (2007, 2008) and Bowker (2006, 2007) are worth mentioning. All highlight the persuasive character of this genre, but only Bamford (2007) acknowledges that this persuasion can also be achieved by means of paralinguistic and kinesic features (although her analysis does not focus on them):

> Apparently it is easier to hide the truth or to guild the pill with a written account than an oral. In fact investors are said to take courses on the interpretation of body language and intonation to be able to understand what these can reveal, not to speak of analysing the actual words used in the presentation to see what speakers are trying to keep hidden. (Bamford 2007: 137)

Another relevant study already mentioned in Chapter 2 is Kast (2008) on the iPhone presentation. He notes, for example how appeals to emotions in this presentation are achieved through the use of rhetorical figures of speech which are supported by non-verbal and paraverbal communication, as well as by visual aids (the catchy projected slogans can be considered the media counterpart of the linguistic figures of speech). He also notes how argumentations are very often implicit or even replaced by multimedia aspects (e.g. an image in a slide can make explicit the conclusion the speaker wants to make, or prove a claim). Persuasion in the presentation of the iPhone is achieved, therefore, by a complex interplay of semiotic resources, or in other words, by a carefully orchestrated multimodal ensemble.

Likewise, Pan, Wong and Scollon (2002) note how frequently in recommendations on how to make a good presentation (e.g. May 2014) emphasis is given to non-verbal communication (e.g. maintaining eye-contact, controlling body movement and gesture, etc.).

From a pedagogical perspective, Palmer-Silveira (2015) supports these claims by showing how students can improve communication skills in product presentations when they learn how to exploit proxemics, head movements, gestures and eye contact. For example, he shows how a good use of space and body posture can project confidence, gently nodding while smiling can help reinforce positive aspects of the product while overusing hand gestures can be distracting. With a similar pedagogical aim, Tamarit and Skorczynska (2014) have focused on the multimodal realisation of argumentative strategies in oral business discourse. They show how the repetitive use of certain gestures and metaphoric images can increase the effectiveness of a specific type of business presentation: the project management training session. They point out, for example how gestures can have a confirming or emphatic effect in assertive moves, or how images can be used to attract attention and facilitate understanding, but also to build rapport with the audience. The results of their analysis have been used to design a Master's course module that aims at teaching students how to use these multimodal argumentative strategies.

The role of multimodality in research dissemination has only been recently dealt with. Caliendo (2014) and Caliendo and Compagnone (2014) and Compagnone (2015) all note how speakers in TED talks resort to a variety of semiotic resources. Some studies that probe into the use of intonation in these talks will be discussed in section 3.3.

3.2 Words, intonation, gestures and head movements as semiotic modes

The present research is based on the assumption that there is much more to communication than just words. Meaning is conveyed through different modes, and words constitute one of these modes, but they are by no means the only one. In adopting a multimodal approach to oral discourse, two aspects reveal themselves as particularly salient and relevant because of their contribution in the meaning-making process. These are *paralanguage* and *kinesics*. In this section I will undertake a review of literature that will highlight their relevance as meaning components, with focus on three specific aspects: intonation, gestures and head movements. The selection of these modes in particular is based on the fact that their presence is virtually unavoidable in oral speech and on their important contribution to the communicative purpose of the genres under discussion in this dissertation, which has been frequently underestimated. The rationale for my selection of modes is further discussed in Chapter 4.

The concept of communication being a complex including more than words was already set forth in 1953 by Hall and Trager. Trager (1958) suggests that communication is achieved by a set of interacting systems. One of these systems is *voice set*, or the physiological and physical characteristics of the voice of an individual. This system constitutes a background for *speech*, which in turn includes *paralanguage* and *language*. Paralanguage can be divided into: 1) *voice qualities*, or recognizable speech events including pitch range, vocal lip control, articulatory control, etc.; and 2) *vocalizations*, or noises different from the sounds of language. Language is defined as a cultural system which employs in a systematic way noises made by organs of speech. Finally, the system of *kinesics* includes body movements and postures, a term he borrows from Birdwhistell (1952). Interestingly, Trager does not consider paralanguage or kinesics as subordinate to language, but rather contributing to the communicative event at the same level.

A similar approach to Trager's proposal is Poyatos (1983). He talks about a *basic triple structure* in communication, made up of three systems: language, paralanguage and kinesics. He points out how these three systems are semantically cohesive, and how any stimulus from the outer world can elicit a response in any of the three systems. Thus, a bad odour can prompt a verbal reaction (e.g. 'That stinks'), a hand-to-nose gesture or a paralinguistic reaction (e.g. 'Eeugh!'). The three systems, he argues, can "operate simultaneously, alternate with or substitute for each other as needed" (Poyatos 1983: 132). Poyatos also argues that the verbal code is not necessarily the main one in a message, and that the meaning transmitted by the three systems of verbal code, paralanguage and kinesics is usually not redundant but complementary.

The term kinesics, as Trager acknowledges, was coined by Birdwhistell (1952) to refer to these non-linguistic aspects of the message. Poyatos (1983: 139) defines kinesics as:

> The systematic study of psychomuscularly-based body movements (gestures and manners) and/or their resulting positions (postures), either learned or somatogenic, of visual, visual-acoustic and tactile or kinaesthetic perception, that, whether isolated or combined with the linguistic-paralinguistic structures and with the situational context, possess communicative value, either in a conscious way or out of awareness.

On the other hand, he describes paralanguage as "a series of vocal/narial voice modifications and independent sounds and meaningful silences" which, together with gestures, head movements, facial expression and gaze (i.e. kinesics), is what "truly gives the spoken words their total meaning" (1983: 129).

According to Poyatos (1975, 1976, 1983), paralanguage is constituted by four aspects. The first is *nonverbal voice qualities* such as pitch or rhythm. The second is *voice modifiers*, including pharyngeal control, labial control, articulatory control, etc. He also considers part of paralanguage *independent sounds* like clicks, hissing sounds, hesitation sounds, narial sounds, etc. Finally, he includes silences as the fourth component of paralanguage. As it can be seen, his description of paralanguage includes most of the aspects considered by Trager, but also adds silence to the equation.

Brown (1977), on the other hand, approaches paralanguage focusing on what Poyatos calls nonverbal voice qualities (pitch span, voice range, direction of pitch, tempo, loudness) and voice modifiers (voice setting, articulatory setting, articulatory precision, timing) but does not consider silences (except for pauses) or independent sounds.

Brown also finds that there is a strong connection between paralanguage and kinesics. In the complex system of communication, he claims that language carries the conceptual meaning, while paralanguage and kinesics carry the affective meaning. He further points out how, in cases in which paralanguage and kinesics contradict the verbal message, we tend to rely on the former, rather than on words. These ideas are also shared by Arndt and Janney (1991) and have been proved experimentally by Argyle et al. (1970, 1971): in their experiment subjects were exposed to friendly words plus unfriendly kinesic features vs. unfriendly words plus friendly kinesic features, and what they remembered was the effect of kinesics, not the words. That is, when the modes were not coherent, the effect of kinesics prevailed. The complex connections between paralanguage, kinesics and words is a topic that I will resume in sections 3.2.1.2 and 3.2.2.4.

3.2.1 Words as a semiotic mode

I have opted for the term *words* to refer the semiotic mode that includes the verbal-linguistic material used by speakers (as opposed to paralinguistic or kinesic aspects). The term “words” is arguably rather unspecific, but I have preferred it over other more technical terms such as "verbal mode", “lexical mode" or “language”, because I consider that it describes more accurately the way I conceive this mode. I find the term “verbal” misleading in the context of the present study, because it can also include material that is produced verbally (i.e. using the articulatory organs) but which is non-linguistic, and therefore would fall into “paralanguage” (see section 3.2 in this chapter). The term “lexical” would be too restrictive, because under the semiotic mode *words* I also consider aspects such as grammatical structure, which is usually considered out of the scope of lexis. Finally, the term “language” would be far too generic, because, as argued in section 3.2 in this chapter, language can be considered a macro-system including paralinguistic and kinesic features. Therefore, I use the term “words” to include *lexis (vocabulary choice), grammar, style* and *register*. I have focused on these four aspects in particular because I believe they can have an important role in the design of a persuasive message, as it can be concluded from the discussion of verbal realisations of persuasion offered in Chapter 2. A discussion of each follows in the next paragraphs.

One of the aspects of the mode *words* is lexical resources, or the particular vocabulary choice that a speaker makes. Drawing on my discussion of verbal persuasion in Chapter 2, I think that the following specific vocabulary choices can contribute to persuasion:

i) evaluative adjectives, adverbs and nouns (Bamford 2007; Bhatia 2005; Pomerantz 1986; Hunston 2000) such as “magic” or “easier to use” (both examples from my product pitch sub corpus);

ii) intense language (Pomerantz 1986; Sopory and Dillard 2002; Perloff 2003) , or words that carry a high emotional semantic content (in Aristotle's terms they would appeal to pathos), like "a <u>revolution</u> needs to happen within each an everyone of us” (an example from a research dissemination talk in my corpus);

iii) use of pronouns: for example inclusive pronouns that can contribute to common ground and rapport (in the previous example, “each and everyone of <u>us</u>”); or first personal pronouns that direct the attention to the speaker and make the text personal.

iv) lexical creativity (Lakoff 1982), or original use of vocabulary, for example using the acronym KISS to mean “keep it simple, but not stupid" (example from one of the research dissemination talks in my corpus); and

v) semantic field of novelty (Lakoff 1982, Bamford 2007, 2008): using words that convey the idea of “new", like “we <u>found out</u> we had to focus more on the outside than the inside of the story” or “it’s a very <u>exciting</u> way to work with undergraduate students and it’s <u>growing in popularity</u> in America” (examples from two conference presentations in my corpus).

vi) references to the event: either through deictic terms (e.g. “here”) or by mentioning the title of the event or the previous presentations.

A second aspect which I consider relevant in the expression of persuasion is the use of specific grammatical structures, in particular when they facilitate understanding (e.g. through ordering or listing) or they engage the audience of prompt a particular response. This is the case, for example, with inferences and presuppositions (see section 1.2.2 in Chapter 1), direct addresses to the audience through questions and imperatives, and the use of direct speech (i.e. somebody else's exact words) instead of reporting. For example, in one of the research dissemination talks in my corpus the speaker begins the presentation by asking the audience "Are you smart?" (DTLI1)[2], with the aim of engaging them and prompting them to answer "yes". Also interesting are hedging devices for their potential to make the message not imposing, and the choice of personal vs. impersonal structures that can have the effect of assigning or taking away responsibility for the action or situation described by the verb, and can also affect the level of objectivity. For example, in one of the conference presentations the speaker refers to previous research in impersonal terms but to her own in personal terms, and at the same time hedges her research decisions: "the position of the researcher in terms of gender, in terms of race, in terms of ethnicity has been discussed a lot, but I haven't seen much attention paid to the more, the other levels here, the local ethnographic categories, the more transitory roles. So I'll try to pay attention to that in my paper." (CPDO1)

The use of stylistic devices can also contribute to persuasion, in particular by making the message more memorable. A key resource here is *repetition*. In some of the product pitches in my corpus the name of a product is repeated throughout the presentation to make the audience remember it. A specific form of repetition is the use of *symmetrical structures, also called parallel structures* or *contrasting pairs* (Bowring et al 2013). A classic example of this is Armstrong's reputed quote as he first stepped on the moon "That's one small step for [a] man, one giant leap for mankind". A second form of repetition that can particularly contribute to memorability is the so- called *circular structure*, which implies repeating the same element at the beginning and end of the excerpt. Another frequent stylistic device is the use of *three-part lists* (Carter 1997; Bowring et al 2013). Three seems to be the magic number for a list, since two makes it look incomplete and four makes it more difficult to retain. Again, a classic example is Cesar's alleged "vini, vidi, vinci". I have identified the same resource in some of the product pitches in my corpus, for example when a product is described as "cheaper, faster and easier to use". A third important stylistic device is the use of *narratives (*Bamford 2007, 2008; Boker 2006, 2007; Thompson 2002) and *examples* (Lichinksi 2008),

[2] See Chapter 5 for a description of the coding system used to identify the presentations.

which can present the information as closer, more vivid and more relevant to the audience than when it is accounted for in abstract terms. For example, in one of the research dissemination talks in my corpus the speaker describes a specific need of doctors which his product in meant to supply for, and he does this by telling a narrative about his own sister (who happens to be a doctor), in this way also providing a specific example.

Finally, I also consider that specific choices in register can contribute to persuasion, in particular when they are unexpected. In pragmatic terms this would be a mismatch, and in SFL this would be a marked option (see Chapter 1), and therefore it signals a conscious communicative intent. For example, when in one of the research dissemination talks in the corpus the speaker says “Cool, I think. Well, if you are geeky enough you should find it cool”, he is using slang that is highly unexpected in the context of a presentation, and for this reason the audience is prompted to pay particular attention to it, eventually laughing. This builds rapport, which in turn contributes to persuasion.

A further terminological distinction regarding the use of the term “speech” in this thesis might be relevant at this point. Throughout this thesis I use the term “speech” with two meanings: 1) either referring to the verbal component of oral discourse, including both linguistic and paralinguistic features; and 2) with reference to a specific genre as in "political speech". It is never used as a synonym of “words” as described in the previous paragraphs.

In this section I have clarified my use of the term *words* to refer to the semiotic mode enclosing verbal-linguistic material, and I have pointed out some specific aspects on which my analysis focuses, due to their potential contribution to persuasion. In the next section I turn to intonation, gestures and head movements as semiotic modes.

3.2.2 Intonation

Some of the aspects that fall within the scope of paralanguage, such as for example Poyato’s nonverbal voice qualities, are the focus of study of *prosody*. Prosody is the discipline that studies features of speech affecting more than one sound segment. These features are frequently called suprasegmentals, and include stress, rhythm and intonation (Dalton and Seidlhofer 1994; Noteboom 1997; Wells 2006). Suprasegmental phonetics, therefore, deal with the study of suprasegmental features. Segmental phonetics, on the other hand, deal with features of single sound segments.

The suprasegmental features above mentioned are defined according to three physical properties of sound, which are *pitch*, *intensity* and *duration*. Pitch is usually defined as the frequency of vibrations of the vocal cords. As we speak, variations in pitch make up a pitch contour, which is usually referred to as intonation. Native speakers are able to recognize the pitch contours which are commonly used in their languages (Noteboom 1997). Intensity is related to the energy employed in the production of a sound, and it is what makes us perceive a sound as loud (Pittam and Scheerer 1993). It can be defined as the “amount of sound energy that passes through a given area over a certain amount of time” (Halliday and Greaves 2008: 30). Duration refers to the length of sounds, and it is measured in seconds. These three features play an important role in stress: stressed syllables are heard louder; they last longer and carry a variation in pitch. Duration also affects the rhythm of speech, since it influences its temporal patterning (Noteboom 1997).

My interest in suprasegmental features lies on their communicative value, that is, on their meaning-making potential. This is an area that has attracted increasing interest with the evolution of language teaching approaches and the development of speech technology. Language teachers have long noticed that many of the difficulties language learners find are not really verbal or even linguistic, and they have suggested that a more holistic approach to communication should be applied to language teaching, taking into account non-verbal and non-linguistic aspects of communication (Pennycook 1985; Kellerman 1992; Hurley 1992). However, focus on suprasegmentals is relatively recent in language teaching. Chun (2002) and Akita (2005) argue in favour of a prosody-oriented approach to improve learners’ outcomes in productive communicative skills. There is also increasing awareness among authors about the role intonation plays in comprehensibility of non-native speakers (Pickering 2001, 2004, 2009; Levis et al. 2004; Hahn 2004; Kang 2010), in the expression of interpersonal meaning (Hewings 1998; Pickering et al. 2012) and in effective communication (Wang 1999). Similarly, research in speech technology, especially speech synthesis, has long acknowledged the crucial role that suprasegmental features play in making speech sound “natural” (Noteboom 1997; Carlson and Granström 1997).

As the previous discussion suggests, the effect of suprasegmental features on communication has not probably been fully acknowledged yet. In the present study, my main interest is in the communicative value of suprasegmental features, especially intonation, and the way they contribute to the meaning of the message and to the expression of persuasion in particular.

Intonation is frequently defined as variations in pitch that make up a pitch contour. It has also been described as the “melody of speech” (Wells 2006: 1) or “the use of patterns of pitch to convey different kinds of meaning in discourse" (Dalton and Seidlhofer 1994: 176). It is precisely this communicative force of intonation that mainly concerns this study, since it is expected to play a significant role in the expression of persuasion. In the following subsection I will provide an overview of different approaches to intonation that are relevant for a multimodal study of persuasive strategies.

3.2.2.1 Approaches to intonation

From the point of view of the methodological approach, Ladd and Cutler (1983) have identified two main trends in the study of prosody: concrete and abstract. The concrete approach is closer to phonetics, and is concerned with accurate measurements of realisations and the physical correlates of prosodic phenomena. The abstract approach is closer to phonology, and is more concerned with developing models that can provide a formal representation of prosodic forms and their functions. They frequently borrow notions and concepts used to describe other aspects of language and apply them to prosody, because they tend to view prosody as an integral part of language.

More recently, Chun (2002) has reviewed how the study of prosody and intonation has evolved. Similarly to Ladd and Cutler (1983) she has identified two main initial trends. One was developed in the UK. It focuses on tunes and contours, and is more impressionistic. This is the trend including SFL approaches that I will describe in detail at the end of this section. The other one developed in the US, and is more based on instrumental analysis and interested in tone, sequencing and parts. Both of them are form-based theories. The study of intonation later developed to generative models with the rise of generative grammar. Finally, later approaches have moved away from grammar to focus on discourse.

Concerning the functions of intonation, in the literature two distinct trends can be recognised. On the one hand there are authors who claim that intonation is primarily *grammatical or syntactical*. Its main function is to structure the message, for example marking syntactical boundaries, disambiguating syntactical structures, and signalling given vs. new information or contrast (Crystal 1969; Halliday 1985).

On the other hand, there are authors who argue that intonation is primarily *attitudinal or emotional*, and that it is used to show attitudes and emotions (O'Connor and Arnold 1973; Cutler 1893; Bolinger 1986). Cutler (1983) used a study of repair to elucidate speakers' conceptions about the functions of prosody, showing how speakers tend to consider prosody as primarily pragmatic. She notes how some repairs are marked with a distinct prosodic pattern, while others keep the same pattern to go unnoticed. She finds that speakers mark repairs to ensure attention to them only when 1) the error can disrupt communication (e.g. lexical stress is not corrected if it does not affect vowel quality), and 2) especially when the prosody is not in itself incorrect, but inconsistent with the speakers intention. This leads her to think that the function of prosody is primarily concerned with the expression of pragmatic meaning. Particularly interesting is Bolinger's (1986) view of intonation and gestures as forming a "gestural complex". Bolinger (1986) talks about the interrelation between intonation and emotion, which is proved by the interrelation between intonation and gesture.

> The fact that intonational configurations are matched by configurations of facial expressions and bodily gestures, that the two operate much of the time in parallel, and that their similarities betoken similarities of function, points to the configurational approach as the most likely to succeed (Bolinger 1986: 337).

However, as Crystal (1969) points out, there is general agreement on the fact that these functions of intonation are not mutually exclusive, so even those who consider intonation as primarily grammatical acknowledge some attitudinal function, and vice versa. Crystal sees the relationship between both functions of intonation as hyponymic: in all cases in which intonation is primarily grammatical it also carries some attitudinal function, but the opposite is not true.

There is yet a third trend within intonation research that sees intonation as a tool facilitating the flow of discourse and helping the negotiation process involved in every communicative interaction. It is called *Discourse Intonation*. Its main claim is that speakers use intonation to project their understanding of the communicative situation, i.e. their assumptions and intentions (Brazil 1985, 1997; Dalton and Seidlhofer 1994; Cauldwell 2002, 2015; Chun 2002). Couper Kuhlen and Selting (1996) propose that all the functions mentioned above can actually be included within this last discourse function (which they call *interactive*). They consider that what intonation does is to signal or contextualise the use of language in a particular situation to accomplish interactional goals.

For the purposes of this thesis two approaches to intonation mentioned in the literature review presented in the previous paragraphs stand out as particularly valuable: Halliday and Greaves' (2008) Systemic Functional Approach to intonation and Brazil's Discourse Intonation (1985, 1997). Both shed light on the way intonation contributes to make meaning according to moment-to-moment decisions of speakers as discourse unfolds as part of an ongoing process of interaction. A discussion of both approaches follows in the next pages.

Halliday and Greaves (2008) offer a comprehensive study of intonation from a systemic functional approach, which, as discussed in Chapter 1, sees language as a semiotic system, made up of different strata. In each stratum speakers have a set of options, and each stratum is related to the next through realization. In addition, the available choices of the language system convey different types of meaning, which makes it possible to differentiate between three metafunctions of language: ideational (experiential and logical), interpersonal and textual.

Intonation would fit in the phonological stratum. Three systems are differentiated within intonation:

i) Tonality: division of oral discourse into stretches or tone units;
ii) Tonicity: tonic syllables within each tone unit (each tone unit contains at least one);
iii) Tone: pitch movement (fall, rise, level-rising, fall-rising, rise falling).

To exemplify how intonation fits within the system of language Halliday and Greaves (2008) provide the example of a statement with a falling intonation. This fall tone (a category at the phonological stratum) realizes a lexico-grammatical category, which is declarative, which in turn realizes a semantic category, i.e. a statement.

As Halliday and Greaves (2008) point out, the systems of tonality and tonicity realize meanings of the textual metafunction while tone contributes to the interpersonal and logical metafunctions. The only function to which intonation does not seem to contribute is the experiential function. They explain the contribution of intonation in the following way:

Tonality is used to mark the information distribution. Usually each *tone unit* at the phonological stratum corresponds to two other parallel units at the lexico-grammatical stratum, which are the *information unit* and the *clause*. When the three units do not match we find a *marked option* that is the consequence of a particular discursive situation and it becomes therefore especially meaningful.

Tonicity is used to mark the information under focus and differentiate new from given information. New information tends to occur at tonic syllables. Usually tonic syllables are placed at the end of the tone unit. This is the unmarked option. The information focus contains new information, and all elements following the information focus are given.

Tone can be used to mark the relation between clauses in longer stretches of discourse, contributing in this way to the logical meaning. Through tone speakers indicate whether the two clauses have equal status (paratactic) or if there is a relation of dependency (hypotactic). Each relation has a typical unmarked tone. Speakers may use a marked version for communicative effect, for example using a tone implying "condition" when in fact the clauses are connected by a paratactic relationship, as in “Give it a try (fall-rise) and you will be convinced (fall)" . Likewise, tone also signals the logical-semantic relation connecting both clauses. For example a clause that is elaborating on a previous one will repeat the same tone pattern. This is called pitch concord, and it shows that the two clauses have the same status and discourse function.

Finally, tone is used to express meanings of the interpersonal metafunction. As mentioned in Section 1.1.3 of Chapter 1, at the lexico-grammatical stratum two systems are particularly relevant for the expression of interpersonal meaning: mood and modality. At the stratum of semantics there is the system of speech function. Within this system speakers can choose between ‘give’ and 'demand’ either ‘goods and services’ or ‘information’, resulting in statements, offers, questions or commands (see Section 1.1.3). These options within the speech system are realized at the lexico-grammar level by particular choices in the systems of mood and modality, and at the phonological level by a selection of tone. As Table 3.1 shows, there are usual, unmarked, tones for each speech function and mood (except for offers which have no particular tone assigned to them):

Table 3. 1 Speech function, Mode and Tone (Adapted from Halliday and Greaves 2008: 110)

Speech function	Mood	Tone
Statement	Declarative	Falling tone
WH question	Interrogative	Rising tone
Yes/no question	Interrogative	Falling tone
Command	Imperative	Falling tone

However, there are also marked options with powerful communicative effects (Halliday and Greaves 2008). An example is a statement with rising tone, which may function as a question, a challenge or a contradiction. Another example is an imperative with fall-rise tone, in which a nuance of "with reservation" is added, or imperatives with rise-fall tone, which add a nuance of insistence.

As regards modality, speakers use intonation to adjust their utterance within a continuum of aspects such as probability, usuality, readiness and obligation. A fall tone indicates a high modal value, while a fall-rise means low modal value. For this reason, a statement with a fall tone sounds more certain or probable than if a fall-rise tone is used.

In short, a systemic functional approach sees intonation as a system of choices with which speakers make meaning. It reveals how intonation contributes to the ideational, interpersonal and textual function of language as much as the verbal component of the message. For this reason I consider it a semiotic mode in its full right. Other approaches to intonation such as Discourse Intonation will also foreground its communicative value and meaning-making potential, but instead of linking choices at the phonological stratum to grammatical entities, they relate them to moment-to-moment situationally motivated decisions (Chun 2002).

As already mentioned, Brazil's communicative approach to intonation (Brazil 1985, 1997, Cauldwell 2015), known as Discourse Intonation, focuses on the communicative value of intonation, and it proves particularly valuable in order to analyse the relevance of paralanguage in a message. The main tenet of this approach is that speakers make meaningful choices in their use of intonation, or, in other words, a speaker can exploit intonation for communicative effects. In addition, the choices speakers make in the use of intonation reveal their assumptions about the interaction in the communicative process.

Discourse intonation divides speech into *tone units*, each one containing one or two *prominent syllables*. Prominent syllables are more emphatic, because they are usually louder and longer, and therefore can be heard more clearly. The first prominent syllable in a tone unit is called *onset*, the second is called *tonic* syllable. Several tone units together make up a *pitch sequence*. Each tone unit is an increment to discourse and reflects the ongoing planning of the speaker. Every time speakers begin a new tone unit they are making a choice in each of the four systems in which meaningful contrasts can be produced: prominence, tone, key and termination (see Table 3.2). A brief description of each system is provided on the following pages.

Table 3. 2 Discourse Intonation systems (adapted from Brazil 1985)

SYSTEM	DEFINITION	OPTIONS
PROMINENCE	Prominent syllables are louder and longer	Prominent vs. non prominent
TONE	Pitch movement	Rise-fall Fall Level Rise Fall-rise
KEY	Relative pitch of onset syllables in relation to the onset syllable of the previous tone unit.	High Mid Low
TERMINATION	Relative pitch of tonic syllables in relation to the onset syllable in the same tone unit	High Mid Low

Prominence

Speakers can decide what syllables they make prominent, and these decisions reveal their understanding of the context of interaction. Usually, we only make prominent parts of the message that allow for other alternatives (i.e. are selective) and can therefore be misunderstood. The rest of the message is treated as shared knowledge. The number of possible alternatives is determined in every communicative situation by the system of language, by the immediate conversational environment and by the wider context (real world, social conventions and culture). Prominence can be exploited for persuasive purposes, because content presented as non-prominent is assumed to be already negotiated between speaker and hearer and part of what they both take for granted. Presenting something as common ground prevents potential counterclaims. At the same time, it can be used to direct attention to certain parts of the message to the speaker’s convenience.

Tone

Tone is the pitch movement that begins in the tonic syllable and continues throughout the tone unit. Speakers have the choice of using fall, rise, rise-fall, fall-rise or level tones. Rise and fall-rise tones are called referring tones. They usually imply that what is being said is shared information. Fall and rise-fall are called proclaiming tones, and they imply that the information provided is new. Finally, the level tone shows that the speaker is focusing on the exact words leaving aside any interactional activity for the moment.

The main communicative value of tone is to show what the speaker assumes is new or shared information. It also gives an indication of the type of response expected from the listener. This is clearly seen in questions. Questions with referring tone seek agreement or confirmation, and expect a positive answer, while questions with proclaiming tones present a possible interpretation of reality with equal chances of being confirmed or refuted by the listener. Information questions usually have proclaiming tone because the speaker has no reason to present the information as already negotiated. Referring tones are only used in this type of questions when a particular answer is expected. According to Brazil (1997), the use of a referring tone projects a wish to have assumptions confirmed with respect to something that has already being negotiated, while proclaiming tones expect a selection (e.g. rejection or acceptance) from a still not negotiated number of options.

Because the use of referring tone suggests agreement and consensus, it also contributes to a closer interpersonal relationship in communication. As Brazil argues, a referring tone projects an understanding of the situation on behalf of all the participants in the communicative event (what "we" as a group believe), while a proclaiming tone projects the individual’s own perception and ideas (what “I” as an individual believe), and therefore can establish a greater social distance. This can be used for persuasive purposes. Persuasive speakers will try to apply referring tones to parts of their message that they do not want to have questioned, because in this way they are likely to be perceived as already negotiated and agreed.

It must also be noted that Avery and Ehrlich (1992) refer to a continuation rise tone that simply indicates that discourse is not finished, rather than assuming shared knowledge or agreement, which is an option that cannot be discarded.

There are also differences between the two referring tones and the two proclaiming tones, and they lie mainly in dominance. Within the referring tones, the rise tone shows more dominance than the fall-rise, and within the proclaiming tones the rise-fall shows more dominance than the fall tone. Likewise, as Brazil (1997) claims, an offer with a rise tone is perceived as more polite, because the overt control exercised by the speaker in choosing that tone is usually interpreted as sincere willingness.

Key

Key refers to the relative pitch of onset syllables in relation to the onset syllable of the previous tone unit. Speakers have a choice of low, mid or high key, each one conveying a different interpretational framework for the listener.

Low key has an equating effect. When a speaker chooses low key he adds the meaning "what I'm going to say is equivalent to what I have just said". This equating effect can have important consequences in the way a message is perceived.

High key, on the other hand, indicates surprise, disagreement with expectations, or little connection to what has just been said. A common use of high key, for example, is contrasting. In short answers, a high key expresses disagreement (i.e. not the answer that the person making the question expects). In the specific case of "Yes" responses, a high key is adjudicating, and means "you're right" (not wrong), as opposed to mid key, which is concurring and means "I agree with you".

Mid key implies addition. When a speaker chooses mid key the implication is that the tone unit is simply adding information to what has already been said, as opposed to equating or contrasting. In conversation, mid key usually accompanies concurring answers.

In persuasive discourse a mid-key (concurring response) is what is obviously expected from the persuadee, and depending on the case also from the persuader, since a high key can potentially disrupt rapport.

Termination

Termination refers to the relative pitch of tonic syllables in relation to the onset syllable in the same tone unit. Termination reveals the speaker's expectations regarding the interlocutor's reaction. As it was the case with key, speakers can choose from high, mid and low termination. Each choice gives an indication of the expected role the interactant is expected to assume.

In general terms, high-termination indicates the interlocutor is expected to have an active role of some sort. In the case of questions a high termination shows that the listener is expected to adjudicate from a number of possible options. High termination is also related to surprise. Questions that anticipate a surprising response use high termination. In the case of statements high termination is also employed when the speaker expects to surprise the listener.

Mid termination, on the other hand, indicates the interlocutor is expected to have a passive role (i.e. concur) and as such it is a sign of dominance in conversation. In extended monologues termination choices represent the speaker's expectations of a particular feedback from the audience. Thus, when a mid termination is used, it signals that the audience is expected to concur, or to accept something without further argument. On the other hand, when a high termination is applied, the audience is being prompted to decide on something, (i.e. to adjudicate from a number of possible options). As Brazil points out, it is a frequent device to start with a high termination and end with a mid termination. The effect of this is that listeners have the feeling of being actively deciding, but in fact at the end they are being prompted to agree.

When a termination choice occurs immediately before giving the turn, it can also restrict the possible options of the listener, because of what Brazil calls *termination-key concord*. This means that a mid termination indicates an expectation of a concurring response with a corresponding mid key, and a high termination expects an adjudicating response with a corresponding high key. Low termination, on the other hand, indicates no expectation at all, so it can be interpreted as a cue for the listener to take the floor (a signal that the speaker has finished speaking) or even a cue to finish the communicative event.

Now the four systems of Discourse Intonation have been introduced, there are three additional important aspects that need to be taken into account when applying Discourse Intonation to an oral text: *pitch sequences, orientation* and *redundancy*.

First, we need to consider a unit above the tone unit to account for longer stretches of speech that make a coherent whole. The unit proposed by Brazil (1997) is the *pitch sequence*, which is a series of tone units ending with a low termination and with no occurrences of low terminations within it. Often in pitch sequences there is a progression from high to low pitch. This is a frequent pattern among newsreaders for example. Sometimes, however, there are also variations in pitch which are referred to as *internal choices*. That is to say, when looking at longer stretches of utterances comprising several tone units, it becomes clear that speakers do not always start their discourse with a high pitch and smoothly fall to a closing low pitch. As Brazil argues these “deviations from the norm” are meaningful. As shown in the previous discussion of key and termination, speakers exploit key and termination systems to project their understanding of the context of interaction, and to exercise some control over their listener's response.

Brazil (1997) refers to the initial key choice and the closing termination choice as *external choices*. The initial key choice in a pitch sequence expresses how it can be connected to the previous pitch sequence, as follows: a) high key (both pitch sequences are unrelated), b) mid key (addition or expansion), and c) low key: equative.

Regarding closure of pitch sequences, two aspects are of particular interest. One is the position of this closure, that is, where the speaker decides to put the boundary. As Brazil points out, this is not always determined by grammar. A speaker often extends a pitch sequence over several sentences. In doing this, he presents the information as a unity, to be considered as a meaningful whole. On the other hand, the speaker may also decide to split a sentence into two pitch sequences, presenting the information in discrete parts to be considered separately. The other remarkable aspect of pitch sequence closure is its relation to dominance. Brazil argues that it is the dominant participant who actually closes a pitch sequence when this extends over several turns in a conversation, and that, in fact, a low termination response can be interpreted as a bid to take the turn (the speaker shows in his response his intention to end the current pitch sequence and start a new one).

In addition to accounting for longer stretches of speech and looking at the relationship between subsequent tone units, it is also necessary to take into account the attitude of the speaker towards the message and whether he is using it to communicate or as metalanguage. In this sense Brazil distinguishes two different focuses or *orientations* when speaking: *direct and oblique*. When speakers adopt a direct orientation, which happens most of the time, their goal is to transmit a particular meaning, that is, their words have communicative value. On the other hand, when speakers adopt an oblique orientation they are focusing on the language used, and presents it as a specimen with no communicative value. This happens frequently in language courses when language is used to practise rather than to really *say* something. It is also found when reciting stretches of speech learnt by heart, and also when the speaker needs to concentrate on the verbal organisation of the message (i.e. the encoding process). In this latter case oblique orientation frequently accompanies fillers (e.g. ‘eeerr’).

Finally, a third aspect that needs to be considered in a DI approach is that the meaningfulness of choices in the different systems of prominence, tone, key and termination fades out as their communicative value becomes evident or recoverable from the context, or in other words, as they become *redundant*.

To finish this review of DI, I would like to stress the fact that the different systems of prominence, tone, key and termination do not work independently from each other. On the contrary, their effect can be cumulative. In the same way, they can –and in fact they do- co-occur with other aspects of paralanguage, such as silences or loudness, and with kinesic features such as gestures and posture that also contribute to the meaning of the message. This study is based on the assumption that intonation does not work independently from other modes, and that the full communicative value of intonation can only be grasped taking into account the whole communicative situation.

3.2.2.2 The contribution of intonation to the multimodal ensemble

As Crystal (1969) claims, one of the main difficulties in undertaking a systematic study of the meaning of intonation is that there are several systems that work simultaneously, such as other paralinguistic features and kinesics. Therefore, isolating the variable that makes the difference in meaning is complicated. To exemplify this, he points out how a falling intonation can vary its meaning depending on whether it is combined with a frown or a smile.

Bull and Conelly (1985) also acknowledge a close relationship between kinesic features and paralanguage. They focus their research on how kinesic features combine with paralinguistic features to express emphasis. As they proved in their study of body movement and emphasis, tonic stresses are usually accompanied not only by hand and arm movements, but also by a wide range of head, trunk, legs and feet movements. They conclude that it is precisely this synchrony with tonic stress that provides body movements with their emphatic character.

On the other hand, Kellerman (1992) claims that kinesic features are closely connected to prosodic features, since embodied actions are coordinated with the rhythm of speech. She agrees with Brown (1977) that it is actually possible to “see” stress by looking at body movements. She points out how some facial expressions are clearly connected to intonation and stress: a raised eyebrow frequently accompanies questions, and eyes wide open are usually synchronous with prominent syllables. Furthermore, she argues that kinesic features can have an organisational function, signalling a change of topic or highlighting foregrounded information. As a consequence, they can direct the addressee’s interpretation, serving as signals that what is being said is particularly relevant, interesting or complex and therefore requires special attention. These ideas are shared by Wrembel (2007), who argues in favour of the teaching of suprasegmentals through kinesics (physically matching pitch and tone with body movements and face expression), because she claims that awareness of the communicative value of gestures and facial expressions can be used as a basis to transfer awareness to the communicative value of intonation (e.g. an angry face can be used as a stimulus for an angry intonation).

Several studies have focused on how different paralinguistic features can combine to produce emphatic speech. Selting (1994) claims that speakers make use of high pitch, greater number of stressed syllables and increased loudness in order to mark discourse as emphatic. These features may appear on their own or they can co-occur with lexical intensifiers, (extreme) evaluations and also syntactically elliptical units (in which only the relevant parts are physically pronounced so as to make the addressee focus on them). This use of emphatic style prompts the addressee to show signs of alignment with what is being said. In fact, this tendency is so strong that listeners usually feel compelled to show alignment with what is being said emphatically even if they have failed to listen or to understand properly. These results are supported by Goodwin and Goodwin (2000), who point out how loud intensity, high pitch and prominence are all signals of a strong emotional stance, and also Yu's (2011) study about the display of frustration in arguments, which also shows how the expression of frustration is aided by gestures, posture and facial expressions.

Finally, within the specific context of oral presentations, Hincks (2005, 2009) has studied how prosody can influence perceptions of liveliness in a presentation. Speech that is directed to a large audience tends to feature more pitch variation than conversational speech (Johns-Lewis 1986). In addition, speech that is delivered with a monotonous intonation makes it more difficult to recall the message (Hahn 2004), and can convey disengagement by missing chances of expressing common ground and ensuring mutual understanding (Pickering 2001). However, a common problem with novice presenters and NNSs is that their speech tends to be monotonous. Hincks measures the pitch variation quotient in a corpus of students' oral presentations, and finds that it correlates well with perceptions of liveliness. She suggests using a computer assisted automatic feedback mechanism in order to provide feedback to students regarding their use of prosody so that they can improve their performance. After all:

> [i]t is vital to use one's voice well when speaking in public. It is the channel of communication, and when used poorly, communication can be less than successful. If listeners stop listening, or fail to perceive what is most important in a speaker's message, then all actors in the situation are in effect wasting time. (Hincks 2009: 46)

3.2.3 Gestures and head movements

Gestures and head movements fall within the scope of kinesics, together with other bodily actions such as posture, gaze and facial expression.

Kendon (2004: 7) defines gesture as “visible action when it is used as an utterance or as part of an utterance". He claims that for a gesture to be regarded as such, it has to be perceived as intentionally expressive, i.e. used to communicate something. This is what differentiates them from other body movements.

All speakers use gestures when communicating, but they do so to a great extent unconsciously. In fact, research suggests that kinesic and paralinguistic features work at a more subconscious level than words and require less intellectual processing (Poyatos 1983). Yet, they closely intertwine with words to create meaning, so much so that it is possible to consider them as integral parts of a single unit. This idea was first suggested by Kendon (1972, 1980) and is also shared by McNeill (1992), who claims that gesture and speech are so closely connected because they spring from the same process of utterance formation. This process of utterance formation brings together two opposite forms of thought: one is imagistic, and makes use of a mode (i.e. gestures) which is global and synthetic. The other one is analytic, and makes use of socially regulated coded symbols (i.e. language) which are segmented, linear and sequential. McNeill considers that gestures are global because, unlike language, they are not made up of discrete meaningful parts that can be combined, each one adding a piece of meaning. In the case of gestures, it is rather the whole that determines the meaning of the parts. They are also synthetic because they condense different meaning elements. Additionally, Kendon (2004) provides further proof that gestures are integral parts of the utterance by pointing out how they tend to be revised or repeated as speech is repeated or rephrased. Furthermore, the same verbal utterance can achieve different communicative effects only by adding or changing accompanying gestures. The relationship between gesture and speech is a topic I will return to in section 3.2.3.4.

Gestures are extremely varied, and for a systematic study of their contribution to the message, a classification may prove a useful tool. For example it can smooth the annotation process, and it can also facilitate the discussion of results with speakers when they lack the required metalanguage (see chapters 4 and 5). The classification of gestures is the focus of the following section.

3.2.3.1 Types of gestures and head movements

Many attempts have been made at classifying gestures. This task has proved to be a difficult one, and several researchers have provided different classifications according to different criteria and scientific purposes. Some examples are Efron (1941), Ekman and

Friesen (1969), McNeill (1992), Kendon (2004) or Krauss et al. (2000), which I will first review in the following paragraphs to later find some points of agreement among them.

Efron (1941) examines gestures according to three perspectives: *spatio-temporal*, *interlocutional* and *linguistic*. The first, spatio-temporal, includes aspects such as range of movement and body parts involved. The second, interlocutional, looks at whether the gesture involves the interactant and how. Finally, the linguistic perspective distinguishes between *discursive* gestures, which refer to the flow of discourse, and *objective* gestures, which make reference to the object of discourse. The latter, in turn, can be: 1) deictic (pointing); 2) physiographic (depicting); and 3) symbolic (they bear no morphological relation with the referent).

Later, Ekman and Friesen (1969) have suggested that in the study of nonverbal behaviour, the following aspects need to be considered: origin, usage and coding. *Origin* refers to whether the gesture is learnt or inborn. *Usage* includes aspects like the context, the association with verbal behaviour, level of consciousness, and type of information conveyed. *Coding* refers to the relationship between gesture and meaning, which can be arbitrary, iconic or intrinsic (gesture and its meaning cannot be separated). According to these criteria, they propose the following classification of gestures:

i) Emblems have a direct verbal equivalent or definition;

ii) Illustrators illustrate what is said with words:

   a. Batons: mark discourse flow;
   b. Ideograhps: show line of thought;
   c. Deictic: point;
   d. Spatial: represent spatial relationships;
   e. Pictographs: draw a picture of the referent;
   f. Kinetographs: depict bodily action;

iii) Affect displays: facial expressions of emotion;

iv) Regulators: control interaction;

v) Adaptors: attempts to adapt to the communicative situation, e.g. control emotions.

The terms ‘emblems’, ‘illustrators’ and ‘batons’ have been widely accepted and are frequently used in the literature on gestures.

More recently, McNeill (1992) has offered the following typology of gestures:

i) Imagistic: depict action, shape, movement. They can be further divided into *iconic* if they represent something concrete and *metaphoric* if they stand for abstract concepts.
ii) Non imagistic, which can be *deictic* (pointing) or *beats* (they mark discourse flow)
iii) Cohesives: they tie thematically related but temporally separated parts of discourse. As I will discuss in the next section, this can be plausibly considered a function rather than a type.

Finally, Krauss et al. (2000) have suggested the following classification:

i) Symbolic: can communicate on their own, without the need of speech because they tend to have a well-established meaning (equivalent to emblems);

ii) Lexical: equivalent to imagistic gestures;

iii) Deictic: used for pointing;

iv) Motor: repetitive, rhythmic movements with no obvious relation to the semantic content of the accompanying speech.

Like McNeill, they also contemplate the distinction between metaphorical and iconic gestures within the lexical group, but they consider this difference to be more a matter of degree than a discrete distinction.

As Kendon (2004) points out, a problem with these classifications is that they mix different criteria. For example some gestures are classified according to their meaning, others according to their function and others according to their relationship with speech. Each classification suits the purposes of the research from which it emerges and is perfectly valid for that, but can hardly claim to be universal. This problem in the classification of gestures is also noted by Mittelberg (2007). As she points out, the most widely used coding parameters in the gesture literature are hand presence, hand shape, palm orientation, movement, and the location in space, but she warns that during the process of annotation and analysis there is always a risk of mixing the physical properties of the gesture with the semantic and pragmatic functions they perform. Kendon's gesture families can be a way of reconciling both criteria, as I will discuss in section 3.2.3.3.

Kendon suggests that for the sake of clarity, it is more accurate to consider gestures alongside a continuum expressing a unique criterion. He first adopted this approach when analysing the different ways in which gestures are used as part of the utterance. He noticed how some gestures generally need to be accompanied with words in order to be understood, while others can stand on their own. The former are idiosyncratic, the latter are conventional and stereotyped. Drawing on this idea, McNeill (1992) suggests an arrangement of gestures along a continuum, which he called Kendon's continuum (see Figure 3.1):

Figure 3. 1 Kendon's Continuum (Kendon 1988, McNeill 1992)

Gesticulation comprises idiosyncratic gestures that occur together with words and generally need the verbal component to be correctly understood. Language-like gestures differ from gesticulation in that they form part of the grammatical structure of the utterance (e.g. as when a gesture completes an utterance). Pantomimes represent objects or actions and are conventionalized enough so as to be able to stand on their own (e.g. directing your thumb to your mouth meaning "drink"). Emblems are similar to pantomimes but they are more standardized, and they can fully be interpreted without words. An example of an emblem is the OK sign. Finally, sign language constitutes a system of its own, with well-defined rules and a community of users. On one extreme of the continuum we find what many authors call conversational gestures, which are idiosyncratic and occur together with speech, and on the other extreme we find stereotyped gestures, which are autonomous. For the present research it is non-conventional gestures co-occurring with speech that are of particular interest, for their potential contribution to the expression of persuasion in a creatively orchestrated multimodal ensemble which includes, among other modes, gestures and words. Goldin-Meadow (2003) suggests that this type of gestures in particular have been classified quite consistently in the gesture literature (see the classification presented in the previous paragraphs), although with different numbers of sub-categories, as Table 3.3 shows.

Table 3. 3 Classifications of gestures, adapted from Goldin-Meadow (2003)

Lexical	Iconic	Kinetographic
		Spatial
		Pictographic
	Metaphoric	Ideographic
Deictic	Deictic	Deictic
Motor	Beats	Baton

So, in spite of divergences in the gesture literature, it can be safely claimed that there is general agreement on three main aspects. The first is that some gestures are used to point out, which are usually called deictic gestures. The second is that some gestures are used to represent an aspect of the content of the message. They are usually called topic gestures (lexical, imagistic). Finally, some gestures refer to the structure of discourse rather than content, and they are usually called non-topic gestures (batons, beats).

Regarding head movements some early studies (Hadar et al.1983, 1984, 1989) have classified them according to their frequency, amplitude and duration, resulting in four types: narrow rapid, ordinary, wide slow and postural shift. There are actually limited ways in which people can move their head: back and forth (i.e. nodding), from side to side (i.e. quick shaking or a slow lateral sweep), tilting it to one side and in circles. Because of this, classifications of head nods tend to be less elaborate than classifications of gestures. For the purposes of this study I focus on nods, shakes and lateral movements (sweeps and tilts) because I hypothesize they can be expected to occur in the context of oral presentations and contribute to persuasion. I also consider it interesting to look at the amplitude of the movement and the number of repetitions, as Hadar's studies do, because I think it can provide insights into pragmatic meaning. For example Hadar et al. (1983) note how a nod can have different meanings depending on the number of repetitions (one nod means polite involvement, two nods real interest, but three nods can convey impatience or desire to take the floor).

3.2.3.2 Functions of gestures and head movements

In order to fully understand the contribution of gestures to a persuasive message it is necessary to analyse the possible functions they can perform.

Kendon (2004) concluded that gestures perform three main functions:

i) Referential: they are part of the referential content;

ii) Pragmatic: gestures that show the attitude of the speaker towards the content and indicate how content is to be interpreted. Three sub-functions can be distinguished:
   a. Modal: show how the utterance is to be interpreted;
   b. Performative: indicate the kind of speech act;
   c. Parsing: equivalent to punctuation in writing;

iii) Interpersonal: regulate interaction.

As Querol-Julián (2011) points out, we could add McNeill's (1992) *cohesive gestures* to this typology of functions, since the term *cohesive* actually describes a function rather than a type of gesture (i.e. connect thematically related but temporally separated parts of discourse).

These functions are by no means clear-cut, and some gestures can perform more than one function at a time. This is the case with *deictic* gestures. They are referential in that they refer to an object under discussion, but pragmatic when they also show the speaker's attitude towards this object or what the addressee is expected to do with this object (e.g. examine it, consider it an example of something, etc.). Another example is *beats*. They can be considered pragmatic in that they serve to guide the understanding of the message. However, Bavelas, Chovil, Coates and Roe (1995) have argued that beats are actually interactive gestures, because they ultimately refer to the addressee in different ways. They highlight how these gestures share a pointing action with fingers or palm to the addressee, and have proved that they are significantly more abundant in conversational mode than in monologues. They point out how a beat can mark the type of information offered. If the beat is performed with both open hands palms facing upwards, the information is presented as new. A beat performed with one hand facing upwards and fingers pointing to the addressee can indicate that the information is shared, as a kinesic equivalent of "as you already know", or it can function as a kinesic marker of citing with which the speaker acknowledges the origin of his contribution (equivalent of "as you said"). Holler (2010) has probed into this type of interactive beats in particular and has experimentally proved that they are more frequent when there is more common ground between speakers. Two types can be differentiated: a deictic pointing with the finger to the addressee, which seems to be perceived as more compelling, and an open hand palm up "offering" gesture. Through these different gestures speakers signal what type of feedback or backchannel (evidence of understanding) they need from the addressees and how urgent this is. If these gestures appear in a presentation they can be hypothesized to help build rapport with the audience, because they create the feeling that the audience is brought in a conversation with the speaker.

It is important to note that, as Goldin-Meadow (2003) or Kendon (2004) suggest, the function or meaning of a gesture is not inherent in the gesture, but determined by the context, and the same gesture can have different functions depending on the context in which it appears. Kress (2009b), on the other hand, points out that we understand gestures by mimicking them and feeling them, we need the bodily experience of it to grasp their full meaning. This implies that meaning potentials of gestures are embodied.

Turning now to head movements, Maynard (1989) distinguished 4 functions of head nods in Japanese conversation, highlighting their importance in conversational dynamics: 1) emphasis, 2) mark clause boundary/turn completion, 3) fill turn transition pause, and 4) claim next turn. More recently, McClave (2000) undertook a study of head movements in American conversation that revealed semantic, narrative, cognitive and interactive functions. Among the *semantic* functions she finds that nods and shakes can serve as *emblems* for affirmation and negation in the absence of speech. In the case of lateral sweeps, they can convey concepts such as “all of us”, “everyone”, “whole”, which can be grouped under the generic semantic theme of "*inclusivity*”. Head shakes can convey appreciation of something out of the ordinary or unbelievable, in which case they would have an *intensifying* function, or they can acknowledge other possibilities, or the fact that a piece of information is missing, in which case they convey *uncertainty*.

The *narrative* functions include marking *direct quotes* (a movement of the head can show a change in footing, or the switch to direct speech), portrait *mental images of characters* (speakers can pretend to look at fictional characters as if they were present), perform deixis (the direction of a head movement can locate a referent in space) or *indexing items in lists* (the head moves which each item).

Head movements can also help *cognitive* processing (cognitive function). Speakers frequently mark self-initiated repairs with a head shake, and they also mark contrasts with preceding utterances of interlocutors. Finally, head movements have *interactive* functions, in particular nods, because they serve as backchannelling requests.

Kendon (2002) focuses specifically on head shakes. He notes how head shakes can occur on their own (as McClave notes, as emblems for “no") or simultaneously with speech.

Furthermore, they can affect the referential content, the illocutionary force of the message or the stance of the speaker. He distinguishes the following functions of shakes:

i) emblem of “no”, with no accompanying speech;

ii) combining with the word “no” to emphasize or clarify verbal production;

iii) kinesic equivalent of a verbal negative operator (e.g. “don’t”);

iv) addition to the meaning conveyed by words, i.e. simultaneous with speech but not semantically synchronous (e.g. making a positive statement and at the same time responding negatively to a previous question or offer with the head: “I have to study” + head shaking as meaning “Thank you for the offer but I can’t go because I have to study”);

v) additional move or expressing stance, similarly to iv) but in this case pragmatically asynchronous with words;

vi) anticipating counterclaims in combination with implied negations (e.g. “the only thing” implies “there is no other”) or with intensifying and extreme expressions that imply “no exceptions allowed”, such as superlatives (e.g. “it was really so good” implying “it couldn’t have been better”). The rationale behind this last interpretation is that this type of expressions is likely to prompt counterclaims due to their categorical nature. With the head shakes speakers are anticipating and refuting these counterclaims.

3.2.3.3 Gesture families

Kendon (2004) suggests that it is possible to group gestures with similar kinesic characteristics into families, each family sharing a common semantic theme, and each gesture within the family adding a particular semantic nuance. Of course, it is the context and the accompanying words that make it possible to grasp the full meaning of a gesture, but Kendon’s approach of gesture families sharing semantic themes makes it possible to make useful generalizations about the meaning expressed by gestures –in particular the pragmatic function they fulfil- and facilitates our understanding of how they contribute to our interpretation of the whole message. Kendon distinguishes the following families, which are explained in the following paragraphs:

i) Precision grip:
   a. G-family;

b. R-family;

ii) Open Hand:

a. Open hand prone (OHP);

b. Open hand supine (OHS).

The *precision grip* gesture families are grouped under this name because the gestures they include are thought to derive from the action of gripping something with precision (Morris 1977). However, the action of gripping can be performed in very different ways, and this is why two different precision grip families are identified: *G-family* and *R-family*.

The gestures within *G-family*, also called “finger bunch”, share the following position: fingers are brought together touching at their tips, forming a sort of bunch. The semantic theme conveyed by these gestures is that of *seizing*, but depending on the specific gesture used this generic meaning is made more specific. For example when the fingers close to form the bunch this means ‘selection of a topic'. When, in addition, the fingers are moved towards the speaker this adds the meaning of ‘please pay attention to this'. This gesture is frequently found in situations in which the speaker demands an explanation or justification for something unexpected. If, on the other hand, the bunch moves vertically, this adds the meaning of "core" or “essential”. Finally, a different option is to start from a finger bunch position and open the fingers. This implies that the speaker is commenting on a topic.

The *R-family* gestures all have in common a position of the fingers emulating a ring. They have the semantic theme ‘making precise’. These gestures are used when this precision seems necessary to gain agreement, persuade or facilitate understanding. There are three main gestures within this family. The first is the *R- to-open*, in which the fingers form a ring and then open up. This gesture is used when the speaker selects a specific topic (the kinesic equivalent of this being the ring shape) and then comments on it (the kinesic equivalent of this being the opening movement). The second gesture is the *R-display*, in which the fingers close to form a ring and this position is held up as in a display for the addressee to notice. This gesture is used when precise information is being offered which in a way contradicts expectations or when something is being quoted. With this gesture the speaker seems to draw attention to the accuracy of his words. The third gesture in this family is the *R-vertical*. This gesture is similar to the R-display, only that a vertical movement of the ring shape is added. With this movement the speaker adds a nuance of insistence. It is frequently used when the speaker wants to make clear an opinion that explicitly contradicts the interlocutor's.

Gestures realised with an open hand and palm facing down are grouped within the *open hand prone family*. They all share the semantic theme of "halt" or "stop". However, there is a difference introduced by the verticality or horizontality of the gesture. If the wrist is bended so that the hands are in a vertical position, the speaker expresses a wish to stop something, which can be his own course of action, the interlocutor's course of action or the course of action shared with the interlocutor. The gesture is therefore performative. On the other hand, if the gesture is realised with palms facing down and a horizontal lateral movement of the hands, the meaning expressed is not performative but descriptive. The speaker is no longer expressing a wish to stop something but describing a process in which something is actually stopped. These gestures are used to express that a course of action is impossible, to describe something as complete (implying that no further action is necessary), together with universal truths (implying that no exception is possible) and with extreme assessments (implying nothing else can be as good, bad, etc.). It is noticeable how, in the last three uses, the gestures make explicit what the words only leave to infer.

The last family suggested by Kendon is the *open hand supine*. These are gestures made with the palm of the hand facing up. Their semantic theme is "offering" or "willingness to receive something". Within this family we find three gestures. The first is the p*alm presentation gesture*, by means of which the speaker presents something for inspection. Examples in which this gesture is found are: explaining the meaning of a word or phrase, setting the scene for a narrative, concluding, summarizing or commenting on something. As Kendon notes, all these examples involve clarifying or commenting on something previously said or going to be mentioned. For this reason, they can be taken as a sign of meta-discursivity. The second gesture in this family is the p*alm addressed gesture*, in which the palm facing upwards is directed towards somebody or something. It is used for example when acknowledging the source of what is being said, or to indicate that something mentioned is correct. This gesture corresponds to what Bavelas et al. (1995) call interactive gestures. Finally, in the *palm lateral gesture* palms move away from the frontal space of the speaker. This gesture differs in the semantic theme of the family, since it conveys withdrawal or non-intervention. It is found in situations in which speakers show unwillingness or inability to intervene or position themselves with regard to something, such as a question for which there is no answer, a situation that cannot be helped, or when something is obvious or redundant (implying that nothing further can be said about it).

Finally, although Kendon does not treat pointing gestures as a family in itself, he remarks how similar differences in meaning can be spotted in different ways of pointing. When pointing with the index finger and palm down the effect is that of individuating the referent, and we indicate that the exact location is relevant. When pointing with the palm up and hand open, we offer the object referred to or present it for inspection. Note how this is coherent with the rest of the gestures in the Open hand supine family. When we point using an open oblique hand we indicate a negative attitude towards the referent, and finally, pointing with the thumb conveys the idea that the exact location is not important.

3.2.3.4 The contribution of gestures and head movements to the multimodal ensemble

In this section I want to highlight the crucial contribution of gestures and head movements to the multimodal ensemble. To this aim, I zoom in on the tight relation between the modes I analyse in this study (words, intonation, gestures and head movement) to show how they closely intertwine in a multimodal ensemble. I undertake a literature review that will show how even approaches that are not multimodal acknowledge this strong connection between modes.

The relationship between gestures and speech remains a debatable topic. Goodwin (1986) investigated this relation and concluded that gestures can be said to have a twofold contribution: being part of the message and a resource for organising orientation and direct attention. He pointed out how speakers can vary the degree to which a gesture is essential to understanding, making some gestures particularly salient, and at the same time listeners monitor for cues that tell them when gazing at the gesture is essential.

McNeill (1985, 1992) argues that gestures and speech present a single cognitive representation, and he bases his claim on three main reasons: the synchrony that can be observed between gesture and speech (a topic that I will discuss in the next paragraph); the fact that gesture and speech are affected in similar ways when there is neurological damage such as aphasia; and finally, the fact the both undergo similar stages in development from concrete during childhood to increasing abstraction in adults. Further empirical proof of this gesture-speech relationship comes from studies that focus on the contribution of gestures for word retrieval (tip-of-the-tongue phenomena, when the speaker cannot find the word), which show that gestures tend to precede speech and actually facilitate speech production (Butterworth and Beattie 1978; Morrel-Samuels and Krauss 1992). Beattie and Shovelton (1999) follow up on this line of research to show that iconic gestures do contribute to the verbal communication of certain types of meaning, in particular position and size of objects in a description, but they do not rule out other potential contributions of these gestures to the message, which further research can eventually reveal.

Analysing the exact contribution of gestures to the message, McNeill (1992) found out that hand gestures can be semantically and pragmatically co-expressive (i.e. they can express the same meaning as the verbal message) or complementary (i.e. be partially overlapping with the verbal component of the utterance), with words and gestures expressing different aspects of meaning. He provides examples in which words indicate an action and gestures specify the instrument used to carry out that action. His study of hand gestures revealed that they are so closely linked to language that they can be synchronised with speech. He distinguishes three phases in the realisation of a gesture: preparation, which usually anticipates the verbal expression, the stroke, which precedes or is synchronous with the tonic syllable, and the retraction. He suggests three synchrony rules:

i) Phonological synchrony rule: stroke precedes or is synchronous with the prominent syllable of the utterance;
ii) Semantic synchrony rule: gestures are synchronous with words that share their meaning;
iii) Pragmatic synchrony rule: gestures are synchronous with words that share their same pragmatic function.

Later, McNeill et al. (2001) conducted a study which concludes that the organization of discourse is inseparable from gesture and prosody: the three components are different sides of a single mental-communicative process. They explain this relation using the concept of *catchment*:

> A catchment is recognized when two or more gesture features recur in at least two (not necessarily consecutive) gestures. The logic is that the recurrence of an image in the speaker's thinking will generate recurrent gesture features. Recurrent images suggest a common discourse theme. In other words, a discourse theme will produce gestures with recurring features. (McNeill et al. 2001: 10)

They argue that, at least in the case of discourse concerning spatial description, gestures seem to have the leading role in the discourse-gesture-prosody triad. For example, they find that fall and rise tones are determined not only by the organisation of discourse but also by iconicity (description of a physically high place prompts rising tone).

Goldin-Meadow (2003) sides with the notion of a tight cognitive relationship between gesture and speech, and provides extensive proof for this. According to her, gestures can convey attitude, regulate interaction and express part of the content of the message. It is in this last function that she is most interested. As proofs of the tight connection of gesture and speech as part of one single cognitive process she points out how we keep doing gestures even when nobody is watching, and how even blind people gesticulate. In addition, gesture is known to halt during stuttering (Mayberry and Jaques 2000). Furthermore, and despite the existence of some universals in gesture that are not affected by language (i.e. use of gesture to help plan speech) (Aboudan and Beattie 1996), gestures do change depending on the language we use. Nicoladis et al. (1999) found that English-French bilingual children change their gestures when they change languages. Özyürek and Kita (1999) also noticed how the type and timing of gesture can be influenced by the language. In particular, the way a language codes the world determines the gestures we use. For example English tends to be more precise with motion verbs regarding path and manner, with the consequence that English speakers make more precise gestures when they use these verbs to include details about manner and direction of moving.

However, as Goldin-Meadow (2003) acknowledges, gestures and words encode information in very different ways and work very differently as representational resources: words are a codified, recognisable system with hierarchically organised segments, while gestures are not; words depend on the linearity of time, while gestures use space and movement, and can present more information simultaneously. To shed more light on how these two very different resources can work together as one single cognitive process, she suggests the idea of a continuum based on the level of information overlap between gesture and speech (similar to Kendon's continuum, and drawing on McNeill's idea of semantic and pragmatic co-expression and complementarity). On the one end, we have gestures that elaborate on a topic that has already been introduced in speech (a gesture speech-match, or co-expression in McNeill's terms) and, on the other end, we have gestures that convey new information that has not been said with words (a gesture-speech mismatch or complementarity according to McNeill). She explains that the information that is expressed only with words can also be expressed through gestures, but the other way round is not always possible. In other words, mismatches usually indicate that the information that is conveyed only through gesture is not easily articulated in words. Goldin-Meadow (2003) actually provides evidence from previous literature showing that both deictic and iconic gestures convey an important part of the message. In the case of deictic gestures Thompson and Massaro (1986) show how they help in the perception of ambiguous sounds, and how listeners tend to rely on the information conveyed by the gesture in cases of mismatches. For iconic gestures, Riseborough (1981) proved that the understanding of descriptions of objects is impoverished when iconic gestures are removed. Beattie and Shovelton (1999), discussed above, is another study contributing to this body of evidence.

Another interesting issue raised by Goldin-Meadow (2003) is the way a mismatch is understood, or, in other words, how do we interpret a message when gestures and words are not doing the same thing. Research shows that when addressees are faced with these situations the cognitive load may be increased: they have two pieces of information to decode coming from different resources. Indeed, it seems that addresses gather less information from speech when it is accompanied by gesture than when it is not, but this does not mean that gesture has a negative effect on communication: when we add the information obtained from gesture with the information we get from speech, the total amount of information remains the same, only that it comes from different modalities. Goldin-Meadow (2003) suggests that gestures not only help understanding, but they also help the encoding process of the message. She hypothesizes that one factor causing gesture is cognitive load, and that speakers use gesture to reduce this cognitive load. As she puts it, “gesture and speech form an integrated and, indeed, synergistic system in which effort expended in one modality can (at times, but probably not always) lighten the load on the system as a whole” (Goldin-Meadow 2003: 166) Evidence shows, however, that gestures do not always help (i.e. sometimes their presence does not make any difference). A reason for this might be that the effectiveness of a gesture depends on the degree and manner of overlap with speech. Taking this argument to the multimodal arena, I would argue that the same applies to the interaction of different modes (not only gesture and speech) in any multimodal ensemble.

Kendon (2004) describes the gesture-speech relationship as two kinds of expressive resources that are components of a final product. He points out how gestures can contribute in three main ways: combine with speech, relate to the discursive situation or constitute another layer of meaning. The three are explained in the following paragraphs.

Gestures can combine with speech. For example gestures can make the meaning more precise, they can become the object of a deictic reference, they can serve as an illustration, they can signal the type of speech act being used (e.g. an expressive, a command, etc.), or they can mark features of the discourse structure (e.g. beats).

Gestures can relate to the discursive situation: there are cases in which gestures do not refer to the verbal component they accompany but rather to the dialogue this verbal component implies, that is, they can anticipate potential moves from addresses and refer to potential responses. For example an OHP gesture can be understood as “I will not accept any counterarguments”.

Finally, gestures can constitute another layer of meaning: a gesture can add an additional layer of meaning to the utterance and constitute a move or a speech act by itself, in this way making explicit something that could only be inferred from the words. This would be the case of an assertion or expressive utterance that is turned into an invitation when the appropriate gestures are added. For example, an open hand supine gesture added to the assertion “Talk to me during the coffee break” can add the meaning “if you want to” or “you are free to do whatever you choose, I am not going to intervene” (my own example).

Kendon also coincides with McNeill in distinguishing three parts in a gesture: a preparation, a stroke and a recovery (equivalent to McNeill’s “retraction”). He uses two basic units of analysis: the *gesture unit*, which goes from relaxation moment to the next relaxation moment, and the *gesture phrase*, which includes the preparation and the stroke (which is the one I apply in this study). It is this stroke that is synchronous with speech, and Kendon notes how, if necessary, pauses are included or gestures are prolonged to force this synchrony. He also points out that gestures tend to be revised or repeated as speech is repeated or rephrased, which serves as further proof that they are integral parts of the utterance. Furthermore, the same verbal utterance can achieve different communicative effects only by adding or changing accompanying gestures.

Zappavigna et al. (2009) add a SFL perspective to this issue of the gesture-speech relation. Although Halliday (1985) does not initially consider gesture as part of the system of language (because it cannot be systematically divided into constituent units), Zappavigna et al. (2009) side with Kendon’s approach and they think it accords with SFL stance. SFL understands this gesture-speech relationship in terms of *coupling*, which is the co-selection of functional features. A recurrent co-selection of features that contributes to a rhetorical strategy is a *syndrome*. In their study they look at how the coupling of gesture and speech contributes to meaning and they identify the following scenarios:

i) Gestural pitching: movement up or down of the gesture as intonation rises and falls;

ii) Gestural beating: kinesic analogue of rhythm, adds salience;

iii) Gestural articulation: analogous to phonological articulation, can realise deixis and representation.

Kress (2009, 2009b), on the other hand, objects to the idea of considering speech and gesture as integral parts of a single cognitive representation. He argues that speech and gestures are two different modes with different materialities and therefore different affordances. If they work together towards a specific communicative goal it is because they are orchestrated in a multimodal ensemble. However, both positions may not be so different if we consider how Kendon (2004) describes the gesture-speech relationship as two kinds of expressive resources that are components of a final product. Both agree on the basic assumptions, but multimodality broadens the concept of *utterance, cognitive representation* or *final product* to include other semiotic resources beyond words and gestures and refers to it as *multimodal ensembles*.

Concerning the contribution of head movements to the multimodal ensemble, previous literature has suggested some interesting perspectives, but seems to have concentrated on head nods. The relation of head nods to prosody has long been acknowledged. According to Kendon (1972) head nods correspond to limits of prosodic phrases and they contribute to move the discourse forward. Hadar et al. (1983, 1984, 1989) note how speakers tend to nod during speech and keep quiet during silences, and how nods tend to be synchronous with peaks in loudness. They further notice that head movements seem to have a motor function, helping to channel articulatory energy. For example, they tend to begin slightly before the speech onset because they help initiate speech, just as a conductor makes a gesture to get the orchestra ready. They also become faster as the pitch increases. More recently, Aoki (2011) has suggested that nods complement and heighten the monitoring and regulating function of gaze. They are a cue that the addressee is expected to do something, but by themselves they cannot clarify exactly what. Therefore, they are almost always used in combination with other modes, but they are not redundant because they are intensifiers when there is misalignment from the recipient. McClave (2000) claims that nods prompt backchannel feedback (verbal or non-verbal). Aoki, on the other hand, claims that they are more than backchannels: they are a request for realignment, a request to complete an adjacency pair (in CA terms). A proof of this is that nods are more frequent when there is mutual gaze. Like Kendon, she finds that most nods occur at the end of tone units, but not always, and depending on their position they can prompt different responses. When they occur in turn final position they convey “please respond accordingly, complete the adjacency pair”. When they occur in turn internal prosodic boundary they are mostly interpreted as “please backchannel or display recipiency”. Finally’ during production of prosodic units they express emphasis.

In conclusion, research on gestures and head movements used in oral discourse reveals how they constitute an integral part of the message, as much as paralinguistic features do. Gestures and head movements ensemble with the verbal component and with paralinguistic features, apart from other semiotic modes, in order to create a coherent meaning. Dismissing the importance of kinesic and paralinguistic features is missing much of what happens in oral communication. The next section of this chapter will review previous studies on oral academic and business discourse, focusing on the extent to which the role of intonation, gestures and head movements has been considered.

3.3 The role of intonation, gestures and head movements in conference presentations, research dissemination talks and product pitches

In this section I provide a review of literature that will help me highlight the relevance of the paralinguistic and kinesic modes I include in my study (i.e. intonation, gestures and head movements). My aim with this review is to support my view that these modes play a crucial role in the expression of persuasion, which makes them equal in importance to words and by no means ancillary to them.

Regarding the use of gestures in CPs, Hood and Forey (2005), in their multimodal study of introductions of plenary presentations, showed how gestures express attitudinal meaning and encourage alignment between audience and presenter. Räisänen and Fortanet-Gómez (2006) tackled the use of nonverbal communication in conference presentations, looking at gestures, gaze, face expression and posture. They concluded that discussion sessions display a greater variety of kinesic features than the monologic presentation. Querol-Julián (2011) and Querol- Julián and Fortanet-Gómez (2012) provide a detailed study of the multimodal expression of evaluation in discussion sessions of specialized conference paper presentations. They find that gestures in Q&A sessions show stance and are frequently face-saving (also noted in Fernandez-Polo 2014). They conclude that kinesic and paralinguistic features have an important role in the meaning-making process during the Q&A session of CPs. They have a pragmatic and modal function, showing the speaker’s attitude or intensifying the evaluative meaning expressed in words.

A particularly detailed multimodal study of the academic presentation, including both kinesic and paralinguistic features is provided by Rendle-Short (2005, 2006). She shows how speakers indicate orientation to the audience through positioning, gaze and action. For example, the body usually faces the audience during topic talk, to show full engagement.

Another way of showing engagement is looking at the audience, sometimes intermittently as attention is also paid to visuals. A third way is gesticulating with hands, in particular iconic and metaphoric gestures that visually represent part of the referential content, and beats that mark the rhythm of discourse and are usually synchronous with stresses. Body orientation, gaze, gesture, also show that the speaker is no longer engaged with the audience during a pause between sections. Usually speakers avert gaze and hands return to their home position (Schegloff 1998: 542). The body may also be averted but only partially, to show that this disengagement is only temporary. Frequently, this is followed by doing some tasks like moving on to the next slide. When speakers show engagement again the audience knows that talk will be resumed. Interestingly, the audience only interrupts during these disengagement moments.

Another relevant finding of these studies is that speakers often use kinesics, prosody (reduced pitch or volume) and their actions (e.g. changing the slides) to mark the macrostructure of their discourse, in particular, to announce a pause or a topic transition. Some markers of a coming pause between sections are:

i) quieter voice;
ii) faster pace;
iii) fall intonation;
iv) words (reformulations, summaries, discourse markers);
v) averted gaze;
vi) change in body position;
vii) start an action (e.g. changing slides).

Rendle-Short (2005, 2006) also identifies the following pattern for resuming talk after a pause:

Pause + (pre-beginning) + orientation + (uuhmm) + (pause) + topic talk

The pre-beginning is a way of showing a new turn will be initiated, e.g. opening the mouth, directing gaze, clearing the throat, using discourse markers, etc. During orientation speakers direct attention again to the topic. The talk is always initiated with louder volume, raised pitch, and slower pace. On the other hand, non-verbal behaviours indicating re-engagement may be slightly delayed until after the orientation and the beginning of the topic talk. A slightly modified home position might be used to signal the speaker is still in the orientation and has not started the topic talk yet, and no visual representations with gestures are done until the speaker begins with the topic talk.

Finally, Rendle-Short (2006) also notes that pointing in academic presentations implies three steps, that are enacted multimodally:

1) Setting an expectation that the image is relevant to the talk;

2) Talk and non-verbal action invite the audience to direct attention to the image (or part of it);

3) The relevant part of the image is incorporated into the talk, and a deictic gesture is performed, not necessarily at the same time but the previous preparation in the two former steps makes this potential time mismatch unambiguous and unproblematic. Sometimes, however, speakers either slow down speech or hold a gesture until the appropriate words arrive to make them closer in time and unambiguous.

To sum up the points made so far, previous research on CPs has proved that speakers fulfil their communicative aims through a skilfully orchestrated use of words and paralinguistic and kinesic means. In particular, gestures and intonation serve as signposts to walk the audience through the presentation, making the message easy to understand. They also help achieve a desirable interpersonal relationship, showing (dis)engagement with the audience and facilitating the expression of affect and evaluation.

To my knowledge, very few studies deal specifically with the role of intonation, gestures and head movements in research dissemination talks or product pitches. Some corpus-based studies on business English consider intonation. This is the case of the studies derived from the Hong-Kong Corpus of Spoken English, developed at the English Department of the Hong Kong Polytechnic University. This audio-corpus is divided into four subcorpora: conversations (natural conversations collected in a wide range of social settings), business (including job interviews, informal office talk, service encounters, presentations and Q&A sessions, among others), academic (students' presentations, lectures, seminars, tutorials, etc.) and public (interviews, speeches, press briefings, etc.). It includes native and non-native speakers, and it is the only corpus that, to my knowledge, has been transcribed orthographically and prosodically applying Brazil's Discourse Intonation (Warren 2004). Cheng (2004) provides an interesting example of the important role of intonation in service encounters based on the study of this corpus. She shows, for example, how a hotel receptionist's use of intonation results in a too insistent and pressing discourse that does not suit the communicative situation nor is in line with the mission statement of the hotel as a company.

Mc Gregor et al. (2016) looks at the use of intonation in TED talks, and finds that intonation strongly contributes to convey the speaker's attitude towards both the content and the audience of the talk, and proposes the use of TED talks as valid material to teach intonation. Tsai (unpublished) compares the prosodic features of TED talks and university lectures. Considering TED talks as examples of effective speaking (judging from their popularity), the study wants to probe into which prosodic features make them so effective. The results show that TED speakers give talks that are more dense (i.e. less silent moments), speak with a deeper voice (i.e. in a low pitch range), and have a more consistent flow of delivery (loudness is more consistent throughout). Since these features are not easy to achieve, they are a sign that TED speakers tend to be well prepared in the art of public speaking, and this probably sets them apart from lecturers. With specific reference to product presentations, Lowe and Haws (in press) is a marketing study that investigates the role of pitch in product perception. They conclude that the pitch in a voice-over or background music can affect the way consumers perceive a product: a lower pitch conveys a larger product size. Finally, Niebuhr et al. (2016) analyse the use of prosodic features in Steve Jobs' product presentations, to find out how intonation contributes to project a charismatic figure. With the aid of specialised software, they reveal important characteristics that make Jobs' public addresses stand out from the rest. For example, Jobs' pitch range is particularly high for a male voice, and he produces a particularly varied pitch contour throughout his presentations. He also varies loudness, and, in comparison with other public speeches, produces shorter tone units in a relatively fast and fluent speech rate. In addition, this use of prosody frequently co-occurs with emphatic words, that are elongated and pronounced louder for emphasis.

Regarding the role of kinesics in Business English, Hendon, Hendon and Herbig (1996) approach it from a cross-cultural perspective. They talk about non-verbal behaviour as a generic term, enclosing body language or kinesics, which in turn comprises gestures, body movement, facial expression and eye contact, and vocalics, which is every vocal activity that is not verbal, such as tone or volume. They point out how interpretation of messages depend more on the non-verbal than on the verbal, and that non-verbal messages seem to have more credibility. Because of the relative importance of non-verbal behaviour, misunderstandings can create what they call "cross-cultural noise" and hinder communication. On the other hand, Chu et al. (2005) studied the perceived importance of the influence of proxemics, kinesics and physical location in the development of cross-cultural negotiations, by means of questionnaires distributed among business actors with negotiation experience. Most of their findings were in line with existing non-verbal communication literature. For example, negotiators tended to agree that they interpret a kinesic sign such as moving closer or leaning forward as a marker that what is being said is important, and acknowledged that they were keener on cooperating with a person who smiles.

Filliettaz (2004) provides a very interesting multimodal study of service encounters. This author argues that in any communicative action there is a praxeological process, made up of collective goal-directed actions, and a communicative process of intercomprehension, and shows how speech and gestures can have different roles in each. He distinguishes four different scenarios. The first is gesture as co-verbal action, in which gestures co-occur with the linguistic utterance contributing to the communicative process. Words still carry the bulk of the interaction. The second is gesture as communicative action, in which gestures are a direct instrumental contribution to the social action, and also have a communicative value, because they are the response to a speech act that enables conversation to go on. The third is speech as cogestural communication. In this case gestures are direct instrumental contributions to the social action and speech co-occurs to make this gesture interpretable, but gestures carry the bulk of the interaction. Finally, the last scenario is gesture as autonomous action. In this case there is a mismatch between gesture as instrumental act and communicative symbol. Gesture and speech can co-occur, but the gesture is not contributing to the communicative process but to the social action (i.e. a seller can hold out his hand to the customer to collect payment and close the selling encounter while saying thank you).

The study of spoken Business English reveals that communication (verbal and non-verbal) is used to construe business relationships (Boden 1994; Poncini 2002). Indeed, interpersonal meaning has been shown to have the same importance as transactional meaning in business contexts (McCarthy and Handford 2004). Because intonation and kinesics play an important role in the expression of interpersonal meaning, their contribution in business contexts needs to be considered. This is especially true of business situations that demand persuasiveness. Intonation, gestures and head movements play an important role in achieving a persuasive message, especially when this persuasion is intended to be subtle and indirect (Bamford 2007). Further proof of the relevance of kinesic and paralinguistic features is the fact that interpretation of messages relies more on them than on verbal features, attaching more credibility to the former (Hendon et al. 1996). Likewise, the negative consequences of an inappropriate use of intonation are exemplified in Cheng (2004), and the important role of gestures in communicative actions both as instrumental acts and communicative symbols is revealed in Fillietaz (2004).

Therefore, the contribution of intonation, gestures and head movements in inherently persuasive genres cannot be ignored. Through intonation, gestures and head movements speakers can direct the interpretation of the message according to their needs. For example, through the choice of tone they can present parts of the message as common ground, ruling out other possible interpretations. Using prominence, they can direct attention to the parts of the message they are most interested in, and imply that the rest of information that is out of focus is common knowledge and does not need to be discussed. Using prominence and tone persuaders can manipulatively "impute particular knowledge or views to the other participant or participants" (Wells 2006: 115). Likewise, through the use of key speakers can lead the listener to certain desired interpretations, for example that something is the necessary consequence of something else. Termination proves a valuable tool to direct reactions: it can signal whether the listener is expected to consider something carefully and make a decision or simply accept something as already negotiated. Likewise, gestures constitute another rich source of information for the addressee, and have a strong influence on the way the message is interpreted. Gestures can convey certain information more economically and efficiently than words, for example position, movement, size or spatial relations (McNeill 1992; Kendon 2004). Gestures and head movements can also contribute to highlight what the speaker considers most important (Kellerman 1992). They also reveal the attitude of the speaker towards what is being said, and serve as cues for expected responses, even when this is not mentioned explicitly with words. This will undoubtedly be exploited in situations in which social conventions prevent interactants from expressing attitude and emotions openly. For example criticizing a competitor product overtly, contradicting an opinion directly or blatantly preventing a counterclaim may be inappropriate in many academic and business contexts, but these meanings can be conveyed through gestures and head movements more safely in social terms (Arndt and Janney 1991).

In this chapter I have highlighted the multimodal nature of the three genres which are the object of this study: CPs, research dissemination talks and product pitches. I have focused on the role that words, intonation, gestures and head movements have as semiotic modes in these genres. I have reviewed different approaches to intonation and several classifications of gestures and head movements than can provide a useful methodological toolbox for a multimodal analysis of these genres. In the following chapter, I draw on the literature that I have discussed in chapters 1, 2 and 3 to distil a methodological framework that will enable me to approach the study of persuasive strategies in these three genres from a multimodal perspective, focusing on words, intonation, gestures and head movements as semiotic modes.

CHAPTER 4. FROM THEORY TO PRACTICE: MY METHODOLOGICAL FRAMEWORK FOR A MULTIMODAL ANALYSIS OF PERSUASIVE STRATEGIES

In this chapter I will link the theoretical construct presented in the previous three chapters with the multimodal analysis of persuasive strategies that I propose in this thesis. I will show how the theories discussed in previous chapters about discourse, genre, persuasion, intonation and kinesics can help me build an appropriate methodological framework to approach persuasive strategies in oral genres from a multimodal perspective. To this aim, throughout this chapter I will explain a range of key terminology drawing on previous literature in the aforementioned fields. The first section of this chapter states my research questions. In the second section I discuss the use of combined methods for the analysis of oral genres. I then turn to the analysis of words, gestures, head movements and intonation as modes. Finally, the chapter concludes with a proposal for an analytical framework that can facilitate the multimodal analysis of persuasive strategies in oral genres.

4.1 Objectives of the analysis and research questions

This thesis presents a contrastive study of the use of persuasive language in three genres, viz. conference presentations, research dissemination talks and product pitches. The rationale behind the selection of these genres in particular is the likelihood of interdiscoursive and intertextual relationships. In other words, I use a multimodal and ethnographic approach to shed light on common traits and differences among these three genres regarding their use of persuasive strategies, which can in turn help refine their definitions in further research and improve the way they are taught in LSP. The three types of presentations included in the study are professional oral genres which are widely used by their respective discourse communities (i.e. academia and the world of business), and this makes them crucial tools for professional success. In addition, the three genres show important parallelisms, as discussed in chapters 1 and 2. They all involve speakers addressing an audience in an extended turn of speaking, and they present some novelty: either scientific knowledge or a product. They also share a similar purpose: to convince the audience that what is presented is valuable. For this reason persuasive language can be expected to play an important role in all of them. I consider this overlapping in communicative purpose as an important, strong link between them. In spite of this, the different communicative situations in which they are used will also prompt important differences that could determine their generic nature.

In this thesis I adopt a multimodal approach for the study of persuasive strategies in the aforementioned presentations, focusing on the way in which different semiotic modes work together to create a persuasive message. In particular, I am interested in identifying the verbal-linguistic, kinesic and paralinguistic features of the various strategies that speakers resort to in conference presentations, research dissemination talks and product pitches in order to design a persuasive message. By verbal-linguistic persuasive strategies I refer to lexico-grammatical, semantic and stylistic resources such as narratives, or lexical evaluation that speakers can use to boost persuasion. By paralinguistic features (Trager 1958) I mean voice qualities (e.g. pitch) and vocalizations (noises different from the sounds of language). The term kinesics, as used by Birdwhistell (1952), refers to body movements and postures. The present research will focus specifically on one paralinguistic feature, i.e. intonation, and two kinesic features, i.e. head movements (nods, shakes and lateral movements) and gestures. These features have been selected because of three reasons. The first reason is that they are rather conspicuous in oral discourse. Unlike other semiotic modes that speakers can choose to use (e.g. visual aids), they are virtually unavoidable. Indeed, it is actually very difficult for speakers to avoid making gestures or using a particular tone contour as they speak without jeopardizing the communicative effectiveness of their messages (Butterworth and Beattie 1978; McNeill 1985, 1992, 2008; Morrel-Samuels and Krauss 1992; Brazil 1997; Beattie and Shovelton 1999). The second reason is that these modes have received little attention in multimodal studies in comparison to other non-verbal modes such as visuals (e.g. the use of power point slides). Finally, the third reason is that research on the pedagogy of oral presentations shows that speakers tend to overestimate visual aids and underestimate delivery skills (which include the use of intonation and kinesics) when they self-assess their presentations (Campbell et al. 2001), which makes a focus on these features something desirable.

In addition to an MDA approach, I also make use of ethnographic techniques such as interviews and observation sheets to gather contextual data and gain insight into speakers' views about the genres and the broader communicative event in which they are embedded, as well as their communicative intentions. The combination of multimodal analysis and ethnography is an increasingly common practice (see for example Cowan 2013). In this particular case it helps me in two main ways: first to interpret the video data and select material for fine-grained analysis, and second to compare the results of my analysis with the speaker's own opinions (see section 4.2).

As discussed in Chapter 2, according to authors such as Fuertes-Olivera et al. (2001), O'Keefe (2002), Perloff (2003), Virtanen and Halmari (2005) and Dafouz-Milne (2008), persuasive messages tend to be more effective when:

i) Speakers have credibility and the audience can identify with them: addressees are more prone to be persuaded by speakers they can trust.

ii) The message is easy to understand: when the effort to understand the message is too great, addressees will tend to stop trying and eventually they will stop paying attention.

iii) The content is relevant for the audience: addressees are less likely to be persuaded by messages that do not appeal to them or they feel as not applicable to them.

iv) The message is made memorable: addressees are more likely to be convinced by a message that they can easily remember without making a conscious effort.

v) The message is innovative and surprising: persuasion becomes more effective the more unexpected it is. Forewarning is a persuasion killer, because it activates the addressee's mind and stirs potential counterarguments in advance, or in other words, activates elaboration-based persuasion (see Chapter 2).

vi) The message is perceived as not imposed, but inferred. Addressees should feel they have reached their own conclusions, even if they have been guided towards them.

In addition, there are other *persuasive features* that are more *specific to each type of presentation*. Research on academic discourse suggests that for researchers to communicate persuasively they need to find "a judicious balance of tentativeness and assertion, and the expression of a suitable relationship to one's data, arguments and audience" (Hyland 1997: 440). Engaging with an audience implies recognition of the

> need to adequately meet readers' expectations of inclusion and disciplinary solidarity (...), the writer pulls readers into the discourse at critical points, predicting possible objections and guiding them to particular interpretations with questions, directives (mainly obligation modals and imperatives) and explicit references to shared knowledge (Hyland 2006: 33).

Academic presentations, therefore, can be considered to become more persuasive when the speaker 1) proves the originality and value of the piece of research described, 2) proves that it has been done according to disciplinary rules, 3) shares practices and admits responsibility for decisions and 4) becomes part of the whole communicative event in which research is being disseminated (Hyland 2006).

On the other hand, research on presentations in the area of business (Bamford 2007, 2008; Boker 2006, 2007) highlights some specific persuasive features that are particularly frequent in this genre, such as 1) facts presented as given and already agreed upon even when there is no clear supportive evidence, 2) minimized negative aspects and highlighted positive aspects, 3) direct addresses to the audience 4) narratives and personal testimonies. When applied to a product presentation, these features have the effect of presenting the product as valuable for the audience and feasible from a marketing point of view.

Previous research has also pointed out some *linguistic* strategies that are not necessarily genre-specific and help provide the message with the above-mentioned characteristics to make it more persuasive. Some of them are particularly relevant for the present study and were discussed in detail in Chapter 2:

i) Framing or orientation refers to the way in which information is presented so as to guide a particular evaluation and interpretation of the content.

ii) Narrative techniques in business presentations can ease interpretation and present information as relevant for the addressee, which is expected to enhance persuasion under conditions of high elaboration. Also narratives in past tense during conference presentations are used as a persuasive appeal that helps create rapport.

iii) Examples vs. statistics: the use of examples can contribute to credibility and persuasion.

iv) Language intensity: the use of metaphors, strong, vivid language and emotionally charged words can also contribute to a persuasive message.

v) Stylistic devices such as three-part lists, symmetric structures, contrasting pairs or lexical creativity.

vi) The use of vocabulary from the semantic field of novelty can also boost persuasion.

vii) Direct addresses to the audience can also become persuasive when they are used to prompt a particular response or to build rapport.

In this thesis, I attempt to demonstrate that paralinguistic and kinesic features also play a crucial role in providing presentations with the above-mentioned characteristics (i.e. easy, memorable, relevant, innovative, etc.) to make them more persuasive. In other words, speakers communicate a persuasive message through the use of different modes and not only through words. The specific research questions addressed are the following:

RQ I	**What type of persuasive strategies can be identified using an MDA approach?**
RQ II	**What is the contribution of the multimodal ensemble orchestrated through words, intonation, head movements and gestures to the creation of persuasion in the genres included in the analysis?**
RQ II. a)	Can persuasion be considered multimodal in these genres?
RQ II. b)	How do modes orchestrate in multimodal ensembles to create a persuasive message?
RQ II. c)	What type of persuasive activity (i.e. persuasive efforts on the part of the speakers) is more likely to be realised through the modes which are the object of this study? When, how and how often does it happen?
RQ III.	**To what extent is the use of persuasive strategies genre-specific and to what extent is there overlap between the three genres?**
RQ III. a)	Are persuasive strategies enacted through the same modes in the three genres?
RQ III. b)	Are there strategies which can be considered genre-specific?
RQ III. c)	How do the three genres compare in terms of their use of persuasive strategies?

Research Question I addresses the methodological challenges of approaching persuasive strategies from a multimodal perspective. This type of study requires the combination of different methods (e.g. multimodal analysis and ethnography) and a constant triangulation of results based on different sources (literature, videos and the informants). It also requires a way of classifying or accounting for the persuasive activity enacted through different semiotic modes.

Research Question II aims at probing into the way specific kinesic and paralinguistic modes combine with words in a multimodal ensemble to create a persuasive message. As pointed out before, the modes selected have so far received little attention in the study of academic and professional discourse compared to, for example, visuals, and yet they are quite conspicuous whenever people speak. Indeed, a presentation featuring a flat intonation throughout or no body movements at all loses much of its communicative power. Rather than focusing on how persuasive strategies are realised through words or gestures, I direct my attention to how speakers orchestrate all modes in a unified multimodal ensemble to achieve persuasion.

I have differentiated intonation as a separate mode rather than including it together with words (e.g. in a wider mode that could be called “speech”) because I consider it can fulfil Halliday’s SFL metafunctions (ideational, interpersonal and textual) explained in chapter 1. In addition, I argue that people in general exploit it communicatively to a great extent, even if not so consciously or systematically as words. We are all familiar with the expression “it's not what you say, it’s the way you say it”. Therefore it meets the two requirements suggested by Kress (2010) to be considered a mode: it is fully capable of expressing meaning and it is socially accepted as a way of doing it.

Gestures will be analysed following Kendon (2004) and McNeill (1992) among others (see section 4.3). Particular attention will be paid to the interaction of the gestures with the other modes studied, and how they work together in the expression of meaning. In the case of head movements, my analysis draws mainly on McClave (2000) and Kendon (2002), focusing especially on amplitude and repetition, as well as interplay with other modes. The approach adopted for the analysis of intonation is Discourse Intonation (DI) (Brazil 1985, 1997). As explained in section 4.3, DI looks at the communicative value of intonation by analysing four systems in which speakers can make a range of meaningful choices. Through these choices speakers project a particular understanding of the communicative situation, steer the interpretation of the message or favour certain responses, all of which can prove to be highly persuasive tools.

Research Question III implies a practical application of the multimodal approach suggested in RQ1 to test its usefulness for a contrastive study of the use of persuasive strategies in the three genres. Although persuasion is an important component of the communicative purpose of the three genres, the different communicative situations in which these genres are used are likely to prompt different persuasive strategies from speakers, in an attempt to adapt to their audiences. In this sense, I expect to find both similarities and differences: cases in which the use of persuasion overlaps across genres and cases in which the use of persuasion is genre-specific and therefore different strategies and modes are used in each genre.

4.2 The use of combined methods: video based multimodal discourse analysis and ethnography

This study analyses a corpus of presentations that has been video recorded. Video-based research presents several advantages. For example, it preserves the real temporal and sequential structure of interaction, allowing researchers to keep talk in the context in which it was produced. It also transforms oral discourse from something ephemeral to something durable and researchable by means of a fine-grained multimodal record (Jewitt 2004).

On the other hand, video-based research implies important decisions as to what and who to include or exclude, and how to do this. This will highly shape and condition the subsequent analysis (Widdowson 1998; Flowerdew 2003, 2004; Jewitt 2004). Among these decisions I would like to highlight the following:

i) *Type* of camera: Researchers have to decide whether to use alien cameras (i.e. intrusive in the environment) that may alter the communicative situation and the interaction, or body mounted, integrated cameras that offer 1st person perspective (1st person generated data). For this study, an alien camera was used, since its intrusiveness in the communicative situation was limited given the physical environment. For example the size of the room and the audience helped reduce this impact, and the presence of the camera was not distracting for speakers.

ii) The *number* of cameras: Multiple cameras give a more detailed record, but also more fragmented; a single camera forces the researcher to focus and select research interests. Only one camera was used for this study, because the focus is on the speaker.

iv) The *shooting* can shape and affect interpretation. For example, a moving camera can make a static environment look dynamic. This can affect reliability of data (see next point). In this study, the camera is constantly focused on the speaker and moves only when it is required to keep this person in focus.

v) *Reliability* of data: Despite common belief, pictures do lie, especially when they are decontextualized from the communicative event where they originated (see also next point). In a sense the video is the researcher's view, and for this reason triangulation is important. Researchers cannot observe reality without being part of it and therefore change it (e.g. the participant may look directly to the camera). There are two extreme positions in this sense. On the one hand, video can be considered as the most accurate representation of reality, and the closest to truth possible. On the other hand, video can be considered as a distorting filter that modifies reality. For this study, I tend to side with the latter position. For this reason I decided to be present in all the events recorded and gathered first-hand information about the event using observation sheets. In addition triangulation of results through interviews was included for the sake of objectivity.

vi) *Contextualisation*: Video itself is raw data. Video preserves the materiality of image and sound, but cannot preserve everything that happens during a complex communicative situation and is beyond this materiality (e.g. how participants perceive and interpret what is happening and respond accordingly). It is the job of the researchers to give meaning to this data through their interpretation. To this aim, in this study observation sheets were used during the recording sessions to gather as much data as possible about the context. These sheets contained information about different aspects that helped interpret the persuasive strategies the speaker resorted to. A detailed explanation of these sheets is offered in Chapter 5.

vii) Video as a *source*: It is also important to consider how video recorded data stands in relation to other sources. In the case of the present study, videos are a primary source of information, which is nonetheless complemented with observation sheets and ethnographic interviews with speakers.

There is yet another risk involved in the analysis of data with such a high level of detail as video provides, which is precisely a tendency to focus on minute detail and micro analysis that can lead to de-contextualised results whose relevance in the global picture can be overestimated (Jewitt 2012).

For all these reasons, video data needs to be cautiously interpreted (Schindler 2009) and the results obtained complemented with data from other methodological approaches. In this thesis video-based multimodal analysis is combined with ethnography, i.e. observation sheets during recordings and semi-structured interviews with speakers (see Chapter 5 and Appendix 3). Ethnography as a methodological approach presents a remarkable advantage for the study of persuasive strategies: it makes it possible to gather first-hand information about the communicative situation and about the communicative intentions of the speakers, which is crucial to interpret the data in a more objective way and cross-check initial hypotheses. On the other hand, it also presents some limitations regarding access and reliability of the data obtained. Observations can be obtrusive, as it is the case with video, and interviews are time consuming and require a good predisposition on the part of the informant. In addition, the reliability of the data gathered depends also on the quality of the interview design, the type of questions included and the interviewer's skills (Rubin and Rubin 1995).

Yet, the use of both methods in combination can help to overcome the limitations of each methodological approach used separately (Jewitt 2012). Moreover, this combination of methods can prove useful to inform important decisions as the analysis unfolds, such as:

i) what to record;

ii) how to screen data (e.g. what should be considered persuasive, what material to select for fine-grained multimodal analysis);

iii) how to interpret the data gathered;

iv) how to triangulate results;

v) how to actively involve participants in the study and take into account their opinion when interpreting results.

For all these reasons, I decided to take on a combination of methods for the present research.

4.3 Analysing words, gestures, head movements and intonation as semiotic modes

The analysis presented in this thesis focuses on four semiotic modes: words, gestures, head movements and intonation. In this section I describe my approach to each of these modes.

As mentioned in Chapter 3, when I refer to "words" as a mode I use the term in a broad sense to include all the verbal-linguistic material (as opposed to paralinguistic or kinesic aspects) used by speakers. Therefore, under the umbrella term "words" I include: 1) lexical resources (e.g. a specific word choice); 2) grammatical resources (e.g. a particular choice in sentence structure); 3) stylistic resources (e.g. the use of a three-part list); and 4) register (e.g. the use of slang).

It must be noted that in most cases there are no clear-cut divisions between these four categories. For example, when one of the speakers in a product pitch describes his product as "safer, faster and easier to use than waxing kits" I find that there is an interesting word choice in the three comparative forms of the adjectives that express evaluation, but I am equally interested in the fact that the speaker is using a three part list that provides the text with memorability.

My approach to words is totally data-driven. I have not searched my corpus for specific instances of lexical choice, stylistic devices or particular types of sentences. This is a conscious research decision to avoid a natural tendency as a linguist to focus my analysis on words and limit the analysis of the other modes to instances of correlations. For example, I might find that a three part list is synchronous with an enumerating gesture, that the intonation used divides the three items of the list in three tone units with parallel intonation, and that three head nods are used to index each item, which would be a prefect correlation between the four modes. However I consider that the effect of gestures, intonation or head movements can be interesting in itself, even when no stylistic device is synchronously used or when there is nothing remarkable in the choice of words.

As my description of the other modes that follows in the next paragraphs will reveal, my approach to words is, in comparison, much less systematic. This is again a research decision that is due to the fact that speakers are more linguistically literate and more conscious of their use of words that the other three modes. This means that in order to discuss their use of intonation or gestures (as I do in the interviews described in Chapter 5) some systematic classification was helpful, but it would probably have been limiting in the case of words (for example by imposing a closed list of options from which to choose to describe and classify something that they can explain themselves much better in their own way).

For the analysis of gestures and head movements I draw on previous research by McNeill (1992), McClave (2000), Kendon (2002, 2004), Bavelas et al. (1995) and Querol-Julián (2011). I focus on the aspects summarized in Table 4.1, which are discussed in detail in Chapter 3.

Table 4. 1 Analysis of gestures and head movements

Synchrony McNeill (1992)	Phonological
	Semantic
	Pragmatic
Interaction with other modes (Kendon 2004)	Offers equivalent or complementary meaning
	Relates to the discursive situation and dialogue implied
	Adds a new layer of meaning

Function	Referential (Kendon 2004)	
	Interpersonal (Bavelas et al. 1995)	
	Pragmatic (Kendon 2004)	
	Cohesive (McNeill 1992, Querol-Julián 2011)	
Type	Iconic (McNeill 1992)	
	Metaphoric (McNeill 1992)	
	Beats (Kendon 2004, Bavelas et al. 1995)	
	Deictics (Kendon 2004)	
Gesture family (Kendon 2004)	Precision grip	G-family (finger bunch) = seizing
		R-family (ring)= making precise
	Open hand	Prone (down/ palms facing addressee)= halt, stop
		Supine (up/ palms facing speaker)= offer, willing to receive
	Closed Fist	
Amplitude and repetition (McClave 2000; Kendon 2002)	Normal, high or low	
	One or more repetitions	

As pointed out in Chapter 3 *synchrony* refers to the timing of the gesture in relation to other modes. For this I draw on McNeill's rules that can be summarized as follows:

i) Phonological synchrony rule: the stroke tends to precede or be synchronous with the prominent syllable of the utterance;

ii) Semantic synchrony rule: gestures tend to be synchronous with words that share their meaning;

iii) Pragmatic synchrony rule: gestures tend to be synchronous with words that share their same pragmatic function.

This is interesting to get further insight into how different modes interact in the creation of a persuasive message, in particular how gestures interact with words and intonation. For example from these rules it is possible to deduct a tendency for gestures to co-occur with prominent words in tone units, and with words that share the same referential meaning and pragmatic function.

Interaction with other modes will indicate the role of the gestures within the multimodal ensemble. A gesture can contribute to the meaning expressed through other modes, for example complementing the verbal referential meaning giving a visual representation of a movement described with words. It can also relate to the discursive situation, referring to the dialogue implied. In this case the speaker is focusing on the potential reaction of the addressee and produces a gesture accordingly, for example anticipating a potential counterclaim and refuting it. Finally, as Kendon (2004) points out, a gesture can add a new layer of meaning as a new speech act.

The *function* of the gesture will indicate what its purpose is, that is, whether it is used to represent an aspect of the content (referential), to regulate interaction (interpersonal), to link parts of discourse (cohesive), or whether it is used pragmatically to show attitude or perlocutionary meaning. These functions are by no means clear-cut, and some gestures can perform more than one function at a time. This is the case with *deictic* gestures. They are referential in that they refer to an object under discussion, but pragmatic when they also show the speaker's attitude towards this object or what the addressee is expected to do with this object (e.g. examine it, consider it an example of something, etc.). Another example is *beats*. They can be considered pragmatic in that they serve to guide the understanding of the message, but can also be cohesive when they highlight the structure of discourse. As mentioned in Chapter 3, Bavelas, Chovil, Coates and Roe (1995) have argued that beats are actually interactive gestures, because they ultimately refer to the addressee in different ways. The function of a gesture can be very revealing about the persuasive strategy used. An interpersonal gesture in a presentation, for example, is a sign that the speaker is trying to involve the audience and probably seeking alignment with the views expressed.

The *type* of gesture will reveal its relationship with the meaning it conveys. An iconic gesture represents concrete objects and events, for example a finger ring that represents a circular instrument. Metaphoric gestures represent abstract ideas. Beats are repetitive gestures that, among other things, mark the discourse flow. They frequently accompany prominent words in tone units. Deictic gestures are used to point to something in the environment.

Finally *gesture families* can help group similar gestures according to their kinesic features and build correlations between recurrent shapes and related meanings. As explained in Chapter 3, Kendon (2004) suggests that gestures with similar kinesic characteristics can be grouped into families, each family sharing a common semantic theme, and each gesture within the family adding a particular semantic nuance. A brief explanation of each family as a way of reminder is offered in the next paragraphs.

Two different precision grip families are differentiated. The gestures within the G-family, also called "finger bunch", share the following position: fingers are brought together touching at their tips, forming a sort of bunch. The semantic theme conveyed by these gestures is that of *seizing.*

The R-family gestures all have in common a position of the fingers emulating a ring. They have the semantic theme 'making precise'. These gestures are used when this precision seems necessary to gain agreement, persuade or facilitate understanding.

Gestures realised with an open hand and palm facing down (if the gesture is horizontal) or facing the addressee (if the gesture is vertical) are grouped within the open hand prone family. They all share the semantic theme of "halt" or "stop". The last family suggested by Kendon is the open hand supine. These are gestures made with the palm of the hand facing up, and their semantic theme is 'offering' or 'willingness to receive something'.

To Kendon's classification I have added the closed fist family, including gestures realised with a closed feast. These closed feast gestures were frequent in my corpus, but could not be fully explained using Kendon's suggested families, hence my addition. Figure 4.1 illustrates the analysis of one specific gesture in a research dissemination talk.

"And we have been taught that everybody should be treated the same"

A "stop" sign with open hands, palms facing the audience (open hand prone, vertical) visually represents how society prevents people from considering other options than what she is expressing with words (i.e. "everybody should be treated the same"). This metaphoric gesture is both referential and pragmatic, and adds a sense of disagreement with this statement that is not explicitly said in words.

Figure 4. 1 Example of gesture analysis

Finally, in the case of head movements I consider two aspects as particularly important for persuasion: *amplitude*, that is how wide the head shake or how long the head nod is, and *repetitions* since they can show the speaker's emphasis and stance. For example a wide head shake repeated several times can indicate strong disagreement to what is being expressed with words (i.e. the speaker is talking about views which are contrary to what he or she supports) or to the expected reaction from the audience (i.e. the speaker predicts the audience will deny his or her claim and refutes this potential denial before it actually happens).

One of my hypotheses in this thesis is that gestures can contribute to express meaning in presentations more effectively and subtly than other modes, and for this reason I consider they can be powerful persuasive tools. On the one hand, they can contribute to creating rapport, a sense of inclusivity and shared ground with the audience, and they can subtly communicate meanings which would otherwise not be considered acceptable or polite (e.g. strong disagreement). In addition, given the time constraints and relationship with the audience established in product and conference presentations, they are particularly effective in these communicative situations. For example, a gesture can help reduce the processing effort for the audience by presenting part of the meaning visually.

The analysis of intonation draws on Brazil's discourse intonation (1985, 1997). As explained in Chapter 3, the main tenet of this approach is that speakers make meaningful choices in their use of intonation, that reveal their assumptions about the interaction in the communicative process. Discourse intonation (DI) divides speech into tone units, each one containing one or two prominent syllables. The first prominent syllable in a tone unit is called onset, the second is called tonic syllable. Several tone units together make up a pitch sequence. DI claims that speakers make choices in four systems with which meaningful contrasts can be produced: prominence, tone, key and termination. As a reminder, they are summarized in Table 4.2, together with their meaning addition and their potential contribution to persuasion:

Table 4. 2 Discourse Intonation (adapted from Brazil 1985)

SYSTEM	DEFINITION	OPTIONS	MEANING	POTENTIAL PERSUASIVE EFFECT
PROMINENCE	Prominent syllables are louder and longer	Prominent vs. non prominent	New vs shared knowledge Selective vs only alternative	Debatable vs already negotiated
TONE	Pitch movement	Rise-fall Fall Level Rise Fall-rise	New vs shared knowledge Type of response expected	Debatable vs already negotiated Seek active response vs prompt to concur
KEY	Relative pitch of onset syllables in relation to the onset syllable of the previous tone unit.	High Mid Low	Equating Addition Contrast	Prompt inferences Build rapport
TERMINATION	Relative pitch of tonic syllables in relation to the onset syllable in the same tone unit	High Mid Low	Type of response expected Surprise	Seek active response vs prompt to concur

In conclusion, by looking closely at intonation it is possible to learn about aspects such as which parts of the message the speaker regards as shared knowledge (and therefore agreed upon) and which parts are presented as new information (and potentially open to discussion). It also reveals the reaction that the speaker expects of the interactant or who is in a dominant position in the communicative process. These are all crucial for the understanding of how persuasion is achieved. It must be noted that in the present study combined tones (fall-rise and rise-fall) have not been distinguished. The reason for this research decision is that they mainly convey information about power relationships and in particular dominance, which in the case of the presentations analysed tend to remain rather stable throughout. It was considered that the potential added value of this distinction did not justify the analysis cost at this stage. Figure 4.2 illustrates my analysis of intonation applying DI:

INTONATION:

(Prominent syllables are capitalized, syllables above or below the line represent higher or lower pitch respectively, tone is indicated with arrows next to each tone unit)

1. but of COURSE I am not here to tell you a sTORY ↘
2. I am NOT a story TELler ↘
3. I'm HERE just to SHOW you →
4. that the possiBIlities with SMART textiles ↗
5. are END () LESS →

"Endless" is prominent and split in two for further emphasis.

In unit 2 the low key in "not" has equating effect (I am not here to tell you a story = I am not a story teller)

The high key in unit 3 suggests the opposite of equating: reversal of expectations, little connection with what was said before. In fact the speaker interprets this use of intonation as conveying that it is more important "what I'm here for than what I'm not here for".

In unit 5 the low key again has equating effect (possibilities= endless), while the low termination suggests discourse final and closes the potential argument as to the truth of this statement.

Figure 4. 2 Example of intonation analysis

4.4 Framework for the multimodal analysis of persuasive strategies

The analysis of persuasive strategies proposed in this PhD is framed as a data-driven and cyclical process that draws from three sources:

i) input from previous literature, including:

a) Previous studies on persuasion (Fuertes-Olivera et al. 2001; O'Keefe 2002; Perloff 2003; Virtanen and Halmari 2005; Dafouz-Milne 2008);

b) Methodological proposals for a MDA approach to genre analysis (Querol-Julián 2011; Querol-Julián and Fortanet-Gómez 2014);

c) Work on kinesics (McNeill 1992; McClave 2000; Kendon 2002, 2004);

d) Work on intonation (Brazil 1997; Chun 2002; Pickering et. al 2012);

e) Previous research on professional and academic discourse (Lakoff 1982; Carter 1997; Hyland 1998, 2006, 2009; Bowker 2006, 2007; Bamford, 2007, 2008).

ii) multimodal data (results from the multimodal analysis); and

iii) ethnographic data (results from interviews and observation sheets).

In other words, I triangulate what I observe in the videos (i.e. the results of the multimodal analysis) with what I learn from previous literature, and with the speakers' own observations and reported intentions. In this way, in an iterative process of literature reviewing, multimodal analysis and ethnographic analysis, I obtain data that enables me to propose a list of persuasive strategies realised multimodally (see Chapter 6). The data retrieved from the interviews with the speakers (in particular information regarding the speaker's intentions and aimed effects) help me to support tentative conclusions or reject them, refining the list of persuasive strategies, which becomes the basis for my analysis of persuasive strategies. Figure 4.3 represents the framework for the analysis.

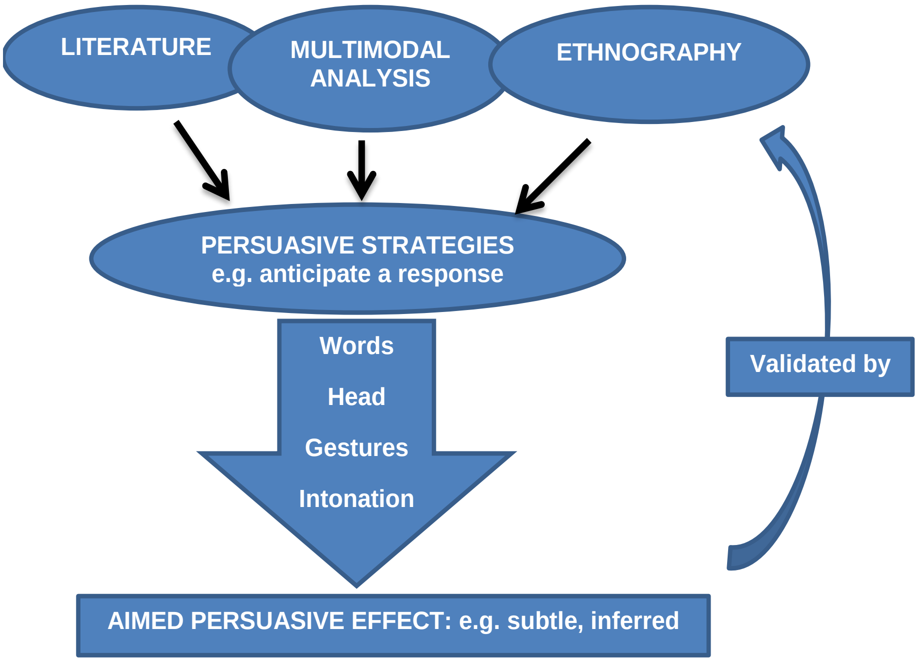


Figure 4. 3 Multimodal analysis framework

A concrete example may help clarify this cyclical process. Input from literature informs me that anticipating responses of interlocutors can contribute to persuasion. For example, when your interlocutor is likely to process a message in detail (i.e. central route to persuasion) it is important to anticipate possible counterarguments and try to prevent them or at least be ready to refute them (O'Keefe 2002; Perloff 2003). The multimodal analysis of my corpus further reveals that speakers seem to anticipate their audience's reactions combining words, intonation, gestures and/or head movements. For example, the speaker in one of the research dissemination talks presented a hypothetical situation: what would the audience respond if she asked a specific question. She anticipated that the audience would be "very silent", and as she said these words (which by the way are prominent) she made a sweeping gesture and nodded her head, in what can be interpreted as an attempt to prevent a counterclaim to this anticipated response she was describing (see Figure 4.4).

Figure 4. 4 Multimodal anticipation of responses

With this information I hypothesize that "anticipation of responses" can be a multimodal persuasive strategy. Later, the interview with the speaker confirmed that she was indeed trying to anticipate responses from the audience, and using also gestures and head nods for this aim: "maybe it was a confirming gesture that actually, yes, it was the right anticipation". During this interview I also discussed my interpretations regarding the aimed effect of this - and other - strategies that the speaker was using. She agreed with me that part of her intention in this case was to subtly guide the audience towards a particular interpretation of the possible response that she is describing as the most likely. A full description of this strategy is provided in section 6.1.1 in Chapter 6.

I believe that one of the main advantages of the analytical framework I suggest is that it makes it possible to study how a particular aspect of the message (in this case persuasion) is jointly expressed through all the semiotic modes included in the analysis, instead of starting from a particular modal realisation and then looking for correlates in other modes. My intention with this is to avoid prioritising any mode in particular and treat all of them as integral parts of a multimodal ensemble. Whether I fully achieve this aim is object of discussion in Chapter 7.

In the following chapter I describe how I have applied the methodological framework presented here to study a corpus comprised of conference presentations, dissemination talks and product pitches.

CHAPTER 5. THE STUDY

In this chapter I will provide an account of the different steps followed in this study. I will first describe the data gathering process and building of the corpus. After that I describe the macro-analysis of the presentations, to continue with an account of the multimodal analysis at the micro-level. Then I turn to the process of creating a taxonomy of persuasive strategies. The final section of the chapter focuses on the comparative analysis within and across subcorpora.

5.1 The data gathering process

In this initial section of the chapter I provide a detailed account of the process of data collection and building of the corpus with the aid of ethnographic methods such as observations sheets and semi-structured interviews.

5.1.1 Recording of the events

For this study three different communicative events were video-recorded, namely, a symposium on business discourse, a TED contest and a networking event for universities and entrepreneurs. The events were selected according to the following criteria:

i) The type of presentation: three events were selected in order to obtain a similar amount of the three different types of presentations that concern this study: conference, research dissemination and product pitch;

ii) Accessibility: the organizers granted access and helped contact individual participants;

iii) Feasible timing and location to facilitate the researcher’s physical presence during the recording of the events.

Table 5.1 provides an overview of the three events recorded for this study, which are explained in more detail in the following paragraphs.

Table 5. 1 The corpus

EVENT	ORGANIZERS	SELECTED/ RECORDED	TOTAL TIME	ROOM AND PROPS	AUDIENCE
The Ins and Outs of Professional Discourse	ABC	5/13	173 min	-large room -podium -lectern -projector -screen -computer -microphone	-14 to 23 persons -members of ABC -professors -teaching assistants -doctoral students -bachelor students
TEDX GENT	Gent University	5/6	45 min	-large room -podium -computer -projector -screen -lectern -desk -lapel microphone	-90 persons approx. -general public
BRIDEE	Flemish government	5/7	40 min	-medium-sized room -jury desk -laptop -screen -projector -pointer -lapel microphone	-34 persons approx. -other members of the participating teams -general public

The first of these events was an international symposium on business discourse organised by ABC (Association for Business Communication) held in Modena, Italy, in March 2014. The title of the event was *The Ins and Outs of Professional Discourse*, and special attention was paid to two main topics: access to business environments and feeding back research results to business communities. These are considered two extreme ends of the scholarly process in business discourse research: the input, or obtaining data to analyse, and the output, or reporting the results of the analysis as feedback to practitioners, hence the *ins and outs*. They are, therefore, a very specific type of conference presentations in the field of humanities, and particularly in the area of Discourse Analysis. In this case they can also be considered examples of a *process genre*, because they reported work in progress and the publication came later. The event took place in a large room with a high ceiling that made the acoustic conditions far from perfect. The room was equipped with a small podium, a lectern, a projector, a screen and a computer. The lectern had a microphone attached, and speakers stood still behind the lectern in order to be able to use the microphone and be heard. From this position, they had to reach out to the computer next to them to move on with the slides in their presentations. The room could seat 80 people, but the audience ranged from 14 to 23 persons, among which there were members of ABC, professors, teaching assistants, doctoral students and bachelor students. Each conference presentation in this symposium was allotted 40 minutes, including a Q&A session. A total of 13 presentations were recorded and from these, 5 were selected on the basis of the following criteria (which apply to the other two events as well):

i) Consent granted for the recording;

ii) Good quality of image and sound of the recording;

iii) Availability of the complete text (the recording included the complete presentation);

iv) Availability of speakers for follow-up.

The total size of the conference presentations dataset is 173 minutes. Two of these presentations were co-presented, so the total number of speakers in this case is 7. Three of these speakers were native English speakers (two of them were co-presenting the same presentation). The rest were all non-native speakers with very high proficiency in English. Regarding expertise in presenting at conferences, only one of the native speakers co-presenting reported being novice, this one being her first conference. Table 5.2 provides an overview of the conference presentations dataset, listed in the order in which they were recorded.

Table 5. 2 Conference presentations dataset

1	**TITLE: Participant observation in a foreign business context with sensitive data. Ethical and methodological reflections from the field.**					
	Speaker's Gender	**Speaker's English proficiency**	**Speaker's Expertise**	**Delivery**	**Support/ devices**	**Duration**
	Female	High	High	Explain	PPT Lectern Microphone	00:28:44

2	**TITLE: The role of the researcher in ethnographic investigations of workplace settings: Expert, consultant or confidante**?					
	Speaker’s Gender	**Speaker’s English proficiency**	**Speaker’s Expertise**	**Delivery**	**Support/ devices**	**Duration**
	Female	High	High	Explain	PPT Lectern Microphone Written notes	00:35:20
3	**TITLE: Disseminating professional discourse research to occupational communities: A case study examination of the mental health**					
	Speaker’s Gender	**Speaker’s English proficiency**	**Speaker’s Expertise**	**Delivery**	**Support/ devices**	**Duration**
	Sp 1 female Sp 2 female	Sp 1 Native Sp 2 Native	Sp 1 High Sp2 Novice	Explain	PPT Lectern Microphone Written notes Sp 1: Handouts	00:25:05
4	**TITLE: From classroom to company to classroom: Adapting classroom contents to the evolving business environment**					
	Speaker’s Gender	**Speaker’s English proficiency**	**Speaker’s Expertise**	**Delivery**	**Support/ devices**	**Duration**
	Female	Native	High	Explain	PPT Lectern Microphone	00:43:54
5	**TITLE: “Mars keeps students going”: From student observer to professional insider**					
	Speaker’s Gender	**Speaker’s English proficiency**	**Speaker’s Expertise**	**Delivery**	**Support/ devices**	**Duration**
	Sp 1 Female Sp 2 Male	High	High	Explain	PPT Lectern Microphone Sp 2: pointer	00:40:53

The next event was an independent TED contest organised by Ghent University. TED started in 1984 as a technology conference that aimed to bring together entertainment and design (hence the acronym standing for "Technology, Entertainment and Design). Its increasing popularity soon gave rise to independent events, now called TEDx, which are aimed at spreading ideas among specific communities[3]. In the case of the event described here, the community at stake is Ghent University. There were 6 presentations of scientific research in which scientists disseminate their activity to a layman audience. The event was open to the general public. The size of the audience was approximately 90 people. Each presentation had a time limit of 6 minutes and was followed by a Q&A session (which means that the total duration including Q&A was always beyond 6 minutes). The presenters were doctoral and post-doctoral students. All of them were non-native fluent speakers of English and were experienced in presenting their research, although in more academic contexts (e.g. in their respective research groups at their universities). The aim of the contest was to make research public and attract interest, but as an incentive the winner also received free tuition in public speaking and free participation in other bigger scale TED events. The event took place in a large room equipped with a podium, a computer, a projector, a screen, a lectern and a desk. The speakers did not stand still on the podium, but came down and moved around the front part of the room to be closer to the audience. They all wore lapel microphones. Four of them controlled their presentations with a pointer, and one used no visual aids at all. From the 6 presentations, one was discarded due to poor quality of the image of the recording. The total size of this dataset is 45 minutes. The details of this dataset are provided in Table 5.3.

Table 5. 3 Dissemination talks dataset

1	**TITLE: Microwave chemistry: time is money**					
	Speaker's Gender	**Speaker's English proficiency**	**Speaker's Expertise**	**Delivery**	**Support/ devices**	**Duration**
	Male	High	High	Explain	PPT Lapel mic Pointer	00:07:15

[3] www.ted.com (last accessed 3/2/2017)

2	TITLE: Why does mathematics count?					
	Speaker's Gender	**Speaker's English proficiency**	**Speaker's Expertise**	**Delivery**	**Support/ devices**	**Duration**
	Male	High	High	Explain	Prezzi Lapel mic Pointer	00:07:56
3	**TITLE: With adaptive systems, we can make the best of our differences**					
	Speaker's Gender	**Speaker's English proficiency**	**Speaker's Expertise**	**Delivery**	**Support/ devices**	**Duration**
	Female	High	High	Explain	Lapel mic	00:11:07
4	**TITLE: Snow White's smart textiles twist**					
	Speaker's Gender	**Speaker's English proficiency**	**Speaker's Expertise**	**Delivery**	**Support/ devices**	**Duration**
	Female	High	High	Explain	PPT Lapel mic Pointer	00:08:59
5	**TITLE: Open transport data**					
	Speaker's Gender	**Speaker's English proficiency**	**Speaker's Expertise**	**Delivery**	**Support/ devices**	**Duration**
	Male	High	High	Explain	PPT Lapel mic Pointer	00:10:17

Finally, the third group of presentations consists of product pitches in an event organised by a government collaborative platform funded by the Flemish government, which brings together entrepreneurs with design and business experts. The event takes place once a year, and it is aimed at facilitating cooperation between (former) university students and enterprises. Students team up with design and business experts, and they work for several months on a business plan to launch a product. The presentations are the final step of the process, and they are organised as a competition where one member of each team pitches the product in front of a jury. As a prize the winning team obtains support for the realisation of the business plan. The event was open to the general public, but required registration in advance (free of charge). The presentations had a time limit of 3 minutes and a Q&A session follows (which means that the total duration including Q&A was always beyond 3 minutes). The total number of presentations in this event was 7. There were 2 native speakers of English and 5 non-native speakers with good command of English. Two presentations (one from a native speaker and another from a non-native speaker) were excluded from the analysis because no consent to use the material was obtained. Only one speaker (native) reported having wide experience in pitching. The event took place in a medium-sized room equipped with a laptop, a screen, a projector and a pointer. There was no podium, and the speakers were mainly standing next to or in front of the screen and facing the audience. The size of the audience was 34 people (other members of the participating teams and general public) who were sitting in front of the screen. In addition, on the right side of the room there was a table perpendicular to the audience and the screen where the four members of the jury were sitting. Speakers wore a lapel microphone and used the pointer to change slides in their presentations. None of them had written notes with them and they all used an explanatory style. The total size of the product pitches dataset is 40 minutes. The details of this dataset are summarised in Table 5.4.

Table 5. 4 Product pitches dataset

1	**TITLE: Waxpert**					
	Speaker's Gender	**Speaker's English proficiency**	**Speaker's Expertise**	**Delivery**	**Support/ devices**	**Duration**
	Male	High	Novice	Explain	PPT Lapel mic Pointer	00:07:31
2	**TITLE: Anapad**					
	Speaker's Gender	**Speaker's English proficiency**	**Speaker's Expertise**	**Delivery**	**Support/ devices**	**Duration**
	Male	High	Novice	Memorized	PPT Lapel mic Pointer	00:07:55

3	**TITLE: I-BAR**					
	Speaker's Gender	**Speaker's English proficiency**	**Speaker's Expertise**	**Delivery**	**Support/ devices**	**Duration**
	Male	High	Novice	Explain	PPT Lapel mic Pointer	00:08:01
4	**TITLE: Wigoh**					
	Speaker's Gender	**Speaker's English proficiency**	**Speaker's Expertise**	**Delivery**	**Support/ devices**	**Duration**
	Male	High	Novice	Memorized	PPT Lapel mic Pointer	00:06:12
5	**TITLE: Uxprobe**					
	Speaker's Gender	**Speaker's English proficiency**	**Speaker's Expertise**	**Delivery**	**Support/ devices**	**Duration**
	Male	Native	High	Explain	PPT Lapel mic Pointer	00:09:33

One of the initial hypotheses of this study is that the three genres selected are intertextually and interdiscursively interrelated. In this sense, it is interesting to note that there is a business element in the CPs included in the corpus, since the event is a symposium on business discourse, and that there is also an academic component in the research dissemination talks, since they speakers are all doctoral students and the event is organised by Ghent University.

In the building of the corpus I followed Sinclair's (2005) guidelines for spoken corpus design. The first of these guidelines claims that content (texts, specimens) must be selected according to communicative functions and regardless of the language they contain. For this reason the presentations were not selected on the grounds of specific linguistic, paralinguistic or kinesic features, but according to a common persuasive communicative purpose. In other words, the rationale behind the selection of these different types of presentations and their organization in three datasets is that the three genres share a broad communicative purpose:

persuading an audience of the validity of a product or a piece of research. From the point of view of genre analysis, it is particularly interesting to study how persuasive language differs or overlaps in them, and it provides a solid tertium comparationis (Chesterman 1998).

The second guideline states that corpus must be as representative as possible. The size of my corpus does not allow for quantitatively-based generalisations, but it is valid for qualitative analysis. It is also in line with previous multimodal studies, which due to their minute level of detail and the lack of automating tools cannot afford the use larger corpora. Querol-Julián (2011), for example, uses corpora of 59 and 71 minutes, from which subcorpora of 15 and 14.5 minutes are then extracted for multimodal analysis. Morell (2014, 2015) uses corpora of 80 minutes.

Third, the corpus must be as homogeneous as possible. To this aim, each dataset is drawn from a single communicative event, and each of the three datasets contains 5 presentations.

Fourth, information about the text must be stored separately from plain text. The use of annotation software, as I will explain in section 5.3.1, makes it possible to store this information in layers or tiers, separately from the video-recorded material.

Fifth, the samples should consist of entire documents or complete speech events, even if this implies differences in size. For this reason presentations were considered in their entirety for the analysis even if this implies a considerable difference in length (conference presentations are notably longer). Finally, the design and composition must be documented including the reasons that support decisions taken.

As mentioned in the previous paragraphs, informed consent (see Appendix 1) was previously obtained from all speakers, who received detailed information about the objectives of the analysis and the potential use of their videos, as a way to comply with ethical standards in research (Flewitt 2014).

For the purpose of this research the events were recorded using only one camera focused on the speaker. The camera was mainly static and only moved if the speaker moved. No close ups of the head, body or the room (e.g. screen behind speaker) were used. These decisions were made on the basis of the research questions that drive the analysis: my interest is mainly in the persuasive activity of the speakers, rather than in the uptake by the audience. No special attention was paid to the visual aids used by speakers, because they are out of the focus of this research.

The video material was finally edited for later use with specialised annotation software, i.e. converted into avi and wav formats using free conversion software (Free Video Converter version 3.2.0.0).

5.1.2 Ethnographic observations and first interviews

During the recording, I used observation sheets to gather contextual information about the event, the presentation and the speaker. These sheets included detailed information about the physical and chronological location of the presentation within the event and the performance of the speaker, in particular:

i) The event (e.g. date, organisers, frequency, etc.);
ii) The presentation (e.g. title, length, order in the program, genre);
iii) The speaker (e.g. name, age, gender, mother tongue, performance);
iv) The physical environment (e.g. the room, the audience, devices, etc.).

An example of this observation sheet is included in Appendix 2.

Shortly after each presentation (at the end of the event, and on the same day) face to face, semi-structured interviews were carried out with the speakers in order to obtain crucial information about their previous knowledge and assumptions about the event and their preparation for the presentation, which in turn helps us interpret their use of persuasive strategies. These interviews probed into aspects such as:

i) The speakers’ motivation to participate in the event;

ii) What they knew about the event beforehand (e.g. size and type of audience, topics expected, types of presentations expected, etc.);

iii) The way they prepared for the presentation;

iv) Their main goal in their presentation;

v) Their satisfaction with their performance;

vi) Their opinion of other presentations as members of the audience.

Items i) and iv) are quite related to each other. I considered that the reason why speakers had decided to participate in the event (i.e. their motivations) would ultimately determine their main goal in their presentations and shape their main message, something crucial to fully understand the communicative intentions lying behind their use of modes. Items ii) and iii) were also important to fully understand how speakers were using the modes, and in particular to determine how much of this was unconscious or spontaneous and how much the result of practice or explicit training. Item v) was a prompt for self-assessment, and was intended to put my own interpretation of what I had seen in perspective and help me focus on particularly persuasive moments. Finally, the last item tries to include the audience uptake in the analysis, although admittedly in a very limited way.

Widdowson (1998) and Flowerdew (2003, 2004) have stressed the importance of context and ethnographic information to interpret video data. In my case the ethnographic data provided by the observation sheets and the interviews proved crucial in two ways. First in the selection of the data to be analysed (which will be discussed in the next section) since it provided a starting point to carry out an analysis of the macrostructure of each presentation and to focus the analysis on potentially persuasive moments. Second, the data provided by the interviews also helped interpret the speakers' intended communicative effect and their use of persuasive strategies, as mentioned in the previous paragraph.

A second round of interviews was carried out after the analysis in order to triangulate results and compare interpretations with the speakers, which will be explained in detail in section 5.3.3.

5.2 The macroanalysis and selection of a subcorpus for subsequent microanalysis: moves and rich points

The next step was an analysis of the macrostructure of the presentations, which provided an insight into their topic flow and moves structure (Swales 1990; Carter-Thomas and Jowley- Rolivet 2003). From this analysis, integrated with the information obtained from the interviews with speakers and observation sheets, a series of potentially persuasive points were identified, which I have called *rich points*. The term *rich point* was coined by Agar (1996) and it has become a common concept in ethnographic studies. It refers to moments of experience that stand out from the rest because they reveal a cultural difference, to the point that they often imply a breakdown in cross-cultural communication. This makes them worthy of attention as a research focus. However, I am using the term in a different way. My rich points are moments in the presentation which are particularly rich in terms of persuasive efforts from the speakers. Their length varies in each case, since the contours were established according to a sense of completeness: they seem to fulfil a communicative intention or objective (e.g. illustrates a claim, prompt the audience to action, etc.). For example, in the following rich point from a product pitch the speaker is describing his product and listing its unique characteristics (see Appendix 3 PPPA2 for the full transcription of the four modes; the coding system for naming rich points will be explained in the next section). The contours of this rich point have been set according to when this description begins and ends (i.e. without interrupting the list with two items).

Example of Rich Point contours:

> Uxpro is an analytic service that integrates in your app and helps to understand and measure your users' experience and how your users feel about that app. It includes two important concepts. First behaviour and satisfaction tracking in real time. What your users are doing, whether they are failing, whether they are succeeding, how long they need to do things, how they feel about that app. And secondly, really important thing, because the combination of these two things is the real magic. It's, it's collection of users' sentiment in a very pragmatic way with feedback and microsurvey. Letting your users tell you how they feel about how they are doing with your app.

It must be stressed that the rich points were not selected because they illustrated a particular use of words or a specific paralinguistic or kinesic feature, but exclusively on the grounds of strong persuasive effort. In the previous example, the rich point was not chosen because it featured an example of a list or intense language like “magic” (this was noted after closer analysis), but because it was considered persuasive.

The identification, delimitation and selection of rich points was what Goldman et al. (2007) would call an *inductive* approach to video data:

> Inductive approaches are considered more apt for sampling when working with 'raw' video data sets that have been collected with broad questions in mind but without a strong orienting theory. The process is usually to view all of the video data repeatedly and in increasing depth where the research team agree on major events, themes and identify key moments of importance and describe the structure of the event. (Jewitt 2012: 19)

The process was supported by three sources. The first one was the information gathered from the observation sheets and the first interviews with speakers (e.g. their explanations of what was their main goal with their presentations). For example in one of the research dissemination talks (see Appendix 3 DTLA1), the rich points identified correspond to moments in which the speaker tries to convince the audience that a better world is possible with a change of attitude. During the interview, the speaker explained that she had been inspired by a TED talk about motivational speeches that presented a contrast between what currently exists and what can be achieved (e.g. Martin Luther King “I have a dream”, Steve

Jobs Keynotes) and she had tried to reproduce this effect in her presentation. I considered that this supported my choice of rich points in this presentation.

The second one was a pilot analysis carried out simultaneously and independently by an external researcher and myself in one of the presentations, in this case a product pitch (see Appendix 3 PPSE1 and PPSE2). The two researchers viewed the presentation and independently selected two rich points using the same criterion: high persuasive effort from the speaker. Interestingly, we obtained the same results (i.e. the same rich points were selected). If two researchers select the same persuasive moments, this indicates that our perception of high persuasive efforts is not totally subjective.

Finally, my selection of rich points was cross-checked with the opinion of the three supervisors of this PhD and consensus was sought in (rare) cases of disagreement. For example, one of the rich points in a product pitch (see Appendix 3 PPKE2) was a dubious case. Two of the supervisors perceived as not particularly persuasive in comparison with other rich points already selected. After deliberation, we decided to keep it for two reasons: 1) it was still more persuasive that the rest of the presentation; 2) even if the persuasive effect was not fully achieved (neither of us were really convinced by the speaker), the persuasive effort could still be perceived (and was later confirmed by closer analysis), which is the real focus of this study. In other words, we considered that it could also be interesting to analyse an example in which persuasion is there, but it is not successfully conveyed.

The use of rich points was adopted to avoid prioritizing any semiotic mode in particular. I did not want to use one mode as the driver of the analysis and limit the analysis of the other modes to instances of co-expression (e.g. starting the analysis looking for lexical expressions of persuasion and then looking for correlates in the other modes). In this sense, the approach has proved useful to keep the focus on the multimodal ensemble as a whole (see section 1.1.4 in Chapter 1) and the way different modes interact to encode a persuasive message, as it will be discussed in detail in chapters 6 and 7. In using the rich point as a unit I side with Adolphs et al. (2011: 320) in their suggestion to analyse corpora including heterogeneous data, in particular with their claim that "the unit of analysis may shift from the word or sequence of words to a contextually defined episode of interaction which may include multiple modes of discourse and which is dynamic in nature". I believe this definition of unit of analysis applies to the rich points I have defined in my study.

For the sake of building a comparable corpus, and as I mentioned in section 1.3.1 in Chapter 1, I excluded the Q&A session of all presentations in the selection of the rich points.

During this section a rather different communicative situation is built, entailing a different relationship between presenter and discussant. Because of this, previous research has actually considered it a closely related, but different, genre (Querol-Julián 2011; Räisänen 1999, 2002). Different communicative and persuasive strategies are, therefore, to be expected, which would further complicate the contrastive analysis of the genres. This will, however, constitute a fascinating topic for further research.

For the sake of feasibility, and in order to keep the corpus to a manageable size for multimodal analysis, I restricted the selection of rich points for fine-grained multimodal analysis to two per presentation, which gave a total of 30 rich points (ten per each of the three events). I did not try to make these rich points share the same length or position in the presentation, but in the two presentations that were co-presented I did select one rich point for each speaker on purpose. The total size of the conference presentations rich points is 7.2 minutes. The rich points in dissemination talks add up to 4.9 minutes and the rich points in product pitches have a length of 3.8 minutes. These differences in corpus size reflect the differences in the overall length of the presentations, conference presentations being considerably longer, followed by dissemination talks and finally product pitches which are the shortest. As will be explained in detail in Chapter 5, the results are presented in the form of a frequency per minute, thus facilitating the comparative analysis across presentations that differed considerably in length.

To keep track of the rich points while preserving the anonymity of speakers, rich points were coded according to the type of presentation, the initials of the speakers and then numbered (e.g. DTPI1 stands for Dissemination Talk, two first letters of speaker’s name and first rich point in the presentation). The transcripts of all of the rich points are compiled in Appendix 3. Table 5.5 offers an overview of the rich points in the three genres.

Table 5. 5 Rich Points subcorpus

CONFERENCE PRESENTATIONS			
The role of the researcher in ethnographic investigations of workplace settings: Expert, consultant or confidante?			
	Begin	**End**	**Duration**
CPDO1	00:05:14	00:06:00	00:00:46
CPDO2	00:17:00	00:17:24	00:00:24

“Mars keeps students going”: From student observer to professional insider			
	Begin	**End**	**Duration**
CPAS1	00:16:05	00:16:35	00:00:30
CPTO1	00:19:26	00:20:13	00:00:48
From classroom to company to classroom: Adapting classroom contents to the evolving business environment			
	Begin	**End**	**Duration**
CPAM1	00:03:13	00:03:37	00:00:24
CPAM2	00:21:10	00:21:58	00:00:48
Disseminating professional discourse research to occupational communities: A case study examination of the mental health			
	Begin	**End**	**Duration**
CPRE1	00:00:38	00:01:36	00:00:58
CPKE1	00:15:24	00:16:14	00:00:48
Participant observation in a foreign business context with sensitive data. Ethical and methodological reflections from the field.			
	Begin	**End**	**Duration**
CPPE1	00:04:43	00:05:22	00:00:39
CPPE2	00:07:07	00:08:12	00:01:05
RESEARCH DISSEMINATION TALKS			
With adaptive systems, we can make the best of our differences			
	Begin	**End**	**Duration**
DTLA1	00:03:13	00:04:04	00:00:51
DTLA2	00:04:04	00:04:52	00:00:47
Snow White’s smart textiles twist			
	Begin	**End**	**Duration**
DTLI1	00:00:00	00:00:24	00:00:24
DTLI2	00:04:40	00:04:50	00:00:10

Why does mathematics count?			
	Begin	**End**	**Duration**
DTRA1	00:01:45	00:1:52	00:00:07
DTRA2	00:05:03	00:05:08	00:00:05
Open transport data			
	Begin	**End**	**Duration**
DTPI1	00:00:33	00:01:03	00:00:30
DTPI2	00:03:13	00:03:53	00:00:40
Microwave chemistry: time is money			
	Begin	**End**	**Duration**
DTJO1	00:00:00	00:00:24	00:00:24
DTJO2	00:00:24	00:01:17	00:00:53
PRODUCT PITCHES			
Wigoh			
	Begin	**End**	**Duration**
PPKE1	00:00:00	00:00:47	00:00:47
PPKE2	00:02:10	00:02:18	00:00:08
Uxprobe			
	Begin	**End**	**Duration**
PPPA1	00:00:00	00:00:23	00:00:23
PPPA2	00:00:31	00:01:17	00:00:46
I-BAR			
	Begin	**End**	**Duration**
PPPI1	00:00:02	00:00:13	00:00:11
PPPI2	00:01:35	00:01:52	00:00:17
Waxpert			
	Begin	**End**	**Duration**
PPSE1	00:02:14	00:02:21	00:00:07
PPSE2	00:02:51	00:02:56	00:00:05

Anapad			
	Begin	**End**	**Duration**
PPTO1	00:00:42	00:01:14	00:00:32
PPTO2	00:01:38	00:02:10	00:00:32

5.3 The microanalysis

In this section I discuss the microanalysis of the rich points with the help of specialised software, and the subsequent triangulation of results by means of ethnographic interviews.

5.3.1 Multimodal analysis by means of annotated transcripts

A multimodal analysis usually entails the use of an annotated transcript to facilitate the understanding and interpretation of a particular communicative situation. Such transcripts must include data regarding all the modes that integrate the analysis. According to Baldry and Thibault (2006), the main aim of a transcription is to relate semiotic modes to each other and to the activities they contribute to, and also to the functions of these activities in their context of situation and context of culture. In addition, they claim that “multimodal transcriptions are ultimately based on the assumption that a transcription will help us understand the relationship between a specific instance of a genre, for example a text, and the genre's typical features" (Baldry and Thibault 2006: 30). I agree with this view and it is with this aim that I use transcripts in my study, although the identification of a genre's typical features would require a larger-scale corpus analysis, as I will argue in more detail in Chapter 8.

According to Bezemer (2014) an annotated transcript has two functions. The first one is epistemological, that is, it is an analytical exercise that provides insight into the collaborative activity people engage in when communicating. The second one is rhetorical: it provides evidence for claims made in a paper or other scientific publication.

He claims that transcripts reveal information about the communicative event they represent, but also about their authors and the way they perceive this event. The choices made in their design reveal the researchers’ intentions and also how they interpret what they are studying. These choices concern aspects such as:

i) who is included in the analysis (i.e. what participants)?

ii) what parts or aspects of the event are included?

iii) what modes are represented (language, gesture, head, gaze, space, etc.)?

iv) what level of analysis is applied (macro, micro)?

v) how are results presented (i.e. what type of transcript)?

Mavers (2011) agrees that a transcript is always partial and reshapes reality, including certain things and excluding others. Its meaning is not independent from its author or from the reader. The transcript is rather part of the analysis, and it does not mean distorting reality, it means turning material into analysable data.

An annotated transcript represents a multimodal communicative event (in this case a presentation), but it is also multimodal in itself, because it uses different semiotic modes for representation (e.g. image and writing). In their representation they reshape the communicative event and what is expressed in one mode becomes expressed through a different mode. For example words can be represented in writing, or a movement can be captured as a sequence of still images. This phenomenon is known as transmediation (Suhor 1984), transduction (Kress 2003) or transmodal remaking (Mavers 2011). Due to different valuation and affordability of modes, a remaking process is never seamless. For example in the case of this study, intonation is represented in the transcript through standard writing and specific symbols, and in this transmodal remaking some meaning might inevitably be lost.

This study is designed as inductive or corpus-driven, which means that the transcription and annotation process was largely determined by the findings obtained from the corpus (Tognini-Bonelli 2001; Adolphs and Carter 2007; Querol-Julián 2011). I have decided to focus this research on the figure of the presenter, and specifically on the use of persuasive strategies. The modes selected are words, gestures, head movements and intonation. The study entails a macro-analysis of the presentations recorded and a microanalysis of selected rich points. The results of the microanalysis are presented by means of an annotated transcription aided by specialised software (see next section). The transcript includes data concerning the four modes included in the analysis, and in the transcription process some transmodal remaking (Mavers 2011), as explained, in the previous paragraph, is inevitable. In particular, words are represented in writing both orthographically and following MICASE[4] conventions (Michigan Corpus of Spoken and Academic English) to represent relevant paralinguistic information such as pauses or false starts. These paralinguistic features are

[4] Quod.lib.umich.edu/m/micase

actually out of the scope of my study, but they were found to be occasionally relevant for the interpretation of the modes analysed. The orthographic transcription was not aided by software, but the transcript was cross-checked with the speakers during the second interviews (see section 5.3.3), and they occasionally made some useful corrections. Intonation is also represented in writing using the following conventions: tone units are sequentially listed, prominent syllables are capitalized, syllables below or above the line represent a lower or higher pitch, arrows indicate rise or fall tone when relevant for the interpretation of persuasion. As it will be explained in the following section, the intonation analysis was aided by the use of specialised software. Gestures and head movements are represented through still pictures with arrows signalling movement when necessary[5]. With the help of the aforementioned software it was possible to time-align the gestures and head movements with the rest of the modes. The transcript includes the accompanying words and the intonation with which they were uttered below each still picture. A final visual schematic representation of the modal configuration or multimodal ensemble (i.e. the modes that are involved in the rich point) is also offered. Different size in balloons indicates the degree to which each mode is being used to fulfil the action expressed in the centre of the diagram, or in other words, how many persuasive strategies they are realising (strategies are counted once per each modal realisation, i.e. there cannot be two instances of the same strategy with the same modal realisation; see section 5.5). An example of a transcript, corresponding to the rich point CPKE1 is offered in Table 5.6:

Table 5. 6 Example of transcript: CPKE1

Orthographic transcription
So the APA Annual Meeting was the best opportunity I could really find to disseminate my research. There are thousands of people there, the whole occupational community gathers to listen to talks. It's largely like this, but for the psychiatric community. I presented a poster in Young Investigator's Poster Session and that's my poster right there, it was actually interesting because I seemed to be the only person that wasn't a psychiatrist there, and a lot of people would walk by and they would double-take and look, like almost that I didn't belong there. But once I was able to talk to them about it, a lot of them were very interested on the impact that rhetoric has on psychiatry and how they could use that practically.

[5] Occasional differences in the colour or the size of these arrows are due to the need to adapt to the background of the picture and provide a better visualisation. They do not indicate particular characteristics of the gesture.

MICASE transcription

so the apa annual meeting was the best opportunity i could really find to disseminate my research eh there are thousands of people there the whole eh occupational community gathers to eh listen to talks, it's largely eh like this but for the psychiatric community. i presented a poster in young investigator's poster session, and, eh that's my poster right there it was actually interesting because, i seemed to be the only person that wasn't a psychiatrist, there, and a lot of people would walk by and they would double-take and look ... like almost that i didn't belong there but, once i was able to talk to them about it a lot of them were very interested on the impact, that rhetoric has on psychiatry and how they could, use that practically.

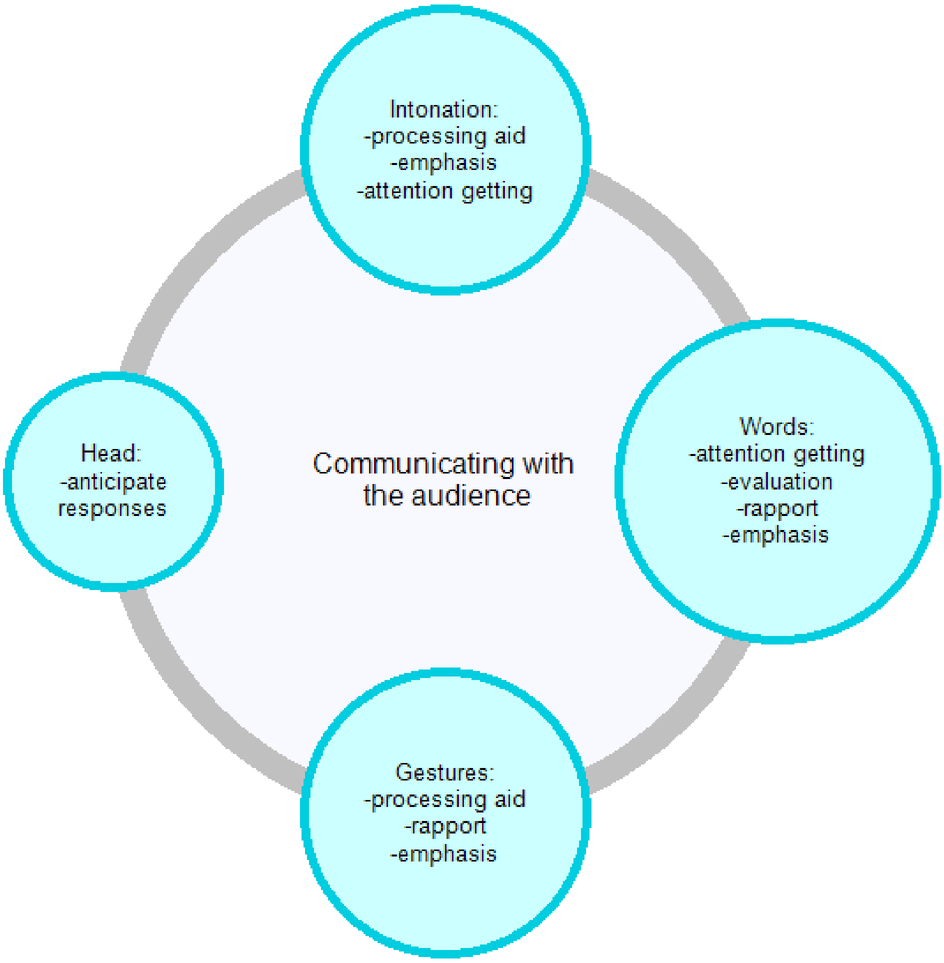
DI (Discourse Intonation) transcription

1. so the Apa annual meeting was the BEST opportunity i could really find
2. to disSEminate my reSEARCH
3. eh →
4. there are THOUSands of PEOple there
5. the WHOLE
6. eh →
7. occupational community GAthers to
8. eh →
9. listen to TALKS
10. it's LARGEly
11. eh →
12. like THIS
13. but for the psyCHIAtric community
14. i presented a POSter
15. in young investigator's POSter session
16. and eh →
17. THAT'S my poster right there
18. it was ACtually INteresting because
19. i SEEMED to be the only person that WASN'T a psychiatrist there ↗
20. and a LOT of people would WALK by
21. and they would DOUble-take and look
22. like almost that i didn't beLONG there
23. but ONCE i was able to TALK to them about it
24. a lot of them were VEry interested
25. on the IMpact
26. that RHEtoric has on psyCHIAtry
27. and how they could USE that PRACtically

Gestures and head movements
 there are THOUSands of PEOple there  it’s LARGEly eh like THIS but for the psyCHIAtric community

a lot of them were VEry INterested

Multimodal ensemble



A pilot analysis of one rich point was first carried out to test the suitability of the transcription and annotation method and the feasibility of combining two programs: ELAN and PRAAT. Both are discussed in detail in the following section.

5.3.2 The use of specialised software in the multimodal analysis of persuasive strategies

The annotation software ELAN[6] (EUDICO Linguistic Annotator) was developed by The Language Archive group in the Max Planck Institute for Psycholinguistics, in Nijmegen, The Netherlands. It is a linguistic annotation tool that makes it possible to transcribe and annotate audio and video files. Transcriptions and annotations are organised on layers (or tiers in the program nomenclature) that relate to each other in different ways. It presents several advantages, for example it is free software, it is flexible in the management of annotations, and it has been improved for time accuracy, which is especially important for the synchronization of the different semiotic modes. It also allows the user to export files as standard formats and import specific formats used by other programs such as PRAAT.

In order to obtain more accurate measurements of prosodic aspects and disambiguate difficult instances I used the program PRAAT[7] (Doing Phonetics by Computer), designed by Paul Boersma and David Weenink, from the University of Amsterdam. This is a tool for phonetic analysis that allows accurate measuring of pitch and intensity. This makes it possible to base a prosodic transcription on objective measurings and avoid to a great extent transcriber's subjectivity.

One of the most important aspects to bear in mind while using annotation tools is the encoding of the information. Kipp (2014) insists that our encoding schemes must be consistent (i.e. data should be set with the same structure), valid (i.e. annotations should correspond to acceptable 'variables' or aspects that are object of study), efficient (i.e. annotations should be easy to use and understand); and reliable (i.e. annotations must clearly refer to one specific aspect).

[6] http://tla.mpi.nl/tools/tla-tools/elan/

[7] http://www.fon.hum.uva.nl/praat (last accessed 3/2/2017)

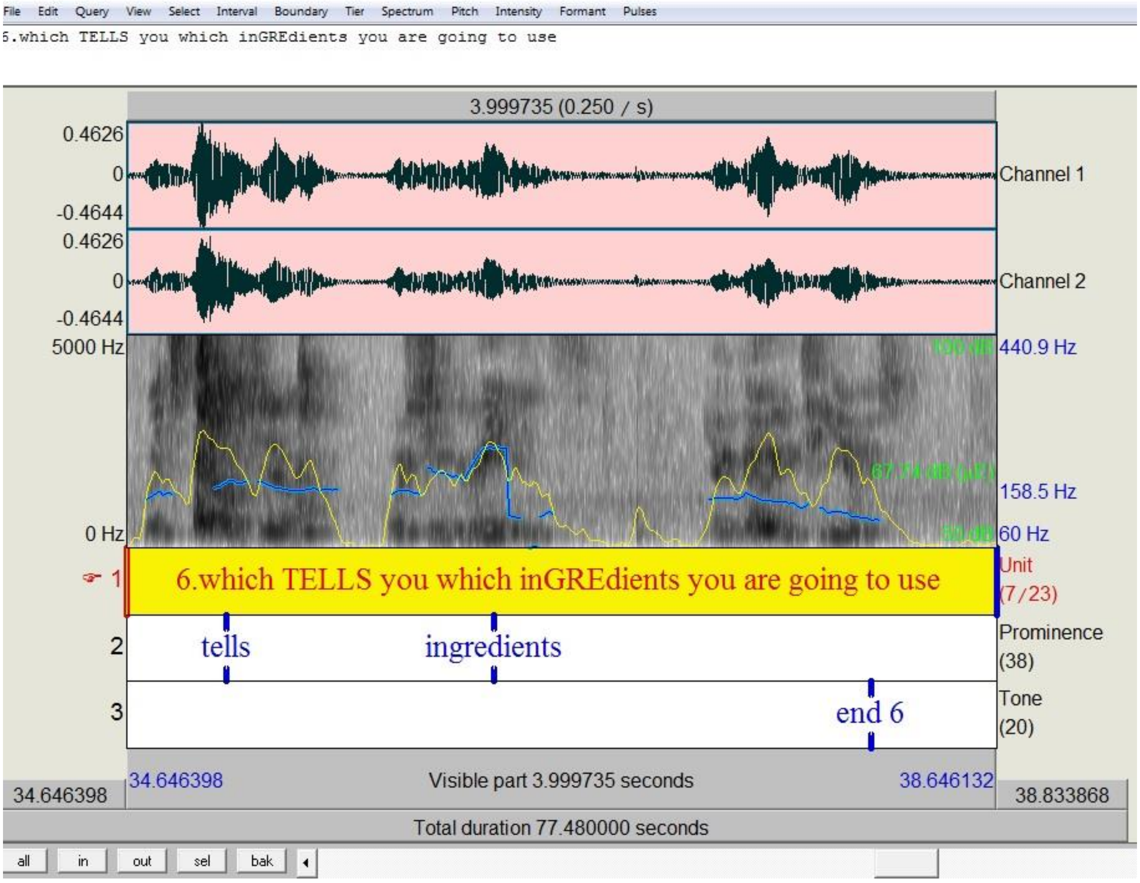


Figure 5. 1 PRAAT tiers encoding prosodic transcription and annotations

For the prosodic analysis in PRAAT I created three tiers. As Figure 5.1 above shows, one tier (Unit) contains the boundaries of the tone units, a second tier (Prominence) contains the position of the prominent syllables and a third tier (Tone) shows the position of the last syllable in the unit. PRAAT enabled me to delimit tone units and locate prominent syllables in a very accurate way. Using a software script developed by Daniel Hirst (2007) I was able to obtain pitch measurements (in blue in figure 5.1) and intensity measurements (in green in figure 5.1) of each prominent syllable, which were useful to identify key and termination. Comparing the pitch levels of the tonic syllable and the last syllable in the unit I obtained a pitch movement or tone. The tiers in PRAAT were later imported for use in ELAN.

For the multimodal analysis I have used the tier structure shown in Figure 5.2 below. An explanation of each tier follows in the next paragraphs.

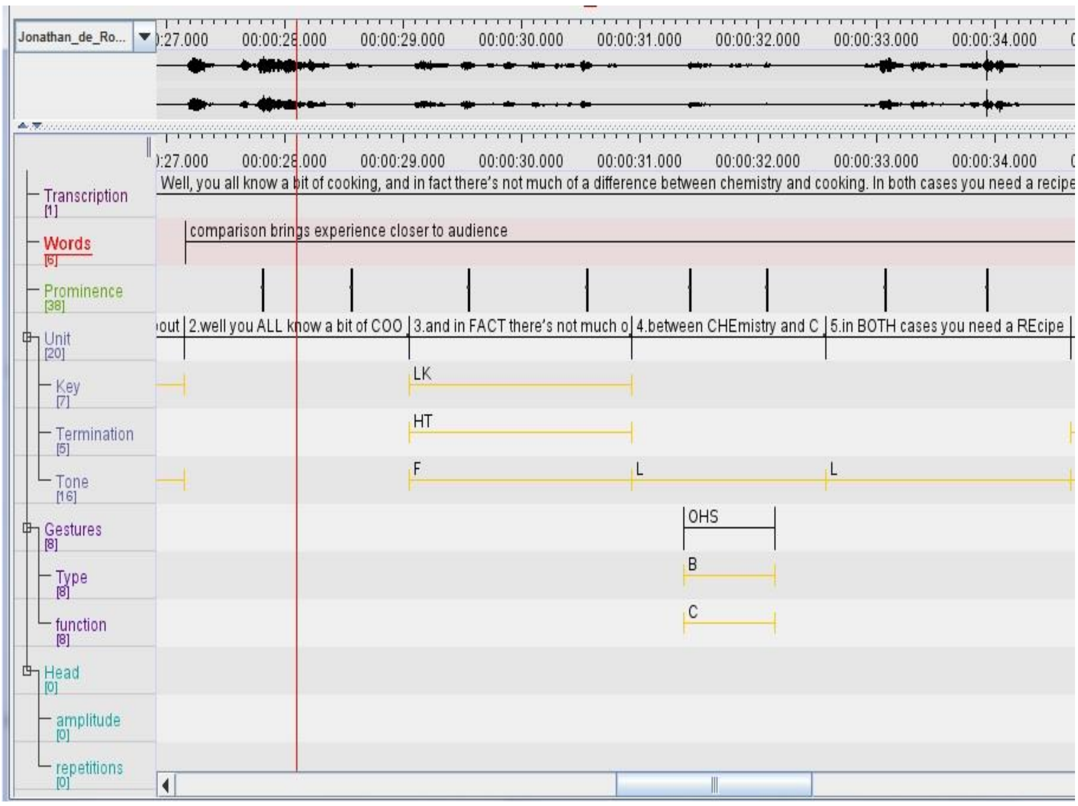


Figure 5. 2 ELAN tiers encoding multimodal transcriptions annotations

The tier *"Transcription"* includes the orthographic transcription, which is synchronized with the audio and video files.

The tier "*Words"* contains annotations about grammatical, lexical and stylistic devices such as evaluative language, three-part lists, symmetrical syntactic structures, inclusive pronouns, examples, comparisons, narratives, among others (Lakoff 1982; Carter 1997; Hyland 1998, 2006, 2009) that can have an effect on persuasion, because they help provide the message with characteristics that according to literature make the text persuasive.

The prosodic transcription tiers ("*Prominence*", "*Unit*", and three dependent tiers: *"Key", "Termination"* and *"Tone"*) include a Discourse Intonation (DI) transcription (Brazil 1997). The tier labelled "*Prominence*" is a tier imported from PRAAT of the type 'point tier' according to the programme's nomenclature, which means that a vertical line marks the time when the prominence occurs. In the tier named "*Unit*" the following standard DI conventions are used: upper case letters for prominent syllables and lower-case letters for non-prominent syllables. Tone, key and termination are indicated in three dependent tiers according to the system shown in Table 5.7:

Table 5. 7 Transcription conventions for Tone, Key and Termination

Tone	Key	Termination
F: fall	HK: high key	HT: high termination
L: level	MK: mid key	MT: mid termination
R: rise	LK: low key	LT: low termination

In the present study I have not differentiated combined tones (fall-rise and rise-fall). The reason for this research decision is that they mainly convey information about power relationships and in particular dominance, which in the case of the presentations analysed tend to remain rather stable throughout. I considered that the potential added value of this distinction did not justify the analysis cost.

The tier *"Gestures"* indicates the gesture family (Kendon 2004). This tier hosts two other dependent tiers (i.e. their division into slots and their input is linked with and determined by the tier they depend on, in this case "*Gestures*") with data concerning the type and function of the gesture. The classification system draws from McNeill (1992), Bavelas et al. (1995), Kendon (2004) and Querol-Julián (2011), and it is shown in Table 5.8:

Table 5. 8 Transcription and annotation conventions for gestures in *Gestures tier*

Gesture family:	Gesture type:	Gesture function:
R: ring (Kendon 2004)	I: iconic (represent concrete objects and events) (McNeill 1992)	R: referential (represent some aspect of the content) (Kendon 2004)
G: grappollo (finger bunch) (Kendon 2004)	M: metaphoric (represent abstract ideas) (McNeill 1992)	I: interpersonal (regulate interaction) (Bavelas et al.1995)
OHP: open hand prone (palms down) (Kendon 2004)	B: beat (repetitive gestures that usually mark the discourse flow) (Kendon 2004,Bavelas et al. 1995)	P: pragmatic (show attitude or perlocutionary meaning) (Kendon 2004)
OHS: open hand supine (palms up) (Kendon 2004)	D: deictic (point to something) (Kendon 2004)	C: cohesive (link parts of discourse) (McNeill1992,Querol-Julián 2011)
CF: close fist		

The tier *"Head"* indicates the kinesic characteristics (lateral movement, shake, nod), and two dependent tiers with information about *Amplitude* and *Repetitions* (McClave 2000; Kendon 2002).

The tiers "*Words*" and "*Gestures-function*" contain annotations, since they provide additional interpretative information about the mode and its potential contribution to persuasion, while the rest of the tiers describe the modes and can be regarded transcriptions (Leech 2005). These transcriptions are used to build the sort of transcript that was shown previously in Table 5.6: for example the different prosodic tiers in ELAN are the basis of the intonation section in the transcript. The annotations, on the other hand, pave the way for the interpretation of the results and help elucidate the contribution of each mode to persuasion. The fact that all tiers are time aligned allows me to keep track of the multimodal ensemble at each moment in the rich point.

It will be noted that the transcription and annotation of gestures and intonation are highly systematic and follow pre-defined conventions that have been incorporated in the software through the use of controlled vocabularies. This limits the input possibilities in specific tiers through a predefined list of options from which the analyst selects. The annotation of words, on the other hand, is corpus-driven rather than pre-defined. The rationale behind this methodological decision, as explained in Chapter 4, is a need to keep the focus on the multimodal ensemble. I took into account the input from previous literature regarding persuasive strategies realised through words (i.e. lexical, grammatical and stylistic devices), and some of these strategies were indeed identified in the corpus (e.g. use of direct speech or evaluative lexicon). However, I avoided having a pre-defined list as the starting point of our analysis, because this would very likely lead to prioritizing words as the main mode and would limit the analysis of the other modes to finding instances of co-expression. On the other hand, the interviews prove that speakers are generally very capable of accounting for their use of words, but find it harder to explain their use of intonation and body movements because they are less conscious of them. Thus, the use of a systematic classification system was necessary to facilitate the identification of particular uses of intonation, gesture and head movements and discuss them during the interviews, but it was not necessary in the case of words.

5.3.3 Triangulation of results: second interviews

Ethnographic interviews were used as a final step in the microanalysis to discuss preliminary results with speakers and cross-check interpretations. This practice is becoming increasingly widespread in multimodal studies, see for example Norris (2004) who calls this practice ‘playback methodology’, Querol-Julián and Fortanet-Gómez (2012) or Morell (2014). In the case of the present study it was also considered a requirement to make the analysis acceptable in ethical terms. It made it possible to turn subjects into participants and share outcomes of the analysis which could benefit both parties (Flewitt 2014). For example, in most cases the speakers thought that the discussion of results and exchange of interpretations was a very useful feedback to improve their public speaking skills. For the researcher, this exchange of interpretations provided the analysis with more reliability. Occasionally, speakers offered alternative interpretations, and their opinions were taken into account and integrated in our discussion, as the examples in Chapter 6 will illustrate.

Nine of these interviews took place face to face, they were audio recorded and I took written notes. Six of them took the form of online video calls and they were equally recorded. Two speakers could unfortunately not be reached. These interviews were open and did not include a predefined battery of questions, but roughly followed the same structure:

i) Explanation of the content and purpose of the interview;

ii) Visualisation of the excerpts analysed with the speakers;

iii) Checking of the orthographic transcript;

iv) Discussion of aspects the analysis had revealed as relevant for the persuasive effect (e.g. a particular use of intonation), including speakers’ opinion about their intention and potential effect on audience;

v) Exchange of interpretations regarding the intent and potential effect of these aspects.

The time gap between the actual presentation and this second interview was considerable (in some cases even a year). For this reason, speakers were asked to watch the video of their presentation (without annotations or transcripts). In every case their own interpretation was prompted before sharing my findings with them.

5.4 Building a taxonomy of persuasive strategies

The MDA approach to persuasive strategies that I propose in this dissertation hinges on the integration of previous literature, multimodal analysis and ethnographic data, as explained in Chapter 4. With the term *strategy* I refer to the different techniques or resources that speakers use to achieve their communicative goal of persuading their audience. Triangulating input from literature with the results of multimodal analysis and ethnographic analysis (i.e. observations and interviews) I have identified a series of persuasive strategies that are realised through different modes (words, intonation, gestures and head movements), which in turn are materialised through different *media* (i.e. the speakers' voice and the speakers' body).

It must be noted that the process of identifying the strategies was cyclical rather than linear. With this I mean that I did not follow the same steps in exactly the same order to reach each strategy. Some examples may help clarify this point. Some strategies such as "anticipation and control of responses" were first suggested by the literature on persuasion (O'Keefe 2002; Perloff 2003) and later revealed to be realized multimodally through the MDA. Other strategies such as "projection of context of interaction" (Brazil 1997) were also first derived from the literature on intonation and not previously associated with persuasion, but the MDA showed how they could contribute to persuasion and how they could also be realised through a multimodal ensemble and not only through intonation. Finally, other strategies such as "processing aids" were first suggested by the MDA analysis of the videos. The analysis revealed that this is a recurrent behaviour of the speakers which was realised through different modes. Later, the literature review supported the claim that messages are more persuasive when they are easy to understand, which confirmed my hypothesis that this multimodal behaviour (i.e. using processing aids) was intended to have a persuasive effect. The role of ethnography in the elaboration of this list is that of validating the strategies through discussion with the speakers about their communicative intentions. For example, in the case of the strategy "anticipation and control of responses", the speaker validated it by describing two head nods as a confirmation that she was indeed anticipating the right response from the audience. In her own words, "maybe it was a confirming gesture that actually, yes, it was the right anticipation" (DTLI1).The full list of strategies will be discussed in detail and illustrated with specific examples in Chapter 6.

5.5 Comparative analysis

The final step in the analysis entailed a comparison of the persuasive strategies used and their modal realisations (i.e. whether they are realised through words, intonation, gestures or head movements), both among presentations within the same event and then across genres. As mentioned in Chapter 4, the rationale behind this comparison is to test the usefulness of my suggested MDA and ethnographic approach to identify common traits and differences in the use of multimodal persuasive strategies in the three subcorpora. The results obtained in this study cannot be considered generalizable for the three genres under analysis. A study using a larger corpus and including entire texts rather than selected rich points would be required for this. Such a study could help refine the definitions of each genre and improve the way they are taught in LSP. I believe the methodology described in the following paragraphs is a step forward in that direction.

Persuasive strategies are not discrete entities that can be easily quantified and co-related with modes in one-to-one relationships. It is not straightforward to decide when one strategy ends and another begins, and some overlap is unavoidable. In fact, strategies are best understood as part of a continuum. What we perceive as members of the audience or interlocutors is always a multimodal ensemble, in which different semiotic modes are orchestrated to fulfil the communicative intentions of the speaker. In other words, persuasive strategies may be realised through different modes, and this may be at the same time or sequentially during the presentation. Emphasis, for example, can be realised through the use of an intensifying adverb (words) plus prominence (intonation) plus a hand beat (gesture) at the same time, but these three modal realisations of the same strategy can also be found sequentially (or even more scattered throughout the presentation) if the prominence is applied to the adjective immediately following rather than to the intensifying adverb and the beat comes somewhat later.

In addition, the reason why a certain gesture or use of intonation can be considered persuasive (i.e. materialising one or several persuasive strategies) needs to be understood in context, taking into account the whole multimodal configuration. A head shake, for example, might not be persuasive in itself, but it acquires this character when it accompanies an evaluation and it can be understood as preventing a counterclaim to this evaluation. Likewise, a pointing gesture can become persuasive when it acquires an implied negative evaluation of what is being pointed at (see examples in Chapter 6).

In spite of these challenges, quantification is required to facilitate the comparative analysis and determine, for example, the frequency of strategies and modes in each presentation. To this aim, the identification of strategies and their modal realisations was carried out in the following way:

i) The four modes included in the analysis (words, intonation, gestures and head movements) were analysed to see how they contributed to persuasion (i.e. what persuasive strategies they were enacting). It is important to note that the same mode can be simultaneously contributing to different strategies. Strategies and their modal realisations are seen as two sides of the same coin: the *what* and the *how*. Strategies are *what* speakers are doing, the modes are *how* they are doing it. Both aspects are quantified, according to the criteria described in the following points.

ii) Each modal realisation of a persuasive strategy was counted once if the speaker uses it in the rich point analysed. This means that in each rich point there can be several instances of the same strategy with different modal realisations, but not two instances of the same strategy with the same mode. Likewise, there can be several instances of the same mode, but always realising different strategies. For example emphasis through words and emphasis through intonation in the same rich point would count as two instances, but I do not count emphasis through intonation (1) and emphasis through intonation (2) even if they are identified at two different moments in the rich point. This is a way of solving the problem of delimiting strategies. It also avoids artificial differences in the way instances of strategies are counted across modes, since not all modes are equally discrete (Iedema 2003). For example, if evaluation is realised by a set of three evaluative words, and of these three words two carry particular prominence that adds to this evaluation, it would be misleading to count this as three instances of evaluation through words (because there are three words) and two instances of evaluation through intonation (because there are two prominences). With the system I have used, they both count as one instance of evaluation in each mode.

iii) The results are formulated as a frequency per minute (i.e. number of instances /duration of each subcorpus of rich points in minutes) to achieve an acceptable *tertium comparationis*. This facilitates the comparison across rich points and genres that vary in length, since 3 instances of emphasis in 10 minutes cannot be considered equal to 3 instances in 20 minutes.

iv) To facilitate comparison across genres, I also quantify the use of modes, regardless of the strategy they are realising, and the use of strategies, regardless of the mode in which they are realised (see sections 6.3 in Chapter 6).

v) No statistics were used in the analysis due to the relatively small size of the corpus. Nevertheless, for future larger scale studies, statistical analysis can shed light on the significance of the differences found among genres, as I will discuss in more detail in Chapter 8.

The following example (figure 5.3) can illustrate the quantification process:

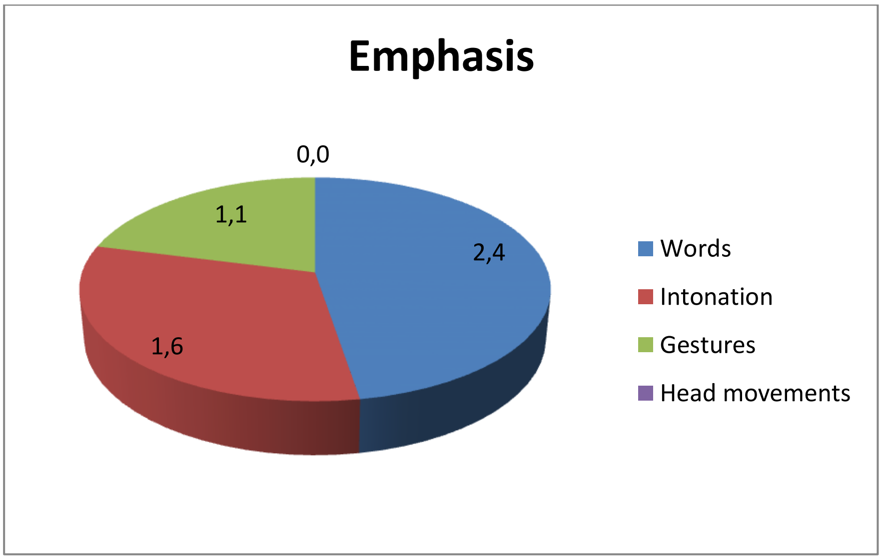


Figure 5. 3 Modal quantification of the strategy "emphasis" per minute (product pitches)

As the graph in figure 5.3 shows, the strategy "emphasis" is realised through the mode "words" (blue) 2.4 times per minute in the product pitch subcorpus of rich points (the whole subcorpus is 3.8 minutes – or 3 minutes and 48 seconds). The same strategy is realised through the mode "intonation" (red) 1.6 times per minute and through the mode "gesture" (green) 1.1 times per minute. Finally, it is never realised through the mode "head".

In this chapter I have explained the different steps I have followed in my study. I have first focused on the data gathering process and building of the corpus. Then I have shown how I have applied the theories discussed in Chapters 1 to 4 to carry out an ethnographic and multimodal analysis of the use of persuasive strategies in my corpus of presentations. A detailed account of the results obtained is the object of the next chapter.

CHAPTER 6. RESULTS

In this chapter the results of the ethnographic and MDA study of persuasive strategies in three types of oral presentations (i.e. conference presentations, dissemination talks and product pitches) will be presented. The chapter is divided into three sections. In the first section I will address my first research question (RQ1), mainly **what type of persuasive strategies can be identified using an MDA approach**. To that aim, I will describe a taxonomy of persuasive strategies identified by applying the framework for multimodal analysis described in section 4.4 of Chapter 4 to the corpus described in Chapter 5. The second section looks into the **contribution of the multimodal ensemble orchestrated through words, intonation, head movements and gestures to the creation of persuasion in the genres included in the analysis**, and answers my second research question (RQ2). To that aim, in this section I present the results of both the ethnographic analysis and the multimodal analysis of each of the three subcorpora. Finally, the last section focuses on a comparison of strategies and modes used across the three subcorpora and closes with a comparative view of the use of multimodal persuasion in them. This provides an answer to my third research question (RQ3): **To what extent is the use of persuasive strategies genre-specific and to what extent is there overlap between the three genres?**

6.1 A taxonomy of persuasive strategies

The first section of this chapter addresses RQ1:

RQ I: What type of persuasive strategies can be identified using an MDA approach?

To answer this question I have applied the framework for MDA described in Chapter 4 to the corpus described in Chapter 5. In other words, I have used a combined multimodal and ethnographic video-based analysis to analyse and classify the persuasive multimodal behaviour of the speakers in my corpus. In view of the outcomes, my answer to this question is that a systematic multimodal approach to persuasive strategies is possible and feasible in practical terms. It is facilitated by the use of video recordings to revisit the material during the analysis. The use of a combined methodology that integrates MDA with ethnographic observations and interviews constitutes a powerful instrument to compare the results with the opinions of the speakers. The analysis also integrates input derived from the literature on persuasion and multimodality. This input has suggested persuasive strategies that have later been revealed as multimodal, or has confirmed that some multimodal behaviours identified in the videos are intended to be persuasive. In the following paragraphs I will provide some specific examples.

Applying the methodology I have outlined in the previous paragraph (and which is explained in detail in Chapter 5) to a corpus of 15 presentations (also described in Chapter 5) I have identified seven strategies, which are realised by four modes (words, intonation, gestures and head movements) as illustrated in Figure 6.1.

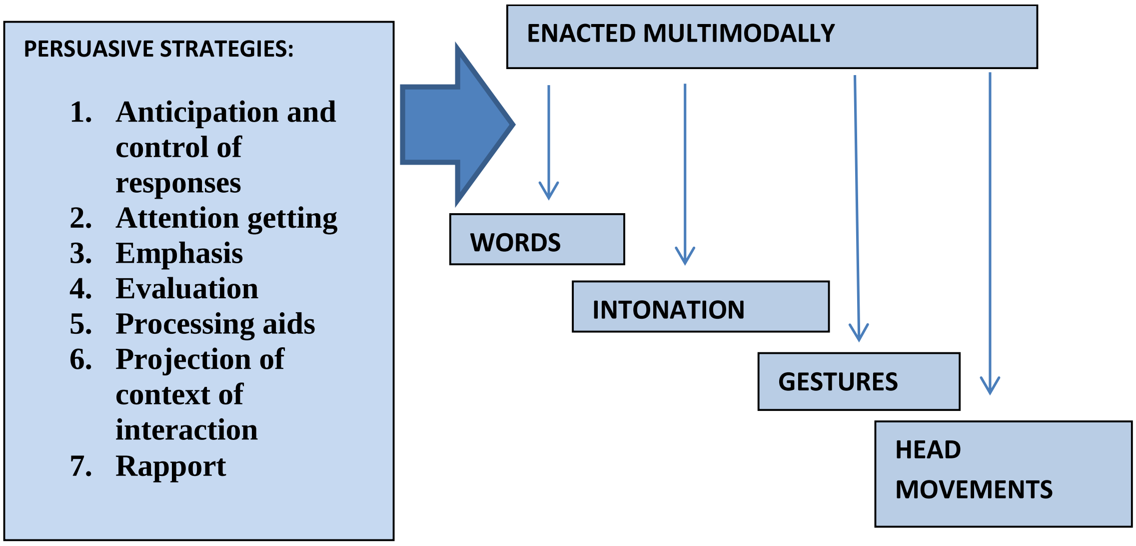


Figure 6. 1 Persuasive strategies

As already mentioned in chapter 5, it must be stressed that the identification of strategies did not start from a predefined list, which was later searched for in the corpus. On the contrary, my approach was data-driven, and based on the analysis of the rich points, which, as explained in Chapter 5, featured a high persuasive activity. The persuasive activity in each rich point was then analysed and classified, taking into account the input from previous literature and the results of the multimodal and ethnographic analysis, to finally yield the taxonomy I present in this study. As explained in Chapter 5, the path to each strategy is not always identical. For example, some strategies were first suggested by literature on persuasion and later revealed to be realized multimodally. Other strategies were first suggested by the MDA analysis of the videos and later the literature review supported the hypothesis that this multimodal behaviour was intended to have a persuasive effect. As a final step in the process I use ethnographic interviews to validate the strategies through discussion with the speakers about their communicative intentions. For example, the first strategy presented, “anticipation and control of responses”, was validated by the speaker’s description of her two head nods (see example 1 in section 6.6.1) as a confirmation that she was indeed anticipating the right response from the audience.

It must be noted that the list of strategies that I am proposing here does not aim at being exhaustive. They are the strategies that I have found in my corpus (described in Chapter 5), but larger scale studies can reveal more strategies and more modal realisations. My intention with this list is to illustrate how the framework described in Chapter 4 can be applied in practice to a corpus of oral presentations in order to study the use of persuasive strategies in them from a multimodal perspective. In doing so, I want to avoid prioritizing any mode in particular, and for this reason I do not assume that any of the modes is the leading one. Therefore, my analysis does not begin with one mode (which would be the main one) and then expands to the others (which would be considered ancillaries). The outcome, as I will explain in section 2 of this chapter is a focus on multimodal ensembles in which modes are revealed to work coherently to achieve a communicative aim: to persuade. The issue of whether I fully manage to avoid prioritizing modes will be resumed in Chapter 7. The rest of this section is devoted to describing and illustrating the persuasive strategies that I have identified.

6.1.1 Anticipation and control of responses

Speakers predict reactions and adapt their behaviour accordingly in order to obtain a desirable response. In other words, they anticipate responses from their audience and also trigger or prevent (i.e. control) certain responses to better suit their communicative intentions. Previous literature has shown how this can be achieved through words, intonation, gestures and head movements. Carter (1997) for example points out how the use of question tags, rhetorical questions and implicit comparisons (only one element in the comparison is explicit and the other one is left to be inferred or presupposed) all trigger a particular response from the audience. For example in “how many people would be happier” (DTLA1) the speaker leaves the second part of the comparison to infer (happier than who or what? implying happier than they are now), and also presupposes that the situation is real by framing the question as an information question with “How many” instead of a yes/no question (“Would people be happier?”).

Anticipation of responses is not only verbal. Brazil (1997) argues that a final high pitch (high termination) can be a cue to show that the speaker expects the listener to be surprised. Gestures and head movements can also be used to prevent potential counterclaims, especially when they co-occur with extreme evaluations that speakers anticipate as debatable (McClave 2000; Kendon 2002, 2004).

Concerning academic discourse in particular, Hyland (2005: 88) has pointed out how anticipating potential counterclaims is crucial in academic writing,

> because writers can only guide readers to a particular interpretation rather than demonstrate proof, readers always have the option of refuting their interpretations. At the heart of academic persuasion then, is writers' attempts to anticipate possible negative reactions to their claims.

The following example (Table 6.1) from one of the research dissemination talks (DTLI1) shows how the speaker is anticipating answers using words, gestures and head movements. The excerpt corresponds to the opening lines of the presentation.

Table 6. 1 Example of anticipation and control of responses (DTLI1)

Orthographic transcription
Are you smart? **I knew that this would come**, but let me rephrase: Is your phone smart? **You would say of course** it is a smart phone so why do you ask. Well then I say, is the clothes you wear smart? The jeans the jacket, the clothes that you have...and **it would be very silent.**
MICASE transcription
are you smart <P03> i knew that this would come but eh let me rephrase, is your phone smart <P02> you would say of course it is a smart phone_phone so why do you ask. well then i say, is the clothes you wear smart, the jeans the jacket the clothes that you have, and it would be very silent.

DI (discourse intonation) transcription[8]	
1.	ARE YOU SMART
2.	i KNEW that this would come
3.	but LET me rePHRASE
4.	is your $^{\text{PHONE}}$ smart
5.	you would say of $_{\text{COURSE}}$ it is a smart $_{\text{PHONE}}$ phone
6.	so $_{\text{WHY}}$ do you ASK
7.	well then I SAY
8.	is the CLOTHES you WEAR smart
9.	the $_{\text{JEANS}}$
10.	the $_{\text{JACKET}}$
11.	the $_{\text{CLOTHES}}$ that you have
12.	and it would be VERY SILENT

Gestures and head movements

and it would be VERY SILENT

The words marked in bold in this transcript show how the speaker explicitly claims to be anticipating answers. Three things stand out in her use of words, 1) the direct, even abrupt, questions for the audience; 2) the high certainty conveyed by the verb “knew” or the adverbial phrase “of course” (which can be considered the opposite of hedging); and 3) the repeated use of conditionals with the auxiliary “would” to describe the hypothetical situation she is anticipating. Interestingly, this anticipation is also noticeable in the gestures she uses, in particular the sweeping gesture with both hands palm down (open hand prone, OHP)

[8] Tone units are sequentially listed. Prominent syllables are capitalized. Syllables below or above the line represent a lower or higher pitch. When relevant, tone is indicated with an arrow next to the tone unit. Frequency measurements in Hertz retrieved from PRAAT

accompanying the expression “very silent”, which includes the intensifier adverb “very”. During the interview the speaker explained that she had chosen a question as a hook to open her presentation because other types of hook (e.g. an image) can distract the attention from the speaker, which is not desirable, as she was informed in specialized courses on presenting skills (see next example for a discussion of hooks). She also confirmed that her communicative intention in this case was to anticipate responses from the audience, and rather than being interested in the information she wanted to get them “activated”. As she described it in the interview, “what I did at the beginning, I imagined a dialogue, but of course I know already the outcome because I know the scientific background, I know they will not be wearing smart clothes” (DTLI1). Regarding the gesture, the speaker accounted for it as a metaphoric representation of silence. On the other hand, I interpret it as a metaphoric, pragmatic gesture having a dual meaning: a) it prevents potential counterclaims to this hypothetical situation that she is describing in a vivid way using the intensifying adverb “very” and b) it represents the scope implied (the whole audience). The speaker found both interpretations plausible and not contradictory with her initial explanation. Interestingly, the gesture is split in two parts to follow the rhythm created by the two prominences in “very silent”, which is also reinforced by subtle head nods synchronous with the prominences. Apart from this reinforcing effect, the speaker interpreted these nods as confirming that she was indeed anticipating the right answers. It is particularly interesting to note that, although words as a mode begin to play a role in anticipating answers earlier in the excerpt, towards the end the four modes are carefully orchestrated so that they work simultaneously to achieve the speakers aim: anticipate a response from the audience.

6.1.2 Attention getting

This strategy is used to raise and maintain the interest of the audience. Andeweg, de Jong and Hoeken (1998) use the term *exordial techniques* to refer to this concept and it is also popularly known as “hook" or “attention-getters”. It can take a variety of forms (i.e. realised in a variety of semiotic modes) including a particularly marked use of intonation (e.g. a high pitch), a marked use of gesture (e.g. wider amplitude or more repetitions) or the use of stylistic devices such as narratives or direct addresses to the audience. Attention getting techniques are especially relevant in the openings of the presentations, but can also appear later on. For example Andeweg, de Jong and Hoeken (1998) or May (2014) claim that having a strong introduction can make the audience more willing to listen, and also more likely to retain the message. The direct questions to the audience featured in the previous example can be considered hooks.

Table 6.2 shows how the speaker in one of the conference presentations (CPDO2) gets the attention of the audience by impersonating a character in her narrative about her research experience.

Table 6. 2 Example of attention getting (CPDO2)

Orthographic transcription
And when I showed up at 4:30 in the morning to follow them for their day of work, they were clearly apprehensive about what I was doing there, was I trying to spy on them cleaning the offices or what? What was the purpose of my being there? And I said I was PhD student, and they caught on to the student part **"Oh so you're going to do a paper. How long is your paper?** Well, it's 300 pages. **Wow, that's a long paper**. Yes it is."
MICASE transcription
and when i showed up at four thirty in the morning to follow them for their day of work they were clearly apprehensive about what i was doing there was i trying to spy on them cleaning the offices or what what was the purpose of my being there. and i said i was phd student and they caught on the student part oh so you're going to do a paper how long is your paper well it's three hundred pages wow that's a long paper yes it is

DI transcription	
1.	and when i showed up at FOUR thirty in the morning
2.	to FOllow them for their day of WORK
3.	they were CLEArly appreHENsive about
4.	what i was DOING there
5.	was i trying to SPY on them cleaning the OFFICES
6.	or WHAT
7.	what was the PURpose of my BEING there
8.	and i said i was PHD STUdent
9.	and they CAUGHT on to the STUdent part
10.	OH
11.	so you’re going to do a PAper
12.	how long is your PAper
13.	well it’s three hundred pages
14.	wow
15.	that’s a LONG Paper
16.	yes it is

Gestures and head movements

or WHAT

well it’s three hundred pages

yes it is

The use of direct speech, first person pronouns and informal speech (e.g. the phrasal verb “showed up”) makes the narrative more vivid and personal. The word choice also reveals a focus on attitudes and emotions of the characters in the narrative (“apprehensive”, “spy”, vocalizations expressing surprise) rather than on the actions, which adds to the intensity of the narrative.

The speaker marks the change of personae by using an overall higher pitch range when assuming the role of the informants in her ethnographic research (i.e. pitch becomes higher when she is not using her own words, but those of her informants, marked in bold in the orthographic transcript), but this higher pitch at the same time conveys that the informants are both surprised at, and interested in, the information they are receiving (Brazil 1997) in this conversational exchange that the speaker is recalling. This is supported by the vocalizations in units 10 and 14, which are stereotypical ways of expressing surprise. At the same time she visually represents the lack of understanding of her informants through a sequence of subtle head shakes, which adds to the vividness of the narrative. The speaker confirmed this interpretation during the interview, but also added an additional one: "maybe I am trying to illustrate that the idea of me being a spy is a bit ridiculous, or maybe illustrating their confusion, like they don't know what I'm doing there". The lateral movements of the head visually represent the researcher's feelings at the reaction of her informants. In combination with a flat unenthusiastic intonation (oblique, in Brazil's terms) they seem to convey a sense of boredom at having to repeat the same thing over and over again. During the interview the speaker offered a different interpretation: it is a way of minimising what she considers a "white lie" in saying she was a student. She did this to gain easier access, but it was not totally accurate.

This example shows three modes (words, intonation and head movements) working together to create an enactment of a research experience. When asked about her communicative intention the speaker reported that she was hoping to make it interesting for the audience: "I was trying to tell a story to try to kind of put them, yeah, in my place. So it was supposed to make it interesting, to spice things up a little. And again, it makes it personal, because I'm telling the story of what happened to me".

6.1.3 Emphasis

Emphasis consists in highlighting parts of the message so that they become more salient. Intonation can contribute through the use of prominent syllables (Brazil 1997). It can also be used in combination with rhetorical devices that make the text more memorable, such as lexical creativity, three-part lists, parallel structures, and words related to the semantic field of novelty (Lakoff 1982; Carter 1997; Bamford 2007, 2008). Emphasis can further be aided by beats (McNeill 1992; Kendon 2004).

The following example (Table 6.3) from a research dissemination talk (DTJO2) shows how the speaker uses words, intonation and gestures to emphasize a part of his message.

Table 6. 3 Example of emphasis (DTJO2)

Orthographic transcription
So think about your kitchen. If you want to heat stuff **really fast** you are not going to use an ordinary oven. You are going to use the microwave.
MICASE transcription
so, think about your kitchen. if you want to heat stuff, really fast, you are not going to use an ordinary oven, you are going to use, the microwave
DI transcription
1. so THINK about your KITchen 2. if you WANT to HEAT stuff 3. REAlly FAST 4. you are NOT going to use an ORdinary oven 5. you are GOING to use the Microwave
Gestures and head movements
 REAlly FAST

It is remarkable how a closed fist beat synchronous with the intensifying adverb “really” adds force to the extreme lexical evaluation in “really fast”, and the prominence in the intensified evaluative adverb sequence "really fast” also contributes to the emphasis. In the interview the speaker described this beat gesture as “more than emphasizing”, as a way of showing emotional involvement, and he pointed out how his facial expression at this point is also consistent with this intention.

6.1.4 Evaluation

Evaluation occurs when speakers assess something and are thus implicitly inviting the listener to accept this opinion (Bamford 2007; Querol-Julián 2011). Hood and Forey (2005) point out how speakers can include multimodal expressions of attitude in their introductions that subtly evaluate their presentations in positive terms and seek alignment with their audience. Along this line, Pomerantz (1986) notes how extreme case evaluations are frequently used to legitimize claims when speakers expect possible counterclaims. Interestingly, these claims are commonly accompanied by head shakes or gestures, which seem to deny in advance a potential counter-argument (Kendon 2002). This was the case in the example shown in Table 6.3 above. Another example of multimodal evaluation is provided in the following excerpt from the research dissemination talk DTRA2 (Table 6.4).

Table 6. 4 Example of evaluation (DTRA2)

Orthographic transcription
Cool, I think. Well. If you're geeky enough you should find it cool.
MICASE transcription
cool, i think well, if you're geeky enough, you should find it cool <AUDIENCE LAUGH> eeeh
DI transcription
1. COOL 2. i THINK 3. WELL 4. if you're GEEky eNOUGH 5. YOU SHOULD find it cool 6. eeehh

Gestures and head movements

if you're GEEky eNOUGH

YOU SHOULD find it cool

In this excerpt there are two head nods synchronous with the adjective “geeky” and the verb “find". The speaker had not initially noticed them until I pointed them out to him when watching the video during the interview, which shows that they were done unconsciously. The speaker interprets them as “emphasizing my agreement, I think" (meaning agreement with the evaluation he had just expressed with the adjectives “cool” and “geeky”). I believe they also prompt to share this positive evaluation and concur with the adjective “cool”. It is also interesting to note that the second nod precedes words, being synchronous with “find" and not with the adjective expressing the positive evaluation itself (“cool”). This is probably a result of the greater encoding effort implied by using words as opposed using gestures or head movements, which are easier and faster to produce (McNeill 1992). Regarding intonation, unit 4 has a high key (+ 124 Hz) that adds the meaning of reversal of expectations. This is consistent with the innovative use of slang (“geeky", “cool”) in this situation. The speaker commented in the interview that he was aware of this use of unexpected vocabulary (which supports claims that persuasion is more effective when it is unexpected, O’Keefe 2002;

Perloff 2003) and agreed with me that his intention with this was to set a humorous tone. Interestingly, unit 5 begins with a low key (-37 Hz) that has an equating effect (being geeky= find cool) and ends with a mid termination that prompts concurrence. Interestingly, the use of a conditional structure with “if” in combination with the mild command “should” also supports this equating effect conveyed by intonation, because it prompts the logical conclusion “if you don’t agree with me then you are not geeky”. In sum, this example shows the orchestration of words, intonation and head nods working together to convey a positive evaluation and prompt concurrence with it.

6.1.5 Processing aids

Hood and Forey (2005) and Rowley-Jolivet and Carter-Thomas (2005) maintain that the immediacy of an audience and the consequent need for real-time processing strongly influence the way a presentation is designed. By *processing aids* I mean resources that help the message to be processed in real time and make it easy to understand, such as visual metaphors materialised through gestures that can help clarify aspects of the referential content of the message, or a use of intonation that clarifies the structure of a message (e.g. contrasting elements, or elements in a list). The following example (Table 6.5) illustrates a processing aid in a product pitch (PPPA2).

Table 6. 5 Example of processing aids (PPPA2)

Orthographic transcription
(…) because the **combination** of these two things is the **real magic**
MICASE transcription
because the combination of these two things, is the real magic.
DI transcription
1. because the COMbination of these two things 2. is the REAL MAgic

Gestures and head movements
 because the COMbination of these two things is the REAL MAgic

The referential gesture with fingers coming together visually conveys the idea expressed in the noun "combination". The gesture is later repeated synchronous with the evaluative adjective + noun sequence "real magic", which is logical if we consider that these two elements are being equated (combination = real magic). Note also how the relevant words ("combination", "real" and "magic") are prominent. Interestingly, the speaker mentioned in the interview that he was not happy with this choice of words because "to me I feel the phrase 'real magic' is almost a little bit too marketing-ish and I might avoid it if I'm superconscious". However, he agreed that his intention with the gesture in synchrony with these words was to represent his meaning visually, which ultimately would make the message easy to follow.

This example shows once more how gestures, words and intonation are orchestrated, in this case to make the message easy to understand for the audience.

6.1.6 Projection of context of interaction

Context of situation is a notion previously used by Malinowski (1923), Odgen and Richards (1923) and Firth (1957). Brazil (1997) uses a similar concept when he talks about *projection of the context of interaction*. According to him, speakers use intonation to project their understanding of the context of interaction as they communicate to negotiate and classify information: what is considered new vs. given information or contrastive (selection among a range of options) vs. non-contrastive (no other options available). In other words, speakers make moment-to-moment decisions concerning whether a piece of information will simply refer to the context of interaction as it currently is, or on the contrary will modify it (because it adds something new). In the context of presentations this means that speakers present some parts of the message as shared with the audience and agreed upon, as opposed to new and open to discussion. In Example 1 (DTLI1) in this chapter when the speaker pronounces

you would say of $_{\text{COURSE}}$ it is a smart $_{\text{PHONE}}$ phone

she is using intonation (i.e. a low key and termination) to mark this idea as obvious and out of discussion. The following example from the conference presentation CPAM1 (Table 6.6) also illustrates this strategy.

Table 6. 6 Example of projection of context of interaction (CPAM1)

Orthographic transcription
So one of the – shall we say, well, the distinctive feature of our Master's as compared to other Master's in International Human Resources Management is the fact that we have a strong concentration on foreign language skills and cultural knowledge.
MICASE transcription
so one of the, shall we say_well the distinctive, feature of our master's as compared to other masters in international operations management, is the fact that we have a strong concentration, on foreign language skills and, cultural knowledge.
DI transcription
1. so ONE of the 2. shall we SAY 3. well $^{\text{THE}}$ distinctive FEAture 4. of OUR master's 5. as comPARED to Other masters in international operations management 6. is the FACT that WE have 7. a STRONG concenTRAtion 8. on FOreign language SKILLS 9. AND 10. CULtural KNOWledge

Gestures and head movements

well $^{\text{THE}}$ distinctive FEAture

Prominence and high key in the determiner “the” (unit 3 in the transcript) work together to achieve a twofold effect: on the one hand it makes “the” contrastive (new information not recoverable from context), and at the same time rules out other options (i.e. there are no other distinctive features). This use of intonation matches the self-repair in “one of the…well the distinctive…”. A sequence of four head nods synchronous with “the distinctive feature” has a reaffirming effect. Since these nods occur after a hedging filler (“shall we say”) plus a false start, they can be interpreted as a way for the speaker to reassure herself that she is formulating the message in the correct way and that she is really saying what she wants to say (i.e. this is the main distinctive feature). The fact that she is not keeping eye contact with the audience at these moments supports this interpretation, and can be interpreted as a sign of encoding effort that is also suggested by the false start and filler. At the same time, with this nod she may be implicitly preventing any potential counterargument to what she is saying. In the interview the speaker agreed this is a plausible interpretation, given the defensive attitude that she perceives she was adopting at this point in her presentation, which she thinks is probably due to the fact that she was unsure about fitting in the whole event.

To sum up, this example shows how the speaker orchestrates words (self-repairs and hedging), intonation (prominence) and head nods to present a piece of information as not debatable and prevent potential counterclaims.

6.1.7 Rapport

Projection of the context of interaction is related to the concept of *rapport*, which in the context of presentations refers to a relationship of sympathy and mutual understanding with the audience. Tickle and Rosenthal (1990) identify three essential components of rapport, and show how they can be expressed non-verbally:

i) mutual attentiveness that can be expressed through spatial configurations and bodily postures that signal accessibility;

ii) positivity, which can be materialised through smiling and nodding that signal liking and approval;

iii) coordination, which can take the form of postural mirroring and interactional synchrony.

Intonation can also achieve these abovementioned effects in a more subtle way than words. It can be used in combination with inclusive pronouns that help create rapport with the audience and enhance the sense of shared knowledge (Rowley-Jolivet and Carter-Thomas 2005; Hyland 2005; Bamford 2007, 2008). Gestures in conference presentations can also help create this sense of inclusivity, bringing the audience into the discussion and establishing common ground (Hood and Forey 2005; Holler 2010).

The following example (Table 6.7) from a research dissemination talk (DTPI1) shows the use of words and gesture to build rapport.

Table 6. 7 Example of rapport (DTPI1)

Orthographic transcription
I’m just trying to get you from A to B, that’s all I want to do.
MICASE transcription
i’m just trying to get you from a to b. that’s all i want to do.
DI transcription
1. i’m JUST trying to GET you from a to b 2. that’s ALL i want to do

Gestures and head movements
 that’s ALL i want to do

In this example, the speaker is using words that minimise his role and are humble. For example he uses the adverb of degree “just”, and he equates his research objective (“all I want to do”) with something apparently unimportant (“take you from A to B”). In doing this he is making himself likable and therefore paving the ground for rapport. The open hand supine (OHS) metaphoric gesture synchronous with “that’s all I want to do” suggests openness, fulfilling a pragmatic function. According to the speaker, he wanted to transmit “there’s nothing more to it”. The gesture can be described as Kendon’s OHS palm lateral gesture, which, according to him, is found in situations in which speakers shows unwillingness or inability to intervene or position themselves with regard to something, such as a question for which there is no answer, a situation that cannot be helped, or when something is obvious or redundant (implying that nothing further can be said about it). Note how this last interpretation is to a great extent consistent with the speaker’s. The meaning of the words and the gesture reinforce each other to convey ‘honest humbleness’ and the whole ensemble is consequently prone to create rapport.

To summarise, I have combined ethnography and multimodality to analyse the persuasive behaviour of speakers in three types of presentations. I have used a data-driven approach in a video-based MDA analysis, which has revealed that speakers orchestrate words, intonation gestures and head movements in multimodal ensembles to convey persuasion. In addition, I have classified this persuasive behaviour into seven persuasive strategies, i.e. anticipation of responses, attention getting, emphasis, evaluation, processing aids, projection of context of interaction and rapport.

As the examples of the seven strategies shown in this section illustrate, it is actually the multimodal ensemble taken as a whole that becomes persuasive, to the point that it is difficult to delimit the contribution of each mode and it is not always possible to establish one to one relationships between modes and persuasive strategies, neither to clearly separate the latter. As the previous examples show, most of the time they co-occur and sometimes the boundaries among them blur. In spite of this, I think the taxonomy of strategies presented here, although by no means exhaustive, can serve as a good starting point to approach persuasive activity in oral presentations from a multimodal perspective and shed light on the contribution of multimodal ensembles to persuasion. I would like to conclude this section by summarising my answer to **RQ I:**

What type of persuasive strategies can be identified using an MDA approach?

I argue in favour of a video-based, data-driven type of study that combines input from previous literature with ethnography and MDA, and is aided by specialised software. The outcome of such a study in my corpus of presentations is the seven strategies presented in this section, but I believe further research can shed light on more persuasive strategies that speakers apply multimodally, or more complex realisations of these strategies in meaningfully orchestrated multimodal ensembles.

6.2 The contribution of multimodal ensembles to persuasion

In the second section of this chapter I will provide an answer to RQ II:

RQ II	**What is the contribution of the multimodal ensemble orchestrated through words, intonation, head movements and gestures to the creation of persuasion in the genres included in the analysis?**
RQ II. a)	Can persuasion be considered multimodal in these genres?
RQ II. b)	How do modes orchestrate in multimodal ensembles to create a persuasive message?
RQ II. c)	What type of persuasive activity (i.e. persuasive efforts on the part of the speakers) is more likely to be realised through the modes which are the object of this study? When, how and how often does it happen?

To answer this research question I have undertaken an ethnographic and multimodal analysis of the speakers' use of persuasive strategies in the 30 rich points selected from the corpus of conference presentations, research dissemination talks and product pitches. Although the results presented lack the representativeness required to be considered conclusive for genre analysis and they cannot be generalised, they are valuable as illustrations of what an ethnographic and multimodal approach can offer to larger-scale genre studies. I have structured section 6.2 in two subsections (6.2.1 and 6.2.2), The first provides a specific answer to the subquestions a) and b) above. The results presented in this section are valid for the three genres. The second answers subquestion c) above, and presents results that are specific for each genre, which are the basis for the comparative analysis presented in the last section of the chapter (6.3). It must be noted that the references to the three genres throughout this section and the subsequent sections in this chapter are not meant as attempts at generalisations: I am using the terms 'genre', 'conference presentations', 'research dissemination talks' and 'product pitches' to refer exclusively to the three subcorpora that I have analysed, and not to the three genres as a whole.

6.2.1 Modal density and modal coherence

RQ II. a)	Can persuasion be considered multimodal in these genres?

The results of the analysis indicate that regardless of the genre, the moments in the presentations that have been identified as showing high persuasive activity and selected as rich points tend to feature *high modal density*. With high modal density I mean that the speakers are using several modes to fulfil persuasive strategies. Tables 6.11, 6.12 and 6.13 in section 6.2.2 offer a detailed rundown of the use of modes in each rich point.

RQ II. b)	How do modes orchestrate in multimodal ensembles to create a persuasive message?

Results also show that, regardless of the genre, the rich points feature *high modal coherence*: the modes work coherently towards the same communicative aim, and they do not contradict each other.

It will be noted that both number of modes and coherence among them is broadly what Norris (2004) calls *modal density*, since according to her, the number of modes in itself is not significant, but rather the complex interactions between them. However, in my analysis I use the term *modal density* to refer to the number of modes and *modal coherence* to refer to the way these modes are *orchestrated* into coherent *ensembles*, as Kress (2010) puts it. I have considered this distinction useful because I think that both aspects (i.e. number of modes and coherence among them) can have an influence in the persuasive effect. For example, some of the rich points selected and analysed in this study may show ensembles with only two modes, which can still be persuasive if they are used coherently.

I would like to illustrate these findings about modal density and modal coherence discussing a rich point from a conference presentation (CPPE1) and a rich point from a product pitch (PPKE2).

In CPPE1 the speaker is particularly skilful in the combination and coordination of modes, as shown in Table 6.8, representing an excerpt from the introduction of her presentation.

Table 6. 8 Orchestration of modes in CPPE1

Orthographic transcription
Specific for my case was that I found tension between organizational culture or practices and what I wanted to do as an ethnographer, I wanted to develop contacts with clients and more personal relationship and have deep and long conversations but the agency, their approach to the female clients is more distanced, commercial, and if there must be personal contact they want to keep it to a limit.
MICASE transcription
specific eh for eh my case was eh that i found_found tension between organizational eh culture eh, or practices and what i wanted to do as an ethnographer eeeh i wanted to eh develop contacts with clients and more personal relationship and have deep and long conversations but the agency their approach to the female clients is more distanced eh commercial and eh if there must be personal contact eh eh they want to keep it to_to a limit.

DI transcription
1. speCIfic eh for eh my CASE 2. WAS 3. eh 4. that i found found TENsion beTWEEN 5. organizational eh CULture eh or practices ↗ 6. and what I wanted to DO as an ethnographer ↘ 7. eeeh 8. i wanted to eh develop CONtacts with CLIENTS 9. and more PERsonal reLAtionship 10. and have DEEP and LONG conversations 11. but the Agency 12. their apPROACH to the female clients is more DIStanced eh 13. comMERcial 14. and eh if there MUST be personal contact 15. eh eh 16. they want to KEEP it to to a LImit.

Gestures and head movements

and what I wanted to DO as an ethnographer

i wanted to eh develop CONtacts with CLIENTS

and have DEEP and LONG conversations

their apPROACH to the female clients is more DIStanced eh comMERcial

they want to KEEP it to to a LImit

The use of first personal pronoun “I” / “my” draws attention to the figure of the researcher, as does the use of the verb “want”, which suggests a personal desire rather than a requirement of the research process. This makes the text closer to a narrative of a personal experience. The lexicon also expresses the two attitudes she is contrasting: on the one hand her stance as a researcher, and on the other hand the position of the company. An example of this are the verbs used to refer to “contact”: “develop” (researcher) vs. “keep to a limit” (company).

Turning to intonation, the high termination in “culture” in tone unit 5 together with the rise tone anticipates and prepares the audience for the contrast with “what I wanted to do” in the following tone unit. Prominence in the words that describe the attitudes of both parties further emphasizes this contrast (e.g. “contacts”, “personal”, “deep”, “long” vs. “distanced”, “commercial”). There are numerous fillers and two false starts in the extract. They reveal encoding difficulties, and the speaker confirmed during the interview that she was struggling to find the right words. In this particular case rather than spoiling the persuasive effect this actually adds to it, because it reinforces the spontaneity and honesty of her text and establishes a closer relationship with the audience.

Regarding gestures, as the speaker pronounces “I” in “what I wanted to do” she points to herself, visually directing the attention and highlighting the figure of the researcher. This is an OHS, deictic, referential gesture in phonological, pragmatic and semantic synchrony with words and intonation. Next, the speaker makes an OHS gesture with alternating hands moving in outward circles (first right hand and then left hand; unfortunately the movement is partly hidden by the lectern). This gesture is repeated twice as she pronounces “contact with clients” and “deep and long conversations”. It is phonologically synchronous with the prominent syllables in “contact”, “clients”, “deep” and “long” (they start slightly before the prominence and continue throughout). It is a metaphoric, referential gesture that seems to represent visually the social interaction that is implied with the lexicon used. The fact that it is repeated also provides it with a cohesive function, since it connects different words that are related (in this case lexicon referring to social interaction). During triangulation of results the speaker added a second interpretation of this gesture as a processing aid for her that helped her encode the verbal component of her message. However, she agreed that both interpretations are not mutually exclusive. The next gesture the speaker makes is an OHP that begins at the prominence in “their apPROACH to female clients is more DIStanced” and continues up to the middle of the next tone unit (“commercial”). It is a metaphoric gesture that symbolises the “distance” referred to (as if saying “stop coming closer”), therefore in semantic synchrony with words. Finally, the speaker makes an OHS gesture that precedes and continues up to the first prominence in “KEEP it to a LImit". This is an iconic, referential gesture that visually represents a barrier the company is imposing to restrict personal contact with clients. It is interesting that the gesture in this case is not in semantic and pragmatic synchrony with words, but rather precedes them. This is likely to respond to encoding difficulties (gestures are easier to encode and therefore faster to produce than words).

To sum up, the speaker uses the following persuasive strategies:

i) Builds rapport trough a personal narrative (words);
ii) Makes her message easy by using different processing aids: visual representations through gestures and marking contrast through intonation and opposite words;
iii) Controls a response, preparing for the contrast she wants to convey through intonation;
iv) Puts the emphasis on the person of the researcher and on ideas that convey the contrast discussed using intonation and gestures.

There are two key aspects in this rich point around which persuasion is built, and the four persuasive strategies mentioned above are used to achieve them. The first is a contrast between what the researcher wants to do vs. what she is allowed to do. All the modes examined contribute to this effect: words (“develop” vs. “keep to a limit”), intonation (rise tone that highlights the contrast at the crucial point in the transition from unit 4 to 5) and gestures (contrast between the rotating hands symbolizing flow of social interaction vs. the barrier gesture at the end). The second aspect is a personal and honest sharing of research experiences, again supported by all modes. It is remarkable how the “I” in “I wanted to do” is prominent and further emphasized by a self-deictic gesture. Other paralinguistic features (e.g. vocalisations as fillers), although out of the scope of this study, also contribute to make her text not rehearsed, honest and natural.

As the speaker reported during the interviews, she was unsure of whether her research would fit in the conference. Her efforts to seek rapport and honestly share her research experience are probably a consequence of this perception and reveal efforts to adapt to the specific audience.

The modal configuration, or the multimodal ensemble that is orchestrated in this excerpt can be represented schematically as Figure 6.2 shows, with all modes working towards the higher level action of persuading the audience by means of different persuasive strategies (difference in size shows difference in the level of use of each mode, i.e. how many strategies the mode is realising):

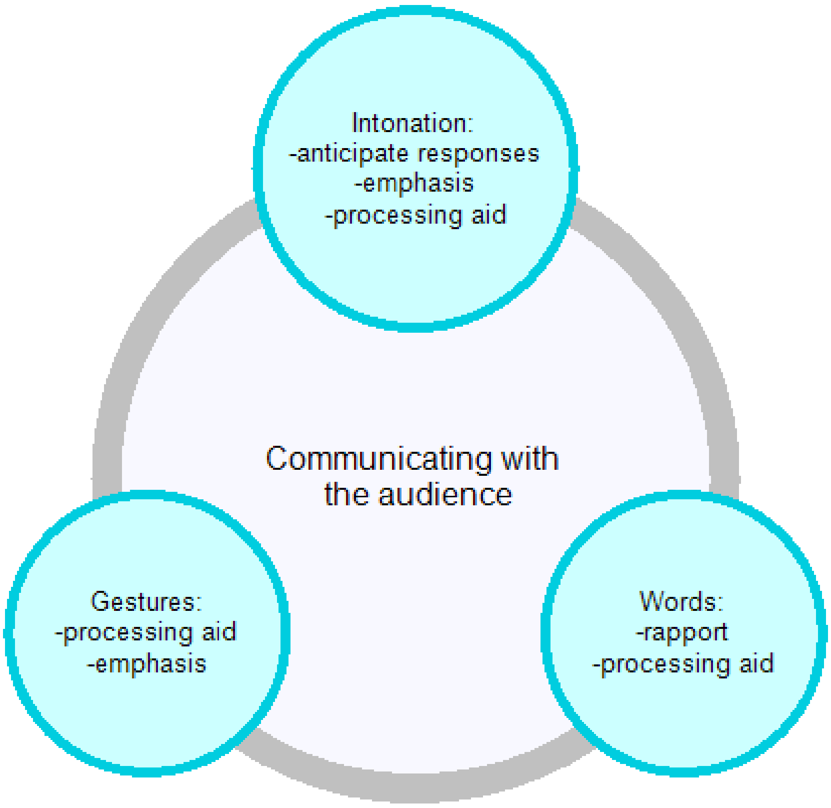


Figure 6. 2 Multimodal ensemble in CPPE1

Figure 6.3 below represents both the modal density (i.e. the number of modes), and the modal coherence, because it shows how different modes work together to realise a persuasive strategy. If several modes are used to realise a single strategy, I consider that they are used coherently. As the bar chart indicates, the speaker uses the strategy ‘processing aids” through gestures, intonation and words, and the strategy ‘emphasis’ through gestures and intonation. Even if ‘anticipation’ and ‘emphasis’ are realised by only one mode each, they do not detract from the overall intended effect, but rather contribute to it, so they are still coherent. In this case the results are specific for this rich point, so they are counted as instances and not as a frequency per minute: each modal realisation of each strategy is one instance (anticipation through intonation = 1 instance)

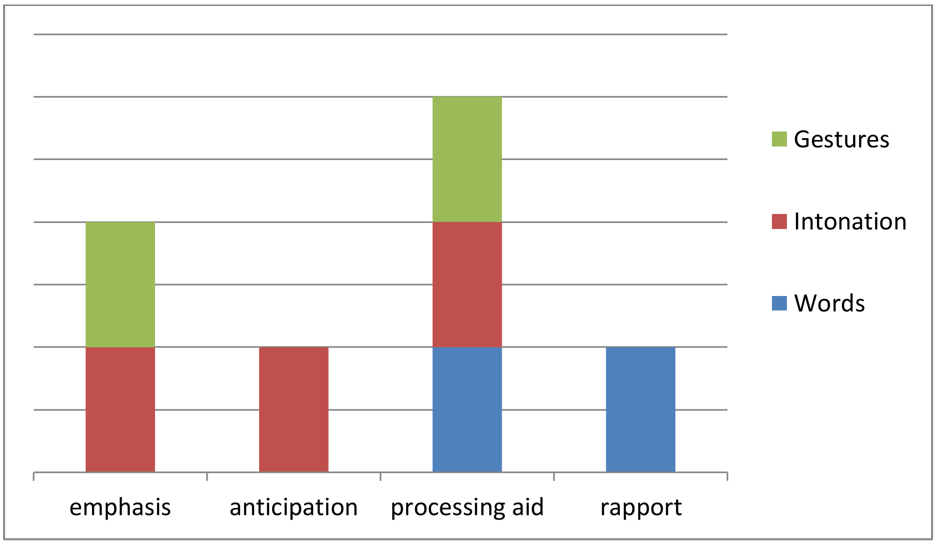


Figure 6. 3 Modal density and modal coherence in CPPE1

In conclusion, the skilful orchestration of 3 modes into a coherent multimodal ensemble to apply 4 different persuasive strategies, together with the speaker's awareness of the need to adapt to her audience, as reported in the interview, both show that the she is engaged in one higher-level action (Norris 2004): convincing the audience by using all the resources available.

In contrast, in the product pitch PPKE2 the speaker's efforts seem to be somehow more divided: although the speaker is trying to convince his audience of the value of the app he is introducing, he is also concentrated on remembering the exact words he is supposed to be saying according to his script. Table 6.2 shows his use of modes in the excerpt, which is the closing of the presentation.

Table 6. 9 Orchestration of modes in PPKE2

Orthographic transcription
So I have just one thing to say, just download the app and you can see what's in it. Thank you for your attention.
MICASE transcription
so, i have just one thing to say, just download the app, and you can see what's in it. thank you for your attention.

DI transcription

1. so i have JUST one thing to SAY →
2. just DOWNload the [APP] ↘
3. and you can SEE what's in it →
4. THANK you for your attention →

Gestures and head movements

just DOWNload the [APP]

A rather flat intonation makes this catchy close go largely unnoticed. There is, however a significant high termination in unit 2 (+ 425 Hertz) that prompts an active role from the audience and matches the imperative form of the verb. The speaker confirmed this was consciously done for this effect of prompting people to action. The effect is, however, somehow counteracted by the level tone and mid termination and key in the last unit that closes the speech.

The only gesture in this extract (pointed out by the speaker himself during the interview) is a subtle shrugging of shoulders synchronous with "download" that actually gives away a sense of disengagement that contradicts the call to action. The speaker interprets it as a sign of feeling tired due to the effort of remembering the exact words in his script, and being willing to finish the stressful experience of the presentation.

Persuasion in most of this presentation is framed as a demand-supply: there is a trendy market niche that we know how to feasibly occupy. It is rational and it works towards highlighting what differentiates the product and the feasibility of the market idea. The closing presented in this rich point, however, is more emotional. The speaker is trying to engage the audience and prompt a response (i.e. try the app) by appealing to sensorial experiences ("see what's in it"). The focus here is not on the practical value of the product, but on its emotional value as something simply enjoyable. The persuasive effect would then probably benefit from interpersonal gestures that would get the audience involved and prompt them to action, which are, however, lacking in this extract. A more marked intonation would also help make this appeal more prominent and likely to be reacted upon. The speaker agreed with this interpretation and actually mentioned that if he were to repeat this presentation he would change two things: first he would not memorize a text (as he was strongly recommended to do during the training he received) because that made him focus on remembering the exact words rather than on communicating with the audience, and second, he would mimic the action of downloading or even demonstrate it with a phone instead of holding a pointer in his hands.

The multimodal ensemble orchestrated in this excerpt and its modal density/coherence are represented in Figures 6.4 and 6.5 below respectively.

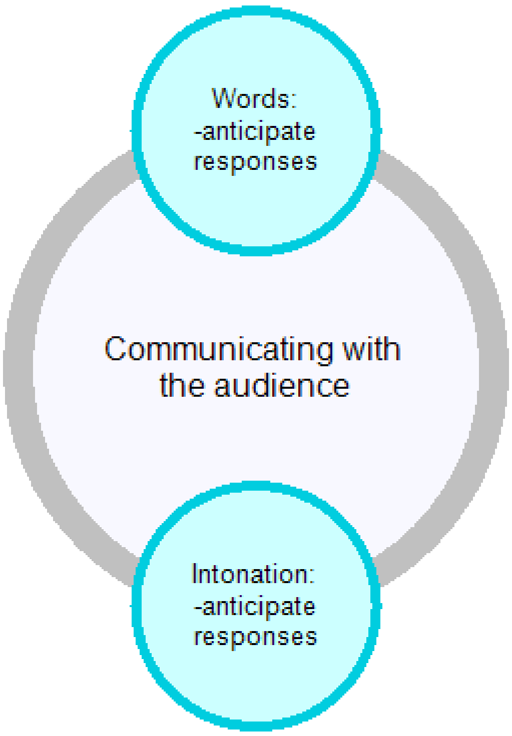


Figure 6. 4 Multimodal ensemble in PPKE2

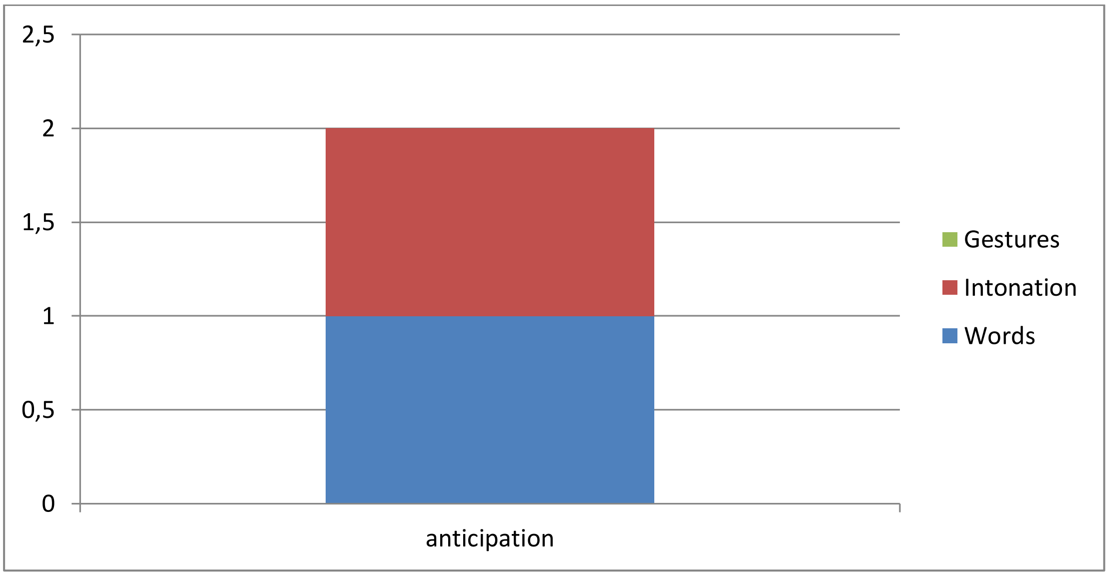


Figure 6. 5 Modal density and modal coherence in PPKE2

It will be noted that the gesture does not contribute to the higher-level action of persuading the audience. On the contrary, it detracts from it. For this reason, it is not part of the persuasive multimodal ensemble. The same applies to the flat intonation with oblique orientation. Words and intonation are only partly coherent in this rich point. They cohere to the extent that an imperative form of a verb is used in conjunction with a high termination showing that the speaker expects an active role from the audience. They do not cohere in the sense that a catchy slogan is used with an oblique orientation, which distracts attention from it.

In conclusion, the first example seems to be more persuasive because there are more modes used (it is more modally dense) and because they are used more coherently (it is more modally coherent) with the clear focus of convincing.

6.2.2 Frequency of strategies and their modal realisations

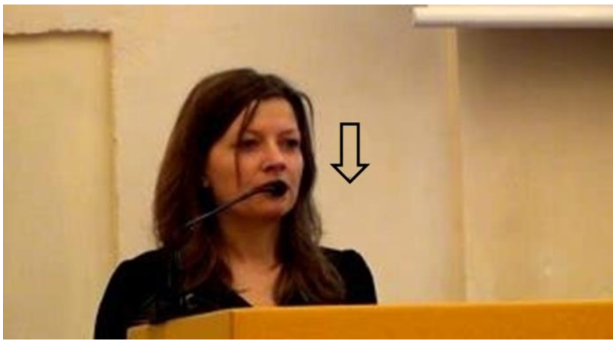
RQ II. c)	What type of persuasive activity (i.e. persuasive efforts on the part of the speakers) is more likely to be realised through the modes which are the object of this study? When, how and how often does it happen?

In this section of the chapter I provide an answer to RQII. c), focusing on how frequently strategies are used in each of the subcorpora analysed and what are the modes that realise them in each case.

An interesting result yielded by the analysis is that the more frequently a strategy is used, the more frequently it is found to be realised by different modes. This is clearly the case of emphasis, which is by far the most common strategy in the three genres. In section 6.1 in this chapter, I provided an example of emphasis in research dissemination talks in which words, gestures and intonation combined to provide emphasis to the message. Table 6.10 shows another example of the use of emphasis, this time in a conference presentation (CPDO1), using words, intonation and head movements:

Table 6. 10 Emphasis in CPDO1

Orthographic transcription
They distinguish between three levels of identity. And in most of the literature that I’ve read in preparation for this paper focuses on these macro, sociological categories. So the position of the researcher in terms of gender, in terms of race, in terms of ethnicity has been discussed a lot, but I haven’t seen much attention paid to the more, the other levels here, the local ethnographic categories, the more transitory roles. So I’ll try to pay attention to that in my paper.
MICASE transcription
they distinguish between, three levels of identity, and, in most of the literature that i’ve read in preparation for this paper focuses on these macro sociological categories so, the position of the researcher in terms of, gender in terms of race in terms of ethnicity has been discussed a lot but, i haven’t seen much attention paid to the more_eh the other, levels here, the local ethnographic categories the more transitory roles so i’ll try to, pay attention to that in my paper.
DI transcription
1. they disTINguish beTWEEN 2. three LEvels of iDENtity 3. and 4. in MOST of the LITerature that i’ve read in preparation for this paper 5. FOcuses on these MAcro sociological categories 6. so 7. the poSItion of the reSEARcher 8. in terms of GENder 9. in terms of RACE 10. in terms of ethNIcity 11. has been disCUSSED a LOT 12. BUT 13. i haven’t SEEN much attention paid to the MORE 14. eh 15. the Other LEvels here 16. the LOcal ETHnographic categories 17. the more TRANsitory roles 18. so i’ll try to pay atTENtion to THAT in my paper

Gestures and head movements
 in terms of GENder
 the Other LEvels here

The speaker identifies a research gap and occupies a niche by making a contrast between “discussed a lot” (previous research) vs. “I haven’t seen much attention” (gap). This contrast is also reinforced by a brief pause immediately following the adversative “but” (which is also prominent) and a head nod immediately following "other”. The speaker interpreted this pause + nod sequence as a momentary focus on her notes. I think it highlights her research gap and reaffirms her occupation of the niche: therefore, apart from being emphatic, it also has a pragmatic function in its reaffirmation that can be considered both a prompt to approve her research topic choice and a reassurance to herself that she had expressed it the right way. In the interview, the speaker thought this is a plausible explanation. In addition, the facial expression with pursed lips seems coherent with this pragmatic function. At the beginning of the excerpt the speaker also uses a three-part list embedding a symmetrical structure (“in terms of”) that makes the message easier to remember, and emphasizes each item with prominence and a head nod.

The use of prominence is particularly interesting in this excerpt. “Most” in unit 4 and “macro” in unit 5 are prominent, because they highlight what has been researched so far and prepare for the identification of the gap. As mentioned earlier, “but” in unit 12 is also prominent and introduces a contrast that shows the existing research gap anticipated before. “Other" in unit 15, “local” and “ethnographic” in unit 16 and “transitory” in unit 17 are also prominent. They constitute the second element in the contrast previous vs. new research. They all show her occupation of the niche and how she is contributing with new research. Interestingly, the speaker commented during the interview that the contrast between previous research and her contribution was something she had consciously planned to do at the word level, but she was not conscious of how her use of intonation and head movements were contributing to it.

In the remaining of this section I present the results of the analysis of the rich points of each of the three genres included in the study: conference presentations, research dissemination talks and product pitches. I begin each subsection with the results of the ethnographic analysis and then I turn to the multimodal analysis. The results of both analyses will be fully integrated during the discussion of results in Chapter 7, where I use the input from the ethnographic analysis to interpret the multimodal persuasive behaviour of the speakers. The analysis focused on identifying the persuasive strategies used and their modal realisations. These results are the basis of the comparative analysis presented in the last section of this chapter (section 6.3). Although the analysis focused on four specific modes (words, intonation, gestures and head movements), these modes are parts of a whole, a multimodal ensemble, and their interpretation is dependent on this whole. For this reason, references to other modes (e.g. facial expression or gaze) are occasionally needed to provide an accurate interpretation. Sometimes these references to other modes are offered by the speakers themselves during triangulation of results, and help to confirm or refute my initial interpretations.

As mentioned in Chapter 5, persuasive strategies are not discrete entities that can be easily quantified and co-related with modes in one-to-one relationships. For the sake of facilitating the subsequent comparative analysis, the identification of strategies and modal realisations was carried out in the following way:

i) The four modes included in the analysis (words, intonation, gestures and head movements) were analysed to see how they contributed to persuasion (i.e. what persuasive strategies they were enacting);
ii) Each modal realisation of a persuasive strategy was counted once if the speaker used it in the rich point analysed. The same strategy with the same modal realisation is not counted twice in the same rich point.

It is also important to note that the same mode can be simultaneously contributing to different strategies. For this reason, it is important to understand each mode in relation to the complete multimodal ensemble.

6.2.2.1 The conference presentations

I begin this account of the results obtained for the genre of conference presentations by offering a summary of the insights provided by the ethnographic interviews.

During the interviews, the speakers described their motivation to present in academic contexts as primarily filling a research gap. For example CPPE explained that her main goal with her presentation was to disseminate ethics in research because she thought it is a gap. Other drivers mentioned were the opportunity to disseminate research, obtaining feedback or networking. For example CPDO described her motivation as “getting new input”.

Regarding expectations and prior preparation, it is interesting that the title of the symposium and the call for abstracts influenced the way participants prepared with a stronger focus on methodology and self-reflection on research process that on results and data. Most speakers started from the abstract they had previously sent, worked on a power point, prepared some written notes, and rehearsed for timing. None of them memorized a text or read from notes and in the interviews they also expressed concern for preparing their presentations in excess. Only one speaker used handouts (CPRE). CPDO, for example, explained that “the aim is kind of to rehearse so much that I don’t need my notes any more, but in the end I did have 6 pages in front of me”. She also mentioned that “you can also overprepare, and you become tired of the topic itself”.

The ABC symposium gathered together a varied group of people, including members of the association, invited keynote speakers, professors from different backgrounds, pre and post-doctoral students, etc. Due to this heterogeneity, some speakers reported that they felt out of their field: “like an outsider” (CPAS), “I expected to feel a little bit like a fish out of water, which I did a little bit”(CPAM) and they adapted their presentations to make them interesting for the particular audience. They pointed out that in more familiar contexts they would have included more detailed data and would have focused on results rather than on methodology. For example, CPDO described this event as “different from the linguistically oriented conferences I usually go to”, and CPAM remarked “I am now presenting at a Business English conference, and I am more used to a more linguistic kind of audience”. She further explained that “from the title of the conference it sounded like my presentation was very in tune with what the theme of the conference was, but I haven't seen any other presentation like mine". They also felt they could assume less shared knowledge and could not take for granted a common interest in their specific fields. Therefore, speakers were concerned about being accessible for the audience. For example, they were worried about using examples that are too culturally specific: "I was afraid that some of the things I said wouldn't translate very well or that people would have a different understanding" (CPKE).

Favourite presentations are described as:

i) Visual (CPAM)

ii) Easy to follow (CPDO)

iii) Related to other presentations (CPDO)

iv) Memorable, engaging (CPDO, CPAM)

v) Honest and humble (CPAS)

vi) Relevant for the audience (CPAM)

vii) Not overloaded with technical content or jargon

viii) Having a sharing tone (CPDO)

Speakers appreciate a small-scale event that allows for more interaction. CPDO, for example explained that "it was presented as a small conference (…) with a focus theme", and CPPE commented "I also like it that it is a small conference, it's very interactive". This interactiveness is related to the idea of sharing research practices, which is one of the motivations reported by the speakers, and to the preference for presentations that try to share rather than teach. In this sense, it is interesting that speakers favour the practice of referring to other presentations and that they look for ways to fit in the communicative event. One speaker, for example, confirmed during the interview that the effort to fit in the event and share experiences was actually conscious, because he was aware that he was "somewhere in the middle" of these ins and outs. He also explained how he kept changing the presentation until the last minute to include references to other presentations and make it more coherent with the whole conference. For him it was really important to show a "clear link with the subject or topic of the conference" and prove that "this is not the tenth time I'm bringing this" (CPTO). A second speaker also mentioned that she particularly enjoyed a presentation that used references to other presentations in the event to show the relevance of her contribution by showing its connection with others: "She was able to draw on what other people had said"

(CPDO). Regarding the concept of "sharing", it is interesting to note that after triangulation of results, one of the speakers (CPPE) admitted that she found confronting to see the overall effect of the modes she was using in the rich points analysed, because it made her realise how explicitly she was positioning herself as an outsider with respect to the object of her study, even judging it as strange. As an ethnographer, this is a tendency that she usually tries to avoid, but in this case her multimodal behaviour was giving herself away. In a way, this shows how the speaker was honestly sharing experiences rather than describing a methodology in an impersonal way, in which case these subjective nuances would not be present. In the same line, another speaker commented that relying on anecdotes from her field work made her presentation more effective (CPDO).

During the interviews speakers also expressed some concerns, such as adding too much text to their slides, which made their presentation less appealing and more difficult to process (CPAM). Next is having a weak conclusion, which implies not being remembered (lack of memorability). Feeling unsecure about technology was also mentioned (CPKE), being a distracting factor that prevents from focusing on the audience. Finally, for some speakers not preparing enough also resulted in insecurity and less credibility (CPPE).

Regarding the multimodal analysis, the following table (Table 6.11) represents the strategies and modal realisations identified in each rich point, offering a summary of the use of each mode in the subcorpus of conference presentations.

Table 6. 11 Rich points in conference presentations

CPDO1	**WORDS**	➢ Emphasis ➢ Processing aid ➢ Rapport	**CPDO2**	**WORDS**	➢ Attention getting ➢ Rapport
	INTONATION	➢ Emphasis ➢ Processing aid ➢ Projection of context of interaction		**INTONATION**	➢ Attention getting ➢ Processing aid ➢ Projection of context of interaction
	GESTURE			**GESTURE**	
	HEAD	➢ Anticipation and control of responses ➢ Emphasis ➢ Processing aid		**HEAD**	➢ Processing aid

CPTO1	**WORDS**	➢ Anticipation and control of responses ➢ Emphasis ➢ Rapport	**CPAS1**	**WORDS**	➢ Attention getting ➢ Emphasis ➢ Projection of context of interaction
	INTONATION	➢ Emphasis ➢ Processing aid		**INTONATION**	➢ Emphasis ➢ Projection of context of interaction
	GESTURE	➢ Emphasis ➢ Processing aid		**GESTURE**	➢ Attention getting ➢ Processing aid
	HEAD			**HEAD**	➢ Processing aid
CPAM1	**WORDS**	➢ Anticipation and control of responses ➢ Attention getting ➢ Emphasis ➢ Evaluation	**CPAM2**	**WORDS**	➢ Rapport
	INTONATION	➢ Emphasis ➢ Projection of context of interaction		**INTONATION**	➢ Emphasis ➢ Projection of context of interaction
	GESTURE	➢ Anticipation and control of responses		**GESTURE**	➢ Anticipation and control of responses ➢ Processing aid
	HEAD	➢ Anticipation and control of responses		**HEAD**	➢ Anticipation and control of responses ➢ Processing aid

CPRE1	**WORDS**	➢ Anticipation and control of responses ➢ Emphasis ➢ Evaluation ➢ Rapport	**CPKE1**	**WORDS**	➢ Attention getting ➢ Emphasis ➢ Evaluation ➢ Rapport
	INTONATION	➢ Anticipation and control of responses ➢ Emphasis ➢ Processing aid ➢ Projection of context of interaction ➢ Rapport		**INTONATION**	➢ Attention getting ➢ Emphasis ➢ Processing aid
	GESTURE	➢ Processing aid		**GESTURE**	➢ Emphasis ➢ Processing aid ➢ Rapport
	HEAD			**HEAD**	➢ Anticipation and control of responses
CPPE1	**WORDS**	➢ Rapport ➢ Processing aid	**CPPE2**	**WORDS**	➢ Rapport ➢ Emphasis
	INTONATION	➢ Anticipation and control of responses ➢ Emphasis ➢ Processing aid		**INTONATION**	➢ Emphasis ➢ ➢ Projection of context of interaction
	GESTURE	➢ Emphasis ➢ Processing aid		**GESTURE**	➢ Processing aid
	HEAD			**HEAD**	➢ Processing aid ➢ Anticipation and control of responses ➢ Emphasis

The number of strategies that each mode realises provides an indication of how extensively each mode is being used. In most cases words is the mode that seems to fulfil the highest number of strategies, a trend that was also noticeable in the other two genres and which will be the object of discussion in Chapter 7.

The following bar chart (Figure 6.6) represents the modal realisations of the persuasive strategies in the conference presentations. Results are presented according to their frequency per minute (number of occurrences /duration of the CP subcorpus of rich points in minutes).

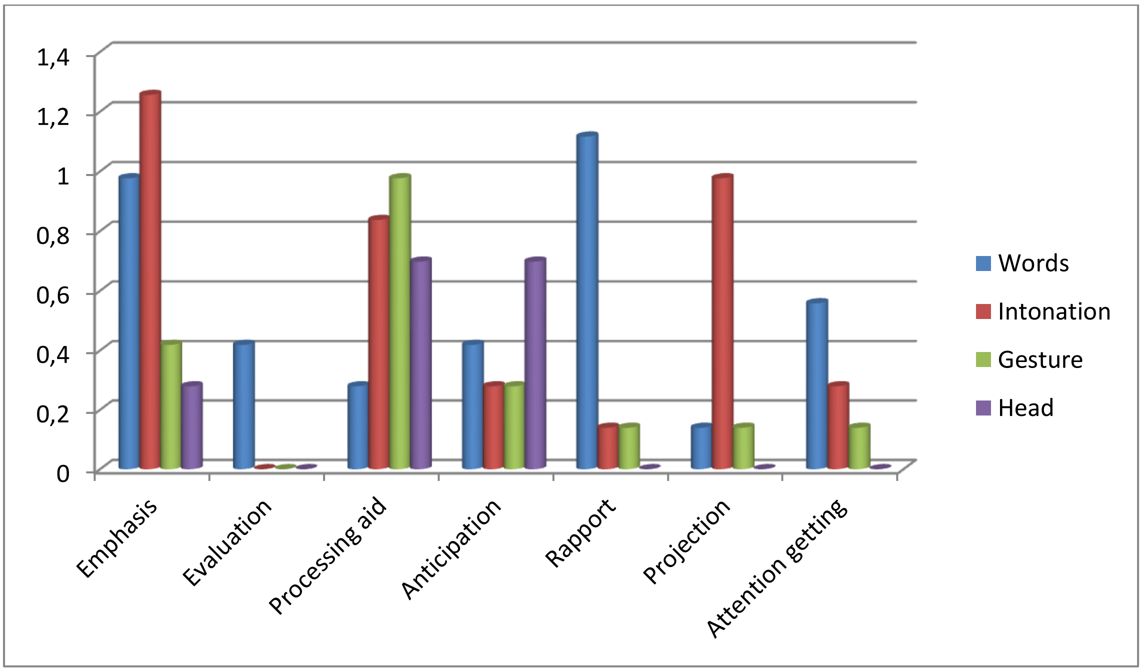


Figure 6. 6 Persuasive strategies and their modal realisations in conference presentations

Some interesting features revealed by these results are worth mentioning here and will be object of further discussion in Chapter 7:

i) Processing aids rarely take the form of words in conference presentations;

ii) Rapport is predominantly realised through words;

iii) Projection is mostly realised through intonation;

iv) Evaluation is the least used strategy and is only realised through words;

v) Intonation is mainly used for emphasis and projection, followed by processing aids;

vi) Ordering the strategies in terms of their frequency (and regardless of their modal realisation) we obtain:

1. Emphasis (21 instances/ 3 per minute)
2. Processing aid (20 instances/ 2,8 per minute)
3. Anticipation (12 instances/ 1,7 per minute)
4. Rapport (10 instances/ 1,4 per minute)
5. Projection (9 instances/ 1,2 per minute)
6. Attention getting (7 instances/ 1 per minute)

7. Evaluation (3 instances/ 0,4 per minute)

Concerning the type of gestures, the most frequently used gestures are deictic referential gestures that point at the speaker or at the audience and metaphoric referential gestures that clarify content. Figure 6.7 offers an example of each:

CPKE1

Deictic gesture pointing at the audience

it's LARGEly eh like THIS but for the psyCHIAtric community

The comparison between two events in "it's largely like this" is clarified by an OHS deictic gesture pointing at the audience.

CPTO1

Metaphoric referential gesture

so they've ALways come BACK to us of course

"Come back" is represented visually trough an inward hand gesture. This is a metaphoric, referential gesture in semantic and phonological synchrony with words that makes the message easier to process.

Figure 6. 7 Examples of gestures in CPs

6.2.2.2 The research dissemination talks

As the interviews with the speakers revealed, in the case of the research dissemination talks, the motivation of speakers is more related to personal satisfaction than to any practical aims such as networking or obtaining funding. One speaker also mentioned receiving feedback from non- experts as an added value. During the interviews, timing came up as a recurrent concern for speakers during preparation. They rehearsed to check and adjust the length of their presentations when necessary; however, in this case they received no formal training for the event. Interestingly, the only two speakers in the whole corpus who reported a certain level of awareness about gestures are DTLI and DTJO, and both reported having received feedback from colleagues and in the case of DTLI also from the organisation Toastmasters[9], devoted to public speaking.

Another recurrent concern in the interviews is the need to adapt to a layman audience making research accessible and showing what its relevance is. According to the speakers, dissemination of research to the general public is different from purely academic presentations because it is less focused on data and specific details, and more sharing and entertaining (DTLI, DTLA, DTJO).

One speaker also commented that in his presentation he was not “evaluating [his] own research directly, that would be too much, it wouldn’t work” (DTRA), but he agrees that he was trying to present it as interesting. The following quote can also serve to summarize most of the views on the genre expressed by the speakers, and suggests that this type of talks share features from both academic presentations and product pitches:

> I hate the way to do research that you are in a cellar, you work for four years and you don't communicate about it. I think it is really important as a researcher that you are able to communicate about it. For me the TEDX Gent was the ideal thing because you have to explain it in layman terms and you have to explain it in a way that’s also persuasive, that people believe there's something in there. It’s almost like a start-up pitch or something where you need to convince people about a certain cause and need to activate them almost to believe in you. Before my research I was active as a, in my own start-up, so I was intrigued by the whole pitch thing and for me TEDX Gent or public speaking about my research is just like this thing from start-ups” [INAUDIBLE] you need to pitch your research all the time and if you cannot do that, for me you are not a good researcher (DTJO)

Finally, the speakers in this event said they value a presentation that is:

i) Easy to follow (DTJO), for example:

a) Made accessible through metaphors;
b) Varied and accurate language (for non-native speakers);
c) A good wrap up for the audience to remember;

ii) Captivating and interesting (DTLI, DTJO), for example:

a) A good hook to capture attention;
b) Enthusiasm of the speaker that is transmitted to the audience;

[9] www.toastmasters.org (last accessed 3/2/2017)

c) Innovation, being surprising (e.g. an invented acronym, an unusual example);

d) Entertaining;

e) Relevant content;

Moving on to the multimodal analysis, the following table (Table 6.12) represents the strategies and modal realisations identified in each rich point of the research dissemination talks. As in the previous case, it provides an overview of the extent to which each mode is being used.

Table 6. 12 Rich points in research dissemination talks

DTLA1	**WORDS**	➢ Anticipation and control of responses ➢ Attention getting ➢ Rapport	**DTLA2**	**WORDS**	➢ Attention getting ➢ Rapport
	INTONATION	➢ Emphasis ➢ Projection of context of interaction		**INTONATION**	➢ Emphasis ➢ Projection of context of interaction
	GESTURE	➢ Anticipation and control of responses ➢ Emphasis ➢ Processing aid		**GESTURE**	➢ Anticipation and control of responses ➢ Emphasis ➢ Processing aid ➢ Rapport ➢ Evaluation
	HEAD	➢ Emphasis		**HEAD**	

DTLI1	**WORDS**	➢ Anticipation and control of responses ➢ Attention getting ➢ Processing aid ➢ Projection of context of interaction ➢ Rapport	**DTLI2**	**WORDS**	➢ Emphasis ➢ Evaluation ➢ Rapport
	INTONATION	➢ Anticipation and control of responses ➢ Processing aid ➢ Projection of context of interaction ➢ Rapport		**INTONATION**	➢ Anticipation and control of responses ➢ Emphasis ➢ Projection of context of interaction
	GESTURE	➢ Anticipation and control of responses ➢ Processing aid ➢ Rapport ➢ Attention getting		**GESTURE**	➢ Anticipation and control of responses ➢ Emphasis ➢ Evaluation
	HEAD	➢ Rapport ➢ Anticipation and control of responses		**HEAD**	
DTRA1	**WORDS**	➢ Emphasis ➢ Evaluation ➢ Processing aid ➢ Projection of context of interaction ➢ Rapport	**DTRA2**	**WORDS**	➢ Anticipation and control of responses ➢ Attention getting ➢ Emphasis ➢ Evaluation ➢ Rapport
	INTONATION	➢ Emphasis ➢ Processing aid ➢ Projection of context of interaction ➢ Rapport		**INTONATION**	➢ Anticipation and control of responses ➢ Emphasis ➢ Projection of context of interaction
	GESTURE	➢ Emphasis ➢ Processing aid ➢ Rapport		**GESTURE**	➢ Rapport
	HEAD			**HEAD**	➢ Anticipation and control of responses ➢ Emphasis

DTPI1	**WORDS**	➢ Evaluation ➢ Rapport ➢ Attention getting ➢ Emphasis	**DTPI2**	**WORDS**	➢ Anticipation and control of responses ➢ Evaluation ➢ Rapport ➢ Attention getting
	INTONATION	➢ Emphasis ➢ Processing aid ➢ Projection of context of interaction ➢ Rapport		**INTONATION**	➢ Anticipation and control of responses ➢ Emphasis ➢ Projection of context of interaction ➢ Rapport
	GESTURE	➢ Emphasis ➢ Processing aid ➢ Projection of context of interaction ➢ Rapport		**GESTURE**	➢ Anticipation and control of responses ➢ Emphasis ➢ Processing aid ➢ Projection of context of interaction ➢ Rapport
	HEAD			**HEAD**	
DTJO1	**WORDS**	➢ Anticipation and control of responses ➢ Attention getting ➢ Emphasis ➢ Rapport	**DTJO2**	**WORDS**	Attention getting Evaluation Processing aid
	INTONATION	➢ Emphasis		**INTONATION**	➢ Evaluation ➢ Projection of context of interaction ➢ Anticipation and control of responses
	GESTURE	➢ Anticipation and control of responses ➢ Attention getting ➢ Emphasis ➢ Processing aid ➢ Rapport		**GESTURE**	➢ Anticipation and control of responses ➢ Emphasis ➢ Processing aid ➢ Rapport
	HEAD			**HEAD**	

The following graph (figure 6.8) provides a visual representation of the modal realisations of the persuasive strategies in the subcorpus of research dissemination talks. As in the case of the conference presentations, the results are presented according to their occurrence per minute (number of occurrences /duration of the research dissemination talks subcorpus of rich points in minutes).

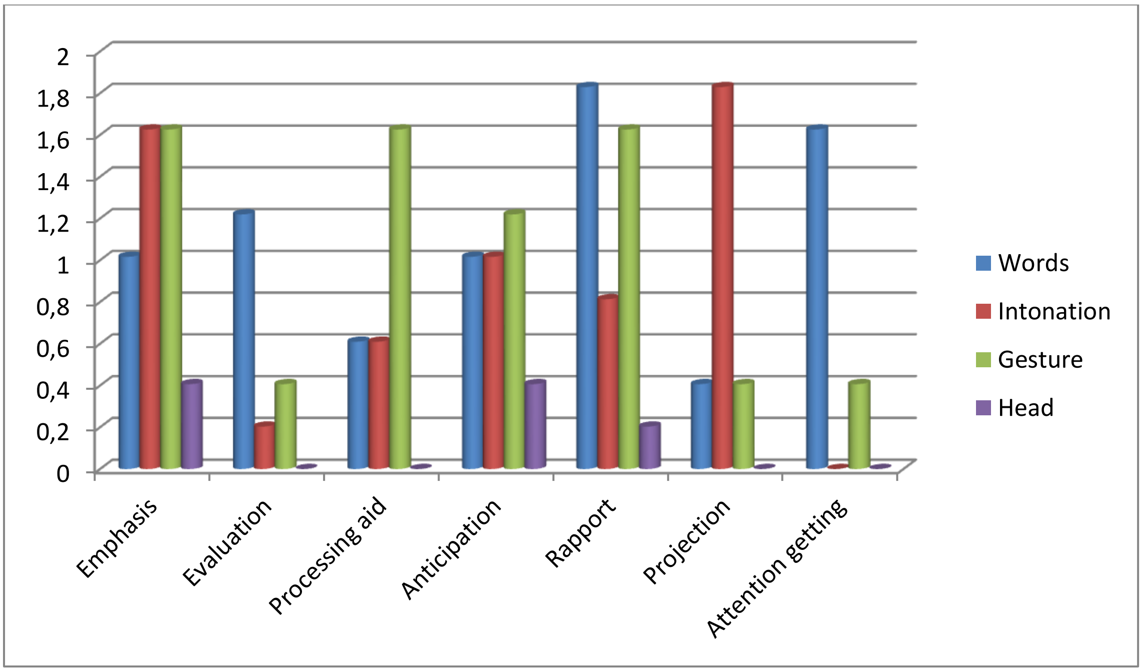


Figure 6. 8 Persuasive strategies and their modal realisations in dissemination talks

Some interesting features revealed by these results:

i) Head is the mode that is less frequently used;

ii) gestures are mostly used for emphasis, processing aids and rapport;

iii) intonation is used mainly for projection and emphasis;

iv) processing aids are mainly realised through gestures;

v) projection is mainly realised through intonation;

vi) ordering the strategies in terms of their frequency (regardless of the modal realisation) yields the following list:

1. Emphasis (23 instances/ 4,7 per minute)
2. Rapport (22 instances/ 4,5 per minute)
3. Anticipation and control of responses (19 instances/ 3,9 per minute)
4. Processing aid (14 instances/ 2,8 per minute)

5. Projection of context of interaction (13 instances/ 2,6 per minute)
6. Attention getting (10 instances/ 2 per minute)
7. Evaluation (9 instances/ 1,8 per minute)

In relation to gestures, the ones that occur more frequently are metaphoric referential gestures that visually clarify aspects of content and beats, that have two main functions: pragmatic and cohesive. It is also interesting to note that there are different types of gestures (metaphoric, deictic and beats) that are used with an interpersonal function. Figure 6.9 offers some examples of beats in research dissemination talks:

DTPI1

Finger ring beats: pragmatic and cohesive

i feel REALly REALly HUMble

The repetition in "really, really humble" is paralleled with 3 finger ring beats that mark the rhythm created by the prominence. The function is therefore partly cohesive (connecting) and partly pragmatic (emphasizing, with the effect of adding sincerity, as if saying "I really mean it").

DTLA1

Open arm beats: interpersonal

ARE YOU SMART

The beats with open arms (OHS) accompanying each prominent syllable in the question "Are you smart?" seem to invite the audience in. They have an interpersonal function and reinforce the conversational style.

Figure 6. 9 Examples of beats in research dissemination talks

6.2.2.3 The product pitches

According to the information that the speakers of product pitches provided in the interviews, they were offered training before the event. During this training, they received advice on how to design their slides and deliver their presentation and they had the chance to rehearse it in advance. Apparently, the training put a lot of emphasis on preparing a written script, and therefore the level of improvisation was not high: most speakers followed the script they had prepared in advance. In two cases, the speakers explained in the interviews that they were advised to memorize a text, which made them focus on remembering the exact words and not on the audience.

Regarding the motivation of speakers to participate in the event, in the interviews they mentioned the opportunity for entrepreneurial collaboration (PPPA), to get hands–on practice and learning by doing (PPPI), and the chance to disseminate a business idea and obtain funding for it (PPTO). I also asked them about what they considered key aspects in a pitch, and they mentioned that for a good pitch it is necessary to i) know your product and what you want to explain about it (PPPA, PPPI), ii) capture attention (PPKE, PPTO), iii) make it personally relevant for the audience (PPPA, PPTO) and iv) show enthusiasm and make it engaging (PPSE).

Concerning the expectations and views about the genre, one of the speakers reported that the main difference between pitching and advertising is the relationship established between the designer of the message and the content (i.e. the product). In pitches, you present your own product, which means it is also relevant to you. This also prompts a closer relationship with the audience than in advertising. Another speaker also explained that “pitching has nothing to do with storytelling" (PPPA), and that the focus has to be on *why* your product might be interesting for the audience. This ‘why’ is what needs to be communicated first, otherwise people will not pay attention to the ‘what’ or the ‘how’. In his own words “if the why doesn’t resonate with them then they won’t care about any other technical aspect” (PPPA). These ideas are supported and complemented by the views of the third speaker who expressed that his intention with his presentation was to “make them focus on a specific problem and the solution” (PPTO).

Turning now to the multimodal analysis, the following table (Table 6.13) represents the strategies and modal realisations identified in each rich point of the product pitches.

Table 6. 13 Rich points in product pitches

PPKE1	**WORDS**	➢ Anticipation and control of responses ➢ Emphasis ➢ Evaluation	**PPKE2**	**WORDS**	➢ Anticipation and control of responses
	INTONATION	➢ Projection of context of interaction		**INTONATION**	➢ Anticipation and control of responses
	GESTURE	➢ Processing aid		**GESTURE**	
	HEAD			**HEAD**	
PPPA1	**WORDS**	➢ Emphasis ➢ Evaluation ➢ Rapport	**PPPA2**	**WORDS**	➢ Emphasis ➢ Evaluation ➢ Processing aid ➢ Rapport
	INTONATION	➢ Emphasis ➢ Projection of context of interaction		**INTONATION**	➢ Emphasis ➢ Evaluation ➢ Processing aid
	GESTURE	➢ Emphasis ➢ Rapport		**GESTURE**	➢ Emphasis ➢ Evaluation ➢ Processing aid
	HEAD			**HEAD**	
PPPI1	**WORDS**	➢ Anticipation and control of responses ➢ Emphasis ➢ Evaluation ➢ Rapport	**PPPI2**	**WORDS**	Emphasis Evaluation
	INTONATION	➢ Anticipation and control of responses ➢ Emphasis ➢ Projection of context of interaction		**INTONATION**	➢ Emphasis ➢ Projection of context of interaction
	GESTURE	➢ Processing aid		**GESTURE**	➢ Emphasis ➢ Processing aid
	HEAD			**HEAD**	
PPSE1	**WORDS**	➢ Emphasis ➢ Evaluation	**PPSE2**	**WORDS**	➢ Attention getting ➢ Emphasis
	INTONATION	➢ Anticipation and control of responses ➢ Emphasis ➢ Projection of context of interaction		**INTONATION**	➢ Attention getting ➢ Emphasis
	GESTURE	➢ Anticipation and control of responses ➢ Evaluation ➢ Processing aid ➢ Rapport		**GESTURE**	➢ Attention getting ➢ Processing aid
	HEAD			**HEAD**	

PPTO1	**WORDS**	➢ Anticipation and control of responses ➢ Attention getting ➢ Emphasis ➢ Rapport	**PPTO2**	**WORDS**	Emphasis Evaluation
	INTONATION	➢ Anticipation and control of responses		**INTONATION**	Projection of context of interaction
	GESTURE	➢ Emphasis ➢ Processing aid		**GESTURE**	Evaluation
	HEAD			**HEAD**	

The following graph (figure 6.10) represents in a visual way the modal realisations of the persuasive strategies in the subcorpus of product pitches, according to their frequency per minute.

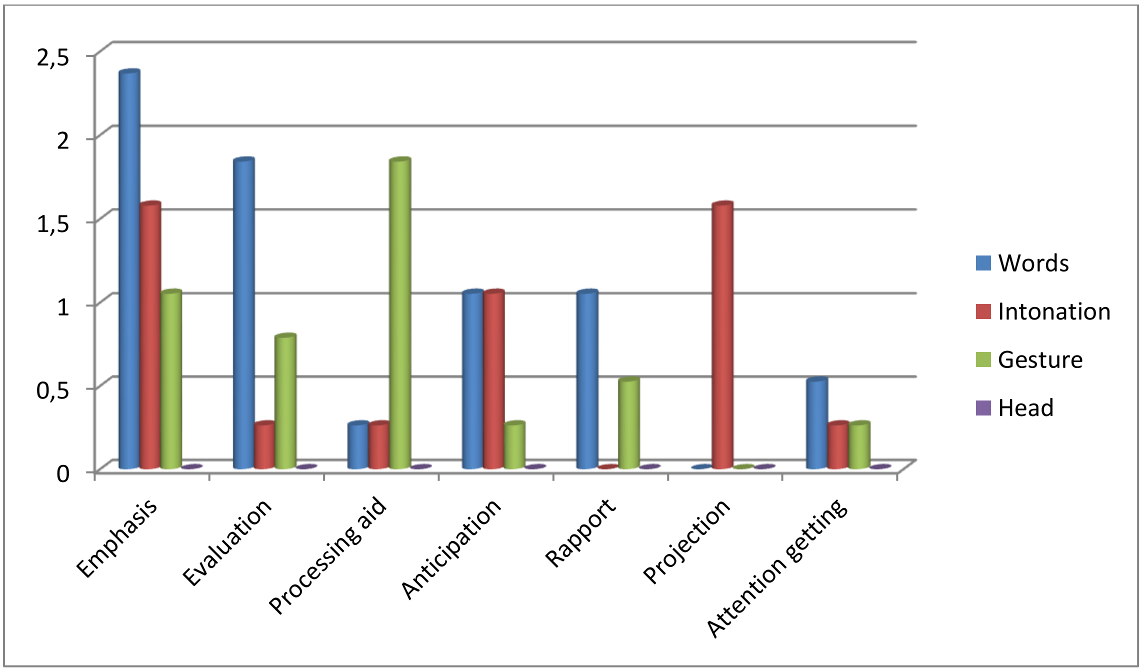


Figure 6. 10 Persuasive strategies and their modal realisations in product pitches

These results reveal some interesting features. Without going into deep interpretation, which will be the focus of chapter 7, it is noticeable that:

i) There is a remarkable lack of head movements used for persuasive purposes in the product pitch subcorpus, regardless of the strategy;

ii) projection of context of interaction seems to be totally dependent on intonation as its only modal realisation;

iii) processing aids frequently take the form of gestures;

iv) ordering the strategies in terms of their frequency (adding all the modal realisations) yields the following result:

1. Emphasis (19 instances / 5 per minute)
2. Evaluation (11 instances /2,9 per minute)
3. Processing aid /Anticipation and control of responses (9 instances / 2,4 per minute)
4. Rapport / Projection of context of interaction (6 instances/ 1,6 per minute)
5. Attention getting (4 instances/ 1 per minute)

It seems that speakers highlight certain parts of their messages and they do it multimodally through emphasis. They also evaluate their products positively and evaluate competitors negatively, and they resort to gestures and intonation –apart from words- to do this.

Rapport also plays a relatively important role, but when speakers use this strategy they do not add intonation or head movements to the multimodal ensemble and restrict their use of modes to words and gestures.

Regarding the type of gestures used, the most frequent gestures are metaphoric, referential gestures that visually clarify some part of the content of the message and cohesive beats that show connections between parts of the message. In both cases, they are processing aids that, as confirmed by speakers during the interviews, are intended to make the message easier to understand. Interpersonal gestures are only used twice, and the two pitches in which they appear (PA and SE) were the two favourites: one was the winner of the event, the other stood out as a favourite when speakers were asked which pitch from the ones they had seen as members of the audience they had liked the best.

I would like to conclude section 6.2 with a summary of my answer to RQ II:

What is the contribution of the modes studied (words, intonation, head movements and gestures) to the creation of persuasion in the genres included in the analysis?

The results of the analysis show that persuasion in the three subcorpora is indeed multimodal. They also indicate that there is a correlation between modal density (i.e. number of modes used) and modal coherence (i.e. using modes consistently) regardless of the genre.

The three subcorpora feature instances of the seven strategies presented in section 6.1, but to different extents and through different modal realisations, something that will be further discussed in section 6.3. The analysis has further revealed that the more frequent a strategy is, the more likely it is to be realised multimodally. Words as a semiotic mode seems to fulfil the greatest number of strategies in the three subcorpora, and emphasis seems to be the strategy that is used the most in the three cases. These similarities and differences across subcorpora are the focus of the next section.

6.3 Comparative analysis of the use of persuasive strategies in conference presentations, research dissemination talks and product pitches

In this third and last section of the chapter I compare the strategies and modes used across the three subcorpora. The aim is this section is not to offer conclusive results of a contrastive genre analysis, for several reasons. First, the corpus size does not allow for generalisations that can be applied to the whole genre. Second, I have only analysed selected excerpts and not entire instances of the genres. Third, my comparison focuses exclusively on the use of persuasive strategies and their modal realisations, which are only two of the many aspects that can define genre. Instead of that, my objective with this section is to test my suggested methodology to see how it can contribute to a contrastive analysis. The results presented are more valuable for their potential as illustrations of the methodology at work than for their representativeness. I believe that gaining more insight into potential differences and similarities across these genres can pave the way to provide better and more accurate definitions in the field of genre studies. Likewise, these differences and similarities may be worth taking into account for the design of LSP materials that make use of these genres as a didactic tool. These are points in the agenda for future genre analysis, and it is my intention to show how the methodology that I propose can contribute to achieve them. Therefore, I will now focus my attention on RQIII:

RQ III.	**To what extent is the use of persuasive strategies genre-specific and to what extent is there overlap between the three genres?**
RQ III. a)	Are persuasive strategies enacted through the same modes in the three genres?
RQ III. b)	Are there strategies which can be considered genre-specific?
RQ III. c)	How do the three genres compare in terms of their use of persuasive strategies?

The following three subsections will provide a specific answer to questions a), b) and c) respectively.

6.3.1 Use of modes in the three genres

RQ III. a)	Are persuasive strategies enacted through the same modes in the three genres?

The results of the analysis reveal differences in the use of the four modes analysed (words, intonation, gestures and head movements) in conference presentations, research dissemination talks and product pitches. Figure 6.11 shows a comparative view of the use of modes in the three subcorpora. In this graph all strategies that are realised by the same mode (i.e. words, intonation gestures or head movements) are added up. Therefore, the graph shows the use of each of the four modes analysed in the study regardless of the strategy which they are realising. Each mode is counted as one instance whenever it realises a strategy in a rich point, and the same mode can enact different strategies within the same rich point, which would account for two or more different instances. To facilitate comparison, the results are again presented according to occurrence per minute (occurrences of the mode/duration of each subcorpus of rich points in minutes), yielding the following frequencies:

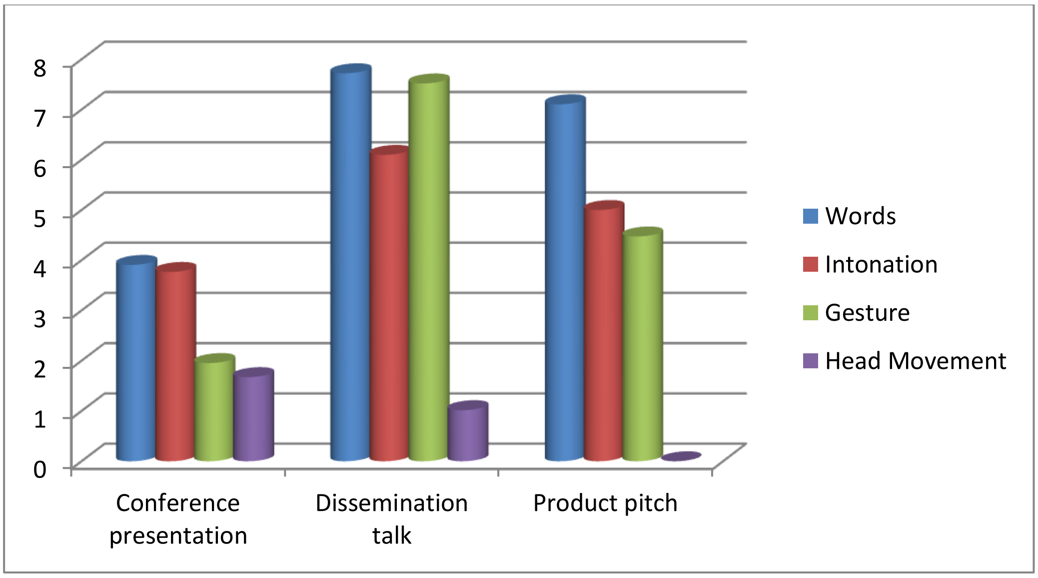


Figure 6. 11 Use of modes in the three subcorpora

A striking result is that, overall, conference presentations show fewer instances of the modes, so they could be described as less modally dense. It is also interesting that, in the three subcorpora, words are used more frequently for persuasion than the other three modes. In conference presentations intonation and words are similar in frequency (intonation being used slightly less), and head movements are used only slightly less frequently than gestures. In research dissemination talks gestures are exploited more than intonation, and the mode that is used significantly less is head movements. In product pitches gestures closely follow intonation and finally head movements are not used at all.

Concerning the type of gestures, in all presentations there is a predominance of metaphoric referential gestures that visually clarify aspects of the content. In dissemination talks and pitches there is also extensive use of beats that acquire a pragmatic or cohesive function. Dissemination talks display the highest use of interpersonal gestures, while conference presentations feature a more extensive use of deictic gestures that point either at the audience or at the speaker.

I will turn back to these results in Chapter 7, where I will discuss them and relate them to the insights offered by previous literature and the ethnographic interviews.

6.3.2 Use of strategies in the three genres

RQ III. b)	Are there strategies which can be considered genre-specific?

According to the results, no strategy is clearly genre-specific, but each genre uses the seven strategies to a different extent. I have also identified two uses of two of the strategies (i.e. attention getting and evaluation) that can be considered genre-specific. In other words, these strategies are used in the three genres, therefore they cannot be said to be genre-specific, but there are remarkable peculiarities in the way they are used in conference presentations and product pitches respectively. Both findings (there are no genre-specific strategies, but there are genre-specific uses of them) are explained in the following paragraphs.

Figure 6.12 shows a comparative view of the use of strategies in the three subcorpora. As in the previous case this figure presents all strategies added up regardless of the mode in which they are realised. Each strategy is counted once per modal realisation, which means that the same strategy with different modal realisations accounts for two or more instances. The results are presented according to occurrence per minute (occurrences of the strategy/duration of each subcorpus of rich points in minutes).

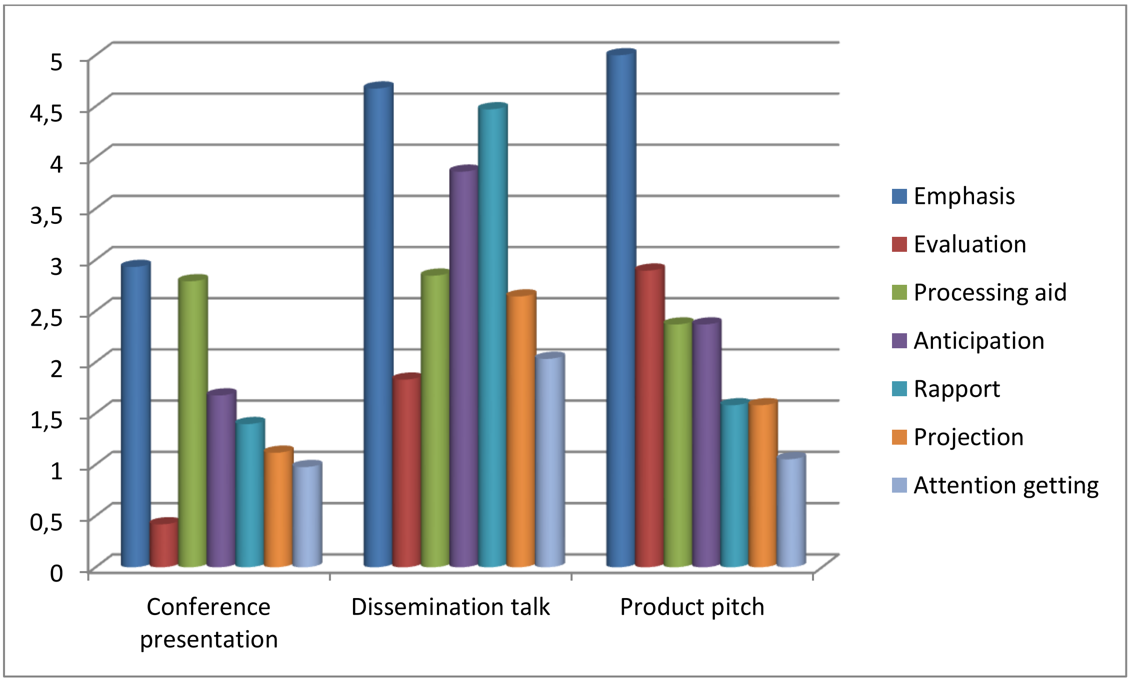


Figure 6. 12 Use of strategies in the three subcorpora

As the graph shows, the seven strategies are found in the three subcorpora in varying degrees. Emphasis is the strategy that is used the most in the three cases, as mentioned earlier in this chapter. Processing aids are used very evenly in the three subcorpora. With the exception of emphasis and evaluation, research dissemination talks feature the highest number of occurrences of the other five strategies. I will come back to this aspect in section 6.3.3.

Table 6.14 shows the persuasive strategies used in each subcorpora ordered according to their frequency:

Table 6. 14 Strategies ordered by frequency

CONFERENCE PRESENTATIONS	**DISSEMINATION TALKS**	**PRODUCT PITCHES**	
Emphasis	Emphasis	Emphasis	**More frequent**
Processing aid	Rapport	Evaluation	
Anticipation	Anticipation	Processing aid	
Rapport	Processing aid	Anticipation	
Projection	Projection	Rapport	
Attention getting	Attention getting	Projection	
Evaluation	Evaluation	Attention getting	**Less frequent**

The results presented in Table 6.14 above show that in conference presentations evaluation is minimal. Speakers use more processing aids and less rapport than in research dissemination talks. Attention getting plays a slightly more important role in research dissemination talks and conference presentations than in product pitches.

In research dissemination talks evaluation occurs significantly less frequently than in product pitches, but there is a higher frequency of rapport and anticipation. Immediately following are processing aids and projection of context of interaction.

Finally, in product pitches evaluating the product in positive terms becomes an important concern, followed by using processing aids and anticipating and controlling responses (both equally frequent), and then by rapport and projection (again equally frequent). Attention getting plays a minimal role according to results. The significance of these results will be further discussed in Chapter 7.

In the remaining part of this section, I would like to focus on two particular uses of strategies in two of the presentations: attention getting in conference presentations and deictic evaluation in product pitches.

I would like to turn my attention first to a particular type of attention getting that I have identified in conference presentations: the enactment of characters. In these cases the speakers act out assuming the roles of people they are referring to in their presentations, such as participants in their research or other researchers. I consider it a way of attention getting, since it turns their account of their research experience into a vivid narrative.

One example of enactment of characters is CPDO2, mentioned earlier in this chapter. As explained in section 6.1.2, the speaker in this rich point does not report what happened, but instead she uses direct speech to bring the experience closer to the audience. Using intonation she indicates who is speaking. At the same time, her head shakes represent visually the misunderstanding that she is relating. Lateral movements of the head also represent how she reacted in the situation she is describing, rather than the way she feels at the moment of speaking. In this way she becomes another character in her narrative.

The next example I would like to discuss is CPAS1. In this excerpt the speaker is discussing how during an internship program where she was involved as a teacher, the students and the company did not manage to build the relationship of mutuality and trust that was required for the project, and this was evident in some communication gaps. Table 6.15 shows the use of words, intonation and head gestures in this rich point.

Table 6. 15 Enacting characters in CPAS01

Orthographic transcription
Also coming back to that concept of freedom, one of their core principles, the students, to them, were very free. When, remember when the students said we weren’t allowed to interview employees, well, in fact it wasn’t because they weren’t allowed, it just did not seem relevant to the company. But in fact, all the resources were there, just if they wanted them, it was there.
MICASE transcription
eh, also, coming back to that concept of, freedom, their _one of their core principles, the students, to them, were very free. eh, when_remember when the students said we weren’t allowed to interview eh employees well, in fact, it wasn’t because they weren’t allowed it just did not seem relevant to the company, but in fact all the resources were there, eh just, eh if they wanted them, it was_it was there.
DI transcription
1. eh 2. also 3. coming BACK to that concept of FREedom 4. their 5. one of their CORE principles 6. the $^{\text{STU}}$dents 7. to $^{\text{THEM}}$ 8. were VEry FREE 9. eh 10. when 11. reMEMber when the STUdents said 12. we weren’t aLLOWED to interview eh emploYEES 13. well in FACT 14. it WASN’T because they weren’t alLOWED 15. it just did not seem $^{\text{RE}}$levant to the COMpany 16. but in FACT all the resources were THERE 17. eh 18. just 19. eh 20. if they WANted them 21. it was it was THERE

Gestures and head movements
 it just did not seem ^{RE}levant to the COMpany

The use of the pronoun “we” to refer to students turns this passage into a direct quote (it is the students talking) and provides a narrative style. The repetition of the adversatives “in fact” and “but” emphasizes the contrast between what students perceived and what the company perceived, and highlights the misunderstanding this passage refers to. High key in “students” (unit 6) and “to them” (unit 7), and later in “relevant” (unit 15) conveys surprise and again highlights the contrast between what the company believes and what the students believed. Repeated head shakes simultaneous with “did not seem relevant to the company” contradict visually the students' wrong interpretation. The speaker is assuming the role of the company at this moment and representing their reaction, paraphrasable as “it’s not that we don’t allow you, we thought it wasn’t useful for you”. Note how the facial expression (although out of the scope of my analysis) also supports head movements in conveying lack of understanding. The speaker agreed with my interpretation and added that this enactment of characters is actually very coherent with the main topic of her presentation: different roles in the interim program. In addition, she also explained that “when I see they [the audience] are engaged it makes me feel more comfortable", which accounts for her efforts at getting the attention of the audience.

In short, in these examples both speakers retain the attention of the audience using intonation, head shakes and direct quotes to build a vivid narrative of what happened, almost providing a re-enactment of the episode. It is remarkable that in both cases the anecdote reported is related to a misunderstanding.

To finish this account of the use of strategies in the three subcorpora I would like to focus on a specific type of evaluation found in product pitches: deictic negative evaluation. Speakers in product pitches have to differentiate their products from competitors, and show their advantage over these competitors. I have found two examples in which speakers do this through a particular way of pointing at an image of a competitor product projected on the screen.

The first example is PPSE1. In this rich point the speaker presents a waxing device for ski equipment, and compares it to competitors to show the advantage it offers. Figure 6.13 shows the use of words, intonation and gestures in this particular moment of the presentation.

and EAsier to use than wAxing kits ↗

Figure 6. 13 Deictic evaluation in PPSE1

The speaker points backwards to the screen as he pronounces “waxing kits” (i.e. competitor products), but keeping his body orientation towards the audience. This is important because by doing this the speaker is making sure the audience does not focus on the image projected but on himself, consequently it seems to suggest that what is being projected does not deserve attention. In addition, the pointing with a thumb suggests that the exact location is not important (Kendon 2004). Therefore, the gesture can be interpreted as disregarding and negatively evaluating what is being pointed at. The effect is supported by the comparative adjective used to describe his product immediately preceding the gesture (“easier to use”), and the referring tone that presents the information as shared with the audience and agreed upon, as if adding "as you already know". During the interview the speaker explained that the gesture is partly a result of his training, because he was taught never to turn his back to the audience, but agrees that the way of pointing is disregarding and entails a negative evaluation that he hopes the audience will share.

The second example is PPTO2. In this pitch the speaker is presenting a tablet with pictures of the body that doctors can use to communicate more easily with their patients. The market idea is to use this product as merchandising that pharmaceutical companies can give to doctors as presents and in this way advertise their products. In this rich point the speaker is explaining that this is already common practice among pharmaceuticals, and is trying to convince the audience that his product can be a much better gift for doctors compared to the merchandising that they usually receive now. As he makes this comparison he uses explicit negative evaluation of his competitors ("a useless gadget"). Simultaneously, he moves his left arm towards the screen behind him where a picture of these competitor products is shown, again without shifting the focus from the audience (figure 6.14). In addition, the gesture is very unspecific (note how the hand stays open and there is no extended finger), so much so that it is not clear what he is pointing at exactly. This conveys lack of interest and suggests to the audience that the image projected is not important and not worth paying attention to. According to the speaker, this gesture was a way of "wiping off" competitors.

a USEless GADget

Figure 6. 14 Deictic evaluation in PPTO2

The two examples of product pitches discussed above show how negative evaluation of competitors is an important element in persuasive product pitches. Furthermore, they prove how this evaluation can be realised multimodally for a more powerful effect.

6.3.3 Multimodal persuasive strategies in the three genres

RQ III. c)	How do the three genres compare in terms of their use of persuasive strategies?

As already introduced in section 6.2, the multimodal analysis of persuasive activity on the part of the speakers reveals that the three genres use the seven persuasive strategies identified, but there are differences regarding the extent to which each strategy is used and the extent to which each mode is used. Likewise, the analysis has also revealed some genre-specific uses of multimodal persuasive strategies, such as specific attention getting in conference presentations and specific evaluation in product pitches. In this section I want to focus on two significant findings of the comparative analysis: 1) differences in modal and persuasive density (level of synchronicity or simultaneity of both modes and strategies), and 2) differences in the speaker-audience relationship.

Regarding the first finding, comparing the frequency with which modes are used in each subcorpus can provide an estimation of their modal density. Results show that, with the exception of head movements (which are more frequent in conference presentations), dissemination talks feature the highest frequency in the use of modes (see figure 6.11), and therefore can be considered more dense on average. Likewise, if we compare the frequency of strategies in the three subcorpora (Figure 6.12), dissemination talks again seem to feature a higher frequency of strategies with the exception of emphasis and evaluation, which are more frequent in product pitches. Therefore, the rich points in the dissemination talks in my corpus can be considered to feature higher *persuasive density*, a concept which I coin in parallelism with modal density to mean “number of persuasive strategies used". Conversely, conference presentations score the lowest in terms of modal density and persuasive density. I will come back to this result and its potential relevance in Chapter 7.

Regarding the second finding (i.e. differences in the speaker-audience relationship), results show that speakers use the seven strategies in varying frequencies to target at a different relationship with the audience depending on the genre: in conference presentations the focus seems to be on fitting in the event and the scientific community, in research dissemination talks, speakers try to show how their research is relevant, while in product pitches the main aim is to present a product as a solution to a problem that the audience needs to solve. This result is coherent with the insights from the interviews presented earlier in section 6.2 in this chapter, and can be clearly illustrated comparing three rich points. The first example I want to discuss is a product pitch (PPA1). In this pitch the speaker is presenting his product, which is an application to assess apps, and he does it identifying a need that this product can cover. Table 6.17 shows the transcript, gestures and intonation used in this excerpt.

Table 6. 16 Getting interpersonal in PPPA1

Orthographic transcription		
You have an app. That app is important to your business, important to your business success. Because of that, you care about how your users experience that app, and you care about how your users feel about that app. But sometimes, you deliver a bad experience. Sometimes, you deliver a real bad experience.		
MICASE transcription		
you have an app. that app is important to your business, important to your business success. because of that, you care about how your users experience that app, and you care about how, your users feel about that app. but sometimes, you deliver a bad experience. sometimes, you deliver a real, bad experience		
DI transcription		
1.	you HAVE an APP	↘
2.	that APP is imPORtant to your business	↘
3.	imPORtant to your business sucCESS	↘
4.	beCAUSE of THAT	↗
5.	you CARE about how your users exPErience that app	↗
6.	and you care about HOW your users FEEL about that app	↘
7.	but SOMEtimes you deliver a BAD experience	↘
8.	SOMEtimes you deliver a REAL bad experience	↘

Gestures and head movements


It is noticeable that the speaker addresses the audience using the personal pronouns "you" and "your" throughout. His consistent use of repetition is also quite remarkable, for example in nouns and verbs that carry key concepts in his message, such as "important", "business", "care", "users" and "app". These repetitions provide the text with a sequence of symmetrical syntactic structures that contribute to memorability. Regarding intonation, he uses fall tones in most of the units, treating the information in this extract as mainly new. This use of intonation draws attention towards the problem or market gap. On the other hand, it also contrasts with the way he is presenting a potentially familiar situation to the audience through his words. Concerning gestures, it is noticeable how beats with an extended finger pointing to the audience get them involved and convey the idea that they are directly affected by the problem presented (your app may be delivering a bad experience for customers). The speaker confirmed during the interview that he was using a potentially familiar situation to create a sense of "shared conspiracy". He also agrees that part of the intention is raising awareness about this potential jeopardy. In fact he described his gestures as "scolding" and agrees with the idea that they are meant to prompt people to consider a potential danger or problem and react to it. He also explained that he decided to adopt this approach of presenting a potential danger, even if it can be considered too negative, because he wanted to differentiate himself from the majority of marketing messages that focus on boosting positive qualities. This is in line with claims in previous literature that persuasion is more effective when it is unexpected (O'Keefe 2002, Perloff 2003). This rich point also provides an example of multimodally realised emphasis, as discussed in section 6.2.2: the intensifier adverb “real” is prominent and is synchronous with an extended finger beat.

The next example is a dissemination talk (DTJO1) in which the speaker also opens his presentation directly addressing the audience. Table 6.17 summarizes the use of words, gestures and intonation in this excerpt.

Table 6. 17 Getting interpersonal in DTJO1

Orthographic transcription
Hi. I’m a chemist, and I really like kissing. I’d like to kiss you, I’d love to kiss you, I’m gonna kiss all of you. Because I really want to keep it simple, but not stupid. Tonight I’m going to show you how your own experience, with a very simple technique, can revolutionise the chemical industry, and finally, your life.
MICASE transcription
hi. i’m a chemist, and i really like kissing. i’d like to kiss you, i’d love to kiss you, i’m gonna kiss all of you because, i really want to, keep it simple, but not stupid. tonight, i’m going to show you how your own experience with a very simple technique, can revolutionise the chemical industry, and finally, your life
DI transcription
1. HI 2. I’M a CHEmist 3. and i REally like KIssing 4. i’d LIKE to KISS you 5. i’d LOVE to KISS you 6. i’m GOnna kiss ALL of you 7. beCAUSE 8. i REAlly want to 9. KEEP it SIMple 10. but NOT stupid 11. toNIGHT 12. i’m going to SHOW you how your OWN experience 13. with a VEry simple technique 14. can revoLUtionise the CHEmical industry 15. and FInally YOUR life

Gestures and head movements
 i’d LIKE to KISS you i’d LOVE to KISS you i’m GOnna kiss ALL of you  i’m going to SHOW you how your OWN experience and FInally YOUR life.

As in the first example of the product pitch, the speaker consistently uses the personal pronouns “you” and “your” to make the audience feel involved. He also selects words that highlight the relevance of his research for the members of the audience, like "your own experience" and uses intense language such as “revolutionise (…) your life". It is also interesting how he gets the attention of the audience using lexical creativity: he plays with the meaning of the noun “kiss” and verb “kissing” to introduce the acronym KISS (keep it simple, but not stupid). In addition, the speaker uses different deictic gestures to point to the audience. First he points with an extended finger, selecting specific members of the audience. As the speaker explained during the interview, his intention was to create “a very personal connection” with the audience. Then he points to the audience in a more generic way with open hand gestures that he considered would not be aggressive and would inspire confidence. Regarding intonation, it is interesting that “own" in "your own experience" and “your” in “your life” are also prominent, contributing to the effect of the synchronous gestures.

The third example is a conference presentation (CPRE1). In the opening of the presentation the speaker strives to prove her satisfaction to be part of the event. The use of words, gestures and intonation is represented in Table 6.18:

Table 6. 18 Getting interpersonal in CPRE1

Orthographic transcription
We’re very excited to be here in this conference because the theme, “Ins and Outs of Professional Discourse,” is very close to our hearts as we’ve been working for the last two years on Kelsey’s very important undergraduate research project. And to give you a little information about undergraduate research in the United States, it’s a movement of engaged learning where undergraduates like Kelsey develop their own research questions, collect their own data and present their own findings under the mentorship of the faculty members. So they are doing the same work that a Master’s or a PhD level student would do with the same kind of mentorship, presenting at conferences, publishing their work, so it’s a very exciting way to work with undergraduate students and it’s growing in popularity in America. Kelsey’s project is really interesting and we’re happy to present it today, in that it deals with both the ins and outs and a different way of thinking about how we enter communities and a broader sense of community that we can address.
MICASE transcription
we’re very excited to be here in this conference because the theme ins and outs of professional discourse is very close to our hearts, as we’ve been working for the last two years on kelsey’s, eh very important undergraduate research project. and to give you a little information about undergraduate research in the united states it’s a_it’s a movement of engaged learning where undergraduates like kelsey develop their own research questions collect their own data and present their own findings, under the mentorship of the faculty members so they are doing the same work that a master’s or a phd level, student would do with the same kind of mentorship presenting at conferences publishing their work, so it’s a very exciting way to work with undergraduate students. and it’s growing in popularity in america kelsey’s project is really interesting and we’re happy to present it today in that it deals with both the ins and outs and a different way of thinking about how we enter communities, and a broader sense of community that we can address.

DI transcription	
1.	we’re very exCIted to be HERE in this conference
2.	because the THEME
3.	INS and OUTS of professional DIScourse
4.	is very CLOSE to our HEARTS
5.	as we’ve been WORking for the last two years on KELsey’s
6.	eh
7.	VEry important
8.	underGRAduate research PROject
9.	and to give you a little information aBOUT underGRAduate research in the united states
10.	it’s a
11.	it’s a MOVEment of engaged LEARning
12.	where underGRAduates LIKE kelsey
13.	deVElop their own reSEARCH questions
14.	colLECT their own DAta
15.	and preSENT their own FINdings
16.	under the MENtorship of the FAculty members
17.	so they are DOING the SAME work
18.	that a MASter’s or a PHD level student
19.	WOULD do
20.	with the SAME kind of MENtorship
21.	preSENting at CONferences
22.	PUblishing their WORK
23.	so it’s a very exCIting way to WORK with undergraduate students
24.	and it’s GROwing in popuLARity in america
25.	$^{\text{KEL}}$sey’s project is really INteresting
26.	and we’re HAPpy to present it toDAY
27.	in that it deals with both the INS AND outs
28.	and a DIFFerent way of thinking aBOUT
29.	how we ENter com$^{\text{MU}}$nities
30.	and a BROADer sense of community that we can adDRESS

Gestures and head movements

in that it deals with both the INS AND outs

The speaker uses intense language that focuses on emotions (e.g. “excited to be here”, “close to our hearts”, “happy to present it today”). This emotionally intense language contributes to rapport. She also evaluates their (her and her co-presenter’s) research in positive terms using words that again appeal to emotions rather than reasoning (“very important research”, “exciting way to work”, “growing in popularity”, “really interesting”). According to the speaker both evaluative and emotional language are connected to a need to make sure that the topic is presented as “valuable, valued, and authentic”. As she explained, she wanted to convey “this is something we care about, so we hope you care about it too”. The speaker also uses two references to the title of the conference both at the beginning and the end of the excerpt (“the ins and outs”). With this she shows how her research fits in the conference. According to the speaker the references to the event are intentional because she was not sure whether the European audience would see the connection of her presentation with the topic of the conference. Concerning gestures and intonation, as she pronounces “ins and outs” both hands move in OHS beats first to her right and then to her left, visually signalling the contrast established. The beats are synchronous with the prominences: “INS", "AND". This use of prominence is interesting, because it emphasizes the coordinating conjunction "and", providing a sense of duality (paraphrasable by “not only ins, but also outs”). The speaker explained in the interview that this is a consequence of the way she interpreted the call for papers, which seemed to indicate expectations of papers focusing on either one aspect or the other, while theirs focused on both and she felt the need to justify or clarify this. A recurring idea that kept coming back in the interview with this speaker is an awareness of *belonging to a different community* where research is done differently. As she puts it, “we come from different academic cultures” and have different “research practices”. This translates into noticeable efforts to fit in, which become evident in the explicit references to the event. The rapport-building also has to do with “knowing I was an outsider in that community, so, I’m just making sure that, it’s building the rapport to show that I was happy to be there even if our work is a little bit different, maybe.” In short, the speaker thinks her way of presenting reveals that “there was a lot about making sure people understood why we were there, that we fit, and that the work we do was valid and rigorous (…)”.

As the three examples shown above have illustrated, in each type of presentation a different relationship is established with the audience. The speaker in the product pitches uses the strategies ‘rapport’ through words and gestures, ‘projection’ through intonation, and ‘evaluation’ through words to present himself and his product as a solution to a problem that the audience needs to solve. The speaker in the research dissemination talk, on the other hand, tries to show how his research is relevant, and can even become life-changing for the audience, and in order to do this he creates rapport through gestures and words and emphasizes parts of his message using gestures, words and intonation. In the conference presentation, finally, the focus seems to be on fitting in the event and the community as legitimate and competent members. To this aim the speaker uses ‘evaluation’ and ‘rapport’ through words, gestures that serve as processing aids, and intonation to emphasize parts of her message.

Summarizing my answer to RQ III: **To what extent is the use of persuasive strategies genre-specific and to what extent is there overlap between the three genres?**

The analysis reveals that the three genres use the seven persuasive strategies identified, but there are differences regarding how much each strategy is used and how much each mode is used. The genre that stands out as more modally dense and more persuasively dense is research dissemination talks.

Likewise, the analysis has also revealed some genre-specific uses of multimodal persuasive strategies (i.e. specific attention getting in conference presentations and specific evaluation in product pitches) and a different speaker-audience relationship being targeted in each genre.

In this chapter I have illustrated how the methodological framework presented in chapter 4 can be applied to the corpus of oral presentations described in chapter 5, yielding a taxonomy of seven persuasive strategies enacted multimodally. I have also analysed the way speakers in each subcorpus orchestrate multimodal ensembles to apply these strategies and convey a persuasive message. Finally, I have compared the use of modes and strategies in each subcorpus, highlighting some similarities and also interesting differences between them. Although these results cannot be generalised to the whole genres under study, they do illustrate the potential of an MDA and ethnographic approach for genre analysis. What these similarities and differences can reveal about each communicative situation will be object of discussion in the next chapter.

CHAPTER 7. DISCUSSION

This chapter discusses the results presented in Chapter 6. The first section discusses the methodological approach used in this study and answers **RQ1: What type of persuasive strategies can be identified using an MDA approach?** Section 2 looks at the contribution of multimodal ensembles to persuasion and deals with **RQ2: What is the contribution of the multimodal ensemble orchestrated through words, intonation, head movements and gestures to the creation of persuasion in the genres included in the analysis?** Finally, section 3 turns to the comparative analysis and focuses on my last research question, **RQ3: To what extent is the use of persuasive strategies genre-specific and to what extent is there overlap between the three genres?**

7.1 A systematic framework to approach persuasive strategies: The use of combined methodologies

This section offers a discussion of the answer that I have provided to RQ1, according to the results presented in section 6.1 in Chapter 6, mainly:

What type of persuasive strategies can be identified using an MDA approach?
I argue in favour of a video-based, corpus-driven type of study that combines input from previous literature with ethnography and MDA. The results of such a study in my corpus of presentations are the seven strategies presented in Chapter 6.

The added value of the methodological framework I propose in this thesis reveals itself in three main aspects: 1) focus on the multimodal ensemble, 2) a potentially more comprehensive view of the use of persuasive strategies and 3) the possibility of triangulation of results.

Regarding the first aspect, I believe that one of the main advantages of the methodology I use for this study is that it places all modes at the same level without prioritising, focusing on the persuasive effect intended by the speaker when designing a multimodal ensemble. Results show that modes contribute to persuasion to a different extent in each rich point and frequently one mode is found to be used more prominently than others. This relationship across modes, or modal configuration, is considered to fluctuate as the communicative situation unfolds, and no mode in particular has been assumed to be playing the leading role by default. This is in line with previous work on multimodal interaction analysis (Norris 2004; Pirini 2015).

The use of rich points as a unit of analysis has greatly contributed to keep the focus on the multimodal ensemble. If the analysis had started from one mode, this selection would imply considering this mode as the leading one, and the rest of the modes would only be probed for instances of co-expression. This is also a perfectly valid approach which can yield extremely interesting results (see for example Querol-Julián 2011 on discussion sessions in CPs, Ruiz-Garrido 2015 on the use of multimodal intensification of adverbials in CPs, or Crawford-Camiciottoli 2015 on OpenCourse Ware humanities lectures), but I believe it destabilizes the balance in terms of attention devoted to different modes.

On the other hand, a focus on the multimodal ensemble can make the analysis more time-consuming and resource-demanding, which in turn can lead to research decisions that artificially limit the scope of analysis for the sake of making it feasible. In the case of the current analysis, only four modes were included, which is by no means the whole multimodal ensemble. In addition, although the methodology promotes a focus on this multimodal ensemble, it is also necessary to acknowledge that both the researcher and the participants are more literate in the use of words as a mode, and more aware of it, which can introduce a bias towards this mode, even if this is unconscious. We are generally more capable of critically reflecting upon the way we use words. This is a tendency that became evident in the follow-up interviews: the speakers were able and willing to provide their own interpretations on their use of words and they provided detailed explanations to account for their choices, but the same process was more difficult when it came to discussing intonation, gestures and head movements, because speakers were less conscious and also lacked the metalanguage (i.e. the speakers knew what an adverb is, but not what an open hand prone OHP gesture means).

Concerning the second aspect, I consider that an MDA approach to persuasive strategies enables researchers to shed light on important persuasive activity that might otherwise go unnoticed if only one mode was studied. The results of my analysis reveal how aspects that might not be considered persuasive in isolation do contribute to persuasion when they are seen as integral parts of a whole. For example a head shake or a beat gesture may not straightforwardly be associated to persuasion if taken in isolation, but a multimodal analysis can reveal these are parts of a multimodal ensemble that do become persuasive, as the examples presented in Chapter 6 amply illustrate.

As for the third aspect, an approach that combines different methodologies can help overcome the limitations inherent in each methodology when they are used independently. As mentioned in Chapter 4, video-based multimodal research has some drawbacks, such as certain subjectivity inherent in the data collection process (e.g. in deciding what and how to record and analyse) and a tendency to microanalysis which leads to a potential risk of overestimating the relevance of specific detail in the global picture. However, the use of ethnographic methods can greatly help overcome these challenges. On the one hand, it informs the data screening process for subsequent analysis. In my case, the observation sheets and the first interviews provided me with data about the events and the speakers' views about the genre that helped me screen my data for moments of rich persuasive activity. On the other hand it facilitates cross-checking of interpretations with speakers concerning their communicative intentions. This playback methodology has previously been used in multimodal studies (Norris 2004; Querol-Julián and Fortanet-Gómez 2012; Morell 2014). In my case it gave the opportunity to my participants to offer their own interpretations before I showed them the results of my analysis. In some cases their opinions supported my interpretations, in others they offered an alternative explanation that frequently helped assign the correct level of importance to each detail. For example in some cases speakers believed I was "reading too much into a head nod" (CPAM), and sometimes they even pointed out aspects that I had overseen but were still relevant to understand the use of persuasive strategies in the excerpt. This was often the case with modes such as gaze or facial expression, which were not included in the analysis, but were clearly important to understand some rich points. A good example is the shoulder shrugging pointed out by the speaker in PPKE2. I consider this a real asset in the study of a concept such as persuasion, which is highly dependent on communicative intentions: nobody knows better what the speakers were actually trying to communicate than the speakers themselves, even if they have not consciously reflected about it before or if they lack the metalanguage to discuss it accurately.

I would like to point out once more that the list of strategies proposed in this thesis is not exhaustive, but rather representative and descriptive of the corpus that has been studied. I believe that this corpus-driven nature of the methodology is in itself another advantage. Instead of being conceived as closed, the list is flexible and adaptable and can be refined and enlarged including more modes and potentially more strategies as they are revealed by further research.

Throughout the preceding discussion I have argued that a video-based, corpus-driven study that combines input from previous literature with ethnography and MDA can be a valid approach to study persuasive strategies from a multimodal perspective. Despite some limitations, it has been proven to work in practice: its application to a corpus of 15 presentations has yielded a taxonomy of 7 persuasive strategies realised multimodally. Focusing on the multimodal ensemble instead of on modes in isolation, this methodology has provided a more comprehensive view of persuasive activity in these presentations. In addition, it has facilitated the triangulation of results.

7.2 Persuasion, modal density and modal coherence

In this section I want to discuss my findings concerning RQ II, mainly:

What is the contribution of the multimodal ensemble orchestrated through words, intonation, head movements and gestures to the creation of persuasion in the genres included in the analysis?
-Persuasion in the three genres is multimodal. -There is a correlation between modal density (i.e. number of modes used) and modal coherence (i.e. using modes consistently) regardless of the genre. -The more frequent a strategy is, the more likely it is to be realised multimodally. -The three genres feature instances of the seven strategies in different modal realisations.

Concerning the first two findings, the results of the analysis suggest that the multimodal ensemble as a whole, rather than isolated modes, make the greatest contribution to persuasion. According to the results, persuasion depends on the number of modes used, since the rich points were found to feature high modal density. However, the coherence across modes seems to be of equal or even greater importance. In other words, it is not only a question of the number of modes, but also a question of how well they are ensembled. The results discussed in Chapter 6 section 6.2.1 show how the persuasive effect can be jeopardized if modes contradict each other. It further illustrates that this can be the result of: 1) over-relying on words as a mode (and consequently neglecting the rest of the modes), and 2) shifting the focus from the higher level action of communicating with the audience to the action of trying to remember the exact words in a written script.

These findings are consistent with previous research on presentations (Rendle-Short 2006) which proves that positioning, gaze and action all signal the level of engagement of the speakers. Speakers, for example, tend to gaze away, shift body orientation away from the audience and start performing other actions such as using a computer or revising notes when they momentarily disengage from the action of presenting. My results show how this disengagement can also be perceived in intonation (e.g. use of a flat, oblique intonation) and gestures (e.g. a shoulder shrug conveying “I don’t care”). Similarly, Querol-Julián and Fortanet- Gómez (2012), Jewitt (2014) and Ruiz-Garrido (2015) notice how speakers are sometimes not fully unaware of their gestures, facial expressions, or intonation, and how these modes occasionally contradict their words, something which eventually also occurred in my corpus.

These results also agree to a great extent with the findings presented by Busà (2010) and Morell (2014, 2015). Busà claims that a greater awareness of kinesic and paralinguistic features as semiotic modes can help students of English become better communicators in general, and when using oral language in particular. With a focus on presentations, Morell finds that a skilful combination of modes makes presentations easier to understand and, in particular, the visual modes can compensate for lack of language proficiency. She also notes how more experienced presenters use body language to interact with the audience. The more aware presenters are of the modes she identifies in presentations (verbal, written, non-verbal material and body language), the more able they are to combine them effectively, foregrounding them at specific moments of their presentations and achieving effective redundancy. Finally, she recommends speakers to “use various modes to refer to the same proposition, and if it is complex it may be a good idea to use the modes consecutively” (Morell 2004: 161).

It is in this last point where my findings differ slightly, since a higher modal density would mean using modes simultaneously and not consecutively. Bringing Goldin-Meadow’s (2003) cognitive perspective on this issue, it is important to consider not only the number of modes used simultaneously, but also whether these modes are expressing the same meaning, or constitute a mismatch, with each mode expressing different aspects of meaning (although not necessarily contradictory). In the first situation, the redundancy achieved can be greatly beneficial, as mentioned before. In the second situation, the cognitive load may be increased and audiences may obtain less information from each mode separately, but the total amount of information gathered from the ensemble will tend to remain the same. The difference is that this information comes from different modes. In addition, I consider that this potential loss of information is not problematic in the case of a persuasive oral presentation, where the focus is on conveying main points rather than on minute detail (which the audience is not likely to remember anyway). Furthermore, if the speaker is confronted with a heterogeneous audience, I would argue that variety in terms of semiotic modes can only ensure that the message gets across in one way or another, catering for the different preferences and cognitive abilities of the individual members of the audience. This, in turn, will make it easier for the audience to understand the message and will pave the ground for persuasion. In other words, using complex multimodal ensembles can have the effect of making the audience miss on details, but will also contribute to make main points more salient and memorable, and the presentation ultimately more persuasive as a consequence.

It also seems that modal coherence (i.e. avoiding contradiction across modes) may be a more important factor for persuasion than the mere number of modes, although further research is needed to confirm this claim. Some progress in this sense has been already done by Tang (2013), who points out how different combinations of modes do not always produce the same meaning even if they deal with the same content, and warns against the "injudicious use of multiple representations" (Tang 2013: 34) just for the sake of motivating and getting the attention of the audience.

The third finding (i.e. the more frequent a strategy is, the more likely it is to be realised multimodally) can be considered a question of statistical frequency: the more frequent, the more likely a strategy is to be realised in different ways. I believe it is also related to the way most speakers reported their preparation for the presentations with a strong focus on the verbal element (in particular in the case of product pitches). The consequence is less awareness and a more restricted use of the other modes, which in turn means that a strategy is not likely to be enacted through other modes unless it is sufficiently frequent.

In the following subsections I will provide a discussion of the fourth finding (i.e. the three subcorpora feature instances of the seven strategies in different modal realisations). To this aim I will come back to the results of the multimodal analysis of each subcorpora and I will draw on the data provided by the speakers in the interviews and the insights from previous literature to provide my interpretations. I would like to stress that I make no attempts at generalisations, and that my results apply strictly to the corpus I have analysed, but further research is needed to check their validity for the genres as a whole.

7.2.1 The case of conference presentations: interpretations of the strategies used and their modal realisations

The frequency with which the persuasive strategies are used in conference presentations seems to be very consistent with the communicative situation of the genre, in which a lengthy and complex message is presented to an audience with whom shared knowledge is expected. The frequent use of anticipation and rapport suggests that speakers are aware that when they present at a conference they are in a way exposing their research for evaluation, and therefore they try to prevent potential counterclaims and build a good interpersonal relationship that will promote a positive evaluation of their research. In this sense, it is interesting to note that evaluation on the part of the speakers is minimal and only realised through words, which is probably a result of this academic context in which objectivity is generally aimed at, and also of the communicative situation established, in which it is rather the audience who is supposed to assume this evaluative role. This is particularly true of process-genres, among which the CPs in my corpus can be included. According to previous research (Querol-Julián 2011, Querol-Julián and Fortanet-Gómez 2012, 2014), evaluation in CPs seems to play a more prominent role in the Q&A session, which is out of the scope of the present study.

Rapport in CPs also takes a predominantly verbal form but it is also frequently realised through gestures, in particular deictic referential gestures that point at the speaker or at the audience. These results are consistent with previous research on academic discourse showing that participants in academic events need to claim a role as legitimate members of the community and integral parts of the event (Hyland 2006, 2009) as a prerequisite to be convincing. The use of gestures in particular is consistent with previous research that proves how gestures in conference presentations can help create a sense of inclusivity, bring the audience into the discussion and establish common ground (Hood and Forey 2005; Holler 2010). It is also interesting that rapport in conference presentations frequently takes the form of an honest sharing of research experiences (with speakers occasionally enacting characters in the narratives that they design to make this exchange more interesting). This agrees with Thompson's (2002) findings about the use of narratives in presentations that relate difficulties and failures along the way in a more honest way than a written paper. It is also in line with Zareva (2013), who in a study of the interplay between written academic genres and oral presentations finds that adherence to rules of the written genres when presenting helps to project a scholarly self, but a successful presentation also needs to show the social and personal self. Interestingly, a similar trend has been noted in business presentations, where the use of direct speech has been found to be a frequent device to boost persuasion (Bowker 2006, 2007).

Speakers did not explicitly comment on the idea of exposing research for evaluation in the interviews, but they did report feelings of being unsure about the suitability of the topic and its potential interest for the audience, and in some cases explained that they were consciously building rapport. This is clearly illustrated by the way the speaker in CPRE1 described her presentation style: “there was a lot about making sure people understood why we were there, that we fit, and that the work we do was valid and rigorous”. The rapport-building also has to do with “knowing I was an outsider in that community, so, I’m just making sure that, it’s building the rapport to show that I was happy to be there even if our work is a little bit different, maybe.” Another example is CPTO1 account of how he and his co-presenter kept changing the presentation until the last minute to include references to other presentations in the event. This practice of citing has been described by Myers (1989) as a potential face threatening act. Actually, Fernandez Polo (2014) shows how speakers in conference presentations use gesture and facial expression (e.g. a smile) to mitigate this threat. My results indicate that in the subcorpus of conferences presentations that I have analysed, the need to fit in the event seems to be higher than the face risk, and speakers do take it, although it is also true that a favoured practice is to refer to the event in general rather than to specific participants.

7.2.2 The case of research dissemination talks

In the research dissemination talks subcorpus, results show that speakers frequently resort to gestures and intonation to highlight parts of their messages, which explains why emphasis is very frequently enacted through a multimodal ensemble that combines words, intonation and gestures (e.g. a beat synchronous with prominence).

Regarding the use of strategies in dissemination talks, the frequent use of rapport and anticipation shows awareness of the need to connect with the audience in a communicative situation in which both parties do not necessarily share the same interests. This is also supported by the relatively frequent use of processing aids (mainly through gestures), which seems to show awareness of a remarkable knowledge gap between an expert speaker and a lay audience. Although the use of projection is not particularly prominent, its presence shows a need to make things clear in a subtle way without being patronising, which again supports my previous claims. Some insights from the interviews further back these interpretations. For example the speakers described dissemination talks as a genre that focuses on entertaining rather than on providing accurate data, and interestingly, the presentation that was voted as the best received positive feedback about extensive use of examples and parallelisms that brought the topic closer to the audience. Recent research on TED talks also supports these findings (Caliendo 2014; Caliendo and Compagnone 2014; Scotto di Carlo 2014).

In relation to gestures, it is interesting that beats are frequently used with two main functions: pragmatic and cohesive. There are also three types of gestures (metaphoric, deictic and beats) that are used with an interpersonal function. The use of pragmatic and interpersonal gestures seems to show a more conscious need to connect with the audience and make the topic relevant, while the use of cohesive gestures can be related to a need to facilitate understanding for a layman audience. Once more this is consistent with the communicative situation described above and the relationship established between speaker and audience.

7.2.3 The case of product pitches

The results of the analysis of the subcorpus of product pitches revealed a remarkable lack of head movements used for persuasive purposes. Although the data gathered from the ethnographic interviews has not provided any plausible explanation for this trend, the physical configuration of the room where the event was held may have had an influence. The audience and the jury were positioned perpendicular to each other, which made the task of keeping eye contact with both parties more difficult for speakers. This may have affected the way they moved their heads.

A second interesting result was that projection of context of interaction seems to be totally dependent on intonation as its only modal realisation. This is probably related to the affordances of this mode, that is, intonation seems to be particularly suitable to convey this meaning. As a matter of fact, projection of the context of interaction is a concept widely used in DI, and it is regarded as one of the more important communicative effects of this mode, as discussed in Chapter 3.

It was also found that, like in the case of research dissemination talks, processing aids frequently take the form of gestures. It seems that speakers try to make their content clear through visual representations which are produced and interpreted faster. This is in line with studies about the contribution of hand gestures to the semantic information conveyed by words (Beattie and Shovelton 1999; McNeill 2008). It is also consistent with the claims of

Hood and Forey (2005) and Rowley-Jolivet and Carter-Thomas (2005) who maintain that the immediacy of an audience and the consequent need for real-time processing strongly influence the way a presentation is designed, imposing a need to make it easy. Although their study was concerned with conference presentations and not product pitches, it can be reasonably assumed that the same constraints apply to product pitches, and even to a greater extent, since they are significantly shorter. On the other hand, this trend can also be a compensation strategy in the case of non-native speakers of English (4 out of 5 in this subcorpus), who may be using gestures to compensate for lack of lexical knowledge. Conversely, the native speaker in the subcorpus (PPPA) may be using these processing aids partly to adjust to the audience's mastery of the language: although he did not mention gestures, when discussing intonation he did explain that he intentionally slowed down because he was aware that he was the only native speaker in the room. This interpretation agrees with the findings of Morell (2015), who suggest that non-native speakers can use non-verbal modes to compensate for limited mastery of the language. However a larger comparative study would be required to confirm this.

Speakers also evaluate their products positively and evaluate competitors negatively, and they resort to gestures and to a lesser extent intonation –apart from words- to do this. It seems that evaluating through gestures and intonation is more subtle than saying it explicitly with words. This subtlety, in turn, has two advantages in the communicative situation of the product pitches. On the one hand it contributes to creating a feeling among the audience of being reaching their own conclusions freely rather than being pushed in a particular direction, which increases chances of this evaluation being accepted according to previous literature (O'Keefe 2002; Perloff 2003). On the other hand, the event was actually a contest and speakers were competing, and it is reasonable to think that an explicit negative evaluation, although occasionally present in the subcorpus (PPTO02 "a useless gadget"), could at some points be considered impolite and disrespectful.

There are two strategies that play a relatively important role in product pitches, viz. anticipation and rapport, but which rely heavily on words and are rarely enacted through a multimodal ensemble including gestures, head movements or intonation. Probably this is because the communicative situation in which the product pitch is used implies an unbalanced power relationship in favour of the audience (and in particular of the jury), and speakers are more cautious in the relationship built with this audience. It is plausible to think that they want to avoid being too conspicuous or patronising, or sounding over-confident. This was not, however, explicitly mentioned in the interviews. In the case of intonation the use of a foreign language is also another plausible explanation, and is consistent with previous studies that show how the fact of using a foreign language can prevent speakers from fully exploiting intonation for rapport building and interpersonal meaning (Pickering 2001, 2009; Pickering et al. 2012). In particular, the finding is consistent with Hewing's (1998) views that when speakers need to focus on grammar, vocabulary, etc. they logically disengage from the interactive activity of speaking to focus on the encoding process. A way to help them would be to sensitize them to the effect of different tone choices while they are listening and while they are freed from the burden of encoding. Hincks (2005, 2009) also suggests the use of computer-assisted automated feedback to help students improve their use of intonation in presentations.

Regarding the type of gestures used, as mentioned above, the most frequent are metaphoric, referential gestures that visually clarify some part of the content of the message and cohesive beats that show connections between parts of the message. They serve as processing aids intended to make the message easier to understand and in turn more likely to be accepted. Interpersonal gestures are only used twice, but the two pitches in which they appear (PA and SE) were two clear favourites according to what speakers said in the interviews. This suggests that audiences appreciate being acknowledged and included in the "fictional" conversation established (Hood and Forey 2005), and is in line with claims that messages are more persuasive when they are both emotionally and rationally appealing (O'Keefe 2002; Perloff 2003). The rest of the speakers in the product pitch subcorpus might have been too cautious in their level of closeness with the audience.

7.3 Conference presentations, research dissemination talks and product pitches as multimodal genres

In this last section I want to highlight some of the similarities and differences that the comparative analysis across the three subcorpora has revealed, and discuss them in the context of the data obtained from the interviews, the observation sheets and the input from previous literature. This offers an interpretation of my answer to RQ III:

To what extent is the use of persuasive strategies genre-specific and to what extent is there overlap between the three genres?
-The three genres use the seven persuasive strategies identified and they realise them through different modes, but there are differences regarding how much each strategy is used and how much each mode is used. -There are some genre-specific uses of multimodal persuasive strategies: specific attention getting in conference presentations and specific evaluation in product pitches. -The genre that stands out as more modally dense and more persuasively dense is research dissemination talks. -In each genre, a different speaker-audience relationship is aimed at.

I will begin by addressing the first two findings, i.e. there are differences regarding how and how much each strategy and mode is used. Concerning the use of modes, it is interesting that in all genres words are used more frequently. This suggests that speakers tend to rely on this mode and consider it their main way of expressing meaning, something that is supported by the data obtained from the interviews regarding the preparation of the presentations. All speakers reported that they had thought about the words they would use to different extents, ranging from writing a script based on their slides and memorising it to having a set of key concepts to expand upon more freely. In contrast, only one of them (DTJO) reported having rehearsed gestures and none of them prepared their intonation.

In the conference presentations subcorpus intonation and words are similar in frequency, and head movements are used only slightly less frequently than gestures. This is interesting, because in the other two events head movements were either not used at all or used significantly less. A possible explanation is the physical configuration of the event: speakers in the conference presentation were standing behind a lectern where a fixed microphone was placed. This invited a resting position of the hands clutching both sides of the lectern and did not facilitate moving around the stage. The lectern also made the body of the speakers less visible for the audience, and it is plausible to think that the more frequent use of head movements was a way of compensating for this constraint in body language.

Concerning the type of gestures, in all presentations there is a predominance of metaphoric referential gestures that visually clarify aspects of content. This reveals the suitability of this mode to express complex ideas in a synthetic, condensed way. Oral presentations are demanding in this sense because they need to be understood in real time, and they also have a limited amount of time assigned (in particular product pitches in this dataset were very limited in time). Therefore gestures become very suitable as the fastest way to facilitate understanding. However, differences in the types of gestures used in the three subcorpora suggest a different focus of attention in each genre. In product pitches and research dissemination talks there is extensive use of beats that acquire a pragmatic or cohesive function. These beats can have the effect of facilitating understanding either by clarifying the structure of the message or by showing how the message is intended to be understood (i.e. what is the pragmatic meaning intended). Research dissemination talks also display the highest use of interpersonal gestures, which reveals a need to connect with the audience and make the topic relevant, as explained in the previous section. Conference presentations, on the other hand, feature a more extensive use of deictic gestures that point either at the audience or at the speaker. This can be considered to show a different communicative situation in which the main concern of the speaker is to legitimize their participation in the event proving that they are competent members of the scientific community and the adequacy of their topic to the other presentations and the event in general. Finally, the use of deictic gestures in product pitches occasionally acquires an evaluative function (negatively evaluating competitors) that has not been found in the other presentations, but which is consistent with the use of strategies in that particular genre (as I will discuss in the following paragraphs) and with the genre’s main communicative purpose.

Regarding strategies, results have shown that in all presentations speakers highlight certain parts through emphasis, being the most widely used strategy and the one that is more frequently realised through a dense multimodal ensemble involving at least three modes. This is probably related to a human limited ability to process messages. Speakers seem to know (probably from their own experience as members of an audience) that attention and reflection is not kept constant throughout a presentation, and emphasis is a way of helping direct this fluctuating attention to the appropriate parts of the message: the takeaways. In fact speakers recurrently described their use of beats and prominence as trying to show the importance of that particular part of their message.

In the conference presentations in my corpus evaluation is minimal, which probably responds to a need to project objectivity. Speakers in my corpus also use more processing aids and less rapport than in research dissemination talks, which suggests that speakers try to make themselves understood and then to make themselves likeable. This might sound contradictory with the greater amount of shared knowledge with the audience that can be expected of these presentations, but is probably a consequence of the communicative situation created in this type of events, where the audience is typically free to select the presentation of their interest. Since it can be assumed that the audience already "likes" the topic, the effort is shifted to making it easy to understand. This can also be further explained by the heterogeneous character of this particular conference, which made a lot of participants unsure about fitting in the event and worried about making their research relevant. The fact that these presentations are longer can also account for the more frequent use of attention getting than in the other genres, and also for the special form of attention getting by means of enacting characters, in which speakers act out assuming the roles of people they are referring to in their presentations, turning the account of their research experience into a vivid narrative. In turn, this narrative can also be interpreted as a way of building rapport with the audience through an honest sharing of experiences, which is very consistent with the views expressed by the speakers and the insecurity about fitting mentioned before. Previous research has also pointed out similar practices in other types of presentations. For example, Ochs et al. (1994) argue that dynamic grammar and gesturing help physicists create a visual representation of experiments so that they are perceived as credible. Morton (2009) also found similar attempts to increase vividness in student architecture presentations through gesture, language, visuals and physical objects.

In the research dissemination talks subcorpus, as shown previously in Table 6.14, evaluation occurs significantly less frequently than in product pitches but more frequently that in conference presentations. There is a higher frequency of rapport and anticipation than in the other two subcorpora. Immediately following in terms of frequency are processing aids and projection of context of interaction. It is probably the case that evaluation in research dissemination talks is present in a less explicit way than in product pitches. This would mean that the speakers in these talks prefer directing their audience towards a positive response rather than evaluating something explicitly and expecting agreement. As one of the speakers commented in the interview, he was not "evaluating my own research directly, that would be too much, it wouldn't work" (DTRA)". This is consistent with the claim that subtlety can boost persuasion (O'Keefe 2002; Perloff 2003), but is not totally in agreement with previous research on TED talks that suggest a much more overt use of evaluation (e.g. Scotto di Carlo 2014). The fact that the presentations in my corpus are part of a small-scale event and not the mass web-streamed talks by reputed scientists that this literature has described may account for this difference. The relative high frequency of rapport also suggests that a main concern in these talks is to make the presentation likeable. This can be related to the need of connecting with the audience and showing them the "what's in it for me". Immediately following is an effort to make presentations easy, which is revealing of the knowledge gap between speaker and audience.

In the product pitches subcorpus evaluating the product in positive terms (and competitors in negative terms) becomes an important concern, which is very consistent with the main communicative purpose of the genre (i.e. sell a product). Following are processing aids and anticipating and controlling responses (both equally frequent), and then rapport and projection (again equally frequent). This seems to suggest that the speakers are trying to make their messages easy to understand and at the same time preventing disagreement and counterarguments by triggering desirable responses, or in other words, they lead the audience to believe that their product can solve their problem, but in a way that does not require a lot of effort from them to understand, since this would be counter-productive (O'Keefe 2002; Perloff 2003). As a next priority the speakers try to make their messages likeable through rapport. This priority of strategies is consistent with the views on the genre provided by PPPA: the focus has to be on *why* your product might be interesting for the audience. This 'why' is what needs to be communicated first, otherwise people will not pay attention to the 'what' or the 'how'. In his own words "if the 'why' doesn't resonate with them then they won't care about any other technical aspect".

Turning now to the third finding (i.e. differences in modal density and modal coherence) an interesting outcome of the comparative analysis is the seemingly greater density of research dissemination talks in comparison with the other two genres, both in terms of modes and in terms of strategies. It could be expected that the shorter presentations (in my corpus the product pitches) would be more dense (since shorter time implies a faster pace), however dissemination talks are more dense than product pitches even if they are twice as long. I believe this can be accounted for in two ways. On the one hand, speakers in research dissemination talks have to overcome a potential knowledge gap and a potential interest gap (the audience may not know or may not be interested in their research), as the speakers actually reported in the interviews. This means that even if their preparation still focuses to a great extent on the words, they need to resort to other means to make their message interesting and accessible, because as previous research proves (Lemke 2009; Kress 2010) not all modes are equally capable of expressing (particular types of) meaning effectively (e.g. an equation is a much more effective way of expressing a mathematical relation than words). This may result in a greater awareness and wider use of the four modes I have analysed, and also in a more extensive use of strategies to keep the audience interested. On the other hand, the speakers in this subcorpus engage in this type of public speaking simply because they genuinely enjoy doing it, and they show an intrinsic greater interest in this activity (e.g. two of them, DTLI and DTLA, reported being members of, or occasional participants in, associations such as Toastmasters, devoted to public speaking, or in their own words "empowering individuals to become more effective communicators and leaders"). It can be argued that while the speakers in the other two subcorpora have more practical aims with their presentations (e.g. networking or obtaining funding), for the speakers in the research dissemination talks the only motivation is the satisfaction of being able to communicate their research effectively to a layman audience, and therefore they devote all their efforts to this aim.

A related outcome is that the conference presentations in the corpus are less modally dense (they feature a lower frequency of modes per minute) and less persuasively dense (they use less strategies per minute) than the other two genres. Unlike the case of research dissemination talks vs. product pitches, in which we find that the latter are more dense despite being longer, I believe that this result is, at least partly, a consequence of the considerably longer duration of the conference presentations (20 minutes, as opposed to 6 in research dissemination talks and 3 in product pitches). Speakers in these presentations, being under no strict time constraints, take more time in their persuasive efforts, and seem to scatter them more throughout their presentations. However, this does not imply that they are less persuasive, because speakers in these presentations still show great ability to seamlessly orchestrate modes into coherent multimodal ensembles to realise persuasive strategies (see example in Table 6.8). In other words, they can use modes very coherently.

Finally, regarding the fourth finding (i.e. different speaker-audience relationship), the different use of strategies and gestures that has been discussed in the previous paragraphs reveals that a different speaker-audience relationship is aimed at in each case. The speakers in the product pitches make an effort to present themselves and their product as a solution to a problem that the audience needs to solve. The speakers in dissemination talks, on the other hand, try to show how their research is relevant, and can even become life-changing for the audience. In contrast, in the conference presentations the focus seems to be on fitting in the event and the community as legitimate and competent members. This finding supports previous research that claims that every presentation is designed for a particular audience, which in CA is called the principle of *recipient design* (Rendle-Short 2006). The fact that speakers in my corpus tend to tailor their presentations to their audiences is a sign that they are trying to abide by this principle of recipient design. I also believe this adaptation to the audience is a pre-requisite to achieve persuasion, and that the multimodal ensembles that speakers orchestrate are the tool to achieve this persuasion. This finding is also very coherent with proposals for genre analysis reviewed in Chapter 1 and their clear emphasis on audience and the design of a message tailored to this audience (Bazerman 1994; Berkenkotter and Huckin 1995; Paltridge 1995).

I would like to close this discussion highlighting three important aspects that the combined multimodal and ethnographic approach to persuasive strategies in the corpus of presentations has revealed. First, it is interesting to note that the particularities of the conference presentations mentioned in the previous paragraph in a way make them closer to research dissemination talks, since the interest and knowledge gap is greater than it could be expected from a very specialised conference. Because of this, the speakers need to pay particular attention to make their research interesting and accessible, which further accounts for the use of story-telling found in the subcorpus. In the same way, the speakers in research dissemination talks are aware that in order to make their research relevant and interesting they need to present it as something that can potentially affect the audience in a direct way, their talks becoming more a call to action than an informative talk. This brings them closer to the characteristic problem-solution structure of product pitches. It can be argued that the design of the corpus may have affected this result. As mentioned in Chapter 5, there is a business element in the CPs included in the corpus, since the event is a symposium on business discourse, and there is also an academic component in the research dissemination talks, since the speakers are all doctoral students. However, I believe the potential transfer that these common elements may have triggered cannot fully account for the interrelationships found. On the one hand, the speakers in the product pitches are not academics, but enterpreneurs (in some cases former students of participating business schools). On the other hand, the speakers in research dissemination talks seem to be assimilating practices from the product pitches (which are non-academic and not related to their research topics) rather than from the CPs (with whom they share an academic background). A clear example of this is DTPI. Therefore, I believe the results of this study support the idea that the three genres can be considered to be intertextually and interdiscursively related. This has been actually suggested by previous studies (Virtanen and Halmari 2005; Bhatia 2005, 2010, 2012, Scotto di Carlo 2014; Caliendo and Compagnone 2014; Compagnone 2015), and could be confirmed by further MDA research on the three genres.

Second, it is interesting to note that the analysis has revealed fewer clearly defined preferences for certain strategies to be realised through particular modes than the researcher had hypothesised, with the exceptions of the pairs ‘projection of the context of interaction’ and ‘intonation’ and ‘processing aids’ and ‘gestures’. A larger-scale analysis using a bigger and more varied corpus could shed more light on this issue. In addition, with the remarkable exception of the types of gesture, the variations in the use of modes seem to be less related to intrinsic characteristics of the genres and more related to the affordances of each mode and the medium through which they are materialised: the physical configuration of the events (i.e. the rooms) and the individual characteristics of the speakers (e.g. native or non-native, presentation style, experience and training, etc.). According to Bateman (2014), this type of differences are attributable to the medium used (not the modes): they do not represent major changes in the rhetorical strategies used to achieve the communicative goals, and are likely to be found in other instances regardless of the genre, so they are generically non-defining. For example, a lapel microphone is likely to grant more mobility to speakers than a stative one, and this can lead to a more extensive use of gestures in the product pitches and research dissemination talks in my corpus, but this does not mean that these two genres are inherently more reliant on gestures than conference presentations (if the speakers in conference presentations had had a lapel microphone or no lectern as media, they would have probably moved more or walked around the stage). For the same reason, the recommendations found in literature about product pitches regarding the use of eye contact, body language and voice can equally apply to the three genres, because they are all related to the media used in oral presentations. However, the differences in the use of strategies, as well as the types of gestures used, are very revealing about the communicative situation that is created in each of them and in particular of the relationship established between speaker and audience, and I believe this can be considered a defining generic trait.

Third, it is also interesting to note that the speakers’ orchestration of modes is generally highly consistent with their motivations, views and expectations about the genre, as the previous discussion has revealed. For example, a greater concern for positioning themselves in the scientific community makes the speakers in conference presentations use deictic gestures to refer to themselves and to the audience, while in the case of product pitches the need to evaluate the product is translated in evaluative deictic gestures. Results also indicate that the dissemination talks in the corpus are more dense, both in terms of number of modes, and in terms of persuasive strategies, which can be considered a consequence of the relationship speaker-audience, but also of the speakers' own attitudes to public speaking. In this sense, results also seem to suggest that training can have an influence on the awareness and correct coordination of modes: more interest and more training in public speaking makes the speakers in dissemination talks better at achieving modal density and persuasive density.

In this chapter I have interpreted the results obtained from the MDA and ethnographic analysis of a corpus of 15 presentations. In the following chapter I will focus on the contribution of this study within the context of discourse and genre analysis, and I will also discuss its potential pedagogical implications and applications.

CHAPTER 8. CONCLUSIONS AND FINAL REMARKS

In this final chapter of my thesis I would like to offer a brief retrospective summary of my research journey. I will then draw some conclusions regarding genre analysis and the pedagogical implications of this research. Finally, I will address some limitations of the study that in turn provide opportunities for interesting further research on the area of MDA, in particular for the study of oral genres.

The motivation behind this multimodal and ethnographic study of persuasive strategies in conference presentations, research dissemination talks and product pitches was to shed light on how these three genres fulfil their main communicative purpose of persuading an audience. I had three objectives with this thesis. First, I wanted to suggest a methodology that can systematically approach persuasive strategies combining multimodality and ethnography. Second, I wanted to test this methodology on a corpus of 15 presentations, which has yielded a taxonomy of seven persuasive strategies that are realized as multimodal ensembles. Third, applying this methodological framework I have analysed the use of persuasive strategies and their modal realisations (focusing on words, intonation, gesture and head movements) in a corpus of the three genres which are object of this study.

In order to meet my objectives I have drawn on different trends within discourse and genre analysis. I have reviewed different approaches to discourse, highlighting two distinct trends: one towards descriptive stances that avoid looking at language as an abstract system independent from its users, and a second one consisting in broadening the understanding of communication, going beyond the purely linguistic and verbal towards a multimodal take on language. I have zoomed in on three disciplines that show potential to shed light on the speakers' intentionality, while taking into account the variety of semiotic resources that characterises oral discourse: Pragmatics, Systemic Functional Linguistics and Multimodal Discourse Analysis. Likewise, I have also reviewed different approaches to genre that highlight the relevance of aspects such as the dynamism that characterizes genres, the need to adapt to social situations and discourse communities, and the importance of the production and interpretation processes, all of which make a multimodal approach to genre particularly appropriate.

Concerning persuasion, my focus has been the use of persuasive strategies by speakers in oral presentations. I have paid particular attention to the fact that it has non-verbal realisations, which justifies a multimodal approach. I have also shown that persuasion has been frequently pointed out in previous literature as a powerful drive for generic change.

Finally, regarding paralanguage and kinesics, I have paid especial attention to several approaches that can account for communicative intent, mainly Brazil's Discourse Intonation and Kendon's classification of gestures. Both have fed my multimodal analysis of intonation, gestures and head movements as semiotic modes.

The research questions that have guided the study are:

RQ I. What type of persuasive strategies can be identified using an MDA approach?

RQ II. What is the contribution of the multimodal ensemble orchestrated through words, intonation, head movements and gestures to the creation of persuasion in the genres included in the analysis?

RQ III. To what extent is the use of persuasive strategies genre-specific and to what extent is there overlap between the three genres?

In order to answer RQI, in this study I have combined different trends of multimodality, i.e. SFL Multimodality and Multimodal Interactional Analysis, with ethnographic methods, i.e. interviews and observation sheets. This combination has been beneficial in several ways. First, it has allowed me to screen my data for potentially interesting moments for analysis (i.e. rich points). Second, it has provided me with a comprehensive view of persuasion, both as a communicative activity that can be systematically studied as an organised set of choices at different levels and within different semiotic modes, and as an interactive phenomenon that is co-construed between speaker and audience and requires constant adaptation as the situation evolves. Third, it has promoted a focus on multimodal ensembles, without assuming that one mode is by default leading the main role. However, as mentioned in Chapter 7, being more literate in the use of words can cause an unconscious bias towards this mode. Finally, it facilitates the triangulation of results with speakers, something which I consider crucial in a study that deeply involves communicative intentions.

Concerning RQII, the results of the analysis have shown that persuasion in the corpus analysed is indeed multimodal and that the three subcorpora feature instances of the seven strategies identified, but to different extents and through different modal realisations. They also reveal that there is a correlation between modal density (i.e. number of modes used) and modal coherence (i.e. using modes consistently) regardless of the genre. This leads me to argue that it is not only the number of modes that matters, but also their consistency: speakers can be more persuasive when they coherently orchestrate multimodal ensembles with a clear aim of persuading.

Turning now to RQ III, the analysis has shown that the genre in the corpus that stands out as more modally dense and more persuasively dense is research dissemination talks, and the less dense is conference presentations. I argue that this is due to two main factors: time constraints and the attitude of the speaker. A longer duration of the conference presentations in the corpus can make them less dense because persuasive efforts are more scattered. In the same way, a positive attitude towards public speaking in the research dissemination talks (which usually implies training and experience) makes these speakers more skilful in their orchestration of modes. In addition, the analysis has also revealed some genre-specific uses of multimodal persuasive strategies (i.e. specific attention getting in conference presentations and specific evaluation in product pitches) and a different speaker-audience relationship being targeted in each genre, which further research could confirm as an important defining generic trait.

In the following two subsections I will discuss the potential contribution and pedagogical implications of these findings for genre analysis and LSP.

8.1 Genre analysis

The research presented in this thesis is a qualitative study that does not aim at representativeness. The results presented in Chapter 6 and discussed in Chapter 7 are valid for the corpus analysed, but a larger-scale study is required to confirm if they are applicable to the three genres as a whole. However, I believe they do show the potential contribution of a combined MDA and ethnographic approach to genre studies, which can be illustrated by the aspects I discuss in the following paragraphs.

One of the main conclusions that can be drawn from this study is that the type of relationship that is established with the audience seems to stand out as a defining trait of each genre, and as a crucial element to pave the way for persuasion. As discussed in the previous chapter, the interpersonal relationship between speaker and audience is different in the three subcorpora analysed, and successful speakers need to be aware of this fact if they want to convince their audiences. The analysis has also revealed that this potential generic difference is revealed in the frequency with which speakers use persuasive strategies and in the type of gestures that they use. Apart from gestures, other differences in the use of modes are attributable to individual differences of the speakers and to the media available in each case, and they are not in themselves a distinctive trait of the genres.

The study also hints at intertextual and interdiscursive connections among the three genres, with CPs drawing on research dissemination talks and these in turn drawing on product pitches to achieve the right interpersonal relationship with the audience. At the risk of oversimplifying and overgeneralising (since further research is needed to confirm this), I would like to suggest that the three genres could be placed along a continuum of what can be called *closeness* with (or degree of *relevance* for) the audience that can be illustrated in the following way:

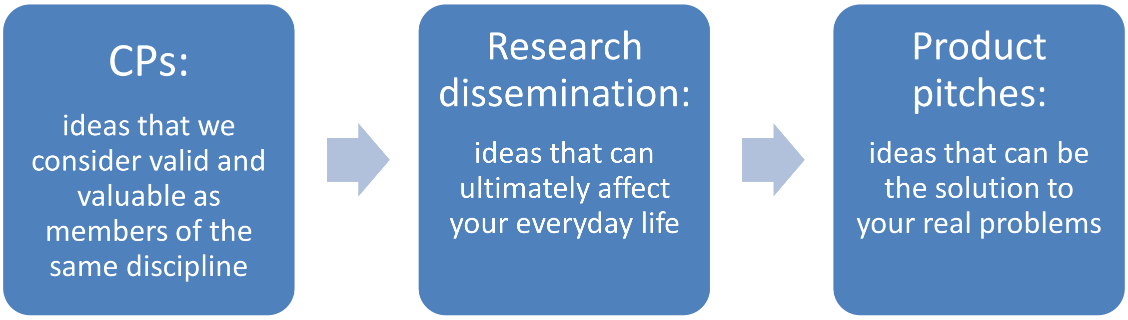


Figure 8. 1 Continuum of oral presentations in academic and business settings

A third important conclusion that I can draw from this research is that there seems to be a correlation between modal density (i.e. the number of modes used), modal coherence (i.e. the way these modes are orchestrated into a coherent multimodal ensemble) and persuasion. In other words, speakers in oral presentations can become more persuasive when they are aware of the affordances and limitations of the modes they can use, and when they are able to use them coherently for a specific communicative purpose. The results of this study also suggest that this consistency across modes is better achieved when speakers foreground the action of communicating a message to the audience. When speakers are distracted by other actions like using props or remembering a script, there is a disengagement from the audience that can hinder persuasion.

Finally, there seem to be two factors affecting modal density and persuasive density (i.e. number of persuasive strategies). One is the total length of the presentations, as suggested by the correlation between a lower density and a longer duration in conference presentations.

The other one is the speaker's own attitude to public speaking. The results obtained in this study suggest that training, experience and a personal interest in public speaking increases speakers' ability to orchestrate powerful persuasive multimodal ensembles.

8.2 Pedagogical implications: Genre, LSP and MDA

Genre has been regarded as a useful tool in the teaching of Language for Specific Purposes (LSP). It is perceived as an important aspect of the knowledge that students acquire as they are enculturated into a professional community, which allows them to become operative professionals (Berkenkotter and Huckin 1995, Fortanet-Gómez, Posteguillo, Palmer and Coll 1998; Luzón 2005; Ruiz Garrido, Palmer and Fortanet-Gómez 2010). In other words, novice researchers who want to become academics or business students who want to become business professionals need to be fluent in the use of the genres that are common in their respective fields. Examples of these genres are product pitches in business and conference presentations in academia.

Following up on the discussion of results presented in the previous chapter, I will now draw some implications for the teaching of LSP, and in particular for the teaching of foreign languages for academic and business purposes.

First of all, based on the results of my study, I believe that prioritizing words (or any other mode) is not always a good strategy. Over-relying on one mode means missing opportunities to communicate effectively (and persuasively) through other modes. Because of this, I tend not to favour the idea of memorising a text word by word. When speakers do this they focus on remembering their memorized text, and neglect other resources to express what they want to convey to the detriment of the overall effectiveness of their messages.

Second, I believe that students should be encouraged to stick to one foregrounded action: communicating with the audience. Focusing on other actions like using props results in projecting disengagement, which hinders persuasion. This is another reason why I claim that memorising a script should be taken with caution. When speakers do this they put all their efforts on remembering correctly, and they consequently neglect their audiences, because attention shifts from them to the words. This results in a recited text that may be very well crafted, but comes across as not genuine, not interesting and not motivating for the audience. I believe this can have a worse effect on persuasion than a possible slip of the tongue, a false start, or a mispronunciation, so it should be given priority.

These two aspects also concern the preparation for a presentation. Most of the participants in my study have shown a tendency to an unbalanced focus on the different aspects that conform the *how* (i.e. the delivery of the presentation). Most of them focus on the words and on the visuals, and have prepared a slide show and rehearsed a script, but without thinking of the many other elements affecting delivery (i.e. other semiotic modes). This is a tendency already spotted by Campbell et al. (2001). They found that speakers tend to overestimate visual aids and underestimate delivery skills (which include the use of intonation and kinesics) when they self-assess their presentations. In my opinion the preparation for a presentation should be more comprehensive. It is necessary to go beyond writing a script and designing a set of slides. With this I do not mean that intonation and body movements should be staged, because this is likely to take away spontaneity and honesty and in turn make the message less persuasive. However, I do argue in favour of rehearsing in a situation which is as close as possible to the real one, for example in front of people who have not been involved in the design of the presentation and have approximately the same shared knowledge with the speaker as can be expected from the real audience. It is also important for novice speakers to gather as much information as possible about the event beforehand, because this can help them fine tune their messages to suit the communicative situation, including the relationship that needs to be established with the audience (e.g. how much shared knowledge can be expected, how much interest in the topic, etc.). The need to adapt to the audience and situation has been previously put forward by Hyland (2005) in relation to academic writing, and I believe it is even more important in the case of oral presentations due to the immediacy of this audience. In my opinion, students should be encouraged to consider this an integral part of their preparation, because it is as important as language accuracy.

A third aspect that could be worth integrating in LSP curricula is an effective and coherent synchronization of modes. In the corpus analysed for this study, inconsistencies across modes have been found to hinder the communicative effect intended. When it comes to using intonation or gestures as a mode, non–native speakers can find it particularly hard to emulate native practices. They might think this is even undesirable, since these modes are to a great extent culturally shaped, and they might feel they are somehow jeopardising their own cultural identity. In this sense, I side with Norris (2015) and Chun (2002) and I believe the aim should not be to master a native model, but to use modes in accordance with the situation. Students should aim at achieving consistency and coherence in their use of modes to convey their meaning as efficiently as possible. Like Chun (2002), I believe that intonation should be judged as “appropriate to the communicative situation” rather than “correct”. No accent is more correct than other. It is rather not contradicting yourself with your use of modes and never losing sight of your audience that plays a role, and this is what I believe students should be trained for.

8.3 Limitations of the study and further research

In this last section of my dissertation I discuss some limitations of the study I have carried out that should be taken into account in order to understand and contextualise the results and conclusions presented in the thesis. I also point to some areas that can provide interesting topics for further research. The study I have described in this dissertation presents limitations in terms of the scope of the multimodal analysis (i.e., number of modes involved) and the size of the corpus. In addition, it leaves out two important aspects of the communicative process, which are the uptake by the audience and the influence of culture and native language. The following subsections will deal with all of these aspects in more detail.

8.3.1 The modes that play a role in oral presentations

Throughout this dissertation I have emphasized the importance of focusing on the multimodal ensemble that is construed in an oral presentation without prioritizing any mode and keeping attention on the joint effect created by the coherent use of all of these modes. However, my MDA analysis of modes used in persuasive presentations is restricted to four specific modes. As pointed out in chapters 4 and 5, I have focused on words, intonation, head movements and gestures for the sake of making the analysis feasible. I have selected those modes in particular because they are difficult to avoid when speaking, because their effect is quite noticeable in oral discourse and because they have received little attention so far.

I believe this research decision has helped to make the analysis more manageable and rigorous, but it implies leaving out other modes that are equally part of the multimodal ensemble and contribute to fulfil the speakers’ communicative intentions. This is the case of facial expression or visuals projected on a screen, to mention just two examples. At some points I actually needed to refer to these modes to interpret the other modes I was analysing. This proves my claim that they are all parts of a whole, and they only make full sense when interpreted as such. A clear example of this is found in the deictic evaluation discussed in Chapter 7. In these cases, a pointing gesture becomes evaluative because the object of pointing is a projection of a competitor product. Further research, in this case, should cater for the integration of more modes and for a more complex approach to the concept of multimodal ensemble.

Another aspect that can be raised in relation to the modes included in the current study is the division into modes. It can be argued that words and intonation are actually part of the same mode, and the same could apply to gestures and head movements. My research decision is based on the criteria suggested by Kress (2010): mainly, a mode becomes a mode when it is capable of conveying the different types of meaning postulated by SFL and society accepts it as such. However, I believe that in the end this issue comes down to a matter of terminology, and whether they are considered modes in themselves or components of a larger mode, or semiotic resources, they all constitute integral parts of a multimodal ensemble that is able to convey meaning effectively when it is well orchestrated.

8.3.2 The variables that affect persuasive behaviour and define (or do not define) genre

One of the major challenges in studying the persuasive behaviour of speakers in oral genres is the large amount of variables that can influence their persuasive activity. These may include individual traits such as age, experience as a presenter, emotional state (e.g. nervousness or stress), among others. Also aspects such as the physical configuration of the room or the timing of the presentation within the event can affect the way speakers perform. Isolating these variables can be in itself a daunting task in a naturalistic environment, but in addition to this there is the question of which of these variables can be considered defining of the genre and therefore worthy of attention in genre studies. From a MDA approach, there is yet a third question to be solved: which of these variables are in themselves ways of conveying meaning, i.e. modes or media? And should they be considered part of the multimodal ensemble?

More advances in multimodal research are definitely needed to overcome these challenges. In the case of the present study, I believe that the combination of MDA and ethnography has proved a useful tool to move a modest step forward in this direction.

8.3.3 Larger comparative studies

As mentioned in section 8.1, the study presented in this dissertation involves 5 instances of each of the 3 genres analysed. Furthermore, from these instances only 2 rich points are selected for fine-grained analysis. This size, although in line with previous multimodal research, does not allow for generalisations applicable to the genres as a whole.

This was actually not the intention of the research, since it was conceived and framed as a qualitative study. However, results suggest that a larger scale quantitative study including entire presentations (and even whole events, including all the presentations in them) would be very useful to carry out a deeper and more rigorous genre analysis. In particular, it would contribute to differentiate personal traits from generic traits, or in other words, what is characteristic of the speaker as an individual and what is characteristic of the genre. It would also help discern what features are a consequence of the media available in a particular event rather than a defining characteristic of the genre (see section 8.3.2). For example, the fact that speakers in my corpus of product pitches do not use head movements to persuade does not necessarily mean that this is a defining characteristic of product pitches, and it may have more to do with the positioning of the speaker relative to the audience and the jury.

Large scale quantitative multimodal analysis presents two main difficulties: it is time consuming and often involves great computational cost. However, new developments in specialised software are promising and can greatly contribute to overcome these difficulties. One example is *Digital Replay System* developed at the University of Nottingham. This type of software is a step forward in the automated identification and coding of gestures and facilitates the analysis of larger corpora. Another example is the software package *Multimodal Analysis*, developed by the Multimodal Analysis Lab at Curtin University, Australia. It features an information storing system that manages to reduce the computational cost of dealing with video, audio and text files avoiding the need for duplicate copies of documents. It also allows the researcher to focus on the interaction between modes through a visualization system that shows relative time for combinations of multimodal choices.

8.3.4 Uptake by audience

In this study I have analysed the use of persuasive strategies in three different genres using a multimodal approach combined with ethnographic methods. However, in a sense, I have only looked at one side of the coin, since my study focuses on the speakers and their communicative intentions. Although the interviews at some points included questions related to how speakers experienced other presentations in the events as members of the audience, the uptake by the whole audience was not addressed thoroughly or systematically. This was a conscious research decision to maintain the analysis scope within feasible limits, but it is definitely an interesting area for further research that could help revisit the researcher's and the speaker's interpretations about their communicative intentions and check if they were in fact conveyed and taken up by the audience as intended.

Different methods could be used to assess the audience uptake. The same type of ethnographic interviews that I have used with speakers could be adapted to test the audience experience of the presentation, and in particular whether they were convinced by it. This would imply a careful design of questions to obtain relevant information regarding the audience comprehension of the main claim, their previous thoughts and opinions about it and whether these thoughts and opinions have been somehow affected by the presentation. Alternatively, questionnaires can substitute for structured or semi-structured interviews in case the size of the audience does not allow for interviewing every member.

Recently developed specialized software can also aid in the analysis of the audience uptake, such as the electroencephalography neuroheadset Emotiv EPOC+, able to register brain activity prompted by certain stimuli. This software has been already applied by the research group DACTIC of the Universitat Jaume I in a pilot analysis of the emotions of professional designers during the process of idea-generation. This software could be applied to register the reactions of selected members of the audience as they watch the presentations, providing real time data about their experience that would probably not be retrievable by ethnographic interviews because audience members would not remember them or probably would not even be conscious of them. This line of research has already been initiated by the GRAPE, Group for Research on Academic and Professional English, at the Universitat Jaume I.

8.3.5 The influence of culture and native language

In this study I have not contemplated the potential influence of the culture and native language of speakers (which could probably be added to the many variables affecting persuasive behaviour mentioned in section 8.3.2). Although it was not one of my research objectives, I have to say that in my corpus I did not find significant differences between English NS and NNS. Nevertheless, previous research points out that culture can be an important factor in both academic discourse (Mauranen 1993; Duzsak 1997; Martin 2003; Hirano 2009; Bennett 2010; Loi 2010; Uysal 2014) and in business discourse (Hendon et al. 1996; Chu et al. 2005). In the case of academic discourse, for example, targeting at the right amount of indirectness and hedging (Uysal 2014) or creating rapport through humour or anecdotes (Morell 2014) is particularly difficult for NNS, and can deeply affect the final effect of the message. Perez-Llantada et. al (2011) have also revealed how Spanish academics report feeling at a disadvantage when they have to speak English in conferences, because they cannot express ideas as clearly and as appealingly as in their mother tongue, and they are unable to express nuances that they would like to convey or modulate strength of claims as accurately as in Spanish. As a consequence, they have problems to clarify, persuade and convince.

In particular, the under-use of intonation for persuasive purposes (e.g. for creating rapport) that I have found in product pitches can be a consequence of using a second language (Hewings 1998; Wrembel 2007), and it can be worth probing into this issue with larger scale studies.

To conclude, the study discussed throughout this dissertation does not claim to be exhaustive, but rather intends to pave the way for larger scale studies that can provide both more qualitative and quantitative data. This, in turn, will help refine our understanding of oral genres in general, and of the way persuasion works in them in particular. Whether further research will ultimately provide a magic recipe to persuasion is still to be seen. This PhD thesis is my modest contribution to this complex field.

REFERENCES

Aboudan, R., and G. Beattie. 1996. "Cross-Cultural Similarities in Gestures: The Deep Relationship between Gestures and Speech Which Transcends Language Barriers." *Semiotica* 111(3–4):269–94.

Adolphs, S., and R. Carter. 2007. "Beyond the Word: New Challenges in Analysing Corpora of Spoken English." *European Journal of English Studies* 11(March 2014):133–46.

Adolphs, S., D. Knight, and R. Carter. 2011. "Capturing Context for Heterogeneous Corpus Analysis: Some First Steps." *International Journal of Corpus Linguistics* 16(3):305–24.

Agar, M. 1996. *The Professional Stranger: An Informal Introduction to Ethnography*. Academic Press.

Akita, M. 2005. "The Effectiveness of a Prosodyoriented Approach in L2 Perception and Pronunciation Training." *Academic Studies, English Language and Literature* 53: 1–22.

Alba-Juez, L. 2009. *Perspectives on Discourse Analysis: Theory and Practice*. Cambridge: Cambridge Scholars Publishing.

Anderson, C. 2013. "How to Give a Killer Presentation." *Harvard Business Review*. Retrieved 3/2/2017 (https://hbr.org/2013/06/how-to-give-a-killer-presentation).

Andeweg, B. A., J. C. de Jong, and H. Hoeken. 1998. "'May I Have Your Attention?': Exordial Techniques in Informative Oral Presentations." *Technical Communication Quarterly* 7(3):271–84.

Aoki, H. 2011. "Some Functions of Speaker Head Nods." in *Embodied Interaction: Language and the Body in the Material World*, edited by J. Streeck, C. Goodwin, and C. LeBaron. Cambridge: Cambridge University Press.

Argyle, M., G. Alkema, and R. Gilmour. 1971. "The Communication of Friendly and Hostile Attitudes by Verbal and Non-Verbal Signals." *European Journal of Social Psychology* 1: 385–402.

Argyle, M., V. Salter, H. Nicholson, M. Williams, and P. Burgess. 1970. "The Communication of Inferior and Superior Attitudes by Verbal and Non-Verbal Signals." *British Journal of Social and Clinical Psychology* 9: 222–31.

Aristotle. 1959. *Ars Rhetorica*. Edited by W. D. Ross. Oxford: Oxford University Press.

Aristotle. 1996. *Poetics* (translated with an introduction and notes by M. Heath) London:

Penguin.

Arndt, H., and W. Janney. 1991. "Verbal, Prosodic and Kinesic Emotive Contrasts in Speech." *Journal of Pragmatics* 15: 521–49.

Askehave, I., and J. M. Swales. 2001. "Genre Identification and Communicative Purpose: A Problem and a Possible Solution." *Applied Linguistics* 22(2):195–212.

Atkinson, M. 1984. *Our Master's Voices. The Language and Body Language of Politics*. London: Methuen.

Austin, J. 1962. *How to Do Things with Words*. Cambridge: Harvard University Press.

Avery, P. W., and S. Ehrlich. 1992. *Teaching American English Pronunciation*. Oxford: Oxford University Press.

Bakthin, M. 1986. "The Problem of Speech Genres." Pp. 60–101 in *Speech genres and other late essays*, edited by W. McGee. Austin: University of Texas Press.

Baldry, A., and P. J. Thibault. 2006. *Multimodal Transcription and Text Analysis: A Multimedia Toolkit and Coursebook*. Equinox Publishing Limited.

Bamford, J. 2007. "Accentuating the Positive. Evaluation and Persuasive Discourse in Business Presentations." Pp. 135–55 in *Business Discourse: language at work*, edited by J. Bamford and R. Salvi.

Bamford, J. 2008. "Telling a Convincing Story: A Corpus Assisted Analysis of Business Presentations." Pp. 277–91 in *Investigating English with corpora. Studies in honour of Maria Teresa Prat*, edited by A. Martelli and V. Pulcini.

Bargiela-Chiappini, F., and S. Harris. 1995. "Towards a Generic Structure of Meetings in British and Italian Management." *Text* 15(4):531–60.

Bargiela-Chiappini, F., and S. Harris. 1997. *Managing Language. The Discourse of Corporate Meetings*. Amsterdam: John Benjamins.

Bateman, J. A. 2008. *Multimodality and Genre: A Foundation for the Systematic Analysis of Multimodal Documents*. London: Palgrave Macmillan.

Bateman, J. A. 2014. "Genre in the Age of Multimodality: Some Conceptual Refinements for Practical Analysis." Pp. 237–69 in *Evolution in Genre Emergence, Variation, Multimodality*, edited by P. Evangelisti Allori, J. Bateman, and V. Bhatia. Bern: Peter Lang.

Bavelas, J., N. Chovil, L. Coates, and L. Roe. 1995. "Gestures Specialized for Dialogue." *Personality and social phychology bulletin* 21(4):394–405.

Bazerman, C. 1988. *Shaping Written Knowledge: The Genre and Activity of the Experimental Article in Science*. Madison, Wisconsin: University of Wisconsin Press.

Bazerman, C. 1994. "Systems of Genres and the Enactment of Social Intentions." Pp. 79–101 in *Genre and the new rhetoric*, edited by A. Freedman and P. Medway. London: Taylor and Francis.

Bazerman, C. 2005. "Emerging Perspectives on the Many Dimensions of Scientific Discourse." Pp. 15–30 in *Reading Science: Critical and Functional Perspectives on Discourses of Science*, edited by J.R. Martin and R. Veel. Routledge.

Beattie, G., and H. Shovelton. 1999. "Do Iconic Hand Gestures Really Contribute Anything to the Semantic Information Conveyed by Speech ? An Experimental Investigation." *Semiotica* 123(1/2):1–30.

De Beaugrande, R., and W. Dressler. 1981. *Introduction to Text Linguistics*. London: Longman.

Bennett, K. 2010. "Academic Discourse in Portugal: A Whole Different Ballgame?" *Journal of English for Academic Purposes* 9(1):21–32. Retrieved 3/2/2017 (http://www.sciencedirect.com/science/article/pii/S1475158509000666).

Berkenkotter, C., and T. Huckin. 1995. *Genre Knowledge in Disciplinary Communication.* N.J.: Erlbaum.

Bezemer, J. 2014. "How to Transcribe Multimodal Interaction? A Case Study." Pp. 155–69 in *Texts, Images and Interaction: A Reader in Multimodality*, edited by C.D. Maier and S. Norris. Mouton de Gruyter.

Bhatia, V. 1993. *Language Use in Professional Settings*. London: Longman.

Bhatia, V. 2002. "Applied Genre Analysis : A Multi-Perspective Model." *Ibérica* 4: 3–19.

Bhatia, V. 2005. "Generic Patterns in Promotional Discourse." Pp. 213–28 in *Persuasion across genres*, edited by H. Halmari and T. Virtanen. John Benjamins.

Bhatia, V. 2010. "Interdiscursivity in Professional Communication." *Discourse and Communication* 21(1):32–50.

Bhatia, V. 2012. "Critical Reflections on Genre Analysis." *Ibérica* 24: 17–28.

Biber, D., S. Conrad, and R. Reppen. 1998. *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge University Press.

Birdwhistell, R. 1970. *Kinesics and Context*. Philadelphia: University of Pennsylvania Press.

Birdwhistell, R. 1952. *Introduction to Kinesics: An Annotation System for Analysis of Body Motion and Gesture*. Washington, DC: Department of State, Foreign Service Institute.

Boden, D. 1994. *The Business of Talk. Organisations in Action.* London: Polity Press.

Bolinger, D. 1986. *Intonation and Its Parts: Melody in Spoken English*. London: Stanford University Press.

Bowker, J. 2006. "Referential and and Affective Force in Oral Business Presentations: The Role of Narration." Pp. 58–71 in *Managing interaction in professional discourse: Intercultural and interdiscoursal perspectives*, edited by J. Bamford and M. Bondi. Rome: Officina.

Bowker, J. 2007. "Uncovering Intentionality and the Process of Persuasion in Oral Business Narratives." Pp. 47–87 in *Business Discourse: Language at work*, edited by J. Bamford and R. Salvi. Rome: Aracne Editrice.

Bowring, M., R. Carter, A. Goddard, D. Reah, and K. Sanger. 2013. *Working with Texts: A Core Introduction to Language Analysis*. Michigan: Taylor & Francis.

Brazil, D. 1985. *The Communicative Value of Intonation in English*. 1st ed. Birmingham: The University Printers, University of Birmingham.

Brazil, D. 1997. *The Communicative Value of Intonation in English*. 2nd ed. Cambridge: Cambridge University Press.

Briñol, P., and R. E. Petty. 2003. "Overt Head Movements and Persuasion: A Self-Validation Analysis." *Journal of Personality and Social Psychology* 84(6):1123–1139.

Brosius, H., and A. Bathelt. 1994. "The Utility of Exemplars in Persuasive Communication." *Communication Research* 21(1):48–78.

Brown, G. 1977. *Listening to Spoken English*. London: Longman.

Brown, G., and G. Yule. 1983. *Discourse Analysis*. Cambridge University Press.

Brown, P., and S. C. Levinson. 1987. *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.

Bull, P., and G. Conelly. 1985. “Body Movement and Emphasis in Speech.” *Journal of Non Verbal Behaviour* 9(3):169–87.

Bull, P. 1986. “The Use of Hand Gesture in Political Speeches: Some Case Studies.” *Journal of Language and Social Psychology* 5(2):103–18.

Burgoon, J., T. Birk, and M. Pfau. 1990. “Nonverbal Behaviours, Persuasion and Credibility.” *Human Communication Research* 17(1):140–69.

Burzynski, J. 2013. “Elements of a Successful Product Pitch.” Retrieved 3/2/2016 (http://38pitches.com/elements-of-a-successful-product-pitch/).

Busà, M.G. 2010. “Sounding Natural : Improving Oral Presentation Skills.” *Language Value* 2(1):51–67.

Butterworth, B., and G. Beattie. 1978. “Gesture and Silence as Indicators of Planning in Speech.” Pp. 347–60 in *Recent Advances in the Psychology of Language*, edited by R. N. Campbell and Phillip T. Smith. New York: Plenum.

Caliendo, G. 2014. “The Popularisation of Science in Web-Based Genres.” Pp. 101–32 in *The language of popularisation: Theoretical and Descriptive models*, edited by G. Caliendo and G. Bongo. Bern: Peter Lang.

Caliendo, G., and A. Compagnone. 2014. “Expressing Epistemic Stance in University Lectures and Ted Talks : A Contrastive Corpus-Based Analysis.” *Lingue e Linguaggi* 11: 105–22.

Campbell, K. S., D. L. Mothersbaugh, C. Brammer, and T. Taylor. 2001. “Peer versus Self Assessment of Oral Business Presentation Performance.” *Business Communication Quarterly* 64: 23–40.

Carlson, R., and B. Granström. 1997. “Speech Synthesis.” Pp. 768–88 in *The handbook of phonetic sciences*. Oxford: Blackwell.

Carter, R. 1997. *Working With Texts: A Core Book for Language Analysis*. London: Routledge.

Carter-Thomas, S., and E. Rowley-Jolivet. 2001. “Syntactic Differences in Oral and Written Scientific Discourse: The Role of Information Structure.” *ASp* 31–33: 19–37.

Carter-Thomas, S., and E. Rowley-Jolivet. 2003. “Analysing the Scientific Conference Presentation (CP): A Methodological Overview of a Multimodal Genre.” *ASp: La Revue* *du GERAS* 40(39–40):59–72.

Cauldwell, R. 2002. *Streaming Speech: Listening and Pronunciation for Advanced Learners of English*. Birmingham: Speechinaction.

Cauldwell, R. 2015. “Discourse Intonation.” Retrieved 3/2/2017 (http://www.speechinaction.org/forteachers/discourse-intonation/).

Celce-Murcia, M., and E. Olshtain. 2000. *Discourse and Context in Language Teaching: A Guide for Language Teachers*. Cambridge: Cambridge University Press.

Chaiken, S., and A. Eagly. 1976. “Communication Modality as a Determinant of Message Persuasiveness and Message Comprehensibility.” *Journal of Personality and Social Psychology* 34(4):605–14.

Chaiken, S., and A. Eagly. 1983. “Communication Modality as a Determinant of Persuasion: The Role of Communicator Salience.” *Journal of Personality and Social Psychology* 45(2):241–56.

Chebat, C., J. Chebat, and A. Vaninski. 1996. “Voice and Advertising: Effects of Intonation and Intensity of Voice on Source Credibility, Attitudes toward the Advertised Service and the Intent to Buy.” *Perceptual and Motor Skills* 83: 243–62.

Cheng, W. 2004. “Did You Took from the Minibar? What Is the Practical Relevance of a Corpus Driven Language Study to Practitioners in Hong-Kong’s Hotel Industry?” Pp. 141–66 in *Discourse in the professions*, edited by U. Connor and T. Upton. Amsterdam, Philadelphia: John Benjamins.

Chesterman, A. 1998. *Contrastive functional analysis*. Amsterdam: John Benjamins.

Chu, Y., W. Strong, J. Ma, and W. Greene. 2005. “Silent Messages in Negotiations: The Role of Nonverbal Communication in Cross-Cultural Negotiations.” *Journal of Organisational Culture, Communications and Conflict* 9(2):113–29.

Chun, D. M. 2002. *Discourse Intonation in L2: From Theory and Research to Practice*. J. Benjamins.

Cialdini, R. B. 1984. *Influence: How and Why People Agree to Things*. William Morrow & Co.

Cialdini, R. B. 2001. “The Science of Persuasion.” *Scientific American* 284: 76–81.

Cialdini, R. B. 2006. *Influence: The Psychology of Persuasion*. HarperBusiness.

Compagnone, A. 2014. “Knowledge Dissemination and Environmentalism: Exploring the Language of TED Talks.” Pp. 7–25 in *Popularisation and the Media.*, edited by E. Chiavetta, S. Sciarrino, and C. Williams. Bari: Edipuglia.

Compagnone, A. 2015. “The Reconceptualization of Academic Discourse as a Professional Prac- Tice in the Digital Age : A Critical Genre Analysis of TED Talks.” *Journal of Language and Communication in Business* (54):49–69.

Constantinou, O. 2005. “Multimodal Discourse Analysis: Media, Modes and Technologies.” *Journal of Sociolinguistics* 9(4):602–18.

Cook, G. 2001. *The Discourse of Advertising*. Routledge.

Coupland, N. 1983. “Patterns of Encounter Management: Further Arguments for Discourse Variables.” *Language in Society* 12: 459–76.

Cowan, K. 2013. “Multimodal Composing in Classrooms-Learning and Teaching for the Digital World 2012 Eds Miller, SM, and McVee MB.” *Writing and Pedagogy* 5(2):375–79.

Crawford Camiciottoli, B. 2007. *The Language of Business Studies Lectures: A Corpus-Assisted Analysis*. John Benjamins.

Crawford Camiciottoli, B. 2011. “Ethics and Ethos in Financial Reporting: Analyzing Persuasive Language in Earnings Calls.” *Business Communication Quarterly* 74: 298–312.

Crawford Camiciottoli, B. 2015. “Elaborating Explanations During OpenCourseWare Humanities Lectures: The Interplay of Verbal and Nonverbal Strategies.” Pp. 144–70 in *Multimodal Analysis in Academic Settings*, edited by B. Crawford Camiciottoli and I. Fortanet-Gómez . London: Routledge.

Crystal, D. 1969. *Prosodic Systems and Intonation in English*. Cambridge: Cambridge University Press.

Cutler, A. 1983. “Speakers’ Conceptions of the Function of Prosody.” Pp. 79–91 in *Prosody: models and measurements*, edited by A. Cutler and D. R. Ladd.

Dafouz-Milne, E. 2008. “The Pragmatic Role of Textual and Interpersonal Metadiscourse Markers in the Construction and Attainment of Persuasion: A Cross-Linguistic Study of Newspaper Discourse.” *Journal of Pragmatics* 40(1):95–113.

Dalton, C., and B. Seidlhofer. 1994. *Pronunciation*. Oxford: Oxford University Press.

Diani, G. 2015. “Visual Communication in Applied Linguistics Conference Presentations.” Pp. 83–107 in *Multimodal Analysis in Academic Settings: From Research to Teaching*, edited by B. Crawford Camiciottoli and I. Fortanet-Gómez. London: Routledge.

Dlugan, A. 2010. “Ethos, Pathos, Logos: 3 Pillars Of Public Speaking.” Retrieved 3/2/2017 (*Http://Sixminutes. Dlugan.Com/Ethos-Pathos-Logos/*).

Duarte, N. 2015. “Great Presenters Find the Perfect Mix of Data and Narrative.” Retrieved 3/02/2017 (http://totalmgmtco.com/great-presenters-find-a-perfect-mix-of-data-narrative/).

Dubois, B. L. 1982. “And the Last Slide Please: Regulatory Function at Biomedical Meetings.” *World Language English* 1(2):263–268.

Dubois, B. L. 1987. “Something on the Order of around Forty to Forty Four: Imprecise Numerical Expressions in Biomedical Slide Talks.” *Language and Society* 16: 527–41.

Dudley-Evans, T. 1994. “Variations in the Discourse Patterns Favoured by Different Disciplines and Their Pedagogical Implications.” Pp. 146–58 in *Academic Listening: research perspectives*. Cambridge: Cambridge University Press.

Dudley-Evans, T., and T. Johns. 1981. “A Team Teaching Approach to Lecture Comprehension for Overseas Students.” Pp. 30–46 in *The Teaching of Listening Comprehension*. London: The British Council.

Duszak, A. 1997. *Culture and Styles of Academic Discourse*. Berlin: Mouton.

Efron, D. 1941. *Gesture and Environment*. Morningside Heights: King’s Crow Press.

Ekman, P., and W. Friesen. 1969. “The Repertoire of Nonverbal Behavioral Categories: Origins, Usage and Coding.” *Semiotica* 1: 49–98.

English, S. L. 1985. “Kinesics in Academic Lectures.” *The ESP Journal* 4: 161–70.

Fairclough, N. 1993. “Critical Discourse Analysis and the Marketisation of Public Discourse: The Universities.” *Discourse & Society* 4(2):133–168.

Fasold, R. 1990. *The Sociolinguistics of Language*. Oxford: Blackwell.

Fernández Polo, F. J. 2014. “Native and Non-Native Speaker Interpersonal Skills at Conferences: Managing Self-Mentions and Humour.” Pp. 163–78 in *Occupying Niches: Interculturality, Cross-culturality and Aculturality in Academic Research*, edited by A.

Lyda and K. Warchal. Heidelberg: Springer.

Filliettaz, L. 2004. "The Multimodal Negotiation of Service Encounters." Pp. 88–100 in *Discourse and technology. Multimodal Discourse Analysis*, edited by P. Levine and R. Scolon. Washinton: Georgetown University Press.

Firth, A. 1995. *The Discourse of Negotiation: Studies of Language in the Workplace*. Oxford: Pergamon.

Firth, J. R. 1987. *Papers in Linguistics 1934-1951*. London: Oxford University Press.

Fitch, K. L. 2003. "Cultural Persuadables." *Communication Theory* 13(1):100–123.

Flewitt, R. 2014. "Interviews." Pp. 136–53 in *Understanding Research with Children and Young People*, edited by A. Clark, R. Flewitt, M. Hammersley, and R. Martin. London: Sage Publications.

Flowerdew, L. 2003. "A Combined Corpus and Systemic-Functional Analysis of the Problem-Solution Patter in a Student and Professional Corpus of Technical Writing." *TESOL Quarterly* 37(3):489–511.

Flowerdew, L. 2004. "The Argument for Using English Specialised Corpora to Understand Academic and Professional Language." Pp. 11–33 in *Discourse in the Professions: Perspectives from Corpus Linguistics*, edited by U. Connor and T. A. Upton. Amsterdam: John Benjamins.

Fortanet-Gómez, I. 2005. "Honoris Causa Speeches: An Approach to Structure." *Discourse Studies* 7(1):35–51.

Freedman, A., and P. Medway, eds. 1994. *Genre and the New Rhetoric*. London: Taylor and Francis.

Frobert-Adamo, M. 2002. "Humour in Oral Presentations: What's the Joke?" Pp. 211–225 in *The language of conferencing*, edited by E. Ventola, C. Shalom, and S. Thompson. Frankfurt: Peter Lang.

Fuertes-Olivera, P. A., M. Velasco-Sacristán, A. Arribas-Baño, and E. Samaniego-Fernández. 2001. "Persuasion and Advertising English: Metadiscourse in Slogans and Headlines." *Journal of Pragmatics* 33(8):1291–1307.

Gili, G. 2013. "Expertise, Justice, Reciprocity: The Three Roots of Teachers‟ Credibility." *Italian Journal of Sociology of Education* 5(1):1–18.

Goffman, E. 1979. "Footing." *Semiotica* 25(1/2):1–29.

Goffman, E. 1981. *Forms of Talk*. University of Pennsylvania Press, Incorporated.

Goldin-Meadow, S. 2003. *Hearing Gesture: How Our Hands Help Us Think*. Belknap Press of Harvard University Press.

Goldman, R., F. Erickson, J. Lemke, and S. Derry. 2007. "Selection in Video." Pp. 19–27 in *Guidelines For Video Research In Education: Recommendations From An Expert Panel*, edited by S. Derry. Data Research and Development Center (NORC at the University of Chicago).

Goodwin, C. 1986. "Gesture as a Resource for the Organisation of Mutual Orientation." *Semiotica* 1/2: 22–49.

Goodwin, M., and C. Goodwin. 2000. "Emotion within Situated Activity." Pp. 33–53 in *Communication: an arena of development*, edited by N. Budwig, I. Uzgiris, and J. Wertsch. Stamford.

Green, G. 1989. *Pragmatics and Natural Language Understanding*. Hillsdale: Lawrence Erlbaum Associates.

Grice, H. 1957. "Meaning." *Philosophical Review* 66: 377–88.

Grice, H. 1975. "Logic and Conversation." Pp. 41–58 in *Syntax and Semantics*, edited by P. Cole and J. Morgan. New York: Academic Press.

Habermas, J. 1981. *Theorie Des Kommunikativen Handeln: Bd1 Handelungsrationalität Und Gesellschaftliche Rationalisierung [The Theory of Communicative Action: Vol. 1 Reason and the Rationalization of Society]*. Frankfurt: Suhrkamp.

Hadar, U. 1989. "Gestural Modulation of Speech Production: The Role of Head Movement." *Language & Communication* 9(4):245–57.

Hadar, U., T. J. Steiner, E. C. Grant, and F. C. Rose. 1984. "The Timing of Shifts of Head Postures during Conservation." *Human Movement Science* 3(3):237–45.

Hadar, U., T. J. Steiner, E. C. Grant, and F. Clifford Rose. 1983. "Kinematics of Head Movements Accompanying Speech during Conversation." *Human Movement Science* 2(1–2):35–46.

Hahn, L. D. 2004. "Primary Stress and Intelligibility: Research to Motivate the Teaching of Suprasegmentals." *TESOL Quarterly* 38: 201–23.

Hall, E., and G. Trager. 1953. *The Analysis of Culture*. Washington D.C.: American Council of Learned Societies.

Halliday, M. 1978. *Language as a Social Semiotic*. London: Arnold.

Halliday, M. 1985. *An Introduction to Functional Grammar*. London: Arnold.

Halliday, M. 1994. *An Introduction to Functional Grammar*. 2nd ed. London: Arnold.

Halliday, M., and S. Greaves. 2008. *Intonation in the Grammar of English*. London: Equinox.

Halmari, H., and T. Virtanen. 2005. *Persuasion Across Genres: A Linguistic Approach*. John Benjamins Pub.

Harklau, L. 2005. “Ethnography and Ethnographic Research on Second Language Teaching and Learning.” Pp. 179–194 in *Handbook of research in second language teaching and learning*, edited by E. Hinkel. Lawrence Erlbaum, Mahwah.

Hendon, D., R. Hendon, and P. Herbig. 1996. *Cross-Cultural Business Negotiations*. Westport: Praeger Publishers.

Hendon, D. W., R. A. Hendon, and P. A. Herbig. 1996. *Cross-Cultural Business Negotiations*. Quorum.

Hewings, M. 1998. “Intonation Choices in the English of Non-Native Speakers.” Pp. 317–31 in *Linguistic choice across genres*, edited by A. Sánchez and R. Carter. Amsterdam: J. Benjamins.

Hincks, R. 2005. “Measures and Perceptions of Liveliness in Student Oral Presentation Speech: A Proposal for an Automatic Feedback Mechanism.” *System* 33(4):575–91. Retrieved 3/2/2017 (http://www.sciencedirect.com/science/article/pii/S0346251X05000679).

Hincks, R., and J. Edlund. 2009. “Promoting Increased Pitch Variation in Oral Presentations with Transient Visual Feedback.” *Language Learning & Technology* 13(3):32–50.

Hirano, E. 2009. “Research Article Introductions in English for Specific Purposes: A Comparison between Brazilian Portuguese and English.” *English for Specific Purposes* 28(4):240–50. Retrieved 3/2/2017 (http://www.sciencedirect.com/science/article/pii/S0889490609000131).

Holler, J. 2010. “Speakers’ Use of Interactive Gestures as Markers of Common Ground.” *Gesture in Embodied Communication and Human-Computer Interaction* 5934: 11–22.

Hood, S., and G. Forey. 2005. “Introducing a Conference Paper: Getting Interpersonal with Your Audience.” *Journal of English for Academic Purposes* 4(4):291–306. Retrieved 3/2/2017 (http://www.sciencedirect.com/science/article/pii/S1475158505000329).

Hornby, P. 1974. “Surface Structure and Presupposition.” *Journal of Verbal Learning and Verbal Behavior* 13(5):530–38.

Hovland, C., I. Janis, and H. Kelley. 1953. *Communication and Persuasion: Psychological Studies of Opinion Change*. New Haven: Yale University Press.

Hunston, S. 2000. “Evaluation and the Planes of Discourse: Status and Value in Persuasive Texts.” in *Evaluation in text*, edited by S. Hunston and G. Thompson. Oxford: Oxford University Press.

Hurley, D. 1992. “Issues in Teaching Pragmatics.” *Applied Linguistics* 13(3):259–79.

Hüttner, J. 2005. “Extended Genre Analysis: Exploring Student Academic Writing.” Vienna University.

Hyland, K. 1995. “The Author in the Text: Hedging Scientific Writing.” *Hong Kong Papers in Linguistics and Language Teaching* 18: 33–42.

Hyland, K. 1997. “Scientific Claims and Community Values: Articulating an Academic Culture.” *Language and Communication* 17(1):19–32.

Hyland, K. 1998a. “Exploring Corporate Rhetoric: Metadiscourse in the CEO’s Letter.” *Journal of Business Communication* 35(2):224–44.

Hyland, K. 1998b. “Persuasion and Context: The Pragmatics of Academic Metadiscourse.” *Journal of Pragmatics* 30: 437–55.

Hyland, K. 2000. *Disciplinary Discourses: Social Interaction in Academic Writing*. London: Longman.

Hyland, K. 2004. *Disciplinary Discourses*. Michigan: University of Michigan Press.

Hyland, K. 2005. “A Convincing Argument: Corpus Analysis and Academic Persuasion.” Pp. 87–114 in *Discourse in the Professions: Perspectives from Corpus Linguistics.*, edited by U. Connor and T. Upton. Amsterdam: Benjamins.

Hyland, K. 2006. *Academic Discourse across Disciplines*. Bern: Peter Lang.

Hyland, K. 2009. *Academic Discourse*. London: Continuum.

Hymes, D. 1972. “Models of Interaction of Language and Social Life.” Pp. 35–71 in *Directions in sociolinguistics.The Ethnography of communication*, edited by J. Gumperz and D. Hymes. New York: Rinehart and Winston Inc.

Iedema, R. 2003. “Multimodality, Resemiotization: Extending the Analysis of Discourse as Multi-Semiotic Practice.” *Visual Communication* 2(1):29–57.

Jewitt, C. 2004. “Multimodality and New Communication Technology.” Pp. 184–95 in *Discourse and technology: Multimodal Discourse Analysis*. Washington D.C.: Georgetown University Press.

Jewitt, C. 2012. *An Introduction to Using Video for Research An Introduction to Using Video for Research*.

Jewitt, C. 2014. *The Routledge Handbook of Multimodal Analysis*. Routledge.

Jucker, A. 1997. “Persuasion by Inference. Analysis of a Party Political Broadcast.” Pp. 121–37 in *Political Linguistics*, edited by J. Blommaert and C. Bulcaen. John Benjamins.

Kang, O. 2010. “Relative Salience of Suprasegmental Features on Judgments of L2 Comprehensibility and Accentedness.” *System* 38(2):301–15.

Kast, B. 2008. “iKeynote-Representation, Rhetoric and Visual Communication by Steve Jobs in His Keynote Ar Macworld 2007.” University of Salzburg.

Kellerman, S. 1992. “‘I See What You Mean’: The Role of Kinesic Behaviour in Listening and Implications for Foreign Language Learning.” *Applied Linguistics* 13(3):239–58.

Kendon, A. 1967. “Some Functions of Gaze Direction in Social Interaction.” *Acta Psychologica* 26: 22–63.

Kendon, A. 1972. “Some Relations between Body Motion and Speech.” Pp. 177–210 in *Studies in dyadic communication*, edited by A. Siegman and B. Pope. New York: Pergamon.

Kendon, A. 1980. “Gesticulation and Speech: Two Aspects of the Process of Utterance.” Pp. 207–27 in *The relationship of verbal and non-verbal communication*, edited by M. Key. The Hague: Mouton.

Kendon, A. 2002. “Some Uses of the Head Shake.” *Gesture* 2(2):147–82.

Kendon, A. 2004. *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.

Kipp, M. 2014. “ANVIL: The Video Annotation Research Tool.” Pp. 420–36 in *Handbook of Corpus Phonology*, edited by U. Durand and G. Kristoffersen. Oxford: Oxford University Press.

Krauss, R., Y. Chen, and R. Gottesman. 2000. “Lexical Gestures and Lexical Access: A Process Model.” Pp. 261–283 in *Language and gesture*, edited by D. McNeill. Cambridge: Cambridge University Press.

Kress, G. 1993. “Against Arbitrariness: The Social Production of the Sign as Foundational Issue in Critical Discourse Analysis.” *Discourse and Society* 4(2):169–93.

Kress, G. 2003. *Literacy in the New Media Age*. London: Routledge.

Kress, G. 2009a. *Multimodality. A Social Semiotic Approach to Contemporary Communication*. 1st ed. Taylor and Francis.

Kress, G. 2009b. “What Is a Mode.” Pp. 61–75 in *The Routledge Handbook of Multimodal Analysis*. London: Routledge.

Kress, G. 2010. *Multimodality: A Social Semiotic Approach to Contemporary Communication.* 2nd ed. London: Taylor and Francis.

Kress, G., and T. Van Leeuwen. 1996. *Reading Images-A Grammar of Visual Design*. 1st ed. London: Routledge.

Kress, G., and T. Van Leeuwen. 2001. *Multimodal Discourse: The Modes and Media of Contemporary Communication*. London: Arnold.

Kress, G., and T. Van Leeuwen. 2006. *Reading Images A Grammar of Visual Design*. 2nd ed. London: Routledge.

Kwasnik, B., and K. Crowston. 2005. “Genre of Digital Documents.” *Information, Technology & People* 18(2):76–88.

Ladd, D., and A. Cutler, eds. 1983. *Prosody: Models and Measurements*. Berlin: Springer.

Lakoff, R. 1973. “The Logic of Politeness: Or Minding Your P’s and Q’s.” Pp. 292–305 in *Papers from the 9th regional meeting of the Chicago Linguistic Society*, edited by C. Corum, T. Smith-Stark, and A. Weiser. Chicago: Chicago Linguistic Society.

Lakoff, R. 1982. “Persuasive Discourse and Ordinary Conversation, with Examples from Advertising.” Pp. 25–42 in *Analysing discourse: text and talk*, edited by D. Tannen. Washinton: Georgetown University Press.

Van Leeuwen, T. 2004. “Ten Reasons Why Linguists Should Pay Attention to Visual Communication.” Pp. 7–19 in *Discourse and technology. Multimodal Discourse Analysis*, edited by P. Levine and R. Scollon. Washington D.C.: Georgetown University Press.

Lemke, J. 2005. “Multiplying Meaning: Verbal and Visual Semiotics in Scientific Texts.” Pp. 87–114 in *Reading Science: Critical and functional perspectives on discourses of science*, edited by J. R. Martin and R. Veel. London & New York: Routledge.

Lemke, J. 2009. “Teaching All the Languages of Science : Words , Symbols , Images , and Actions.” Pp. 1–13 in *Conference on science education in Barcelona*.

Lepers, C. 2015. “How To Pitch an Idea in 7 Steps.” Retrieved 3/2/2017 (http://www.pptpop.com/how-to-pitch-an-idea/).

Levin, I., S. L. Schneider, and G. J. Gaeth. 1998. “All Frames Are Not Created Equal: A Typology and Critical Analysis of Framing Effects.” *Organisational behaviour and Human Decision Processes* 76: 149–88.

Levis, J., and L. Pickering. 2004. “Teaching Intonation in Discourse Using Speech Visualization Technology.” *System* 32: 505–24.

Lichinski, A. 2008. “Examples as Persuasive Arguments in Popular Management Literature.” *Discourse and Communication* 2(3):243–69.

Loi, C. K. 2010. “Research Article Introductions in Chinese and English: A Comparative Genre-Based Study.” *Journal of English for Academic Purposes* 9(4):267–79. Retrieved 3/2/2017 (http://www.sciencedirect.com/science/article/pii/S1475158510000573).

Longacre, R. 1976. *An Anatomy of Speech Notions*. Lisse: Peter de Ridder Press.

Luzón, M. J. 2005. “Aplicación Del Concepto de ‘ Colonia de Géneros ’ a La Enseñanza de Inglés Para Fines Específicos.” *Iberica* 10: 133–44.

Malinowski, B. 1923. “The Problem of Meaning in Primitive Languages.” Pp. 146–52 in *The Meaning of Meaning*, edited by C. K. Ogden and Ian A. Richards. London: Routledge.

Martin, J. 2000. “Beyond Exchange. Appraisal Systems in English.” Pp. 142–75 in *Evaluation in text: authorial stance and the construction of discourse*, edited by S. Hunston and G. Thompson. Oxford: Oxford University Press.

Martin, J., and D. Rose. 2003. *Working with Discourse. Meaning beyond the Clause*. London

& New York: Continuum.

Martin, J., and P. White. 2005. *The Language of Evaluation: Appraisal in English*. London: Palgrave Macmillan.

Martín, P. 2003. "A Genre Analysis of English and Spanish Research Paper Abstracts in Experimental Social Sciences." *English for Specific Purposes* 22(1):25–43. Retrieved 3/2/2017 (http://www.sciencedirect.com/science/article/pii/S0889490601000333).

Martinec, R. 2001. "Interpersonal Resources in Action." *Semiotica* 1(4):117–45.

Matthiessen, C. 2007. "The Multimodal Page: A Systemic Functional Exploration." Pp. 1–14 in *New directions in the analysis of multimodal discourse*, edited by T. Royce and W. Bowcher. Mahwah, NJ: Lawrence Erlbaum Associates.

Mauranen, A. 1993. *Cultural Differences in Academic Rhetoric: A Textlinguistic Study*. Frankfurt: Peter Lang.

Mavers, D. 2011. *Children's Drawing and Writing: The Remarkable in the Unremarkable*. New York: Routledge.

May, G. L. 2014. *Persuasive Business Presentations*. New York: Business Expert Press, LLC.

Mayberry, R., and J. Jaques. 2000. "Gesture Production during Stuttered Speech: Insights into the Nature of Gesture-Speech Integration." Pp. 199–214 in *Language and Gesture*, edited by D. McNeill. Cambridge: Cambridge University Press.

Maynard, S. K. 1989. *Japanese Conversation: Self-Contextualization through Structureand Interactional Management*. Norwood, NJ: Ablex Publishing.

McCarthy, M. 1998. *Spoken Language and Applied Linguistics*. Cambridge: Cambridge University Press.

McCarthy, M., and M. Handford. 2004. "'Invisible to Us': A Preliminary Corpus-Based Study of Spoken Business English." Pp. 167–201 in *Discourse in the Professions: Perspectives from corpus linguistics*, edited by U. Connor and T. Upton. John Benjamins.

McClave, E. 2000. "Linguistic Functions of Head Movements in the Context of Speech." *Journal of Pragmatics* 32: 855–78.

McGregor, A., B. Zielinski, C. Meyers, and M. Reed. 2016. "An Exploration of Teaching Intonation Using a TED Talk." Pp. 143–49 in *Proceedings of the 7th Pronunciation in*

Second Language Learning and Teaching Conference, edited by J. Levis, H. Le, I. Lucic, E. Simpson, and S. Vo. Iowa State University.

McNeill, D. 1985. "So You Think Gestures Are Non-Verbal?" *Psychological Review* 96: 350–71.

McNeill, D. 1992. *Hand and Mind*. University of Chicago Press.

McNeill, D., F. Quek, K. McCullough, S. Duncan, N. Furuyama, R. Bryll, X. Ma and R. Ansari. 2001. "Catchments, Prosody and Discourse." *Gesture* 1(1):9–33.

McNeill, D. 2005. *Gesture and Thought*. Chicago: University of Chicago Press.

Messaris, P. 1996. *Visual Persuasion: The Role of Images in Advertisements*. Thousand Oaks: Sage Publications.

Miller, C. 1984. "Genre as Social Action." *Quarterly Journal of Speech* 70: 151–67.

Mittelberg, I. 2007. "Methodology for Multimodality: One Way of Working with Speech and Gesture Data." Pp. 225–48 in *Methods in Cognitive Linguistics*, edited by M. Gonzalez-Marquez, I. Mittelberg, S. Coulson, and M. J. Spivey. Amsterdam, Philadelphia: John Benjamins.

Morell, T. 2014. "Communicating Research at International Conferences: A Multimodal Analysis of an Intercultural or a Disciplinary Specific Genre?" Pp. 153–62 in *Occupying Niches: Interculturality, Cross-culturality and Aculturality in Academic Research*, edited by A. Lyda and K. Warchal. Heidelberg: Springer.

Morell, T. 2015. "International Conference Paper Presentations: A Multimodal Analysis to Determine Effectiveness." *English for Specific Purposes* 37: 137–50. Retrieved 3/2/2017 (http://dx.doi.org/10.1016/j.esp.2014.10.002).

Morrel-Samuels, P., and R. M. Krauss. 1992. "Word Familiarity Predicts Temporal Asynchrony of Hand Gestures Amd Speech." *Journal of Experimental Psychology: Learning, Memory and Cognition* 18: 615–23.

Morris, C. 1938. *Foundations of the Theory of Signs*. Chicago: University of Chicago Press.

Morris, D. 1977. *Manwatching: A Field Guide to Human Behaviour*. London: Jonathan Cape Ltd.

Morrison, J., and D. Vogel. 1998. "Impacts of Presentations Visuals on Persuasion." *Information and Management* 33: 124–35.

Morton, Janne. 2009. “Genre and Disciplinary Competence : A Case Study of Contextualisation in an Academic Speech Genre.” *English for Specific Purposes* 28(4):217–29. Retrieved 3/2/2017 (http://dx.doi.org/10.1016/j.esp.2009.04.005).

Muntigl, P. 2004. “Modelling Multiple Semiotic Systems: The Case of Gesture and Speech.” Pp. 31–48 in *Perspectives on multimodality*, edited by E. Ventola, C. Charles, and M. Kaltenbacher. Amsterdam, Philadelphia: John Benjamins.

Myers, G. 1989. “The Pragmatics of Politeness in Scientific Articles.” *Applied Linguistics* 10(1):1–35. Retrieved 3/2/2017 (http://applij.oxfordjournals.org/content/10/1/1.abstract).

Nicoladis, E., R. Mayberry, and F. Genesee. 1999. “Gesture and Early Bilingual Development.” *Developmental Psychology* 35(2):514–26.

Niebuhr, O., J. Voße, and A. Brem. 2016. “What Makes a Charismatic Speaker? A Computer-Based Acoustic-Prosodic Analysis of Steve Jobs Tone of Voice.” *Computers in Human Behavior* 64: 366–82.

Norris, S. 2004a. *Analysing Multimodal Interaction: A Methodological Framework*. London & New York: Continuum.

Norris, S. 2004b. “Multimodal Discourse Analysis: A Conceptual Framework.” Pp. 101–15 in *Discourse and Technology. Multimodal Discourse Analysis*, edited by P. Levine and R. Scollon. Washington D.C.: Georgetown University Press.

Norris, S. 2015. *Relevance of Multimodal (Inter)action Analysis to Linguistics*. Castellón.

Noteboom, S. 1997. “The Prosody of Speech: Melody and Rhythm.” Pp. 640–73 in *The handbook of phonetic sciences*, edited by W. Hardcastle and J. Laver. Oxford: Blackwell.

NT&T, 2011. "Towards a linguistics of news production". *Journal of Pragmatics* 43: 1843-1852)

O’Connor, J., and G. Arnold. 1973. *Intonation of Colloquial Engish*. Bristol: Longman Group Ltd.

O’Halloran, K., Sabine T., and B. A. Smith. 2016. “Multimodal Approaches to English for Academic Purposes.” Pp. 256–69 in *The Routledge Handbook of English for Academic Purposes.*, edited by K. Hyland and P. Shaw. London & New York: Routledge.

O’Keefe, D. 2002. *Persuasion: Theory and Research*. London: Sage Publications.

Ochs, E., P. Gonzales, and S. Jacoby. 1994. “Interpretative Journeys: How Physicists Talk and Travel through Graphic Space." *Configurations* 1:151–171.

Ogden, C. K., and I. A. Richards, eds. 1923. *The Meaning of Meaning*. London: Routledge.

Ozyurek, A., and S. Kita. 1999. "Expressing Manner and Path in English and Turkish: Differences in Speech, Gesture, and Conceptualization." Pp. 507–12 in *Proceedings of the Twenty-first Annual Conference of the Cognitive Science Society*, edited by M. Hahn and S. C. Stoness. London: Erlbaum.

Palmer-Silveira, J. C. 2015. "Multimodality in Business Communication: Boody Language as a Visual Aid in Student Presentations." Pp. 171–92 in *Multimodal Analysis in Academic Settings: From Research to Teaching*, edited by B. Crawford Camiciottoli and I. Fortanet-Gómez. London: Routledge.

Paltridge, B. 1995. "Working with Genre " A Pragmatic Perspective." 24: 393–406.

Pan, Y., S. Wong, and R. Scollon. 2002. *Professional Communication in International Settings*. Oxford: Blackwell.

Paré, A., and G. Smart. 1994. "Observing Genres in Action: Towards a Research Methodology." Pp. 146–54 in *Genre and the new rhetoric*, edited by A. Freedman and P. Medway. London, Bristol: Taylor and Francis.

Pennycook, A. 1985. "Actions Speak Louder than Words: Paralanguage, Communciation and Education." *TESOL Quarterly* 19(2):259–82.

Pérez-Llantada, C., Ramón P., and G. R. Ferguson. 2011. "' You Don ' T Say What You Know , Only What You Can ' : The Perceptions and Practices of Senior Spanish Academics Regarding Research Dissemination in English." *English for Specific Purposes* 30: 18–30.

Perloff, R. 2003. *The Dynamics of Persuasion*. London: Lawrence Erlbaum Associates.

Perloff, R., and T. Brock. 1980. "'And Thinking Makes It So': Cognitive Responses to Persuasion." Pp. 67–99 in *Persuasion: New directions in theory and research*, edited by M. Rollof and G. R. Miller. Beverly Hills: Sage Publications.

Petty, R., and J. Cacioppo. 1977. "Forewarning, Cognitive Responding and Resistance to Persuasion." *Journal of Personality and Social Psychology* 35: 645–55.

Petty, R., and J. Cacioppo. 1986. "The Elaboration Likelihood Model of Persuasion." Pp. 123–205 in *Advances in experimental social psychology*, edited by L. Berkowitz. New

York: Academic Press.

Petty, R., and D. Wegener. 1999. "The Elaboration Likelihood Model: Current Status and Controversies." Pp. 41–72 in *Dual process theories in social psychology,* edited by S. Chaiken and Y. Trope. New York: Guilford.

Petty, R., Pablo Briñol, and Z. L. Tormala. 2002. "Thought Confidence as a Determinant of Persuasion: The Self-Validation Hypothesis." *Journal of Personality and Social Psychology* 82(5):722–41. Retrieved 3/2/2017 (http://doi.apa.org/getdoi.cfm?doi=10.1037/0022-3514.82.5.722).

Pickering, L. 2001. "The Role of Tone Choice in Improving ITA Communication in the Classroom." *TESOL Quarterly* 35(2):233–55.

Pickering, L. 2004. "The Structure and Function of Intonational Paragraphs in Native and Nonnative Speaker Instructional Discourse." *English for Specific Purposes* 23: 19–43.

Pickering, Lucy. 2009. "Intonation as a Pragmatic Resource in ELF Interaction." *Intercultural Pragmatics* 6(2):235–55.

Pickering, L., G. Hu, and A. Baker. 2012. "The Pragmatic Function of Intonation: Cueing Agreement and Disagreement in Spoken English Discourse and Implications for ELT." Pp. 199–218 in *Pragmatics and Prosody in English Language Teaching*, edited by J. Romero-Trillo. Dordrecht: Springer Netherlands.

Pirini, J. 2015. "Analysing Business Coaching: Using Modal Density as a Methodological Tool." P. Vol. IV 237-254 in *Multimodality: Critical Concepts in Linguistics*, edited by S. Norris. London: Routledge.

Pirini, J. 2016. *Multimodal (inter)action analysis: Introduction, applications and developments.* Castellón.

Pittam, J., and K. Scheerer. 1993. "Vocal Expression and Communication of Emotion." Pp. 185–97 in *Handbook of emotions*, edited by M. Lewis and J.M. Haviland. New York, London: Guildford Press.

Poggi, I., and C. Pelachaud. 2008. "Persuasion and the Expressivity of Gestures in Humans and Machines." Pp. 45–67 in *Embodied communication in humans and machines*, edited by I. Wachsmuth, M. Lenzen, and G. Knoblich. Oxford: Oxford University Press.

Pomerantz, A. 1986. "Extreme Case Formulations: A Way of Legitimizing Claims." *Human*

Studies 9: 219–29.

Poncini, G. 2002. "Investigating Discourse at Business Meetings with Multicultural Participation." *International Review of Applied Linguistics* 40: 345–73.

Posteguillo, S., J. Palmer, and J. Coll, eds. 1998. *Genre Studies in English for Academic Purposes/ Estudios de Genero En Ingles Para Propositos Academicos*. Digitalia.

Poyatos, F. 1975. "Cross-Cultural Analysis of Paralinguistic 'Alternants' in Face-to-Face Interaction." Pp. 285–314 in *Organisation of behaviour in face-to-face interaction*, edited by A. Kendon, R. Harris, and M. Key. The Hague: Mouton.

Poyatos, F. 1976. *Man beyond Words: Theory and Methodology of Nonverbal Communication*. Oswego: New York State English Council.

Poyatos, F. 1983. "Language and Nonverbal Sign Systems in the Structure of Face to Face Interaction." *Language and Communication* 3(2):129–40.

Pozzer-Ardenghi, L., and W. M. Roth. 2005. "Photographs in Lectures: Gestures as Meaning-Making Resources." *Linguistics and Education* 15(3):275–93.

Punch, M. 2009. *Introduction to Research Methods in Education*. London: Sage Publications.

Querol-Julián, M. 2011. "Evaluation in Discussion Sessions of Conference Paper Presentations: A Multimodal Approach." LAP Lambert Academic Publishing GmbH & Co. K.G., Saarbrücken.

Querol-Julián, M., and I. Fortanet-Gómez. 2012. "Multimodal Evaluation in Academic Discussion Sessions: How Do Presenters Act and React?" *English for Specific Purposes* 31(4):271–83.

Querol-Julián, M., and I. Fortanet-Gómez. 2014. "Evaluation in Discussion Sessions of Conference Presentations: Theoretical Foundations for a Multimodal Analysis." *Kalbotyra* 66: 77–97.

Rackham, N. 1998. "From Experience: Why Bad Things Happen to Good New Products." *Journal of Product Innovation Management* 15: 201–207.

Räisänen, C. 1999. "The Conference Forum as a System of Genres . A Socio-Cultural Study of Academic Conference Practices in Automotive Crash-Safety Engineering." Acta Universitatis Gothoburgensis, Gothenberg.

Räisänen, C. 2002. "The Conference Forum: A System of Interrelated Genres and Discursive

Practices." Pp. 69–93 in *The language of conferencing*, edited by E. Ventola, C. Shalom, and S. Thompson. Frankfurt: Peter Lang.

Räisänen, C. 2015. "There Is More to Multimodality than Discourse Features and Nonverbal Behaviors!" Pp. 133–43 in *Multimodal Analysis in Academic Settings: From Research to Teaching*, edited by B. Crawford Camiciottoli and I. Fortanet-Gómez. London: Routledge.

Räisänen, C., and I. Fortanet-Gómez. 2006. "Do Genres Have Body Language? Nonverbal Communication in Conference Paper Presentations." in *Conference in honour of John Swales*. Michigan: Ann Arbor.

Rendle-Short, J. 2005. "Managing the Transitions between Talk and Silence in the Academic Monologue." *Research on Language and Social Interaction* 38(2):179–218.

Rendle-Short, J. 2006. *The Academic Presentation: Situated Talk in Action*. Ashgate.

Renkema, J. 1993. *Discourse Studies: An Introductory Textbook*. Amsterdam: John Benjamins.

Riseborough, M. G. 1981. "Physiographic Gestures as Decoding Facilitators: Three Experiments Exploring a Neglected Facet of Communication." *Journal of Nonverbal Behavior* 5: 172– 183.

Rosenberg, A., and J. Hirschberg. 2009. "Charisma Perception from Text and Speech." *Speech Communication* 51(7):640–55.

Rowley-Jolivet, E. 1999. "The Pivotal Role of Conference Papers in the Network of Scientific Communication." *ASp* 26(1999).

Rowley-Jolivet, E. 2002. "Visual Discourse in Scientific Conference Papers A Genre-Based Study." *English for Specific Purposes* 21(1):19–40.

Rowley-Jolivet, E., and S. Carter-Thomas. 2005. "The Rhetoric of Conference Presentation Introductions: Context, Argument and Interaction." *International Journal of Applied Linguistics* 15(1):45–70.

Rubin, H. J., and I. S. Rubin. 1995. *Qualitative Interviewing. The Art of Hearing Data*. London: Sage Publications.

Ruiz Garrido, M., J. Palmer, and I. Fortanet-Gómez, eds. 2010. *English for Professional and Academic Purposes*. Amsterdam: Rodolfi.

Ruiz-Garrido, M. 2015. “Intensifying Adverbs in Academic Spoken Discourse: A Contrastive Study between English and Spanish.” Pp. 61–82 in *Multimodal Analysis in Academic Settings:From Research to Teaching*, edited by B. Crawford Camiciottoli and I. Fortanet-Gómez . London: Routledge.

Ruiz-Garrido, M., and I. Fortanet-Gómez. 2008. “El Lenguaje de Las Comunicaciones.” Pp. 125–57 in *Hablar inglés en la universidad: Docencia e Investigación*, edited by I. Fortanet-Gómez. Oviedo: Septem ediciones.

Ruiz-Madrid, N., and I. Fortanet-Gómez. 2015. “Contrastive Multimodal Analysis: Conference Plenary Lectures in English and in Spanish.” Pp. 39–60 in *Multimodal Analysis in Academic Settings: From Research to Teaching*, edited by B. Crawford Camiciottoli and I. Fortanet-Gómez. London: Routledge.

Schegloff, E. 1998. “Body Torque.” *Social Research* 65: 535–96.

Schiffrin, D. 1994. *Approaches to Discourse*. Oxford: Blackwell.

Schindler, L. 2009. “The Production of ‘vis-Ability’: An Ethnographic Video Analysis of a Martial Arts Class.” Pp. 135–54 in *Video Interaction Analysis*, edited by U. Kissmann. Frankfurt: Peter Lang.

Scotto di Carlo, G. 2014a. “Ethos in Ted Talks : The Role of Credibility in Popularised Texts.” *Facta Universitatis* 12(January 2015):81–91.

Scotto di Carlo, G. 2014b. “New Trends in Knowledge Dissemination : TED Talks.” *Acta Scietiarum* 36(2):121–30.

Scotto di Carlo, G. 2014c. “The Role of Proximity in Online Popularizations: The Case of TED Talks.” *Discourse Studies* 16(5):591–606.

Scotto di Carlo, G. 2015. “Stance in TED Talks : Strategic Use of Subjective Adjectives in Online Popularisation.” *Iberica* 29: 201–22.

Searle, J. 1969. *Speech Acts*. Cambridge: Cambridge University Press.

Searle, J. 1971. *Philosophy of Language*. Oxford: Oxford University Press.

Searle, J. 1975. “Indirect Speech Acts.” *Syntax and Semantics* 3: 59–82.

Searle, J. 1976. “A Classification of Illocutionary Acts.” *Language in Society* 5(1):1–23.

Selting, M. 1994. “Emphatic Speech Style-with Special Focus on the Prosodic Signalling of Heightened Emotive Involvement in Conversation.” *Journal of Pragmatics* 22(2/3):375–

408.

Selting, M., and E. Couper-Kuhlen. 1996. *Prosody in Conversation: Interactional Studies*. Cambridge University Press.

Shalom, C. 1993. “Established and Evolving Spoken Research Process Genres: Plenary Lectures and Poster Session Discussions at Academic Conferences.” *English for Specific Purposes* 12: 37–50.

Sinclair, J. 2005. “Corpus and Text –Basic Principles.” Pp. 1–16 in *Developing linguistic corpora: a guide to good practice*, edited by M. Wynne. Oxford: Oxbow Books.

Slembrouck, S. n.d. “What Is Meant by Discourse Analysis?” Retrieved 3/2/2017 (http://www.umsl.edu/~wilmarthp/mrpc-web-resources/discourse-analysis.pdf).

Sopory, P., and J. Dillard. 2002. “Figurative Language and Persuasion.” Pp. 407–26 in *The persuasion handbook: Developments in theory and practice*, edited by J. Dillard and M. Pfau. Thousand Oaks: Sage Publications.

Sparks, J., C. Areni, and K. Cox. 1998. “An Investigation of the Effects of Language Style and Communication Modality on Persuasion.” *Communication Monographs* 65(2):108–25.

Street, B., K. Pahl, and J. Rowsell. 2014. “Multimodality and New Literacy Studies.” Pp. 227–37 in *The Routledge handbook of multimodal analysis*, edited by C. Jewitt. London: Routledge.

Suhor, C. 1984. “Towards a Semiotics-Based Curriculum.” *Journal of Curriculum Studies* 16: 247–57.

Swales, J., and P. Rogers. 1995. “Discourse and the Proyection of Corporate Culture: The Mission Statement.” *Discourse and Society* 6(2):223–42.

Swales, J. 1990. *Genre Analysis: English in Academic and Research Settings*. Cambridge: Cambridge University Press.

Swales, J., and A. Burke. 2003. “‘It’s Really Fascinating Work’: Differences in Evaluative Adjectives across Academic Registers.” Pp. 1–18 in *Corpus Analysis: Language Structure and Language Use*, edited by P. Leistyna and C. F. Meyer. Amsterdam and Atlanta: Rodopi.

Tamarit, I., and H. Skorczynska. 2014. "Argumentative strategies and multimodality in oral business discourse: From theory to practice". *Procedia: Social and Behavioral Sciences* 116: 2330-2337.

Tang, K.. 2013. “Instantiation of Multimodal Semiotic Systems in Science Classroom Discourse.” *Language Sciences* 37: 22–35. Retrieved 3/2/2017 (http://dx.doi.org/10.1016/j.langsci.2012.08.003).

Tannen, D. 1989. T*alking voices: Repetition, Dialogue and Imagery in Conversational Discourse*. Cambridge: Cambridge University Press.

Thompson, G. 1998. “Resonance in Text.” Pp. 29–46 in *Linguistic choice across genres*, edited by A. Sanchez-Macarro and R. Carter. Amsterdam, Philadelphia: Routledge.

Thompson, G., and P. Muntigl. 2008. “Systemic Functional Linguistics: An Interpersonal Perspective.” Pp. 107–32 in *Handbook of interpersonal communication*, edited by G. Antos and E. Ventola. Berlin: Mouton de Gruyter.

Thompson, S. 1997. “Presenting Research: A Study of Interaction in Academic Monologue.” University pf Liverpool.

Thompson, S. 2002. “‘As the Story Unfolds’: The Uses of Narrative in Research Presentations.” in *The language of conferencing*, edited by E. Ventola, C. Shalom, and S. Thompson. New York: Peter Lang.

Tickle-Degnen, L., and R. Rosenthal. 1990. “The Nature of Rapport and Its Nonverbal Correlates.” *Psychological Inquiry* 1(4):285–93. Retrieved 3/2/2017 (http://www.jstor.org/stable/1449345).

Trager, G. 1958. “Paralanguage: A First Approximation.” *Studies in Linguistics* 12(1/2):1–12.

Tsai, TJ. 2015. “Are You TED Talk Material? Comparing Prosody in Professors and TED Speakers.” in *Proceedings of the 16th Annual Conference of the International Speech Communication Association*. Dresden.

Uysal, H. 2014. “A Cross-Cultural Study of Indirectness and Hedging in the Conference Proposals of English NS and NNS Scholars.” Pp. 179–95 in *Occupying Niches: Interculturality, Cross-culturality and Aculturality in Academic Research*, edited by A. Lyda and K. Warchal. Springer.

Ventola, E. 1999. “Semiotic Spanning at Conferences; Cohesion and Coherence in and across Conference Papers and Their Discussions.” Pp. 101–25 in *Coherence in Spoken and*

WrittenDiscourse. How to create it and how to describe it, edited by W. Bublitz, U. Lenk, and E. Ventola. Amsterdam: Benjamins.

Ventola, E. 2002. “Why and What Kind of Focus on Conference Presentations?” Pp. 15–27 in *The language of conferencing*, edited by E. Ventola, C. Shalom, and S. Thompson. Peter Lang.

Virtanen, T., and H. Halmari. 2005. “Persuasion across Genres: Emerging Perspectives.” Pp. 2–24 in *Persuasion across genres*. Amsterdam: John Benjamins.

Warren, M. 2004. “So What Have You Been Working on Recently? Compiling a Specialised Corpus of Spoken Business English.” Pp. 115–40 in *Discourse in the professions*, edited by U. Connor and T. Upton. Amsterdam, Philadelphia: John Benjamins.

Webber, P. 2005. “Interactive Features in Medical Conference Monologue.” *English for Specific Purposes* 24(2):157–81. Retrieved 3/6/2016 (http://www.sciencedirect.com/science/article/pii/S0889490604000213).

Wells, J. 2006. *English Intonation*. Cambridge: Cambridge University Press.

Widdowson, H. G. 1998. “The Theory and Practice of Critical Discourse Analysis.” *Applied Linguistics* 19(1):136–51.

Woods, R., K. Rafferty, J. Murphy, and P. Hermon, eds. 2014. *Engineering Innovative Products*. John Wiley.

Woodwall, W., and J. Burgoon. 1981. “The Effects of Nonverbal Synchrony on Message Comprehension and Persuasiveness.” *Journal of Nonverbal Behaviour* 5(4):207–33.

Wrembel, M. 2007. “Metacompetence Based Approach to the Teaching of L2 Prosody: Practical Implications.” Pp. 189–210 in *Non native prosody:Phonetic description and teaching practice*, edited by J. Trouvain and U. Gut. Berlin: Walter de Gruyter.

Wulff, S., J. Swales, and K. Keller. 2009. “‘We Have about Seven Minutes for Questions’: The Discussion Sessions from a Specialized Conference.” *English for Specific Purposes* 28(2):79–92.

Yates, J., and W. Orlikowski. 2002. “Genre Systems: Structuring Interaction through Communicative Norms.” *Journal of Business Communication* 39: 13–35.

Yates, J., and W. Orlikowski. 1992. “Genres of Organizational Communication: A Structurational Approach to Studying Communication and Media.” *The Academy of*

Management Review 17(2):299–326.

Ylänne-McEwen, V. 1997. "Relational Process within a Transactional Setting: An Investigation of a Travel Agency Discourse." University of Wales, Cardiff.

Yu, C. 2011. "The Display of Frustration in Arguments: A Multimodal Analysis." *Journal of Pragmatics* 43: 2964–81.

Zappavigna, M., C. Cléirigh, P. Dwyer, and J. R. Martin. 2009. "The Coupling of Gesture and Phonology." Pp. 219–36 in *New Discourse on Language: Functional Perspectives on Multimodality, Identity, and Affiliation*, edited by M. Bednarek and J. R. Martin. London: Continuum.

Zareva, A. 2013. "Self-Mention and the Projection of Multiple Identity Roles in {TESOL} Graduate Student Presentations: The Influence of the Written Academic Genres." *English for Specific Purposes* 32(2):72–83. Retrieved 3/2/2017 (http://www.sciencedirect.com/science/article/pii/S0889490612000646).

Zhang, Z. 2015. "Disagreements in Plenary Addresses as Multimodal Action." Pp. 17–38 in *Multimodal Analysis in Academic Settings: From Research to Teaching*, edited by B. Crawford Camiciottoli and I. Fortanet-Gómez. London: Routledge.

Zillman, D., J. W. Perkins, and S. Sundar. 1992. "Impression Formation Effects of Printed News Varying in Descriptive Precision and Exemplification." *Medienpsychologie: Zeitschrift für Individual- und Massenkomunikation* 4: 168–85.

APPENDIXES

Appendix 1. The informed consent

UNIVERSITAT JAUME I (UJI)

Departament d'Estudis Anglesos

GRAPE

Castelló, Spain

CONSENT FORM

Multimodal Academic and Spoken Language Corpus (MASC)

Project Director: Inmaculada Fortanet Gómez Ph.D.

+34 964 729592 Email. fortanet@ang.uji.es

http://www.grape.uji.es

Description and Purpose of the Study:

The goal of this research is to collect additional data to add to the existing holdings of the Multimodal Academic and Spoken Language Corpus (MASC), which is a collection of audio or video tapes, transcripts, and visuals employed in a wide range of presentations in different communicative events. The tapes, transcripts and visuals will be used to study the typical characteristics of naturally occurring speech in these various settings. The primary reason for collecting and analysing this data is to improve the training of present and future university instructors by analysing characteristics of academic speech and interaction. For a more accurate interpretation of the data gathered, interviews with speakers and some members of the audience could also be carried out.

If you have any questions about the project or your participation in it, please do not hesitate to ask the research assistant or to contact the project director named above.

____

My signature below indicates that I understand the nature of the study, and that I agree to participate in this research project as described above.

____

Participant's Name

_____________________________ _____________________________

Participant's Signature Investigator's Signature (MASC)

_____________________________ _____________________________

Date Date

I hereby give my consent to be videotaped. Yes_____ No____

I agree to allow these tapes to be transcribed, but understand that neither my last name nor any other explicit information as to my identity will be included in the transcripts (unless I give explicit permission to the contrary).
Yes___ No___

Portions of the tapes, transcripts or visuals (ppp, OHT, handout, etc.) employed during the session may be used in a variety of ways, and I understand that I may specify (after the recorded session) segments of the tape or transcript that I do not want to be used in these ways.

I hereby consent to the use of the recordings, transcripts or visuals in the following ways:

• For the preparation of **tests or instructional materials** Yes___ No___

• As part of academic **conference presentations or journal articles** Yes___ No___

• As part of **textbooks, CD-roms or other print or electronic media** Yes___ No___

• I would like to give **specific consent for materials to be used on the World Wide** and be consulted every time they are going to be used Yes___ No___

Optional: (This information is only used by researchers for possible follow-up.)

I agree that the researchers may contact me again if they have follow-up questions (e.g., in the event any portions of the recording are unclear to transcribers, or the researchers have additional questions about the recorded interaction):

Yes___ No___

Address:____

Phone (home):_____________________(office): ______________________ Email:_________________________________

Appendix 2. The observation sheet

CODE :			
OBSERVATION DATE:			
OBSERVER:			
1. EVENT:			
Workshop □	Conference □	Symposium □	Other:
Name of the Event:			
Dates of the event:			
Frequency:			
Organized by:			
2.GENRE			
Product Presentation □	Conference Presentation □		
Title:			
Start Time:			
Duration:			
Order in the program:			
3. THE SPEAKER			
Name:			
Nationality:			
Mother tongue:			
Sex:	Male □	Female □	
Age:			
Other relevant information:			
4. THE ROOM:			
Location:			
Lecture room □ Main Lecture Theatre □ Seminar Room □ Laboratory □ Other:			
Sketch distribution of participants, recording devices, furniture and props:			

5. AUDIENCE				
• Type of audience:				
• Total Number:				
• Number of MEN: Number of WOMEN:				
6. RESOURCES/ PROPS				
• PPT □				
• OHP □				
• Handouts □				
• Video □				
• Photographs □				
• Microphone □				
• Whiteboard □				
• CD □				
• DVD □				
Other:				
7. SPEAKER'S PERFORMANCE				
Posture:				
Move □	Stand up □	Sit □	Other:	
Mode of presentation:				
Explain □	Read □	Read and explain □		Other:
Additional Information:				

Observation sheet adapted from GRAPE (Group for research on Academic and Professional English), Universitat Jaume I, PROJECT HUM2004-02599/FILO

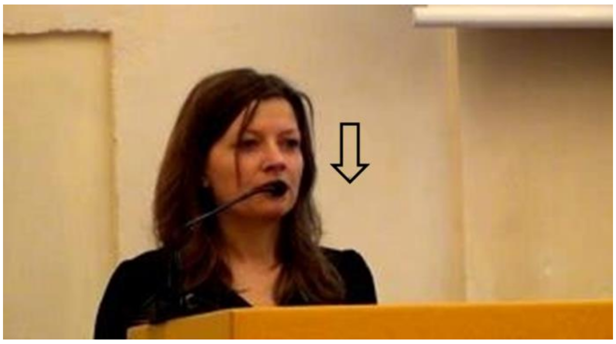
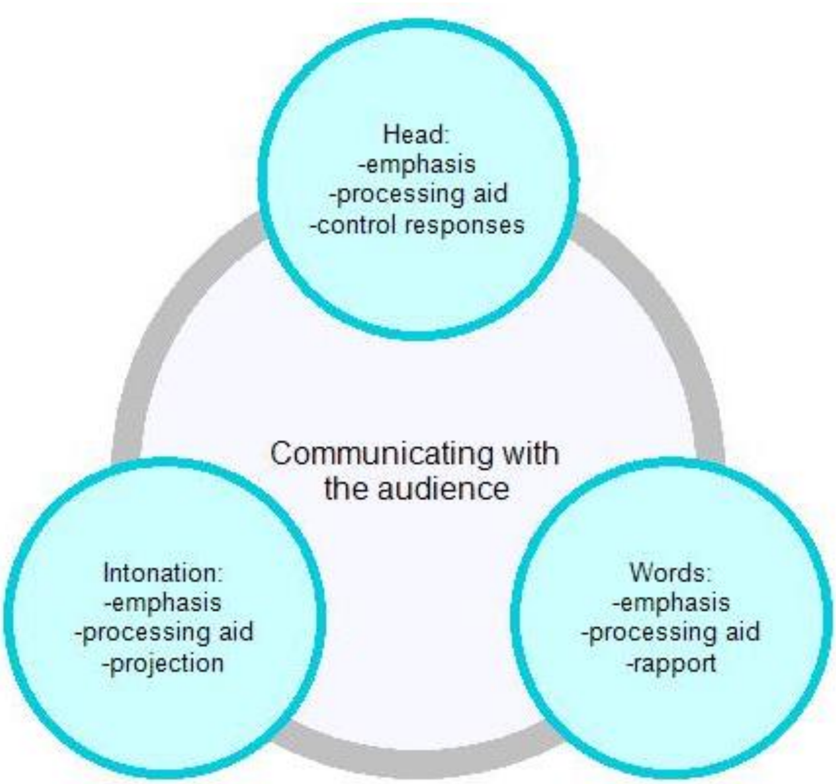
Appendix 3. The rich points[10]

CPDO1

This rich point corresponds to the introduction of the presentation: the speaker is introducing the topic.

Orthographic transcription
They distinguish between three levels of identity. And in most of the literature that I've read in preparation for this paper focuses on these macro, sociological categories. So the position of the researcher in terms of gender, in terms of race, in terms of ethnicity has been discussed a lot, but I haven't seen much attention paid to the more, the other levels here, the local ethnographic categories, the more transitory roles. So I'll try to pay attention to that in my paper.
MICASE transcription
they distinguish between, three levels of identity, and, in most of the literature that i've read in preparation for this paper focuses on these macro sociological categories so, the position of the researcher in terms of, gender in terms of race in terms of ethnicity has been discussed a lot but, i haven't seen much attention paid to the more_eh the other, levels here, the local ethnographic categories the more transitory roles so i'll try to, pay attention to that in my paper.
DI transcription
1. they disTINguish beTWEEN 2. three LEvels of iDENtity 3. and 4. in MOST of the LITerature that i've read in preparation for this paper 5. FOcuses on these MAcro sociological categories 6. so 7. the poSItion of the reSEARcher 8. in terms of GENder ↗ 9. in terms of RACE ↗ 10. in terms of ethNIcity ↗ 11. has been disCUSSED a LOT 12. BUT 13. i haven't SEEN much attention paid to the MORE 14. eh 15. the Other LEvels here 16. the LOcal ETHnographic categories 17. the more TRANsitory roles 18. so i'll try to pay atTENtion to THAT in my paper

[10] The video material corresponding to these rich points is included as a DVD.

Gestures and head movements
 in terms of GENder  the Other LEvels here
Multimodal ensemble


Analysis of words:

Concerning the use of grammar, the speaker makes a contrast between "discussed a lot" vs. "I haven't seen much attention" to highlight the research gap and present it as original. This contrast is also reinforced by a brief pause immediately following the adversative "but" (which is also prominent) and a head nod immediately following "other" (see next sections). It is also interesting how she describes previous research in impersonal terms ("has been discussed and lot"), but turns more personal when occupying the niche ("I haven't seen much attention"… "so I'll try to pay attention to that"), assuming responsibility for this research decision. Note also how she hedges this selection of her focus in "I'll try to pay attention to that" to sound non-categorical.

Regarding stylistic devices the speaker also uses a three part list embedding a symmetrical structure ("in terms of"). This makes the list easier to understand and more memorable. This effect is reinforced through head nods that mark the rhythm of discourse in each item of the list (see analysis of gestures).

Analysis of intonation:

"Most" in unit 4 and "macro" in unit 5 are prominent. They highlight what has been researched so far and prepare for the identification of the gap. "But" in unit 12 is also prominent (and followed by a brief pause) and introduces a contrasts that shows the existing research gap anticipated before. "Other" in unit 15, "local" and "ethnographic" in unit 16 and "transitory" in unit 17 are also prominent. They constitute the second element in the contrast previous vs. new research. They all show her occupation of the niche and how she is contributing with new research.

Parallel rising intonation in each item of the three part list (units 8-10) presents these elements as shared knowledge in contrast to the "new" she is adding. At the same time the parallel intonation mirrors the parallel grammar, making the text more memorable.

Analysis of gestures and head movements:

Head nods are used to index each element in a list of three aspects dealt with in previous research. These nods are synchronous with every prominence in the list (i.e. "gender", "race", "ethnicity"). They have a cohesive function, bringing together the different items in the list, and also pragmatic in their emphatic effect that prepares for the coming contrast (i.e. the research gap).

A nod during a brief pause immediately following “other” in “the other levels here” highlights the research gap she is presenting and reaffirms her occupation of the niche. Apart from being emphatic, the nod also has a pragmatic function in its reaffirmation that can be considered a prompt to approve her research topic choice. The facial expression with pursed lips seems coherent with this pragmatic function.

Analysis of persuasive activity:

Three modes are contributing to persuasion in this excerpt: head movements, words and intonation. The three of them are used to facilitate understanding and emphasize certain parts of the message (e.g. contrast previous vs. new). In addition each mode has an extra specialized function: words are used to create rapport with the audience personalizing the text and making it non-categorical, head nods are used to subtly prompt approval and intonation conveys an assumption of shared knowledge when presenting previous research.

There is minimal shift of attention from communicating to encoding the message that reveals itself in the use of words (i.e. false start in “the more_eh the other, levels here”) and is mirrored in an oblique intonation matching the filler (“eh”). This temporary focus on encoding, though minimal, can give an added function to the nod that follows “other” as self-reassuring, i.e. a confirmation that the way the message has been encoded is correct.

This is consistent with the way the speaker reported to have felt during the presentation: nervous at the beginning (where this excerpt belongs), more relaxed as the audience backchannelled with approval signs such as smiles and nods.

CPDO2

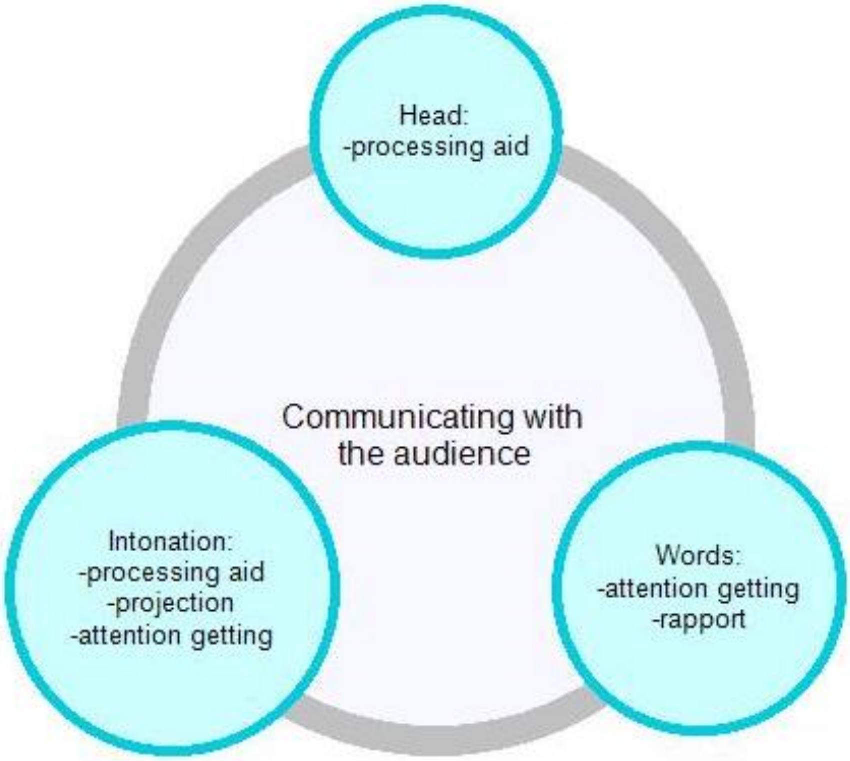
In this rich point the speaker is explaining the reaction and attitude of her informants during her research project to illustrate her positioning as student.

Orthographic transcription
And when I showed up at 4:30 in the morning to follow them for their day of work, they were clearly apprehensive about what I was doing there, was I trying to spy on them cleaning the offices or what? What was the purpose of my being there? And I said I was PhD student, and they caught on to the student part "Oh so you're going to do a paper. How long is your paper? Well, it's 300 pages. Wow, that's a long paper. Yes it is."
MICASE transcription
and when i showed up at four thirty in the morning to follow them for their day of work they were clearly apprehensive about what i was doing there was i trying to spy on them cleaning the offices or what what was the purpose of my being there. and i said i was phd student and they caught on the student part oh so you're going to do a paper how long is your paper well it's three hundred pages wow that's a long paper yes it is

DI transcription
1. and when i showed up at FOUR thirty in the morning 2. to FOllow them for their day of WORK 3. they were CLEArly appreHENsive about 4. what i was DOING there 5. was i trying to SPY on them cleaning the OFFICES 6. or WHAT 7. what was the PURpose of my BEING there 8. and i said i was PHD STUdent 9. and they CAUGHT on to the STUdent part 10. OH 11. so you're going to do a PAper 12. how long is your PAper 13. well it's three hundred pages 14. wow 15. that's a LONG Paper 16. yes it is

Gestures and head movements
 or WHAT  well it’s three hundred pages  yes it is

Multimodal ensemble



Analysis of words:

The speaker is talking about her research experience using a narrative. The use of direct speech, 1st person pronouns and informal style (e.g. in the use of the phrasal verb "showed up") in this narrative makes language more vivid and brings the characters closer to the audience. The word choice also reveals a focus on attitudes and emotions of the characters in the narrative ("apprehensive", "spy", vocalizations expressing surprise) rather than on the actions, which adds to the intensity of the narrative.

Analysis of intonation:

The use of intonation in this excerpt tries to convey the attitude of each character in the narrative, i.e. the participants show surprise and interest when they learn that she is also a student, while the researcher is somehow tired of this reaction, or of having to adopt the role of a student to gain access to her data. The speaker corrected this interpretation in what concerns her own attitude: it is not so much boredom or tediousness that she is trying to convey, but an attempt to minimise a white lie.

The lack of pauses in between turns and the rather speedy pace during the re-enactment of the dialogue (units 10- 16) shows that the speaker is somehow skipping through it as if the audience was already familiar with it, presumably to present it as a choir, something she has found too many times and is tired of. This speeding up is consistent with the almost oblique use of intonation in her own turns in the dialogue (units 13 and 16), which is rather flat and with no prominences, revealing disengagement from the communicative interaction (i.e. in this re-enacted dialogue she is answering mechanically). Taking into account the speaker's interpretation, it seems that she is rather trying to prevent this topic from extending further and move on to something else, because she considers it a white lie she is not comfortable with. As she put it in the interview, she is conveying "let's talk about something else".

In contrast, the intonation applied to the participants conveys surprise and interest. The speaker marks the change of personae by using an overall higher pitch range when enacting her participants, but this higher pitch at the same time conveys that the participants are surprised at the information they are receiving and interested in it. This is supported by the vocalizations in units 10 and 14, which are stereotypical ways of expressing surprise.

Analysis of gestures and head movements:

The speaker explains her informants' reaction when they first saw her using direct or semi-direct quotes (exactly the questions they were asking themselves). At the same time she visually represents this lack of understanding through a sequence of subtle head shakes, as if impersonating her informants, which adds to the vividness of the narrative. These shakes become clearly perceptible as she says "or what".

Regarding the lateral head movements my initial interpretation is that they visually represent the researcher's feelings at the repetitive reaction of her informants. In combination with a flat unenthusiastic intonation they seem to convey a sense of boredom at having to repeat the same thing over and over again. The second time, synchronous with the last tone unit ("yes it is"), this lateral movement becomes more prominent. The speakers interpretation, however, is somehow different: it is a way of minimising what she considers a "white lie" in saying she was a student. She did this to gain easier access, but it wasn't totally accurate.

Analysis of persuasive activity:

In this excerpt the speaker is really focused on communicating with the audience and uses words, intonation and head movements (to a lesser extent) to build an interesting narrative of her research experiences. Words are used to get the attention of the audience and create rapport by focusing on emotions. Head movements ease understanding (e.g. clarify which character in the narrative is at stake in each moment and how they felt) while intonation, apart from also contributing to get the attention and clarifying who is speaking, projects an assumption of common ground with the audience and conveys that the experience described was somehow undesirable.

It is interesting to note that when asked about the strongest points in her presentation, the speaker answered this was the examples she had included from her interviews and field research, which is in line with her efforts to present these examples in the form of a vivid narrative, actually impersonating characters.

CPAS1

The topic of the presentation is the complex interplay between key players in a university master internship project, and the tension between academic and professional profiles. In this rich point the speaker is discussing one of their main findings, i.e. students and company did not manage to build the relationship of mutuality and trust that was required for the project and this was evident in some communication gaps.

Orthographic transcription
Also coming back to that concept of freedom, one of their core principles, the students, to them, were very free. When, remember when the students said we weren't allowed to interview employees, well, in fact it wasn't because they weren't allowed, it just did not seem relevant to the company. But in fact, all the resources were there, just if they wanted them, it was there.
MICASE transcription
eh, also, coming back to that concept of, freedom, their _one of their core principles, the students, to them, were very free. eh, when_remember when the students said we weren't allowed to interview eh employees well, in fact, it wasn't because they weren't allowed it just did not seem relevant to the company, but in fact all the resources were there, eh just, eh if they wanted them, it was_it was there, eh
DI transcription
1. eh 2. also 3. coming BACK to that concept of FREedom 4. their 5. one of their CORE principles 6. the $^{\text{STU}}$dents 7. to $^{\text{THEM}}$ 8. were VEry FREE 9. eh 10. when 11. reMEMber when the STUdents said 12. we weren't aLLOWED to interview eh emploYEES 13. well in FACT 14. it WASN'T because they weren't alLOWED 15. it just did not seem $^{\text{RE}}$levant to the COMpany 16. but in FACT all the resources were THERE 17. eh 18. just 19. eh 20. if they WANted them 21. it was it was THERE

Gestures and head movements

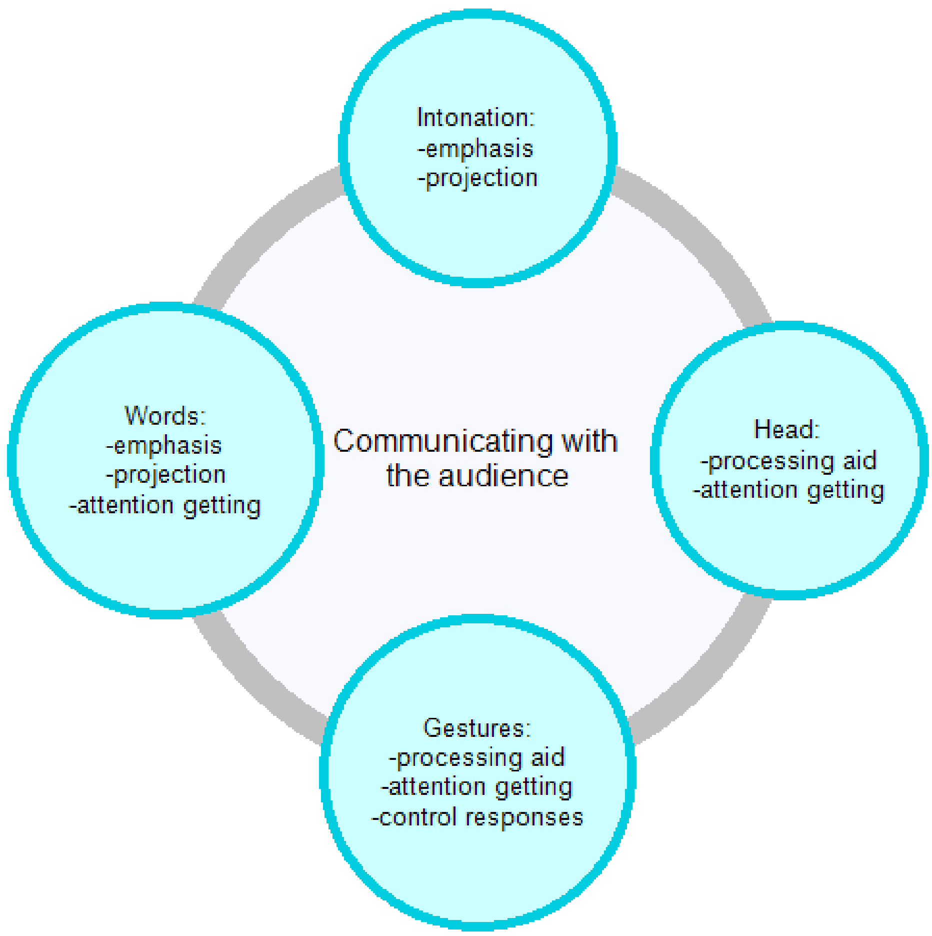
reMEMber when the STUdents said

it just did not seem [RE]levant to the COMpany

if they WANted them

it was it was THERE

Multimodal ensemble



Analysis of words:

The speaker directly addresses the audience and asks them to remember a previous moment of the presentation, which actively engages them in the argument. The speaker reported during the interview this “touch-basing with the audience” is very likely a habit acquired through teaching practice, and it helps her in the design of the message as much as it helps the audience.

The repetition of the adversatives “in fact” and “but” emphasizes the contrast between what students perceived and what the company perceived, and highlights the misunderstanding this passage refers to.

The use of “we” to refer to students turns this passage into a direct quote, which is more vivid and provides a narrative style. The speaker also reported this role switching is a way of connecting with the main topic of the presentation: a student internship project.

Analysis of intonation:

The beginning of the passage is broken into several short tone units (units 6-7-8: the students//to them// were very free). This slows down the rhythm and provides extra emphasis through prominence to words like “students”, “them” and “very free”. This effect is also supported by pauses (see below). According to the speaker it can also be a sign of insecurity, since this is a side project and not her core business, so she didn’t feel as an expert.

High key in “students” (unit 6) and “to them” (unit 7), and later in “relevant” (unit 15) conveys surprise and highlights the contrast between what the company perceived and what the students perceived.

Fillers and false starts can be attributed to encoding difficulties and distraction caused by the chair's coughing fit. There are also other intentional pauses (highlighted yellow in the micase transcript that have the effect of slowing down the rhythm and providing emphasis to the words preceding or following, as in “coming back to that concept of, freedom, their _one of their core principles, the students, to them, were very free”. According to the speaker it can also be a sign of insecurity, since this is a side project and not her core business, so she didn’t feel as an expert.

Analysis of gestures and head movements:

A gesture close to a finger ring beginning at her temple and moving outwards symbolizes the action of remembering and is synchronous with the imperative “remember”. This is a metaphoric, interpersonal gesture that prompts a mental process as a response from the audience, and therefore engages them in the discussion. The speaker added that this gesture can also be a way of anticipating the two-sidedness implied by “interview”, which comes later in the speech, and conveying the two parties involved in the misunderstanding she is reporting. This would explain the shape of the gesture, with two parallel fingers extended.

Repeated head shakes simultaneous with “did not seem relevant to the company” contradict visually the students' wrong interpretation (as if the speaker were assuming the role of the company at this moment and representing their reaction, paraphrasable as “it’s not that we don’t allow you, we thought it wasn’t useful for you). Note how the facial expression also supports head movements in conveying lack of understanding.

Repeated open hand supine gestures as she says "eh just, eh if they wanted them, it was_it was there" visually represent the availability conveyed, through a referential, metaphoric gesture. Note, however that for this speaker OHS gestures are extremely common, and probably this referential function is not intended, being simply a beat. According to the speaker, it can be interpreted as “just take it”.

Analysis of persuasive activity:

Using words and intonation, the speaker emphasizes key parts of her message and highlights two contrasting perceptions (that of the company and that of the students). This lack of alignment between the parties involved in the project she is describing is actually one of the main findings of the study, and therefore it is presented as original and valuable. Furthermore, she retains the attention of the audience using gestures, head shakes and direct quotes to build a vivid narrative of what happened (almost providing a re-enactment of the misunderstanding), which is easy to understand, interesting and memorable. This effort to make the text interesting can be linked to the speaker's reported perception on the event: she explained that she was not familiar with the potential audience as she prepared, and once there she felt “like an outsider”, which can partly explain her concern to make her presentation match their interests and get their attention.

CPT01

The topic of the presentation is the complex interplay between key players in a university master internship project, and the tension between academic and professional profiles. In this rich point the speaker is discussing one of the main findings of their study, i.e., during an internship project, the role of the teacher must be adapted so that students can evolve to become professionals.

Orthographic transcription
So, we found out we had to focus more on the outside than the inside of the story and we found that we should in fact both diminish and expand our role as participants in this whole process. Diminish, why? That means we should move further away from the projects because naturally, students come to us for guidance with questions and we in fact, hold their hands till the end of the project, which is maybe not good, which hinders them in getting involved more with the company, with the consultants. So they've always come back to us of course. So, we saw in the visualization that we should stop at a certain point at the in-phase and not guide the students all along the way until the final output of the project.
MICASE transcription
so, we found out we had to focus more on the outside than the inside of the story. and we found that we should, in fact both diminish and expand our role, as participants in this whole process. diminish, why that means, we should move further away from the projects because naturally students come to us for guidance, with questions and we, in fact hold their hands till the end of the project, which is maybe not good. which hinders them in, getting involved more with the company, with the consultants so they've always come back to us of course. so, we saw in the, eh, visua_visualization, that we should stop at a certain point at the in-phase, and not guide the students all along the way, eh until the final output of the project.

DI transcription
1. SO 2. we found out we had to focus more on the OUTside than the INside of the story 3. and we FOUND that we SHOULD 4. in FACT both diMInish 5. and exPAND our ROLE 6. as parTIcipants in this WHOLE process 7. diMInish WHY? 8. that MEANS we should move further aWAY from the projects 9. because NAturally, students come to US 10. for GUIDance 11. with QUEStions 12. and WE in FACT 13. hold their HANDS till the end of the PROject 14. which is MAYbe not GOOD 15. which HINders them IN 16. getting inVOLVED MORE 17. with the COMpany 18. WITH the conSULtants 19. so they've ALways come BACK to us of course 20. SO we SAW in the eh visua visualization 21. that we should STOP at a certain point 22. at the IN-phase 23. and NOT GUIDE the students 24. ALL along the WAY 25. eh 26. until the final OUTput of the project

Gestures and head movements

we found out we had to focus more on the OUTside than the INside of the story

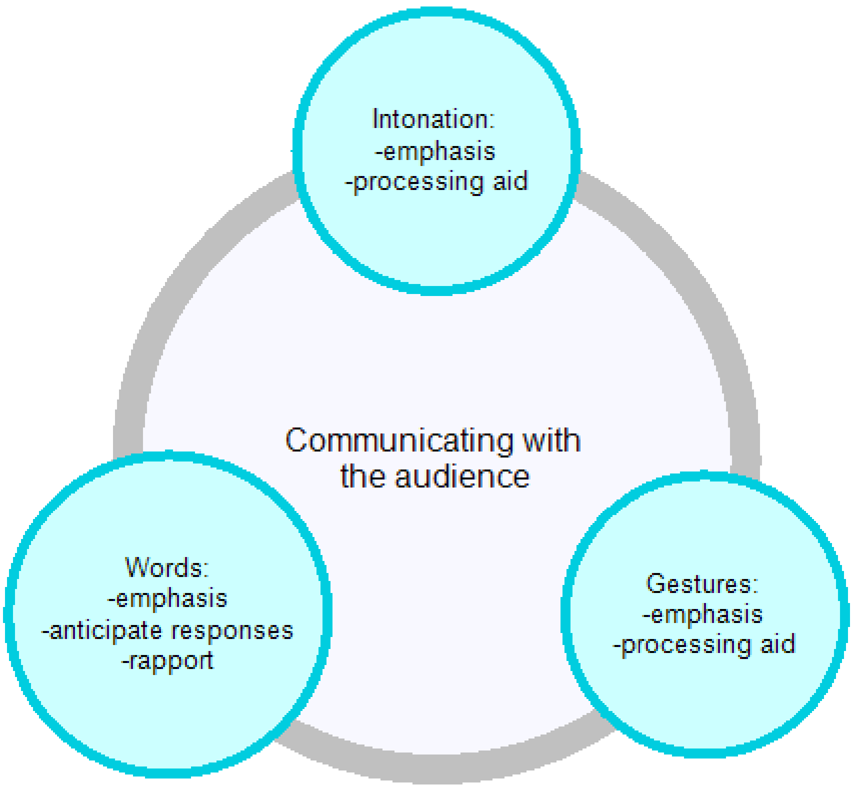
in FACT both diMInish

and exPAND our ROLE

so they’ve ALways come BACK to us of course

that we should STOP at a certain point

Multimodal ensemble



Analysis of words:

There are several indirect references to the title of the event, i.e. “The ins and outs of professional discourse” (outside-inside, in-phase). The speaker expressed during the interview that for him it is really important to show a “clear link with the subject or topic of the conference” and prove that “this is not the tenth time I’m bringing this”.

Two parallel contrasts in outside vs inside and diminish vs expand build a symmetrical grammatical structure that makes the text memorable. The speaker pointed out during the interview that the first contrast was planned, in an attempt to highlight the connection with the overall topic of the conference and also to show their positioning or approach in their research (somewhere in the middle of these “ins and outs”). The second contrast was more spontaneous, but nonetheless it was a way to share their experience, since they, as researchers, were really presented with these two options during the project and they had to discuss them. As the speaker puts it, it is as if they are asking the audience "what should we do, what is your point of view in our discussion?”

The question "why" anticipates a reaction of the audience and at the same time postpones the explanation of the finding (i.e. the need to diminish the role of the teacher), creating surprise and intrigue. The speaker added this was also a way to get the attention of the audience and involve them.

The use of "maybe" is hedging and makes the message not imposing. According to the speaker this non-categorical way of speaking is also a way to share their research experience and bring the audience in to participate in the ongoing discussion: nothing is presented as already settled, but as possible options to which the audience can give feedback.

Analysis of intonation:

The contrasts referred to above are further emphasized through prominence in "outside", "inside", "diminish" and "expand". The words "back" and "stop" are also made prominent through intonation in addition to being visually represented trough gestures. Therefore, in this example, gestures and prominence are in phonological synchrony.

Analysis of gestures and head movements:

Two sets of beats to both sides (left-right in the first case and right-left in the second case) emphasize the contrasts "outside vs. inside" and "diminish vs. expand" respectively. The repetition of the gesture further reinforces the parallelism in the grammar. The function of the gesture can be considered cohesive, since it brings together two related parts of the message.

"Come back" is represented visually trough an inward hand gesture. This is a metaphoric, referential gesture in semantic and phonological synchrony with words and intonation that makes the message easier to process.

As the speaker is pronouncing the utterance "we should stop at a certain point" he makes an open hand supine gesture that symbolizes a wall or barrier and is in semantic and phonological synchrony with the word "stop". This is an iconic, referential gesture with a dual function: it is a visual aid to help understanding and a way of emphasizing an important finding in the study. The speaker explained that in the visuals he was using he had actually drawn a diagram of the process of the project including a red wall representing this barrier. He was probably influenced by this visual input and represented it gesturally as he presented.

Analysis of persuasive activity:

In this example the speaker is using the following persuasive strategies:

-emphasizes crucial parts of the message through intonation, gestures and stylistic devices (words)

-anticipates a response and creates intrigue (words) through the “why” question

-rapport (words): uses non categorical language through hedging, indirectly refers to the event where he is presenting, and shares research experiences involving the audience.

-uses gestures as processing aids and also shows structure through intonation and gestures

With these strategies the speaker makes his text memorable, non-categorical and easy to understand, which makes it more likely to be accepted and remembered by the audience. At the same time he is showing how his presentation fits within the whole communicative event and also highlights the value and originality of the results of the study, while sharing research experiences, which further contributes to the persuasive effect.

The speaker confirmed during the interview that the effort to fit in the event and share experiences was actually conscious, because both speakers were aware that they were “somewhere in the middle” of these ins and outs. He also explained how they kept changing the presentation until the last minute to include references to other presentations and make it more coherent with the whole conference. During the interview the speaker also expressed his strong belief in the importance of visuals in presentations. This can be related to his use of gestures as visual representations of his referential content, for example in the barrier gesture.

CPAM1

In this rich point the speaker is differentiating her Master's from other Master's and indirectly evaluating it in positive terms.

Orthographic transcription
So one of the – shall we say, well, the distinctive feature of our Master's as compared to other Master's in International Human Resources Management is the fact that we have a strong concentration on foreign language skills and cultural knowledge, okay? And that we think is – shall we say the, winning point, although when I come to the results, it'll be interesting to see what they say.
MICASE transcription
so one of the, shall we say_well the distinctive, feature of our master's as compared to other masters in international operations management, is the fact that we have a strong concentration, on foreign language skills and, cultural knowledge. okay and that we think is, eh, shall we say the_the _the winning point, although, when i come to the results it'll be interesting to see what they say.
DI transcription
1. so ONE of the 2. shall we SAY 3. well ^{THE} distinctive FEAture 4. of OUR master's 5. as comPARED to Other masters in international operations management 6. is the FACT that WE have 7. a STRONG concenTRAtion 8. on FOreign language SKILLS 9. AND 10. CULtural KNOWledge 11. okay 12. and THAT we think is 13. eh 14. shall we SAY 15. the the the WINning POINT 16. although when i COME to the reSULTS 17. it'll be inteREsting to see what they SAY

Gestures and head movements

well THE distinctive FEAture

although when I COME to the reSULTS

Multimodal ensemble

The excerpt presents two higher level actions occurring simultaneously:

1. interacting and communicating with the audience

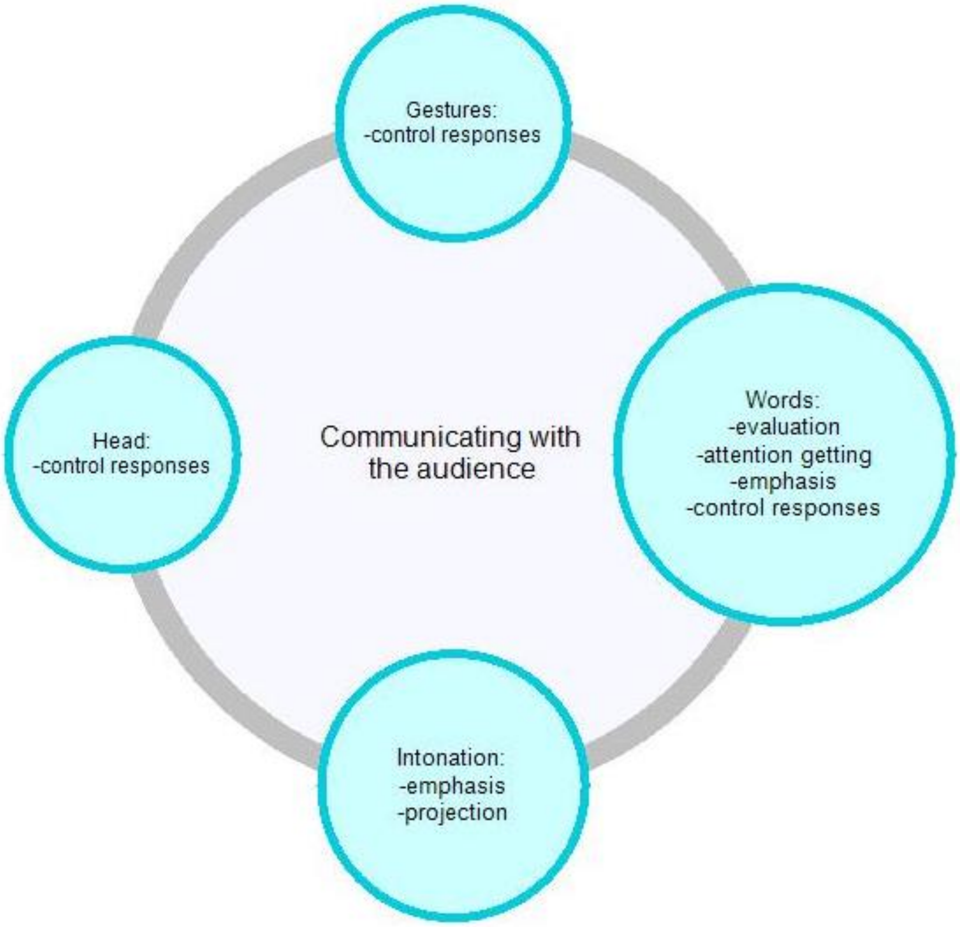
2. designing/planning the message

It is interesting that at points when the action of "planning the message" becomes more salient, gestures and head movements are mainly used to fulfil this purpose. This is clear in the way the speaker avoids eye contact with the audience and makes head movements that are intended also for herself as much as for the audience.

If we compare this to moments in which the action of communicating is fully foregrounded we see how the non-verbal resources are again used to meet this goal, and they become fully oriented to the audience, even becoming interpersonal.

The two following diagrams represent these two scenarios:

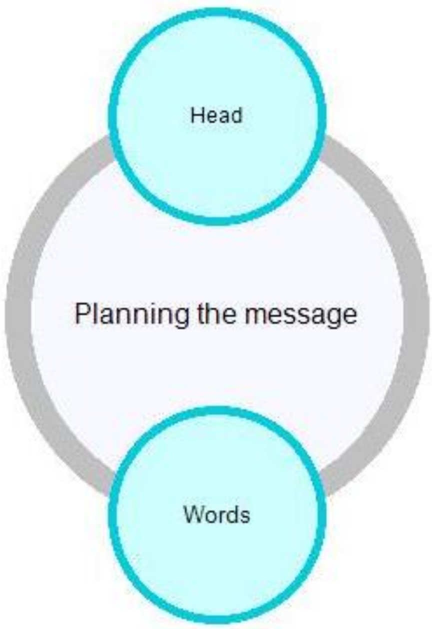
Scenario 1: interacting and communicating with the audience



As the speaker focuses on communicating, she makes extensive use of words and occasional use of intonation to get her message across in a persuasive way. For example she selects intense vocabulary and carefully selects which parts to make prominent to present her master as better than others.

She also uses head movements and gestures to trigger a desirable response from the audience (a positive evaluation), though to a lesser extent.

Scenario 2: planning the message



As the speaker also devotes part of her efforts to design her text, she uses head movements (self-reassuring nods) and words (fillers, self-corrections). These modes also show less synchrony than in the previous scenario, since head nods almost always follow words. The speaker commented during the interview that this is a natural consequence of not having a scripted text: planning the message is done on the spot and not beforehand

Analysis of words:

The speaker evaluates her master describing a unique feature that makes it preferable to others. This positive evaluation is quite intense and noticeable in her lexical choices (e.g. "winning point"). Although it is framed subjectively at the beginning (“we think") the speaker moves to a more objective framing towards the end ("when I come to the results"). The way she anticipates results also creates expectations and raises the interest of the audience.

The speaker commented in the interview that the use of “we” is a way to show she is part of a team and that the master is not only her own. Regarding the anticipation of results, she agrees it was a way of raising interest, but also a consequence of being a data -driven person that creates an "anxiety to show facts". She also added that she perceives a general "defensive" attitude in her speech, which is probably due to the fact that she was unsure about fitting in the whole event.

Analysis of intonation:

The use of prominence is particularly interesting in this rich point. Prominence and high key in “the” (unit 3) makes it contrastive (new information not recoverable from context), and at the same time rules out other options. This use of intonation matches the self-correction in “one of the…well the distinctive…”. Prominence in “our” (unit 4) emphasizes the differentiating effect (our master as opposed to others). The brief pause after “distinctive” (unit 3) also serves the same purpose, as well as the prominence in "we" in unit 6. Prominence in “and” (unit 9) adds a sense of accumulation (=”not only foreign language but also…”) and emphasizes the added value being conveyed.

Analysis of gestures and head movements:

A sequence of four head nods synchronous with “the distinctive feature” has a reaffirming effect. The first three nods immediately follow “the”, which is also prominent (therefore it is not exactly synchronous with words and intonation), the fourth one is synchronous with “feature”.

Following “point” in “the winning point” and filling a micro pause there is a second sequence of two nods that has the same confirmation effect as the previous ones. Since these nods occur after a sequence of a hedging filler “shall we say” plus a false start, they can be interpreted as a way for the speaker to reassure herself that the way she is formulating the message is the correct one and that she is really saying what she wants to say. The fact that she is not keeping eye contact with the audience at these moments supports this interpretation. At the same time she may be implicitly preventing any potential counterargument to what she is saying. The speaker agrees this is a plausible interpretation given the defensive attitude she was adopting at this point in her presentation. In the picture the nod synchronous with “feature” is represented. Note how the speaker also closes her eyes at the same time, maybe as a sign of encoding effort that is also suggested by the false start and filler.

As she moves on to say "although” she makes a slightly lateral head movement and raises her open hand in what seems a deictic gesture pointing vaguely at the audience. This combination of gestures and head movements seems to indicate the shift from subjective to objective evaluation, suggesting that it will be the audience who will decide when presented with the results. Therefore, it can be considered interpersonal. The speaker agrees with this interpretation and perceives an element of hedging in this gesture, a way of making clear to the audience that "you will be involved in the discussion”.

Analysis of persuasive activity:

In this example the speaker is using the following persuasive strategies:

-Emphasis (intonation, words)

-Projection of speaker’s understanding of context of interaction (intonation, i.e. making “the” contrastive)

-Evaluation (words)

-Attention getting technique (words, i.e. postponing results that will be “interesting”)

-Control responses (words, gestures and head movements prepare the audience to objectively conclude that her initial subjective evaluation is actually correct once results are presented)

With these strategies the speaker makes her text interesting and relevant for the audience highlighting the innovation in it (i.e. the unique characteristic of her master). She presents an initially subjective evaluation that she later supports with objective data (the results of her questionnaire) and subtly guides the audience to share this evaluation. Actually, the way evaluation is used in this excerpt is consistent with an important idea the speaker reported in the interview: her main concern was to avoid making the presentation sound like marketing. This explains her efforts to support her evaluation with objective data.

CPAM2

In this rich point the speaker is presenting unexpected results from a questionnaire that she designed and used to assess a master programme.

Orthographic transcription
I asked about most frequent writing typologies. Okay, being in human resources, a lot of them are in recruitment and selections so they're doing descriptions and evaluations of people a lot. Emails, 38%; reports, press releases, web updates and other, the generic other. Here I was, I mean there's been a lot of talk about email and stuff. I was interested, that – I mean, I was sort of anticipating the ultimate results here. I expected to get more novelty from them. For example, if companies have got a website, okay, and they are using the people who know foreign languages to update them. I would've expected a bigger percentage of people actually, you know, writing on the web but in fact, it was not very high, only 7%.
MICASE transcription
i asked about most frequent writing typologies okay being in human resources a lot of them are in recruitment and selections, so they're doing, descriptions and evaluations of people a lot. eh, emails, thirty eight per cent reports press releases web updates, and other the generic other. here i was, em_ I mean there's been a lot of talk about email and stuff eh i was interested, that_ i mean i was sort of anticipating the ultimate results here. i expected to get more novelty, from them for example eh, if companies have got a website okay and they are using the people who know foreign languages, to update them i would've expected a bigger, eh percentage of people actually, you know writing on the web but in fact it was not very high only seven per cent.

DI transcription	
1.	i asked about MOST frequent writing tyPOlogies
2.	okay
3.	BEING in human reSOURces
4.	a lot of them are in reCRUITment and seLEctions
5.	so they’re doing desCRIPtions and evaluAtions of people a lot
6.	eh
7.	Emails, thirty eight per CENT
8.	rePORTS
9.	PRESS releases
10.	WEB updates
11.	and Other the generic other
12.	$^{\text{HERE}}$ i was
13.	em
14.	i mean there’s been a lot of TALK about Email and stuff
15.	eh
16.	i was INterested
17.	THAT
18.	i mean i was sort of anTIcipating the ultimate reSULTS here
19.	i ex$^{\text{PEC}}$ted to get MORE NOvelty from them
20.	for eXAMple
21.	eh
22.	if COMpanies have got a WEBsite
23.	okay
24.	and they are Using the PEOple
25.	who KNOW foreign languages
26.	to UPdate them
27.	i would’ve exPECted a BIGger
28.	eh
29.	perCENtage of people
30.	ACtually
31.	you know WRITing on the WEB
32.	but in FACT it was NOT very high
33.	only SEVEN per cent

Gestures and head movements

i exPECted to get MORE NOvelty from them

i was sort of anticipating the ultimate reSULTS here

i mean there’s been a lot of TALK about Email and stuff

i would’ve exPECted a BIGger

eh

perCENTage of people

you know WRITing on the WEB

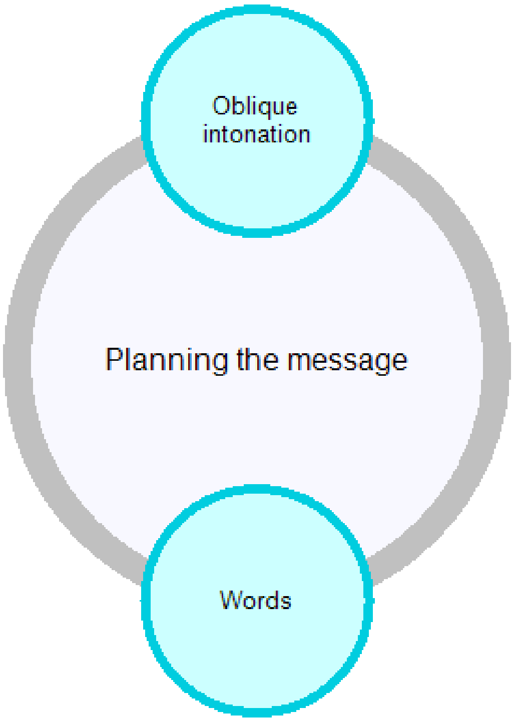
but in FACT it was NOT very high

perCENtage of people

ACtually

you know WRITing on the WEB

Multimodal ensemble



Analysis of words:

The speaker is using a conversational tone with fillers ("okay", "you know") and false starts that provide the text with spontaneity. The frequent references to herself also turn the text into a personal narrative of her research experience. There is also a reference to other presentations in the event.

Analysis of intonation:

A high key suggests surprise in "here" (unit 12) and "I expected" (unit 19). This use of intonation matches the content of the message, that at this point introduces an unexpected result. The use of prominence in unit 19 is also interesting because it emphasizes "more" and "novelty" (precisely the words that describe the unexpected result), and is irregular in the use of three prominences in the unit. The brief pause after "novelty" also adds to the emphasis.

Analysis of gestures and head movements:

A finger bunch beat moving inwards simultaneous with "I expected to get more novelty" represents specific results that she expected to get and did not find. The inward movement suggests receiving something. The gesture is referential, since it visually represents the idea of "obtaining" or "receiving".

Interestingly, the same type of beat in the form of a finger bunch is repeated before and after the previous excerpt, but in the opposite direction (i.e. outwards) when she is not focusing on receiving input from questionnaire as in the previous example, but rather on anticipating results and explaining why she had that expectations based on her informants and the work they do. The outward movement matches the intention of offering an explanation, making the gesture pragmatic.

A deictic pointing to the audience and synchronous with "talk" clarifies the reference in "there's been a lot of talk" specifying who has been doing the talking. The gesture is referential, and partly interpersonal in the way it involves the audience in the discussion. As the speaker herself described it, this gesture was a way of acknowledging the audience.

A head nod synchronous with "percentage" suggests approval and emphasizes the plausibility of her expectations. The speaker did not fully agree with this interpretation, and considers this nod a kind of kinesic stress, to emphasize the lexical stress in "percentage". In her opinion, the hand gesture and the use of gaze to engage the audience are more significant at this point.

A subtle shake synchronous with “web” and immediately preceding “in fact" represents the contradiction between what she expected and what she got. This contradiction is reinforced by a shoulder shrug synchronous with "actually" that seems to convey "I can’t explain this result”. Shrugging shoulders can be considered an emblem that is widely interpreted as "I don't know". The speaker offered a slightly different interpretation of the shoulder shrug as meaning: “if you were in my shoes, wouldn’t you have expected the same?", that is, a request for alignment from the audience.

Analysis of persuasive activity:

The speaker uses the strategies indicated in the diagram to present her research as original and valuable through highlighting unexpected results. Interestingly, she reported that she found these results rather disappointing, but she still acknowledges the value of unexpectedness and novelty in research. Her use of strategies makes her message sound not imposing (since the audience is subtly guided towards an evaluation) and easy to understand (through visual representation). She also shares research practices (how she felt when analysing unexpected results) and by referring to other presentations also tries to fit in the event.

Regarding the modal configuration, or multimodal ensemble orchestrated, words (false starts and fillers) are occasionally used to win time to encode the message and most of these coincide with an oblique orientation in intonation that suggests momentary disengagement. However, this also has the secondary effect of providing the text with spontaneity, which contributes to rapport. All of the four modes analysed contribute to the higher level action of communicating with the audience in a rather specialized way, although in close synchrony:

-Intonation directs attention to certain aspects of the message and presents them as surprising

-Gestures and head movements have a dual function: they ease understanding and also trigger a particular response

-Words create rapport

CPRE1

In this rich point, located at the beginning of the presentation, the speaker presents the concept of undergraduate research, providing the specific example of Kelsey's (her co-presenter) project as particularly relevant for the topic of the conference.

Orthographic transcription
We're very excited to be here in this conference because the theme, "Ins and Outs of Professional Discourse," is very close to our hearts as we've been working for the last two years on Kelsey's very important undergraduate research project. And to give you a little information about undergraduate research in the United States, it's a movement of engaged learning where undergraduates like Kelsey develop their own research questions, collect their own data and present their own findings under the mentorship of the faculty members. So they are doing the same work that a Master's or a PhD level student would do with the same kind of mentorship, presenting at conferences, publishing their work, so it's a very exciting way to work with undergraduate students and it's growing in popularity in America. Kelsey's project is really interesting and we're happy to present it today, in that it deals with both the ins and outs and a different way of thinking about how we enter communities and a broader sense of community that we can address.
MICASE transcription
we're very excited to be here in this conference because the theme ins and outs of professional discourse is very close to our hearts, as we've been working for the last two years on kelsey's, eh very important undergraduate research project. and to give you a little information about undergraduate research in the united states it's a_it's a movement of engaged learning where undergraduates like kelsey develop their own research questions collect their own data and present their own findings, under the mentorship of the faculty members so they are doing the same work that a master's or a phd level, student would do with the same kind of mentorship presenting at conferences publishing their work, so it's a very exciting way to work with undergraduate students. and it's growing in popularity in america kelsey's project is really interesting and we're happy to present it today in that it deals with both the ins and outs and a different way of thinking about how we enter communities, and a broader sense of community that we can address.

DI transcription	
1.	we’re very exCIted to be HERE in this conference
2.	because the THEME
3.	INS and OUTS of professional DIScourse
4.	is very CLOSE to our HEARTS
5.	as we’ve been WORking for the last two years on KELsey’s
6.	eh
7.	VEry important
8.	underGRAduate research PROject
9.	and to give you a little information aBOUT underGRAduate research in the united states
10.	it’s a
11.	it’s a MOVEment of engaged LEARning
12.	where underGRAduates LIKE kelsey
13.	deVElop their own reSEARCH questions
14.	colLECT their own DAta
15.	and preSENT their own FINdings
16.	under the MENtorship of the FAculty members
17.	so they are DOING the SAME work
18.	that a MASter’s or a PHD level student
19.	WOULD do
20.	with the SAME kind of MENtorship
21.	preSENting at CONferences
22.	PUblishing their WORK
23.	so it’s a very exCIting way to WORK with undergraduate students ↘
24.	and it’s GROwing in popuLARity in america ↘
25.	$^{\text{KEL}}$sey’s project is really INteresting
26.	and we’re HAPpy to present it toDAY →
27.	in that it deals with both the INS AND outs
28.	and a DIFFerent way of thinking aBOUT
29.	how we ENter com$^{\text{MU}}$nities
30.	and a BROADer sense of community that we can adDRESS ↘

Gestures and head movements

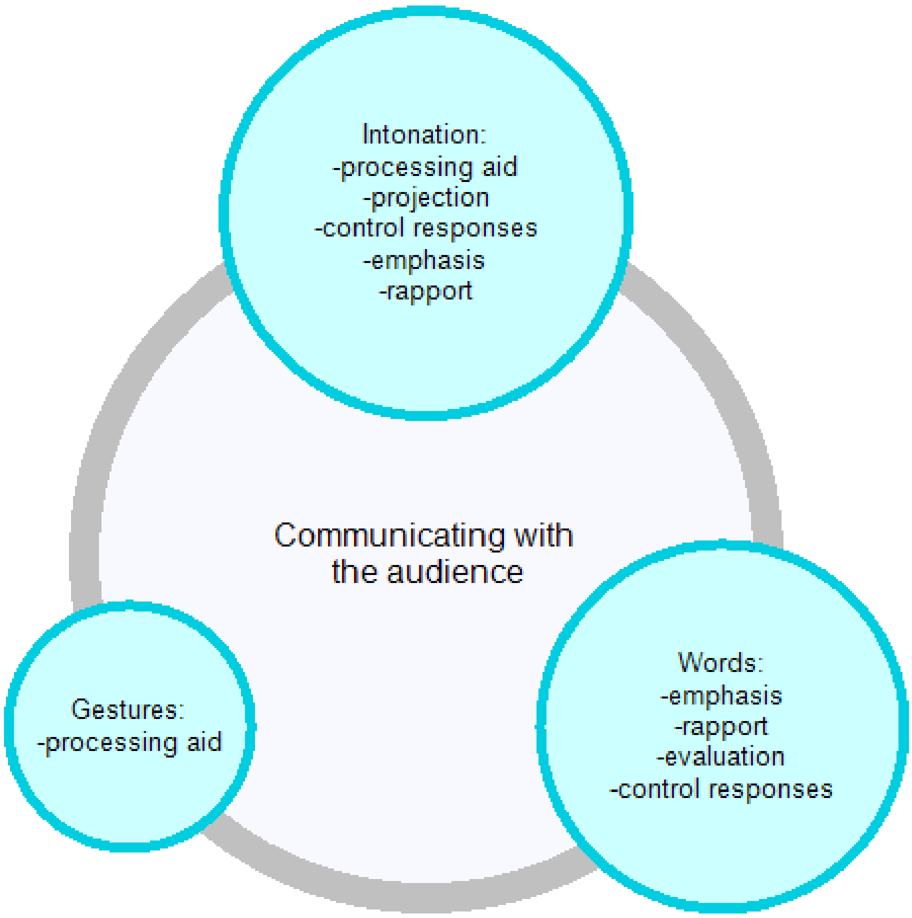
preSENting at CONferences

$^{\text{KEL}}$sey’s project is really INteresting

in that it deals with both the INS AND outs

and a BROADer sense of community that we can adDRESS

Multimodal ensemble



Analysis of words:

In her word choice the speaker uses intense language that focuses on emotions (e.g. “excited to be here”, “close to our hearts”, “happy to present it today”). This emotionally intense language contributes to rapport. She also evaluates their research in positive terms using words that again appeal to emotions rather than reasoning (“very important research”, “engaged learning”, “exciting way to work”, “growing in popularity”, “really interesting”).

This way of describing undergraduate research makes it sound likeable, even if the audience has just learnt about it and most probably does not have enough information to judge it rationally. The speaker explained that ‘engaged learning’ is a term they use widely in their university, so it is not evaluative in this case. The rest of evaluations are meant to describe undergraduate research as ‘rigorous’ and ‘worthy of respect’. To that aim she made a parallelism with the mentor-graduate student relationship, which she expected most people in the conference to be familiar with.

The use of a three part list when she describes what undergraduates do (develop research question, collect data, present findings) also contributes to make the text memorable. The speaker also added that the choice of elements in the list has to do with presenting a familiar analysis framework that would give validity to undergraduate research.

According to the speaker both evaluative and emotional language are connected to a need to make sure that the topic is presented as “valuable, valued, and authentic”. They had been engaged in it for a long period of time and it was something that sincerely mattered to them. The speaker commented in the interview that she is not usually that emotional, and thinks her use of intense language is the result of genuine excitement in this case. She agrees with the intended effect of creating rapport. As she explained, she wanted to convey “this is something we care about, so we hope you care about it too”.

The speaker uses two references to the title of the conference both at the beginning and the end of the excerpt. With this she provides a circular structure that contributes to memorability, and also shows how her research fits in the conference. According to the speaker the references to the event are intentional because she was not sure the European audience would see the connection with the topic of the conference.

Analysis of intonation:

The high key in unit 25 indicates discourse initial and is a signposting device to show that she is now providing a specific example of what she has been describing (i.e. undergraduate research).

A high termination in “communities” in unit 29, reinforced by an immediately following pause, presents this concept as innovative, contrasting with the fall tone and mid key of the final utterance which prompts concurrence and presents the suggestion of understanding communities in a broader sense as perfectly reasonable and out of the question.

The fall tone in unit 23 combined with a parallel fall tone in unit 24 suggests the information in unit 24 is an addition of a detail the speaker almost forgot to include, or is mentioning incidentally. Also the intonation in unit 26 is parenthetical (flat tone and faster speed) and makes this unit sound like an aside. Both provide the text with spontaneity. The speaker agrees with the ‘last minute’ addition interpretation, and in fact remembers it is something she was advised to add, not something she designed herself, which can explain why she initially forgot about it.

The use of prominence in unit 27 is also interesting, because it emphasizes "and", providing a sense of duality (paraphrasable by “not only ins, but also outs”). Interestingly, this contrast “in vs out” is visually represented through gestures, as explained in the following section. The speaker explained in the interview that this is a consequence of the way she interpreted the call for papers, which seemed to indicate expectations of papers focusing on either one aspect or the other, while hers focused on both and she felt the need to justify or clarify this.

Analysis of gestures and head movements:

As she refers to undergraduates presenting at conferences she points to Kelsey as an example. This is a deictic referential gesture synchronous with the prominence in “conferences”. Note how gaze is also directed to Kelsey at this moment. The speaker explained in the interview that she wanted to make sure “the attention was on her, I was just context”. She agrees with the visual representation interpretation and thinks this was a way of personalising undergraduate research, saying “here you have an example, she will be talking to you in a minute”. She repeats the same deictic gesture as she introduces Kelsey’s project. The gesture is synchronous with the prominence in "Kelsey", which is also pronounced with high key.

As she pronounces “ins and outs” both hands move in OHS beats first to her right and then to her left, visually signalling the contrast established. The emphasis on AND reinforces the meaning, announcing the second element in the comparison. The beats are synchronous with the prominences: “INS", "AND". It is a cohesive gesture that makes the text easy to follow.

A circular movement with open hands that expands outwards as she pronounces “broader” visually represents the expansion implied by the word. It is an OHS, metaphoric, referential gesture that clarifies part of the content.

Analysis of persuasive activity:

Regarding modal configuration, there is minimal disengagement from the action of communicating when momentarily focusing on encoding revealed through words (fillers and false start) and oblique intonation in units 6 and 10. The rest of the time the speaker is focused on communicating and uses intonation to a great extent to make her text easy to understand, her descriptions likable, and her evaluations acceptable (at least at an emotional level). Words are used to a slightly lesser extent to make her text memorable and emotionally appealing, as well as to fit in the event. Gestures are used much less in this excerpt, and they are specialised in facilitating understanding.

There were a couple of recurring ideas that were raised during the interview and which are worth mentioning here:

One is an awareness of *belonging to a different community* where research is done differently. The speaker expected differences regarding the way of talking about professional and business communication. As she puts it, "we come from different academic cultures" and have different "research practices". In Europe research is more linguistic-based, and rhetorical strategies are analysed in a different way. This translates into noticeable efforts to fit in. They become evident in the references to the event, which the speaker reports to be intentional because she was not sure the European audience would see the connection with the topic of the conference. The rapport-building also has to do with "knowing I was an outsider in that community, so, I'm just making sure that, it's building the rapport to show that I was happy to be there even if our work is a little bit different, maybe."

The other is a focus on *"paving the ground" for her undergraduate student,* something which starts from the preparation of the presentation. According to the speaker, this preparation was more thorough than what she is used to now, because she adapted to the needs of her student co-presenting. As the speaker explained, she helped her student prepare her content and PPT and then "wrapped mine [the speaker's] around it". They held a mock presentation with their colleagues and incorporated some feedback. It is interesting that this feedback which was incorporated at a later stage and is not part of the original material prepared by the speaker comes through with distinct intonation patterns. These patterns reveal that these pieces of information were almost skipped, and the speaker says them only after her notes remind her of them. The focus on her student also translates into deictic gestures that shift the attention to her.

In short, the speaker thinks her way of presenting reveals that “there was a lot about making sure people understood why we were there, that we fit, and that the work we do was valid and rigorous, even though she is an undergraduate student. So a lot of it wasn’t, a lot of it is not about me, this is me setting up Kelsey, so that people would respect Kelsey when she speaks”.

CPKE1

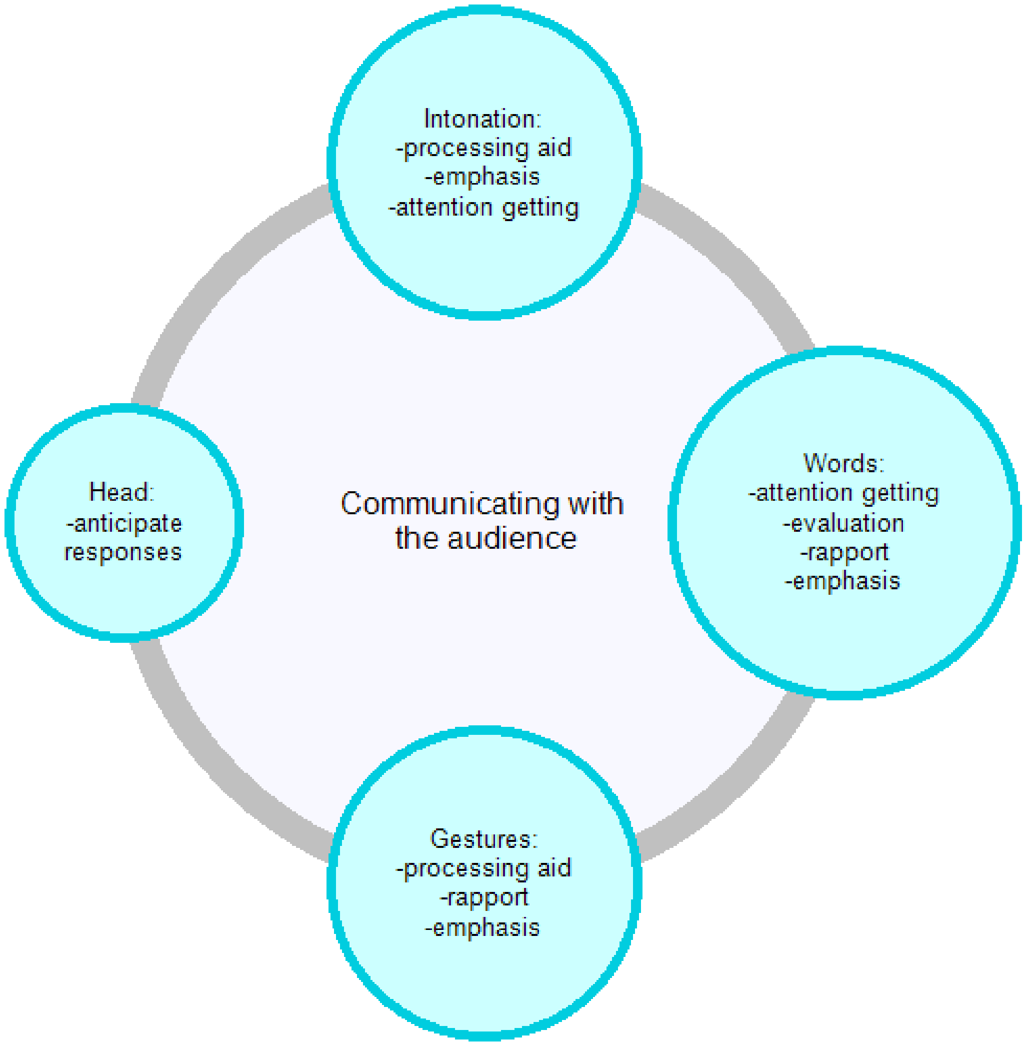
In this rich point the speaker describes her experience disseminating her research project.

Orthographic transcription
So the APA Annual Meeting was the best opportunity I could really find to disseminate my research. There are thousands of people there, the whole occupational community gathers to listen to talks. It's largely like this, but for the psychiatric community. I presented a poster in Young Investigator's Poster Session and that's my poster right there, it was actually interesting because I seemed to be the only person that wasn't a psychiatrist there, and a lot of people would walk by and they would double-take and look, like almost that I didn't belong there. But once I was able to talk to them about it, a lot of them were very interested on the impact that rhetoric has on psychiatry and how they could use that practically.
MICASE transcription
so the apa annual meeting was the best opportunity i could really find to disseminate my research eh there are thousands of people there the whole eh occupational community gathers to eh listen to talks, it's largely eh like this but for the psychiatric community. i presented a poster in young investigator's poster session, and, eh that's my poster right there it was actually interesting because, i seemed to be the only person that wasn't a psychiatrist, there, and a lot of people would walk by and they would double-take and look ... like almost that i didn't belong there but, once i was able to talk to them about it a lot of them were very interested on the impact, that rhetoric has on psychiatry and how they could, use that practically.

DI transcription	
1.	so the Apa annual meeting was the BEST opportunity i could really find
2.	to disSEminate my reSEARCH
3.	eh →
4.	there are THOUSands of PEOple there
5.	the WHOLE
6.	eh →
7.	occupational community GAthers to
8.	eh →
9.	listen to TALKS
10.	it's LARGEly
11.	eh →
12.	like THIS
13.	but for the psyCHIAtric community
14.	i presented a POSter
15.	in young investigator's POSter session
16.	and eh →
17.	THAT'S my poster right there
18.	it was ACtually INteresting because
19.	i SEEMED to be the only person that WASN'T a psychiatrist there ↗
20.	and a LOT of people would WALK by
21.	and they would DOUble-take and look
22.	like almost that i didn't beLONG there
23.	but ONCE i was able to TALK to them about it
24.	a lot of them were VEry interested
25.	on the IMpact
26.	that RHEtoric has on psyCHIAtry
27.	and how they could USE that PRACtically

Gestures and head movements
 there are THOUSands of PEOple there  it's LARGEly eh like THIS but for the psyCHIAtric community  a lot of them were VEry INterested

Multimodal ensemble



Analysis of words:

Personal pronouns (I) highlight the figure of the researcher. She presents her research as an interesting narrative, occasionally using intonation and facial expression to momentarily "enact" characters in her story. There is a reference to the event (in the comparison APA vs. current event) that seeks to fit in. There is also positive evaluation in her lexical choice "best opportunity", "interesting", "very interested", occasionally reinforced through beats.

Analysis of intonation:

High termination in unit 19 adds a sense of surprise and represents the reaction of the people seeing her poster. At the same time a very marked rise tone projects the communicative dominance of more experienced researchers questioning the validity of her research. A micro pause before “there” in unit 19 emphasizes the deictic reference to the event she is describing (which she has also previously compared to the event where she is giving this presentation). There is a two second pause after unit 21, which is filled by a facial expression that mimics the reaction of surprise of the participants in the narrative that she is creating about her experience of presenting a poster. The prominences in units 23-27 emphasize the final positive reaction towards her research (after initial surprise and lack of understanding). Throughout the excerpt there are occasional instances of oblique intonation synchronous with fillers.

Analysis of gestures and head movements:

The extreme quantification “thousands of people" as she describes the APA event is accompanied by a head shake that denies any potential counter arguments. It is a pragmatic gesture that makes her narrative more vivid.

The reference in “largely like this” is clarified by an OHS deictic gesture to the audience. This gesture is repeated three times, and it is meant to be synchronous with the prominences in “largely”, “this” and “psychiatric”. However, this is only achieved in the first and last case, because the hesitation and filler preceding "like this" slows down the flow of words. The gesture is not withheld to match "this" and actually occurs synchronous with the filler. The gesture is partly referential and partly pragmatic in its inclusive effect.

Finally, two closed hand beats moving downwards, the first synchronous with “very” and the second with “interested” in “very interested", emphasize the extreme evaluation.

Analysis of persuasive activity:

As the speaker is focused on communicating with the audience, she makes use of gestures to clarify her content, emphasize certain elements and include the audience. She also uses a head shake to anticipate a response. Words help her create a vivid narrative that sincerely shares her research experiences and evaluates it positively. Intonation is used to vividly convey the reactions of the characters in her narrative.

As some of the attention of the speaker is devoted to encode the message, intonation becomes oblique and fillers are used to gain time, so we can say that the action of encoding temporarily uses the resources of words and intonation.

CPPE1

In this rich point the speaker is explaining challenges in her research through an honest sharing of her experiences.

Orthographic transcription

Specific for my case was that I found tension between organizational culture or practices and what I wanted to do as an ethnographer, I wanted to develop contacts with clients and more personal relationship and have deep and long conversations but the agency, their approach to the female clients is more distanced, commercial, and if there must be personal contact they want to keep it to a limit.

MICASE transcription

specific eh for eh my case was eh that i found_found tension between organizational eh culture eh, or practices and what i wanted to do as an ethnographer eeeh i wanted to eh develop contacts with clients and more personal relationship and have deep and long conversations but the agency their approach to the female clients is more distanced eh commercial and eh if there must be personal contact eh eh they want to keep it to_to a limit.

DI transcription

1. speCIfic eh for eh my CASE
2. WAS
3. eh
4. that i found found TENsion beTWEEN
5. organizational eh $^{\text{CUL}}$ture eh or practices
6. and what I wanted to DO as an ethnographer
7. eeeh
8. i wanted to eh develop CONtacts with CLIENTS
9. and more PERsonal reLAtionship
10. and have DEEP and LONG conversations
11. but the Agency
12. their apPROACH to the female clients is more DIStanced eh
13. comMERcial
14. and eh if there MUST be personal contact
15. eh eh
16. they want to KEEP it to to a LImit.

Gestures and head movements

and what I wanted to DO as an ethnographer

i wanted to eh develop CONtacts with CLIENTS

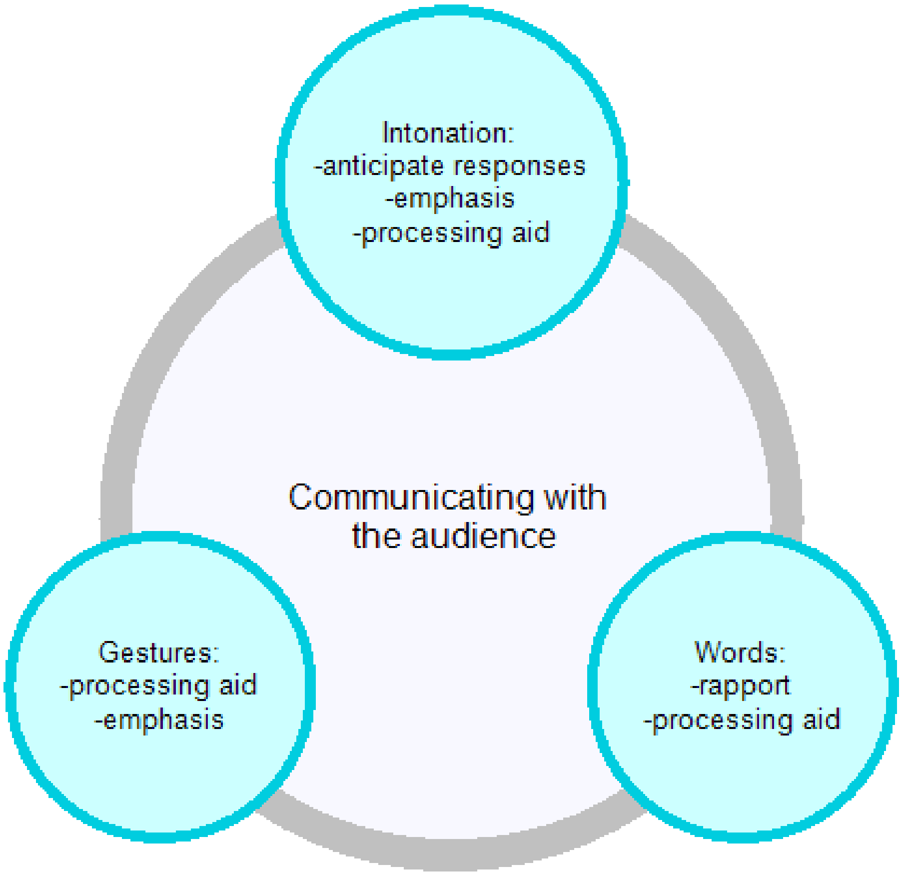
and have DEEP and LONG conversations

their apPROACH to the female clients is more DIStanced eh

comMERcial

they want to KEEP it to to a Limit

Multimodal ensemble



Analysis of words:

The use of first personal pronoun draws attention to the figure of the researcher, as does the use of the verb “want”, which suggests a personal desire rather than a requirement of the research process. This makes the text closer to a narrative of a personal experience. The vocabulary choice also expresses the two attitudes she is contrasting: on the one hand her stance as a researcher, and on the other hand the position of the company. An example of this are the verbs used to refer to “contact”: “develop” (researcher) vs. “keep to a limit” (company)

Analysis of intonation:

The high termination in “culture” in tone unit 5 together with the rise tone anticipates and prepares the audience for the contrast with “what I wanted to do” in the following tone unit. Prominence in the words that describe the attitudes of both parties further emphasizes this contrast (e.g. “contacts”, “personal”, “deep”, “long” vs. “distanced”, “commercial”)

There are numerous fillers and two fall starts in the extract. They reveal encoding difficulties, and the speaker confirmed during the interview that she was struggling to find the right words. In this particular case rather than spoiling the persuasive effect this actually adds to it, because it reinforces the spontaneity and honesty of her text and establishes a closer relationship with the audience.

Analysis of gestures and head movements:

As the speaker pronounces “I” in “what I wanted to do” she points to herself, visually directing the attention and highlighting the figure of the researcher. It is an OHS, deictic, referential gesture in phonological, pragmatic and semantic synchrony with words.

An OHS gesture with hands alternating and moving in outward circles is repeated twice as she pronounces “contact with clients” and “deep and long conversations”. It is phonologically synchronous with the prominent syllables in “contact”, “clients”, “deep” and “long” (they start slightly before the prominence and continue throughout). They are metaphoric, referential gestures that represent visually the social interaction that is implied with the lexicon used. The fact that it is repeated also provides it with a cohesive function, since it connects different words that are related (in this case lexicon referring to social interaction). During triangulation of results the speaker also interpreted this gesture as a processing aid that helps her encode the verbal component of her message.

There is an OHP gesture that begins at the prominence in “their apPROACH to female clients is more DIStanced” and continues up to the middle of the next tone unit (“commercial”). It is a metaphoric gesture that symbolises the “distance” referred to (as if saying “stop coming closer”), therefore in semantic synchrony with words.

Finally, the speaker makes an OHP gesture that precedes and continues up to the first prominence in “KEEP it to a LImit". It is an iconic, referential gesture that visually represents a barrier the company is imposing to restrict personal contact with clients. It is interesting that the gesture in this case is not in semantic and pragmatic synchrony with words, but rather precedes them. This is likely to respond to encoding difficulties (gestures are easier to encode and therefore faster to produce than words).

Analysis of persuasive activity:

The two key aspects in this passage around which persuasion is built are:

- A contrast between what the researcher wants to do vs. what she is allowed to do. All the modes examined contribute to this effect: words (develop vs. keep to a limit), intonation (rise tone that highlights the contrast at the crucial point in the transition from unit 4 to 5) and gestures (contrast between the rotating hands symbolizing flow of social interaction vs. the barrier gesture at the end)

-A personal and honest sharing of research experiences, again supported by all modes. It is remarkable how the "I" in "I wanted to do" is prominent and further emphasized by a self-deictic gesture. Other paralinguistic features (e.g. vocalisations as fillers) also contribute to make her text not rehearsed, honest and natural.

As the speaker reported during the interviews, she was unsure of whether her research would fit the conference. Her efforts to seek rapport and honestly sharing her research experience are probably a consequence of this perception.

CPPE2

In this rich point the speaker is explaining challenges in her research through an honest sharing of her experiences. In this case the focus is on ethically acceptable ethnographic research practices. In fact, during the interview the speaker explained her main goal with this presentation was to disseminate this particular research area (ethical research practices in ethnography), which she considers a gap.

Orthographic transcription
So when I first arrived in Saint Petersburg, I really had the impression that I was in other ethical context. So here regarding the informed consent, I was also advised not to tell about my research which was really strange because in Europe, it was the first question that I got when I was talking with people in the university and outside the university, the people know about your research. In Russia it was different, so first I was surprised by the contrast, but later on these contrasts like faded away and you become aware of your own- yes, values and your own professional and personal moral values, and you have to find a negotiation and a way to deal with these different frames and these different ways of interpretation.
MICASE transcription
so when i first arrived in, Saint Petersburg eh i_i really had the impression that i was in other ethical context. so here eh regarding the informed consent i was eh also advised not to tell about my research, which was really strange because in, europe eh it's_ it was the first question that I, got when i was talking with people in the university and outside the university, the people know about your research, in russia it was, different so, first i_ i was, surprised by the contrast but later on these contrasts like faded away and eeeeh you become aware of your own eh, yes, values and your own eh professional and eh personal moral values, and you have to find a negotiation and, a way to to deal with these different frames and these eh different eh ways of of interpretation

DI transcription
1. so when i FIRST arrived in in saint petersburg 2. eh i i really had the impression that i was in Other ethical CONtext 3. so HERE 4. eh 5. regarding the informed consent 6. i was eh also advised NOT to TELL about my research 7. which was really STRANge 8. because in EURope 9. eh 10. its it was the FIRST question that I GOT 11. when i was talking with people in the university and outside the university 12. the people KNOW about your REsearch 13. in RUssia it was DIfferent 14. so FIRST i i was surprised by the CONtrast 15. but later on these contrasts like FAded away 16. and eeeeh 17. you become aWARE of your OWN 18. eh 19. yes VAlues 20. and your OWN eh professional and eh personal moral VAlues 21. and you have to find a negoTIATION 22. and a WAY to to DEAL with these different frames 23. and these eh different eh WAYS of of interpretation
Gestures and head movements
 eh i i really had the impression that i was in Other ethical CONtext

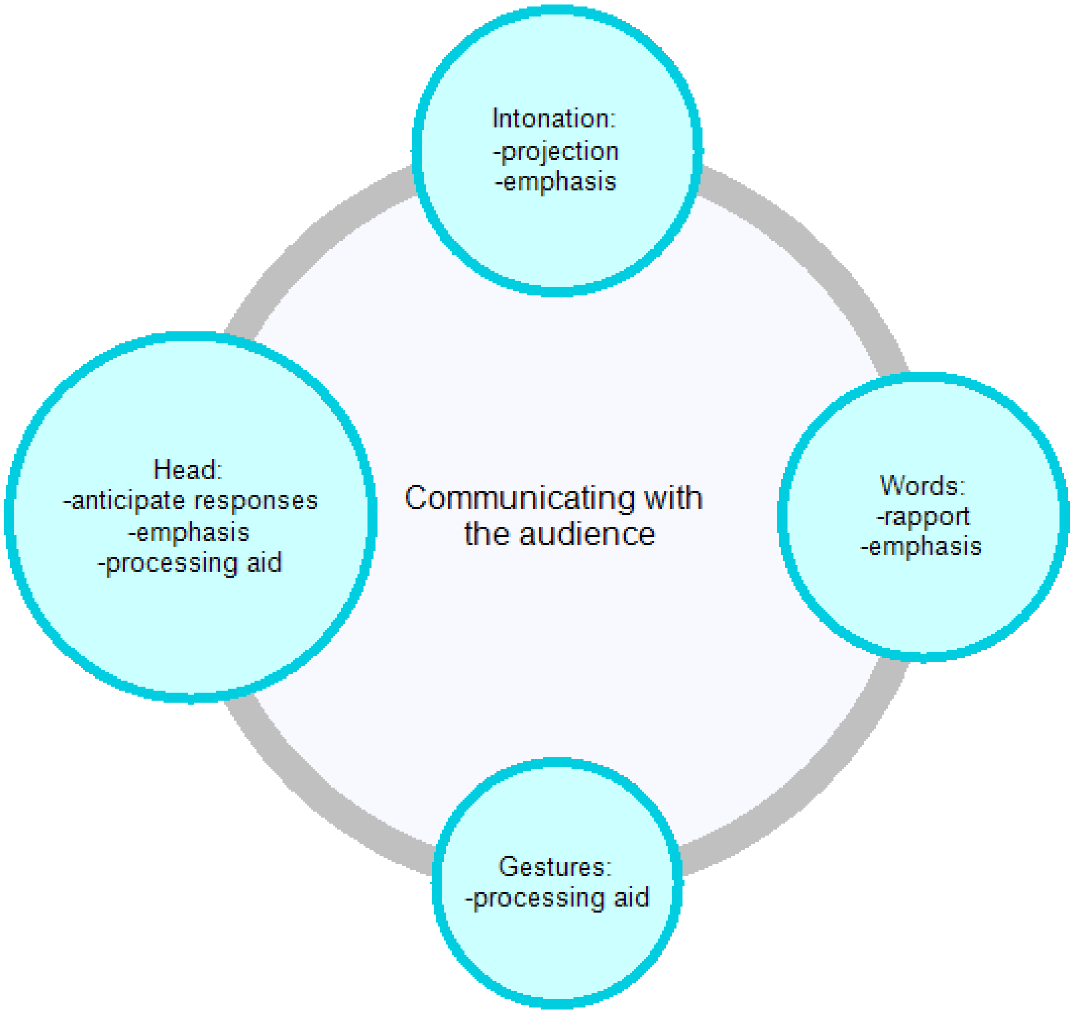
i was eh also advised NOT to TELL about my research

because in EURope

the people KNOW about your REsearch

you become aWARE of your OWN

Multimodal ensemble



Analysis of words:

The lexicon used is personal and subjective, mainly because of the experiential processes referred to by the verbs (“I had the impression”, “I was surprised”, “fade away”, “become aware of your own values”), bringing to the fore the figure of the researcher and her mental/sensory experiences. The speaker is describing her research as a narrative of her personal experience and not as an objective account of the methodology used.

Analysis of intonation:

Prominence in “not to tell” (tone unit 6) “know” (tone unit 12) and rise tone in "surprised by the contrast" (tone unit 14) emphasize the comparison of both scenarios (European ethical standards vs. Russian ethical standards) and anticipates surprise. This can be related to an effort to prove that research is valid (i.e. done according to accepted practices)

Analysis of gestures and head movements:

As she pronounces “other ethical context” she represents the shape of a ball with both hands, metaphorically showing the scope of the context she is referring to. This is a referential OHS gesture that begins with the first prominence in tone unit 2 (“other”) and continues till the second prominence in “context”

The speaker also nods as she pronounces “not to tell”. It is interesting that the nod is synchronous with the verbal negative, which might seem contradictory a priori. It actually seems to anticipate and compensate for disbelief from the audience, as if saying “believe it or not this is common practice in Russia”. In fact, the speaker reported during the interview that she did not expect her audience to be familiar with this research practice.

Next the speaker shakes her head subtly as she pronounces “because in Europe" (difficult to perceive in a still image). The movement is synchronous with the prominence in “Europe”, and seems to visually represent the rejection of the practice she is referring to (i.e. hiding objectives of research to participants) in the European context. The speaker actually confirmed this interpretation during the interview. A head nod (again not easily seen in a still image) also reinforces “know” (which is also prominent) in "people know about your research". This is a way of further emphasizing the contrast she is making: Europe vs. Russia.

An OHS metaphorical, referential gesture (rotating hands inwards) precedes the first prominence and continues up to the second prominence in the tone unit “and you become aWARE of your OWN”. It symbolizes the process of adaptation to a new scenario, and the fact that it is inwards is a reference to the person adapting (the researcher). The fact that is precedes words can be explained by encoding effort (it takes longer to encode words). In fact, during triangulation of results the speaker also interpreted this gesture as a processing aid because she was “wrestling with getting it said”.

Analysis of persuasive activity:

In this example the speaker shares research practices in an honest way that may prompt identification from the audience and additionally proves that research has been done according to accepted practices. Words highlight the figure of the researcher through the references to experiences, while intonation and head movements emphasize the contrast she is establishing. In addition, gestures and head movements visually represent part of the content and make it easier to understand.

Paralinguistic features such as pauses, fillers false starts and repetitions reveal encoding effort. This provides the text with a spontaneous character that makes it sound honest. It is interesting to note that after triangulation of results the speaker admitted that she found confronting to see the overall effect of the modes she was using, because it showed how explicitly she was positioning herself as an outsider with respect to the object of her study and its cultural context, while in fact it is something she usually tries to avoid. In a way this shows how the speaker was honestly sharing experiences rather than describing a methodology in an impersonal way.

DTLA1

In this rich point the speaker is describing the potential applicability and benefits of her research using the example of adapted jobs. It corresponds to minute 3-4, towards the middle of the presentation.

Orthographic transcription

And how about companies or institutions as adaptive systems? How many people would be happier if instead of being forced into predetermined roles and a predetermined career progression, they could have a job that was customised to them, to what they are best at, to their passions, to what motivates them? And this vision might seem a bit far-fetched to you now, but people are working on some of these things, and little by little we can get to this reality.

MICASE transcription

and how about, companies, or institutions as adaptive systems. how many people would be happier, if instead of being forced into predetermined roles, and a predetermined career progression, they could have a job, that was customised to them, to what they are best at, to their passions, to what motivates them. and this vision, might seem a bit far-fetched to you now, but people are working on some of these things, and little by little, we can get to this reality.

DI transcription

1. and HOW aBOUT
2. COMpanies or instiTUtions
3. as aDAPtive SYStems
4. how many PEOple would be HAppier
5. if insTEAD of being FORCED
6. into PREdetermined ROles
7. and a PREdetermined caREER progression
8. THEY could have a JOB
9. that was $^{\text{CUS}}$tomised to THEM
10. to what their BEST at
11. to their PASsions
12. to what MOTIvates them
13. and this $^{\text{VI}}$sion might seem a bit far-FETChed to you now
14. but PEOple ARE working on some of these things
15. and LITtle by LITtle
16. we can GET to this reAlity

Gestures and head movements

that was [CUS]tomised to THEM

to what their BEST at

to their PASsions

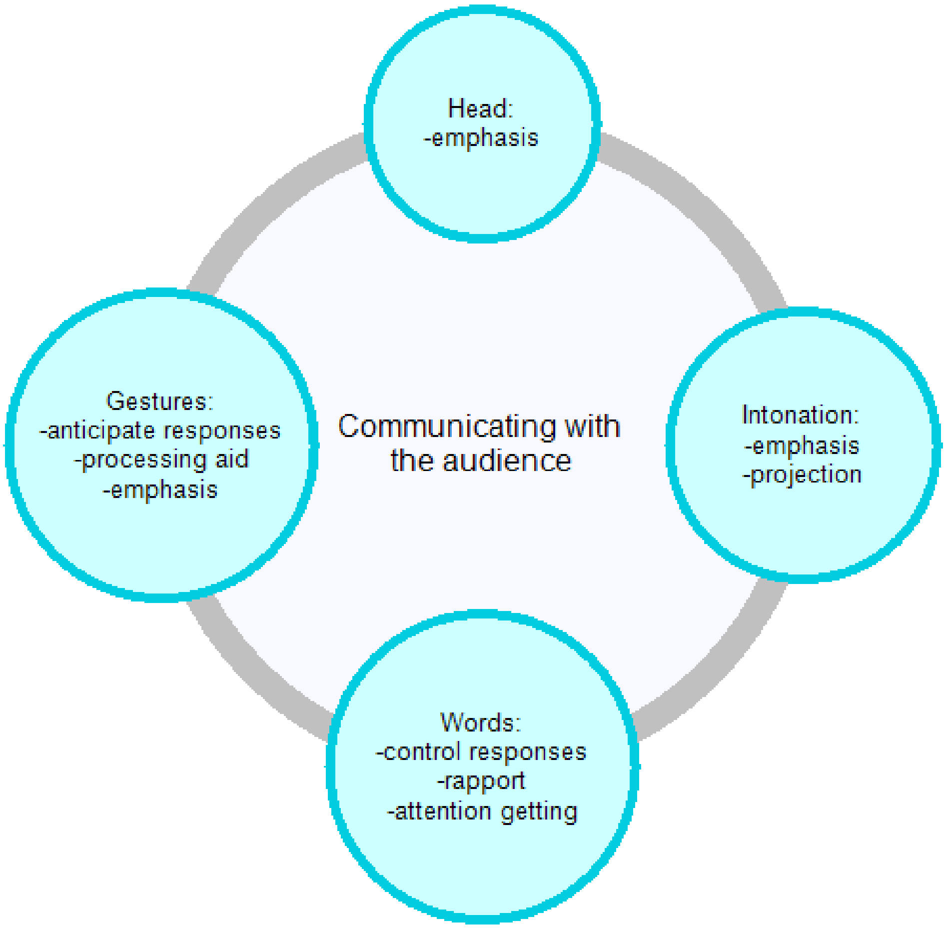
to what MOTIvates them

and this [VI]sion might seem a bit far-FETChed to you now

but PEOple ARE working on some of these things

we can GET to this reAlity

Multimodal ensemble



Analysis of words:

The rhetorical questions guide the audience to a particular line of though and also bring the topic closer to their experience. The questions are framed to prompt acceptance of the idea the speaker is suggesting and taking this potential benefit for granted (note how she says "How many people would be happier" and not "People would probably be happier”). The hedging modal and adversative combination in "Might seem" + “but" mark that she is anticipating a reaction from the audience and refuting it. Also the use of the personal pronoun "we" creates inclusiveness.

Analysis of intonation:

The use of high key in "customised to them" (+ 41 Hertz) and "this vision" (+ 114 Hertz) suggests reversal of expectations and adds to the contrast she is presenting (what we now have vs what we could achieve). "Are" in "people are working" is prominent (and further emphasized by a closed fist beat and a slight nod as mentioned in the gestures section). This stresses the feasibility of this line of research and the plausibility of the option she is suggesting.

Analysis of gestures and head movements:

"Customised" and "them" are emphasized through open hand beats in an attempt to highlight the innovative contrast they present with the current scenario (i.e. predetermined roles). Their function can be considered pragmatic and they show phonological and pragmatic synchrony with words and intonation (the words are also prominent and therefore emphasized).

Next the speaker enumerates with fingers as she lists the aspects that can be customised in a job. This cohesive gesture makes this part of the message easy to process.

Then, as she says "far- fetched", a vertical hand gesture with palms open and facing to the audience (open hand prone) meaning "stop" visually represents the reaction she is anticipating from the audience (paraphrasable by "stop, you are getting carried away"). At the same time, according to the speaker, this gesture also works as a visual representation of the concept of far. This is a metaphoric, referential-pragmatic gesture that is meant to facilitate understanding of the message, and is semantically and pragmatically synchronous with words (far-fetched).

As she pronounces "people are working" she emphasizes "are" with a close fist beat and a slight head nod. This stresses the feasibility of her idea. As in the previous example of beats their function can be considered pragmatic, and they show phonological and pragmatic synchrony with words and intonation ("are" is also prominent and therefore emphasized).

Finally, as she pronounces "get" in "get to this reality"· she extends her right hand forward with palm facing down (open hand prone), as if visually representing this progress and facilitating the processing of the message. This is an iconic, referential gesture that shows pragmatic, semantic and phonological synchrony with “get” (“get" also receives prominence).

Analysis of persuasive activity:

This presentation achieves persuasion by prompting the audience to reflect and consider options they had not contemplated before. The speaker reported in the interview that she was inspired by a TED talk about motivational speeches that presented a contrast between what currently exists and what can be achieved (e.g. Martin Luther King “I have a dream”, Steve Jobs Keynotes). She tried to reproduce this effect in her presentation.

This speaker is subtly prompting the audience to share her positive evaluation of her proposal (i.e. what can be achieved with her research). Note how this evaluation does not reveal itself in an explicit way (i.e. through evaluative language), but comes across in a more subtle way through the inferences created and the expected responses of the rhetorical questions. In doing this she is highlighting the relevance of her research and its potential to become life-changing. At the same time she is also preventing a potential counterclaim that her suggestion might not be feasible. All the modes analysed contribute to this persuasive effect. Rhetorical questions are carefully framed to trigger a particular response from the audience. The anticipation of the counterclaim implied by the use of the hedging modal “might seem” and the adversative "but" is supported by the hand prone gesture that metaphorically represents this reaction of the audience and makes the message easier to process. Beats and also head nods serve to emphasize key parts of her message. At the same time, intonation highlights a contrast between real vs. desired situation and the feasibility of what she is proposing.

DTLA2

In this rich point the speaker states the unlimited potential of her research. She also describes a change in attitude that is required for her research to be beneficial. It is a call for action. It immediately follows rich point 1, during minute 4-5 in the presentation.

Orthographic transcription
And these are just some ideas. The possibilities behind adaptive systems are only limited by our own imaginations. And we have been taught that everybody should be treated the same. So, a revolution needs to happen within each and every one of us before it can be extended to the rest of the world. We need to believe that this can happen. And we need to demand as a society to be acknowledged as individuals and be treated as such.
MICASE transcription
and these are just some ideas. the possibilities behind adaptive systems are only limited, by our own imaginations. and we have been taught that, everybody should be treated the same. so, a revolution needs to happen, within each and every one of us, before it can be extended to the rest of the world. we need to believe, that this can happen. and we need to demand as a society, to be acknowledged as individuals, and be treated as such.
DI transcription
1. and THESE are just some IDEAS 2. the possiBILities behind adaptive systems are only LImited 3. by OUR $^{\text{OWN}}$ imaginations 4. and we have been $^{\text{TAUGHT}}$ that 5. everybody should be TREATED the same 6. SO 7. a revoLUtion needs to HAppen 8. within EACH and Every one of us 9. before it can be exTENded to the rest of the WORLD 10. we NEED to beLIEVE 11. that this can HAppen 12. and we need to deMAND as a soCIETy 13. to be ackNOWledged as indiVIduals 14. and be TREAted as SUCH

Gestures and head movements

by OUR OWN imaginations

and we have been TAUGHT that

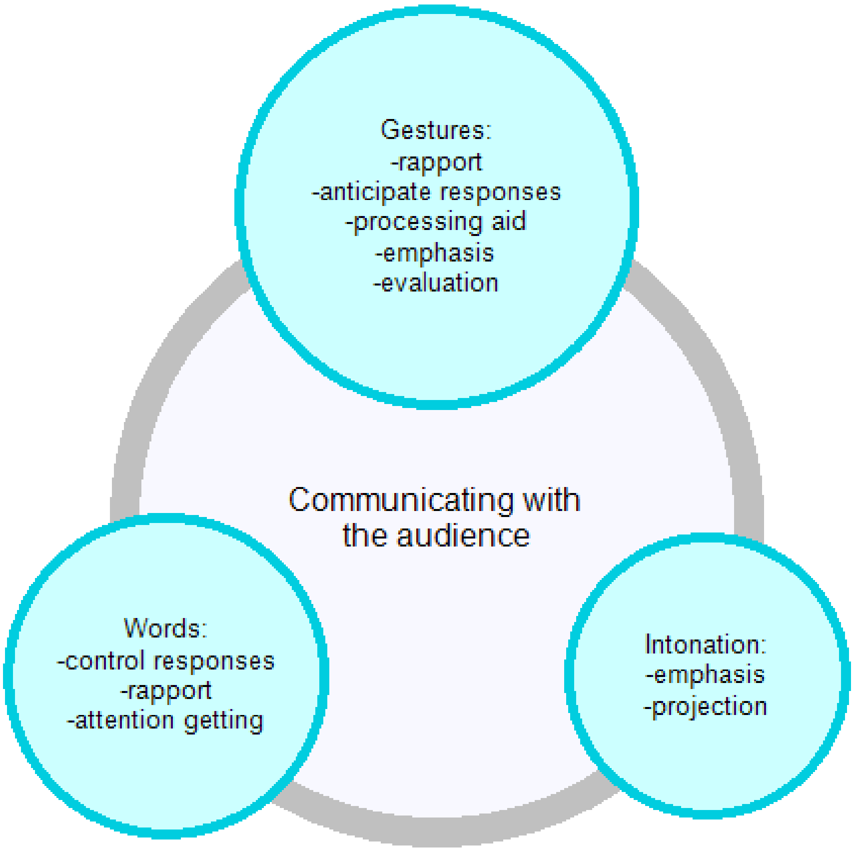
within EACH and Every one of us

everybody should be TREATED the same

before it can be exTENded to the rest of the WORLD

we NEED to beLIEVE

Multimodal ensemble



Analysis of words:

The word choice shows intense language (“revolution”, “demand”) and inclusive language (“we”, “us”) used throughout. The direct addresses to the audience in imperative form add to the intensity and urgency e.g. "we need".

Analysis of intonation:

Prominence in “our and “own" contributes to the sense on inclusiveness created by the use of pronouns. The emphasis attached to “need" (“we need to believe”) and “demand” (“we need to demand”) contributes to the intensity of the language. The speaker pointed out in the interviews that the prominence in “taught” was conscious because she wanted to convey the idea that “taught does not imply true”. The consistent use of fall tones in the first part of the extract (up to “before it can be extended to the rest of the world”) suggests that the speaker treats this part of her message as mostly new information, or rather shows an assumption that this is something the audience has not stopped to consider yet (the speaker actually confirms this interpretation). The rise tones in the following tone units seem to indicate that the message is not finished rather than assuming shared knowledge. Finally, the high termination attached to “our own" (+ 51 Hertz) and “we have been taught that" (+ 146 Hertz) shows the speaker expects the audience to find this pieces of information (i.e. the fact that possibilities are only limited by imagination and the fact of having being socially enculturated into certain views) as surprising or contrary to expectations. This is consistent with the pragmatic content of her message as a whole, which is prompting to reject traditional views and adopt a more innovative stance.

Analysis of gestures and head movements:

Self- deictic gestures are used to refer to both herself and the audience at three different points: as she says “our own imaginations”, “we have been taught” and “within each and every one of us”. This has an inclusive effect, since she groups herself and the audience together. It is a referential and pragmatic gesture that shows pragmatic synchrony with the use of inclusive pronouns that aim at engaging the audience.

Next, a “stop” sign with open hands, palms facing the audience (open hand prone, vertical) visually represents how society prevents people from considering other options than what she is expressing with words (i.e. "everybody should be treated the same"). This metaphoric gesture is both referential and pragmatic, and adds a sense of disagreement with this statement that is not explicitly said in words.

Then, an open hand prone gesture with both hands moving apart suggests the idea of spreading. This is an iconic, referential gesture that helps processing the message, and is semantically synchronous with "extended to the rest of the world".

Finally, as she pronounces “We need to believe” the word “believe” is prominent and further emphasized through a closed fist beat in pragmatic and phonological synchrony with words and intonation. This is a way of emphasizing her suggestion and adding to the intensity of the language.

Analysis of persuasive activity:

The modes analysed contribute to the following persuasive effect:

Through prominence and beats the speaker emphasizes parts of the message she wants the audience to focus on, and adds to intensity, which makes the message more memorable (as the speaker herself reports in the interviews she emphasized the “takeaways”). Using prominence in pronouns she builds an inclusive effect that creates rapport, which is further aided by self-deictic gestures. The speaker is also subtly and indirectly presenting the current situation as undesirable and consequently positively evaluating her suggested change and the benefits it might bring, thus showing the relevance of her research (but note how this evaluation is not explicit, but rather an inference). The use of gestures allows her to do this in a subtle and non-imposing way. Finally, visual representations also contribute to make the message easy to understand and as a consequence more likely to be accepted.

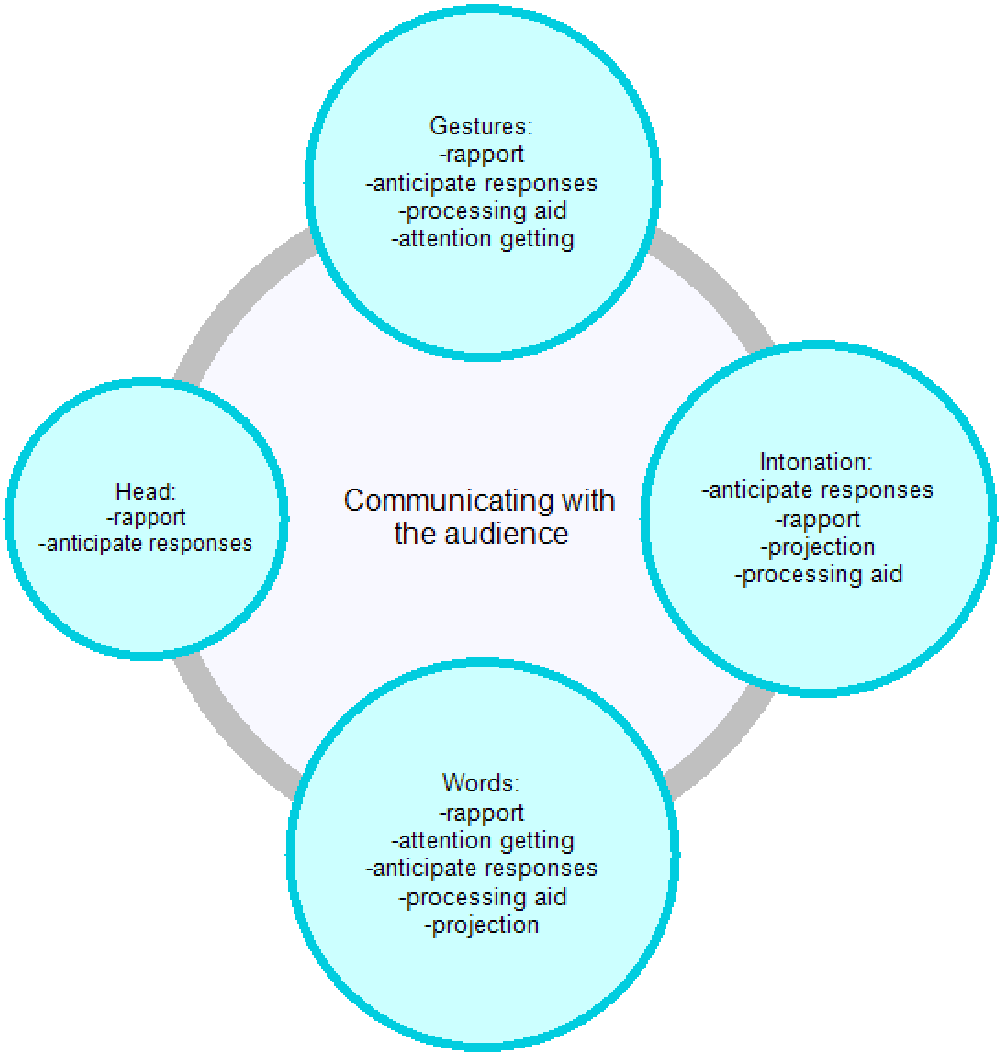
DTLI1

This rich point includes the starting lines of the presentation: a hook to get attention.

Orthographic transcription
Are you smart? I knew that this would come, but let me rephrase: Is your phone smart? You would say of course it is a smart phone so why do you ask. Well then I say, is the clothes you wear smart? The jeans the jacket, the clothes that you have...and it would be very silent.
MICASE transcription
are you smart <P03> i knew that this would come but eh let me rephrase, is your phone smart <P02> you would say of course it is a smart phone_phone so why do you ask. well then i say, is the clothes you wear smart, the jeans the jacket the clothes that you have, and it would be very silent.
DI transcription
1. ARE YOU SMART 2. i KNEW that this would come 3. but LET me rePHRASE 4. is your PHONE smart ↘ 5. you would say of COURSE it is a smart PHONE phone 6. so WHY do you ASK 7. well then I SAY 8. is the CLOTHES you WEAR smart ↘ 9. the JEANS 10. the JACKET 11. the CLOTHES that you have 12. and it would be VERY SILENT ↘

Gestures and head movements
 ARE YOU SMART  the JEANS the JACKET  and it would be VERY SILENT

Multimodal ensemble



Analysis of words:

Three things stand out in this speakers' use of words, 1) the direct, even abrupt, questions for the audience; 2) the high certainty conveyed by the verb "knew" or the adverbial phrase "of course" (which could be considered the opposite of hedging); and 3) the repeated use of conditionals with the auxiliary "would" to describe the hypothetical situation she is anticipating. The speaker uses direct questions to the audience presumably expecting a real answer. In doing so she engages in a conversation that creates rapport. This rapport is reinforced by her anticipation of the answers, something which she explicitly states, and the last line (it would be very silent) that adds a touch of humour. Interestingly, she explained during the interview that she chose a question as a hook because other types of hook such (e.g. an image) can distract the attention from the speaker, as she was informed in specialized courses on presenting skills. She also confirmed that her communicative intention in this case was to anticipate responses from the audience, and rather than being interested in the information she wanted to get them “activated”. The sequence of questions with symmetrical structures provides the text with parallelism and makes it memorable, and the topic is skilfully introduced by a pun with the meaning of the word “smart”. The use of humour can be considered a mismatch in terms of the expected register in a presentation, but it serves a persuasive purpose contributing to rapport.

Analysis of intonation:

In the first tone unit each word receives prominence, making each of them very salient. This is probably a way of marking this utterance as a real question for the audience to consider instead of rhetorical, and the pause following also contributes to this effect, and was reportedly consciously done. Interestingly, the speaker explained in the interview that this question, as well as the rest, was not a real question but simply a way of getting the audience "activated”. Since it was the first sentence in the presentation, she consciously pronounced it louder to get attention.

Contrastive prominence in phone and clothes mark the structure of discourse as a series of parallel questions that smoothly introduce the audience into the topic of smart textiles, playing with the meaning of smart.

The questions in unit 4 and 8 feature a fall tone (-79, -96 respectively) that suggests the speaker does not really know the answer. This marks them as real information questions and adds to the interactive style, but it is somehow contradictory with what she actually says with words and her use of key and termination, as explained below. From the feedback received from the speaker, it seems that the real intention is not to obtain information, but to obtain a reaction (i.e. make the audience respond and keep an active role). This reaction would not be possible if these questions were not perceived as real ones. This is where intonation plays a role in this case.

Unit 4 has a high pitch on “phone” (+49), which adds a nuance of surprise and matches the contrastive prominence with unit 1 (“you” vs. “your phone”). The speaker reported that she wanted to direct attention to the idea of a phone to build a bridge between a concept they know well (a smart phone) and her own topic (smart textiles). Unit 5 has low key and termination (gradual descent -32, -52,-65) and unit 6 also has low key (-65). This has the effect of presenting this potential answer she is anticipating as obvious and out of discussion. The speaker also chooses a low pitch as she lists the items of clothing in units 9-11 (-69,-24,-21 respectively), which is a way of equating them and including them within the same wider group of "clothes". This is a way of easing understanding.

The final fall tone in the last tone unit marks this utterance as a pun and prompts laughter. This effect is reinforced by the head nods referred to in the next section.

Analysis of gestures and head movements:

The beats with open arms (OHS) accompanying each prominent word in the first question (unit 1) seem to invite the audience in. They have an interpersonal function and reinforce the conversational style.

Deictic gestures as the speaker lists the items of clothing have a referential function and serve as visual examples, facilitating understanding. The speaker explained that she consciously decided to do this instead of showing pictures to retain the attention of the audience on herself.

The sweeping gesture with both hands palm down (OHP) accompanying “very silent” (unit 12) is a metaphoric, pragmatic gesture that prevents potential counterclaims to this hypothetical situation and/or represents the scope implied (the whole room). The speaker finds both interpretations quite plausible but finds the second one more likely, and adds a third potential meaning: metaphorical representation of silence. Interestingly, the gesture is split in two parts to follow the rhythm created by the two prominences in “very silent”. This rhythm is also reinforced by subtle head nods synchronous with the prominences. Apart from this reinforcing effect, the speaker interpreted these nods as confirming that she was indeed anticipating the right answers.

Analysis of persuasive activity:

In this extract persuasion is built mainly by means of creating rapport with the audience, raising their interest and anticipating their responses. The audience is also subtly guided towards a particular interpretation of a possible response as obvious. In addition, the text is made easy to follow.

All the modes analysed work towards these effects:

-The initial gestures with open arms contribute to create rapport while intonation marks the utterance as a real question inviting an answer.

-The parallelism in the wording of the questions is reinforced through intonation, in particular contrastive prominence.

-The sense of “obvious” suggested by the expression “of course” when anticipating the answer is also reinforced by intonation.

-The speaker makes the text easy to follow by pointing at visual examples of what she is referring to as she lists items of clothing while intonation contributes to bringing these items together within the same group. The deictic gestures also help keep the focus on herself.

-The final sweeping gesture, head nods and fall tone create a sense of definiteness that provide the last utterance with humour. The gesture is carefully coordinated with the rhythm created by the prominent syllables, which in turn is also reinforced through head nods.

The effort to raise the audience’s interest and make the presentation entertaining is congruent with the views the speaker reported during the interview. In particular she believes that dissemination talks need to focus more on entertainment and less on data, and commented that, as a member of the audience, she preferred a particular presentation that she found especially entertaining.

DTLI2

In this rich point the speaker is discussing the applicability of her research and stating its potential.

Orthographic transcription

But of course I am not here to tell you a story, I am not a storyteller. I'm here just to show you that the possibilities with smart textiles are endless.

MICASE transcription

but of course i am not here to tell you a story i am not a storyteller. i'm here just to show you, that the possibilities with smart textiles are end,less.

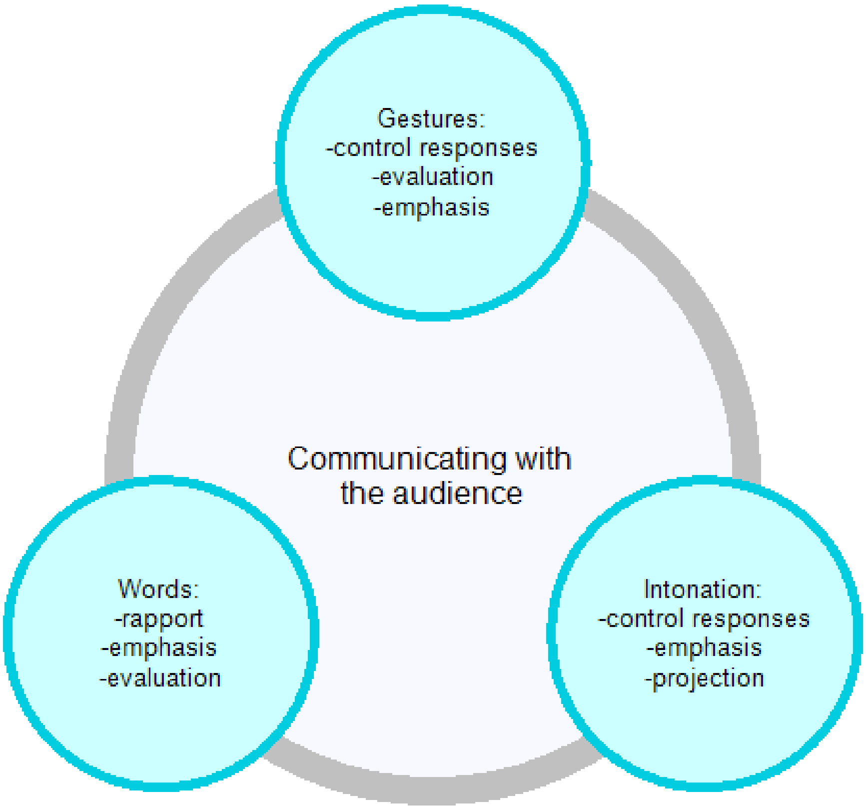
DI transcription

1. but of COURSE I am not here to tell you a sTORY
2. I am $_{\text{NOT}}$ a story $^{\text{TEL}}$ler
3. I'm $^{\text{HERE}}$ just to SHOW you
4. that the possiBIlities with SMART textiles
5. are $_{\text{END}}$ () $_{\text{LESS}}$

Gestures and head movements

are $_{\text{END}}$ () $_{\text{LESS}}$

Multimodal ensemble



Analysis of words:

This speaker's use of words shows extreme positive evaluation of the applicability of smart textiles ("possibilities are endless"). There is also a deictic reference to the event ("here"), which is emphasized through repetition and contributes to rapport.

Analysis of intonation:

It is remarkable that "endless" is prominent and split in two for further emphasis. In unit 2 the low key (-96 Hertz) has equating effect (I am not here to tell you a story = I am not a story teller). The high key in unit 3 (+ 148 Hertz) suggests the opposite of equating: reversal of expectations, little connection with what was said before. In fact the speaker interprets this use of intonation as conveying that it is more important "what I'm here for than what I'm not here for". In unit 5 the low key (-122 Hertz) again has equating effect (possibilities= endless), while the low termination suggests discourse final and closes the potential argument as to the truth of this statement.

Analysis of gestures and head movements:

There is an interesting open hand circular gesture drawing a sphere with both hands that suggests scope of “endless”. It is a metaphoric, pragmatic gesture that guides an interpretation of the word “endless” as something positive, since the concept of “big” conveyed with the gesture (as confirmed by the speaker, she was, as she puts it, “showing the world”) implies “many applications" of her research. It is synchronous with speech, because it coincides with the two prominences in “endless”. The word “endless” is split in two, and so is the gesture, performed in two steps, which further reinforces this synchrony.

Analysis of persuasive activity:

The speaker in this rich point emphasizes important parts of her message through gestures, intonation and words. She also evaluates her research in positive terms and triggers acceptance of this evaluation through intonation and gestures. Words are also used to build rapport by referring to herself and her role in the event.

It is interesting to note that this speaker is the only one who reported having previous experience in this type of event, and her motivation was not only to make her research known, but also to obtain practice and feedback to improve her presenting skills. In fact she also reported being a member of Toastmasters, an organisation devoted to public speaking, which shows her concern and interest in this topic.

DTRA1

The speaker begins his presentation highlighting the relevance and applicability of his research.

Orthographic transcription

We use mathematics to build bridges. To build dinosaur traps, you know, and other things we use on a daily basis.

MICASE transcription

we use mathematics to build bridges. ah to build eh dinosaur traps you know and other things we use on a daily basis

DI transcription

1. we USE mathematics to build BRIDges ↘
2. ah to BUILD eh DInosaur traps ↗
3. you KNOW ↗
4. and Other things we use on a daily BAsis ↗

Gestures and head movements

we USE mathematics to build BRIDges

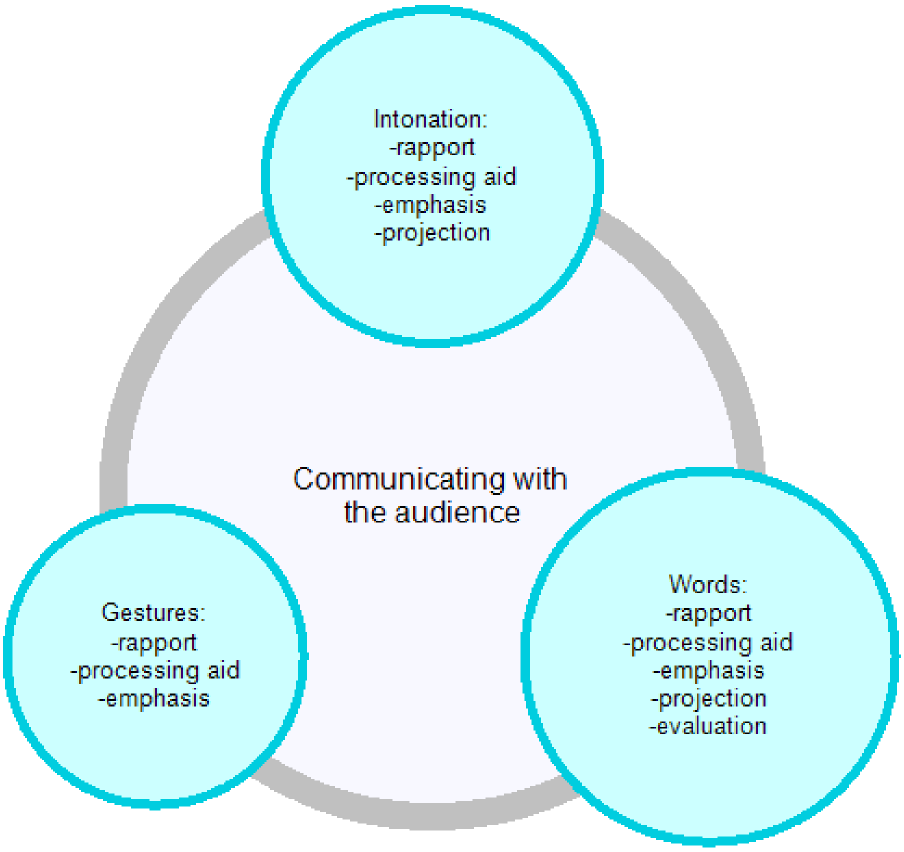
we USE mathematics to build BRIDges

ah to BUILD eh DInosaur traps

you KNOW

and Other things we use on a daily BAsis

Multimodal ensemble



Analysis of words:

The expression “build bridges” was first interpreted as a metaphor that intensifies the applicability of the research described and implicitly evaluates it in positive terms. After consultation with the speaker, he explained that it was actually meant literally (e.g. a practical application to engineering), but the intention of emphasizing the applicability of maths is correct. He also specified how the objective of his research is not exactly to show that maths are applicable, but to understand how we find abstract principles useful to solve practical problems. In this excerpt his is showing the applicability of maths to indirectly show the rationale behind his research topic (why is it that maths are so useful?).

In terms of register, it is noticeable the humour in the example of “dinosaur traps" as a daily routine, that gets attention and builds rapport. The informal expression “you know” is also a way of building rapport and seeking concurrence.

There is a parallel grammatical construction (or symmetrical structure) “to build”…”to build” that adds symmetry and makes it easy to understand and remember. The speaker also added in the interview that his intention with this was to show that this is part of a longer list, which is congruent with the interpretation of making the text easy to understand.

Analysis of intonation:

The first tone unit has a fall, proclaiming tone (-48) suggesting new information, but the rest tend to a rise referring tone that matches the intention of giving familiar examples. This referring tone at the end attempts to assume common ground and creates rapport, and is consistent with the expression “you know”. The speaker interprets his use of tone, specially his use of fall tone, in a more cohesive way: the initial fall tone was his way of showing that this is the beginning of a list, and not a way of projecting new information because he assumes people know maths are used to build a bridge.

The high key in unit 4 (+ 47) presents this information as surprising or denying expectations, and it prepares for the final humorous pun that is achieved by the mismatch between content and tone (an extravagant example presented as daily basis). The speaker agrees with this interpretation, but adds it was also a way of emphasizing the multiplicity conveyed by “other things”.

Analysis of gestures and head movements:

As the speaker says “we use mathematics", rotating hands in a metaphoric, referential gesture seem to symbolise the process of physically *using* something (in this case mathematics). The speaker offered a different interpretation of this gesture, which is the cohesive function of showing that there is a longer list of things coming. Taking into account the next gesture (enumerating with fingers as he actually begins the list) this is probably a more plausible interpretation.

Indeed he starts enumerating with his fingers as he says “build bridges”. This is a cohesive gesture that marks the item it accompanies ("bridges") as the first one in a list. Even if it is not followed up with a number two or three, it does help understanding.

Finally, beats with open hands palm up (OHS) starting with “dinosaur traps” and continuing until the end of the extract can be interpreted as an encoding aid for the speaker. They create the impression that the speaker is improvising his list of examples and provide the text with spontaneity.

Analysis of persuasive activity:

Persuasion in this presentation is built around two key aspects:

- Humour that builds rapport with the audience.
- Highlighting relevance of maths for everybody and not only for experts

The speaker is highlighting the applicability of maths, and indirectly the rationale behind his research question (why and how is it so useful?).To make his message more likely to be accepted he makes it interesting, easy to understand, not imposing and humorous.

Three persuasive strategies are realised through the three modes: processing aid, emphasis and rapport. In addition, both words and intonation are used to project context of interaction. Words alone are used to evaluate, but this evaluation is supported and made more likely to be accepted by the emphasis, rapport and clarity provided by the modal ensemble.

DTRA2

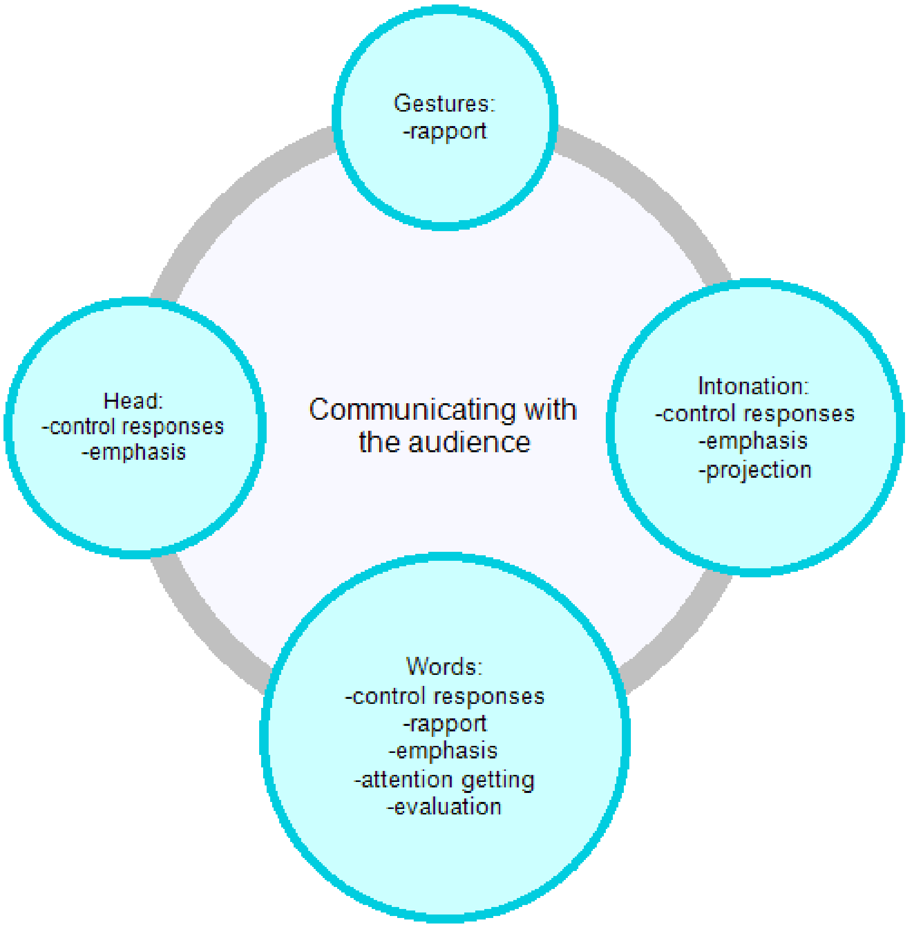
In this rich point the speaker is highlighting the relevance and applicability of maths: something “cool” you can do with maths, i.e. derive several axioms from one single principle.

Orthographic transcription
Cool, I think. Well. If you’re geeky enough you should find it cool.
MICASE transcription
cool, i think well, if you’re geeky enough, you should find it cool <AUDIENCE LAUGH> eeeh
DI transcription
1. COOL ↘ 2. I THINK ↘ 3. WELL ↘ 4. if you’re GEEky eNOUGH → 5. you SHOULD find it cool ↘ 6. eeehh →
Gestures and head movements
 I THINK

if you're GEEky eNOUGH

YOU SHOULD find it cool

Multimodal ensemble



Analysis of words:

There is humour in the statement, as well as use of intense language in the adjectives ("cool", "geeky"). The word "geeky" can also be considered slang. This informal use of language can be considered against expectations in this particular setting.

The speaker also addresses the audience with a mild command (you should) and implicitly prompts agreement with his positive evaluation ("cool") by using a conditional and implying "if you don't agree with my view it means you are not geeky". As the speaker pointed out, this is more than an implicature, becoming a logical consequence.

Analysis of intonation:

The speaker uses fall, proclaiming tones that present the information as new. During the interview he actually confirmed he didn’t expect people to be familiar with what he was explaining at that moment and he wasn’t expecting them to find it cool, either.

There is a high termination in unit 2 (+ 158 Hertz) that prompts an active role from the audience (engage in considering his evaluation of maths as cool). This is consistent with the use of the OHS gesture explained above and also matches the interpretation of the speaker. Unit 4 also has a high key (+ 124 Hertz) that adds the meaning of reversal of expectations. This is consistent with the innovative use of slang in this situation. Indeed the speaker commented he was aware that this was unexpected vocabulary for the audience. Interestingly, unit 5 begins with a low key (-37 Hertz) that has an equating effect (being geeky= find cool) and ends with a mid termination that prompts concurrence.

Analysis of gestures and head movements:

Synchronous with the two prominences in “I think” the speaker makes an open hands palm up (OHS) gesture. It is a metaphoric, interpersonal gesture that visually offers his opinion and engages the audience. The speaker actually explained this gesture as a way of “opening up".

Two head nods synchronous with “geeky” and “find” prompt to share a positive evaluation and concur with “cool”. It is interesting that the second nod precedes words, being synchronous with “find" and not with the evaluation itself (“cool”).

Analysis of persuasive activity:

Persuasion in this excerpt is built around highlighting relevance of maths, but in this case reversal of expectations plays a greater role. The speaker is using a wide variety of persuasive strategies:

-Positive evaluation of maths (words)

-Rapport through humour (words) and inclusive gesture that brings audience in

-Projection of context of interaction (intonation): marks information as new

-Control a response (words, intonation, head): prompt to share an evaluation and infer he implicature (or rather understand the logical consequence) “don’t find it cool= you are not geeky enough”

-Emphasis on (words, intonation, head) certain aspects of his message that do not conform to the rules (e.g. use of slang).

-Attention getting (words) to certain aspects of his message that do not conform to the rules (e.g. use of slang).

All modes analysed contribute to the intended persuasive effect. For example the words "I think" are synchronous with a high key and an interpersonal gesture that engage the audience in this judgement. Also the use of slang together with a high key and a head nod adds the nuance of "reversal of expectations". This makes the piece particularly effective, because persuasion works better as it is not foreseen.

DTPI1

In this rich point the speaker presents his research downplaying its relevance, which is something unexpected in this situation. This reversal of expectations contributes to make persuasion more effective.

Orthographic transcription

I must say all of a sudden I feel really, really humble. I've, just, These these five presentations that came before me they all want to change the world and they are doing it in an incredible way, they are doing superconductors in microwaves, there are things I've never heard of textiles, it's, it's, I'm just trying to get you from A to B, that's all I want to do. It's, it's, it's not rocket science, it's not superconductors, it's just trying to inform you, and this is really difficult.

MICASE transcription

i must say all of a sudden i feel really really humble. i've_just_ these these five presentations that came before me they all want to change the world and they are doing it in an in an incredible way they are doing superconductors in microwaves, there are things i've never heard of textiles, it's_it's, i'm just trying to get you from a to b. that's all i want to do. it's_ it's_it's not rocket science it's not superconductors, it's just trying to inform you. and this is really difficult.

DI transcription

1. i must SAY ALL of a sudden
2. i feel REALly REALly HUMble ↘
3. i've just
4. these these FIVE presentations that came beFORE me
5. they all want to CHANGE the WORLD
6. and they are DOING it in an in an inCREDible way ↘
7. they are doing SUperconductors in MICrowaves
8. there are THINGS i've never HEARD of textiles
9. it's it's
10. i'm $^{\text{JUST}}$trying to GET you from a to b ↘
11. that's ALL i want to do ↘
12. it's it's
13. it's NOT ROcket science ↗
14. it's not SUperconductors ↗
15. it's JUST trying to inFORM you ↘
16. and $^{\text{THIS}}$is really DIfficult

Gestures and head movements:

i feel REALly REALly HUMble

i'm $^{\text{JUST}}$ trying to GET you from a to b

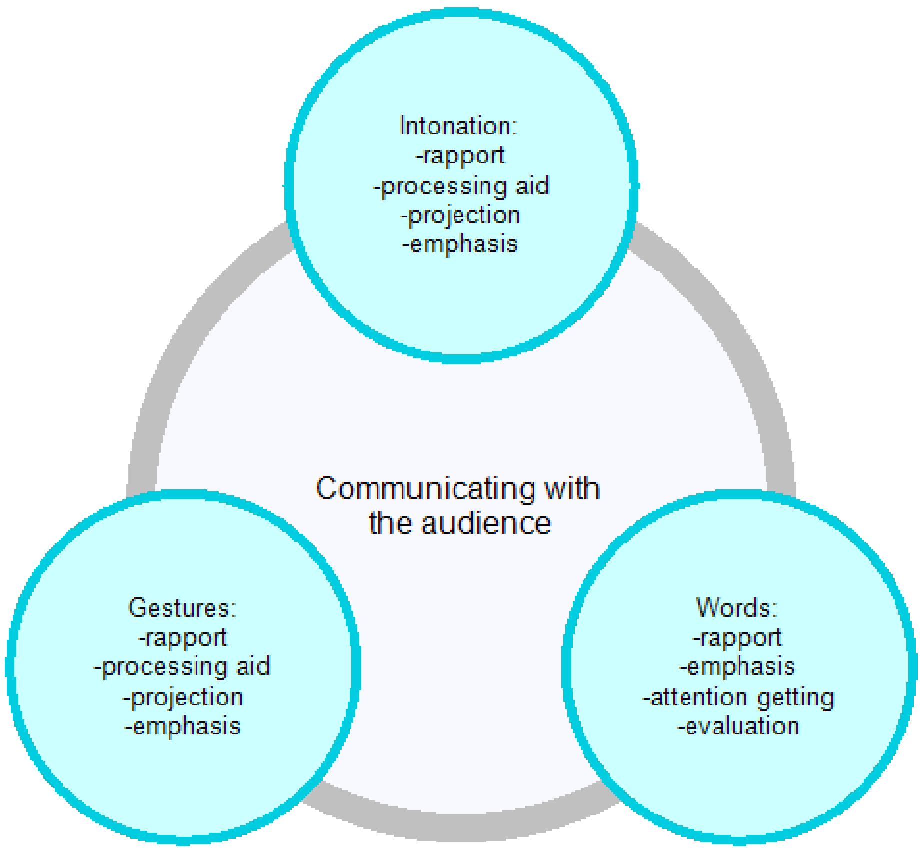
that's ALL i want to do

it's not SUperconductors

it’s JUST trying to inFORM you

and THIS is really DIfficult

Multimodal ensemble



Analysis of words:

The references to previous presentations help to fit in in the event, and intense positive evaluation of the previous presentations establishes a surprising contrast as he humbly presents his own research (“really, really humble”). According to the speaker this is a technique he often resorts to when he is not well prepared for a talk: it is a "captatio benevolentiae" resource that shows interest for other presenters and their presentations and at the same time serves as an attention getting technique.

The speaker is using hedging techniques that minimise his role and suggest humbleness. For example he uses the adverb of degree “just”, he equates his research objective (“all I want to do”) with something apparently unimportant (“take you from A to B”) and explicitly uses the adjective “humble”. In doing this he is making himself likable and therefore paving the ground for rapport. This rapport is also reinforced by the reference to the audience in “inform you”.

The consistent use of “I” and “me” throughout the excerpt also brings to the fore the person behind research and personalises it. This could be interpreted as an indirect way of assuming responsibility of research decisions, which is a common feature in conference presentations. However, the speaker explained he was not at all influenced by conference presentations. He explains he uses these self-references because "It's me on the stage. People want to understand why I am there and my relation to others".

The three part list created by “it’s not…it’s not…it’s just” makes the text memorable.

Interestingly, at the end of the excerpt he evaluates his own research as “difficult”. With this he seems to be implying “my research might not be as impressive, but it is still worth it”. According to the speaker this evaluation of his research as "difficult" is not what he actually wanted to convey. Instead, he wanted to present his research as something that is potentially more familiar, and “down to the level” of the audience (as compared to rocket science or smart textiles). The use of “difficult” is probably a sign of planning ahead, since his next topic is challenges in his research.

Analysis of intonation:

The false starts (e.g. unit 3, 9, 12) are due to encoding difficulties, as the speaker confirmed during the interview.

Concerning prominence, it is interesting that unit 2 presents three prominent syllables. This adds emphasis to the words and this effect is further reinforced by synchronous beats as explained previously. “Inform” (the objective of his research, unit 15) is also made particularly prominent. The speaker added in the interview that slowing down the rhythm and articulating very carefully is another technique he usually resorts to when he wants to sound confident, and this is also part of the intended effect in these examples.

The high key in “I'm just trying to get you…” (+ 43 Hertz, unit 10) conveys a sense of surprise and unexpectedness, and adds a humorous tone. The same sense of unexpectedness is achieved by high key in “this is really difficult” (+ 140 Hertz, unit 16). Interestingly, apart from a high key “this” is also emphasized through a beat.

Regarding tone, the speaker uses fall tone in units 2 and 6 (-51 and -53 Hertz). This adds a nuance of sincerity and differentiates it from polite praising of colleagues/ contest competitors. Units 10 and 11 also feature a fall tone (-54, -105 Hertz) that suggests the speaker is treating this as new information. Units 13 and 14 have a rise tone (+61, +34 Hertz) that indicates a list still not finished, and the fall tone (-60 Hertz) in unit 15 marks the closing element of the list.

The low termination (-214 Hertz) in the final unit indicates this is discourse final and a coming change of topic.

Analysis of gestures and head movements:

The repetition in "really, really humble" is paralleled with 3 finger ring beats that mark the rhythm created by the prominence. The function is therefore partly cohesive (connecting) and partly pragmatic (emphasizing, with the effect of adding sincerity, as if saying “I really mean it”).

The two imaginary locations referred to in “I’m just trying to get you from A and B” are visually represented by two beats with parallel extended hands, the first to the left side and the second to the right side. The gesture is both referential (provides a visual representation) and cohesive (connects two elements). This gesture is not in phonological synchrony: it occurs after the prominent syllables (“just” and “get”).

An OHS metaphoric gesture, synchronous with “that’s all I want to do” suggests openness, fulfilling a pragmatic function. According to the speaker he wanted to transmit “there’s nothing more to it”.

The speaker makes a discarding gesture as he pronounces “superconductors”. It is a metaphoric, pragmatic gesture performed with parallel arms and hands stretched to the left side, meaning “this is not my topic”.

A finger ring beat emphasizes “inform” (the focus of his research). The same gesture is repeated synchronous with the first prominence in “This is really difficult”**.** They have a cohesive function, since they connect “this” with its referent “inform”. They can also be considered to have a pragmatic function, since they emphasize the words they accompany and single them out as particularly relevant. These words are also prominent as shown in the intonation analysis. Interestingly, the shape of a ring suggests specificity, which matches the selection of his specific topic of research: inform about transport data. If we understand this gesture in connection with the previous one, we can see how the speaker is listing potential topics, discarding the ones that do not apply, and selecting the one that applies to him. This effect is also reinforced by the fall tone as discussed in the intonation analysis.

Analysis of persuasive activity:

The speaker uses a wide range of persuasive strategies enacted multimodally, and it is also noticeable that the three modes are used very evenly:

- rapport: words, gesture and intonation

- evaluation: words

-projection of context of interaction: intonation (marks contrast what they do vs what I do and projects surprise) and gestures (mark the selection of research topic)

-emphasis: intonation (prominence), gestures (beats), words (repetition and intense language)

-processing aids: gestures, intonation (marks structure of discourse)

-attention getting: words

The combination of these persuasive strategies is intended to achieve two powerful effects in this extract: create rapport and reversal of expectations. In addition, the message is made memorable through intense language.

Interestingly, there are two features in this excerpt that are more characteristic of the conference presentations in the corpus. One is an attempt to fit in the event and the second is assuming responsibility for research decisions. At a later point on the presentation the speaker also shares research experiences, which is also characteristic of academic presentations. Interestingly, though, the speaker explicitly mentioned in the interview that he was not influenced by academic presentations. In particular he dislikes these presentations when speakers simply assume that the audience will pay attention and do not make an effort to sound interesting. He also explained that at the time he was involved in a start-up project and was actually more interested in product pitches, as the following quote exemplifies:

> “I hate the way to do research that you are in a cellar, you work for four years and you don't communicate about it. I think it is really important as a researcher that you are able to communicate about it. For me the TEDX Gent was the ideal thing because you have to explain it layman terms and you have to explain it in a way that’s also persuasive, that people believe there's something in there. It’s almost like a start-up pitch or something where you need to convince people about a certain cause and need to activate them almost to believe in you. Before my research I was active as a, in my own start-up, so I was intrigued by the whole pitch thing and for me TEDX Gent or public speaking about my research is just

> like this thing from start-ups" [INAUDIBLE] you need to pitch your research all the time and if you cannot do that, for me you are not a good researcher"

This suggests that different disciplines lead to different academic presentations and different persuasive strategies. As the speaker explained, in his field (engineering) persuasion is achieved more by scientifically based claims than by foregrounding the person behind research and this same talk in a conference in his field would have been rejected as "unacademic".

DTPI2

In this rich point the speaker is presenting the benefits of sharing transport data.

Orthographic transcription

Now what do I want to do about this? Well, I think everyone should have the right to reuse data. It's, ok, not all kind of data, but certainly this transport data. Why you ask? well, I have a lot of reasons, too many for just six minutes but I'm going to answer a different question: Why not? Well, the data is not privacy, privacy, there are no privacy issues with this kind of data, and, well, this can be very beneficial for your company because more people are going to take the train or going to take the bus, so, why not?

MICASE transcription

now what do i want to do about this. well, i think everyone should have the right to reuse data. it's, ok not all kind of data but certainly this_this transport data. why you ask well, ehh <SIGH> i have a lot of reasons, too_too many for_for just six minutes but i'm_i'm going to answer a different question why not. well, the data is not privacy, ehh privacy ehh there are no privacy issues with this kind of data, and, well, this can be very beneficial for your company because more people are going to take the train or going to take the bus. so, why not

DI transcription

1. NOW what do I want to do about this ↘
2. well
3. i think everyone should have the RIGHT to reUSE data ↘
4. it's
5. ok not all kind of DAta ↗
6. but CERtainly this this TRANSport data ↘
7. WHY you ASK
8. well ehh
9. i have a LOT of reasons ↘
10. too too MAny for for just six minutes ↘
11. but i'm i'm going to ANswer a different QUEStion
12. WHY NOT
13. well
14. the DAta is not PRIvacy ↗
15. ehh PRIvacy ehh
16. there are no PRIvacy issues with this kind of DAta ↗
17. and well
18. this can be very beneFIcial for your COMpany
19. because MORE people are going to take the train or going to take the BUS
20. so
21. WHY NOT ↗

Gestures and head movements:

NOW what do I want to do about this

i have a LOT of reasons

but i'm i'm going to ANswer a different QUEStion

WHY $_{\text{NOT}}$

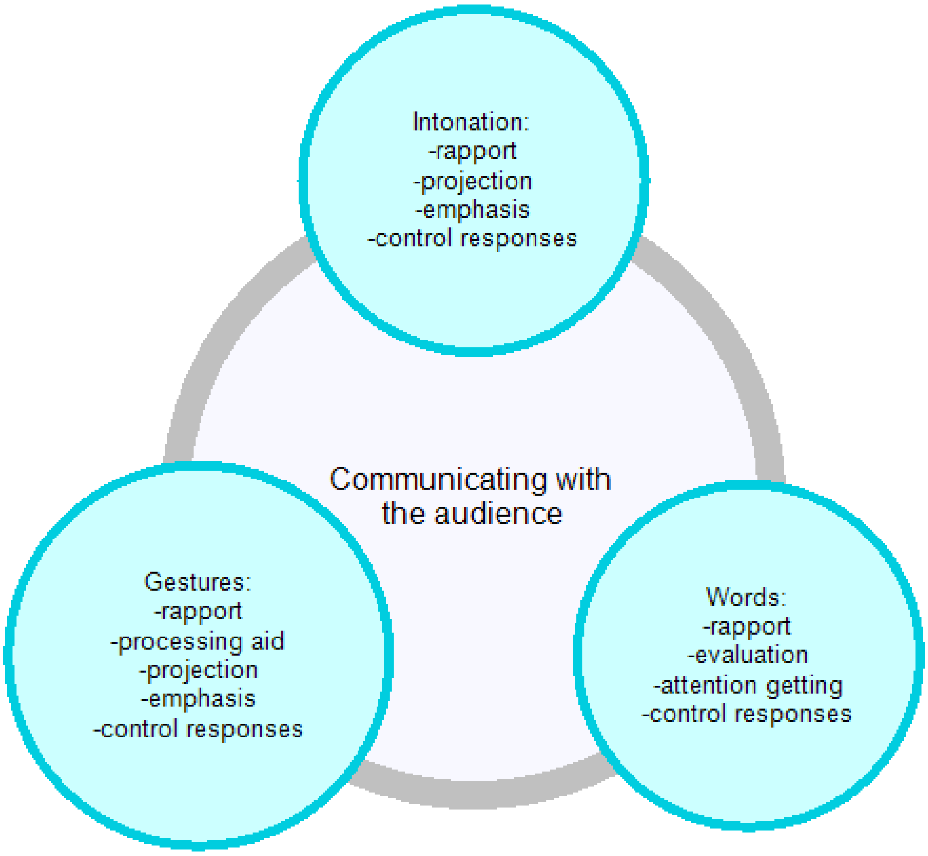
ehh PRIvacy ehh

there are no PRIvacy issues with this kind of $^{\text{DA}}$ta .

this can be very bene^{FI}cial for your ^{COM}pany

WHY ^{NOT}

Multimodal ensemble



Analysis of words:

The rhetorical questions involve the audience in the argument and seem to anticipate their reactions and control their responses. The speaker explained he was actually advised to do this during his preparation for the talk, and in his own words they serve to “activate the audience and keep their attention”.

There is frequent use of first person pronoun that makes the text personal and brings to the fore the figure of the researcher. Interestingly, he presents the motivation behind his research as a moral or ethical principle (“everyone should have the right”) that he personally endorses, instead of providing more rational arguments.

Regarding register, there is a conversational tone with reformulations and self-corrections throughout that provide the text with a spontaneous character and builds rapport.

Finally, the excerpt closes with a positive evaluation ("beneficial").

Analysis of intonation:

“I” in unit 1 is prominent drawing attention to the researcher. “A lot” in unit 9 is also prominent and adds to the emphasis in justifying his opinion (also emphasized through gestures).

Proclaiming tones at the beginning (especially in units 1, 3, 6, 9, 10) serve to state his point of view treating information as new. On the other hand, the referring tones in units 5, 14, 16 and 21 present this line of argument as already shared with the audience: something they actually know (but maybe have not stop to consider yet).

Regarding key and termination, the speaker uses high key to present information as potentially surprising, unexpected or contrasting with previous discourse, for example in unit 6 (+82 Hertz) it contrast "all kind of data" vs. "certainly this transport data", and in unit 18 (+58 Hertz) it seems to imply something that the audience maybe hasn’t stop to consider. The speaker uses low termination to suggests “discourse final” and close the argument, e.g. in units 6 (-87 Hertz) and unit 12 (-28 Hertz). A high termination is also used in units 7 (+86 Hertz), 16 (+24 Hertz), 18 (+70 Hertz) 19 (+28 Hertz) and 21 (+147 Hertz) to prompt an active role from the audience (asking them to consider what he is saying and act or adapt their thinking accordingly). Note that this active role does not imply framing the information as new or even open to discussion: rather the speaker seems to be mentioning things most people know and agree upon (e.g. transport data is not private) and asking them to take this into account to form their responses.

Analysis of gestures and head movements:

A pragmatic beat with OHS hands emphasizes "I" in “what do I want to do”. This brings to the fore the person behind research.

An open hand prone sweeping gesture is synchronous with “I have a lot of reasons”. The gesture is repeated twice, the second time during a brief pause following “reasons”. It is a metaphorical pragmatic gesture that conveys the scope of "a lot" and prevents a potential counterclaim. Interestingly, the first time the gesture is done it ends with arms wide apart that also convey the idea of something "big" (the space taken up by the bulk of reasons). The speaker describes this gesture as saying “you’ll have to take it from me” and interprets it as compensating for the lack of scientific grounding in this argument.

The lines “I’m going to answer a different question: Why not?” are supported by an OHP gesture with an extended pointing finger moving from right to left, synchronous with "answer". It is a metaphoric, pragmatic gesture that conveys the change of perspective that the speaker is proposing.

The speaker makes OHS rotating movements with both hands, the first time immediately following the prominence in "privacy" (unit 15), then synchronous with the prominence in "privacy issues" (unit 16). It is a pragmatic gesture that metaphorically represents "issues" or "problems" that the use of these data may bring, and interestingly occurs both times with the word "privacy", conveying that privacy is indeed a concern when dealing with data. This gesture can also be an encoding aid as he finds the right words, something the speaker finds plausible, and actually both interpretations are not exclusive

An OHS parallel hands gesture to the right simultaneous with "beneficial for your company" visually "offers" the implied benefit of sharing transport data. It is a pragmatic, metaphoric gesture, which begins before the prominence in beneficial and is hold until the beginning of the next tone unit (as he says "more").

A metaphoric, pragmatic open hand supine OHS gesture synchronous with "why not?" and shrugged shoulders convey a sense of "I don't understand". The facial expression, although out of the scope of this analysis, also contributes to this meaning. The open hands suggest openness and visually present this lack of understanding for the audience. The shoulder shrug can be considered a more conventional or stereotyped emblem standing for "I don't know".

Analysis of persuasive activity:

The key to persuasion in this extract seems to be directing a desirable response and building rapport, which interestingly show the highest multimodal density, both being realised through words, intonation and gestures. The speaker seems to be reminding the audience of a situation they already now (mainly that transport data is not private and therefore can be shared) and then prompting them to share an evaluation of this situation: this can be beneficial for both transport companies and their users. At the same time, he is foregrounding the person behind research, and he actually presents the rationale behind his research as a personal mission based on his own moral principles. This makes the message more personal and builds rapport, but it is also a way of assuming responsibility for the decisions taken in his research project. This last feature of this excerpt (assume responsibility for research decisions through a foregrounding of the figure of the researcher) is more characteristic of academic presentations.

DTJO1

In this rich point the speaker is introducing himself and his research and using a hook to get the attention of the audience.

Orthographic transcription
Hi. I'm a chemist, and I really like kissing. I'd like to kiss you, I'd love to kiss you, I'm gonna kiss all of you. Because I really want to keep it simple, but not stupid. Tonight I'm going to show you how your own experience, with a very simple technique, can revolutionise the chemical industry, and finally, your life.
MICASE transcription
hi. i'm a chemist, and i really like kissing. i'd like to kiss you, i'd love to kiss you, i'm gonna kiss all of you because, i really want to, keep it simple, but not stupid. tonight, i'm going to show you how your own experience with a very simple technique, can revolutionise the chemical industry, and finally, your life
DI transcription
1. HI 2. I'M a CHEmist 3. and i REally like KIssing 4. i'd LIKE to KISS you 5. i'd LOVE to KISS you 6. i'm GOnna kiss ALL of you 7. beCAUSE 8. i REAlly want to 9. KEEP it SIMple 10. but NOT stupid 11. toNIGHT 12. i'm going to SHOW you how your OWN experience 13. with a VEry simple technique 14. can revoLUtionise the CHEmical industry 15. and FInally YOUR life

Gestures and head movements

and I REally like KIssing

i'd LIKE to KISS you

i'd LOVE to KISS you

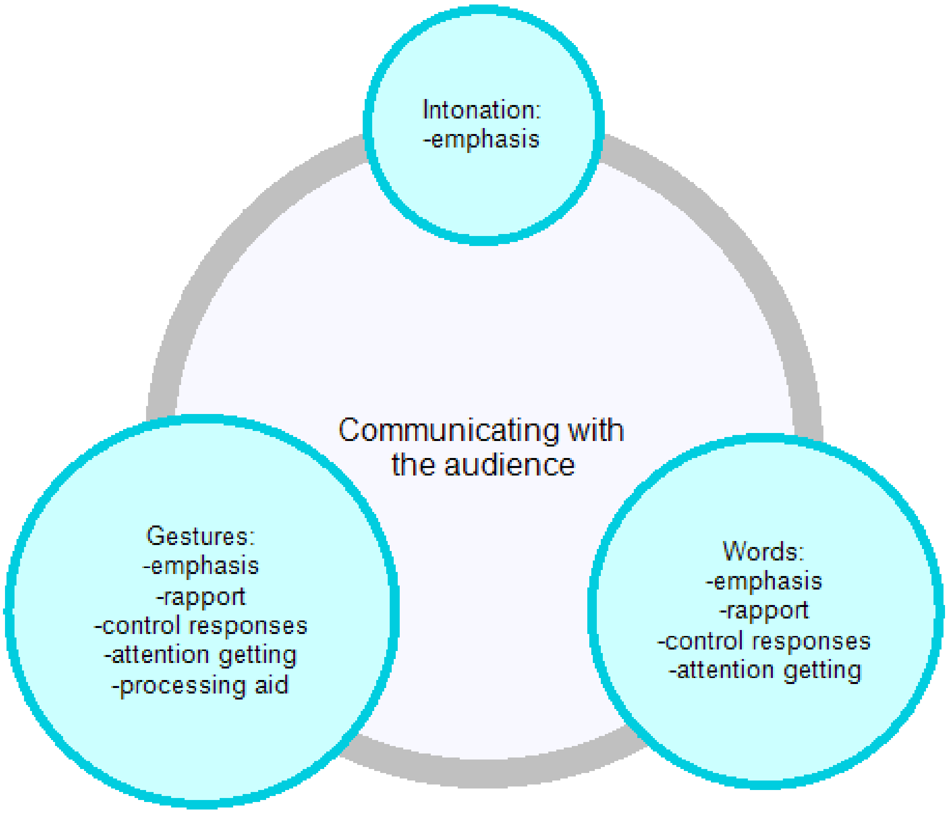
i'm going to SHOW you how your OWN experience

and FInally YOUR life

i’m GOnna kiss ALL of you

with a VEry simple technique

Multimodal ensemble



Analysis of words:

There is lexical creativity in the use of the acronym “KISS” (Keep it short not stupid) and use of parallel structures (“I'd like to kiss you/ I’d love to kiss you”). In addition the speaker uses intense language (“revolutionise your life”). All this contributes to make the message memorable. According to the speaker the acronym was a well-planned hook to draw the attention of the audience, and he consciously repeated it several times to let it “sink in”. The use of “revolutionize" was also conscious, and the intention was to "sell his story"

The repetitive use of "I" also highlights the figure of the researcher. The speaker reported that he was not trying to personalise his research on purpose, because it is the result of team work, but he was consciously trying to make a connection with the audience.

The speaker also refers to the audience's personal experience without actually specifying, which postpones crucial new information and builds suspense. This effect is actually something the speaker was not aiming at consciously, but considers plausible.

Analysis of intonation:

“Really”, “kissing” (unit 3) and “revolutionise" (unit 14) are prominent, which adds to the emphasizing effect of the closed fist beats. “Own" in "your own experience" (unit 12) and “your” in “your life” (unit 15) are also prominent, contributing to the specifying effect of the synchronous gestures that point to the audience.

Analysis of gestures and head movements:

Closed fist beats in phonological synchrony with words and intonation emphasize prominent words in “I really like kissing” and “revolutionise the chemical industry”

Deictic gestures to point to specific members of the audience, in semantic synchrony with words (i.e. synchronous with "you") are used to get the audience involved in a way that is not expected in this communicative situation (speakers do not usually point to members of the audience at least during the monologue section). This “breaking of the rules” raises interest and sets a humorous and relaxed tone. The speaker actually explained during the interview that he wanted to create “a very personal connection” with the audience and “make them awake”, and he was aware of the surprising effect it had on them.

"Your own experience" and “your life” are accompanied by open hand supine gestures subtly pointing to the audience that visually specify whose experience/life he is referring to. It is a referential, metaphoric gesture in semantic and phonological synchrony with words and intonation (synchronous with “own” and “your” respectively, which are also prominent) that helps processing of the message through visual representation. The speaker reported in the interview that he wanted to point to the audience in a way that was not aggressive, and above all, he also wanted to inspire confidence. For that reason he consciously used open hand supine gestures that suggest honesty. His interpretation of this gesture is more pragmatic than referential, although he agrees he was also clarifying the referent in "your own experience" and “your life”.

“Very simple” is synchronous with a sweeping gesture that rules out potential counterclaims. The same gesture is also used previously in “kiss all of you” with the same effect of denying counterclaims to this extreme quantification. The speaker claims he tends to associate this gesture with the concept of “simple” rather than with the extreme evaluation, but on second thoughts he also became aware of the fact that he also uses this gesture with extreme evaluations that imply other qualities as well.

Analysis of persuasive activity:

A good hook (Kiss) gets attention of audience and is later also used at the close giving a circular structure. This makes the presentation memorable. The speaker creates rapport with the audience bringing in his personal experience (e.g. anecdote) and providing familiar examples that make the presentation accessible and brings the topic closer to audience.

The speaker is using the modes analysed to emphasize important parts of his message, help easier processing and subtly direct a response without being imposing (in this case raise curiosity). The message is made memorable and engaging, and the speaker builds rapport with the audience and makes himself likeable as a presenter, which is likely to make the audience more prone to accept his ideas. This is consistent with several ideas the speaker reported during the interview:

-It is important to show enthusiasm in this type of talks.

-He particularly enjoyed one presentation in the event because the speaker was able to engage him from the very beginning. Both points relate to his attempts to create rapport and make his presentation memorable.

-In dissemination talks speakers have to assume that the amount of shared knowledge is limited, which relates to his use of processing aids through visual representations.

DTJO2

In this rich point in particular the speaker is introducing a problem he encountered as he carried out his research and explaining how he solved it.

Orthographic transcription

Now which experience am I talking about? Well, you all know a bit of cooking, and in fact there's not much of a difference between chemistry and cooking. In both cases you need a recipe which tells you which ingredients you are going to use. Now in my case this was hafnium chloride and benzyl alcohol. I just needed to mix these (oops) I just needed to mix these and put them in an oven at 220 degrees. And then I had to wait. For three days. And this was a bit of a personal crisis, because I really can't wait. I like fast reactions. Who doesn't like fast? So think about your kitchen. If you want to heat stuff really fast you are not going to use an ordinary oven. You are going to use the microwave.

MICASE transcription

now which experience am i talking about. well, you all know a bit of cooking, and in fact there's not much of a difference between chemistry and cooking. in both cases you need a recipe which tells you, which ingredients, you are going to use. now in my case, this was, hafnium chloride and benzyl alcohol. i just needed to, mix these oops <P:07>_ i just needed to mix these and put them in an oven at 220 degrees. and then i had to wait, for three days, and this was a bit of a personal crisis, because, i really can't wait. i like fast reactions. who doesn't like fast. so, think about your kitchen. if you want to heat stuff, really fast, you are not going to use an ordinary oven, you are going to use, the microwave

DI transcription

1. now WHICH experience am i TALking about
2. well you ALL know a bit of COOking
3. and in $_{\text{FACT}}$ there's not much of a $^{\text{DI}}$fference
4. between CHEmistry and COOking
5. in BOTH cases you need a REcipe
6. which TELLS you which in$^{\text{GRE}}$dients you are going to use
7. now in my case this was HAFnium chloride and BENzyl alcohol
8. i just needed to MIX these oops
9. i just needed to MIX these and PUT them in an oven at 220 degrees
10. and $^{\text{THEN}}$ i had to WAIT
11. for THREE $^{\text{DAYS}}$
12. and this was a bit of a PERsonal CRIsis
13. because i really CAN'T WAIT
14. i like FAST reactions
15. who $^{\text{DOESN'T}}$ like FAST
16. so THINK about your KITchen
17. if you WANT to HEAT stuff
18. REAlly FAST
19. you are NOT going to use and ORdinary oven
20. you are $_{\text{GOING}}$ to use the MIcrowave

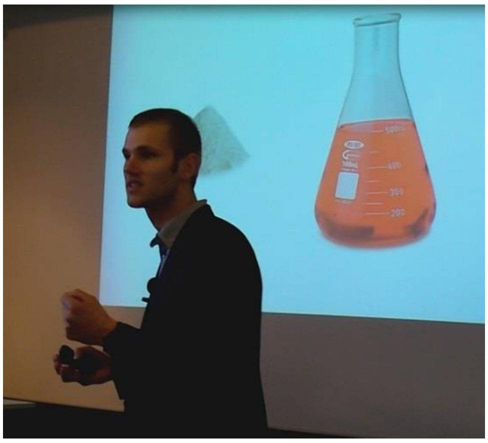
Gestures and head movements

between CHEmistry and COOking

i just needed to MIX these and PUT them in an oven at 220 degrees

who DOESN'T like FAST

so THINK about your KITchen



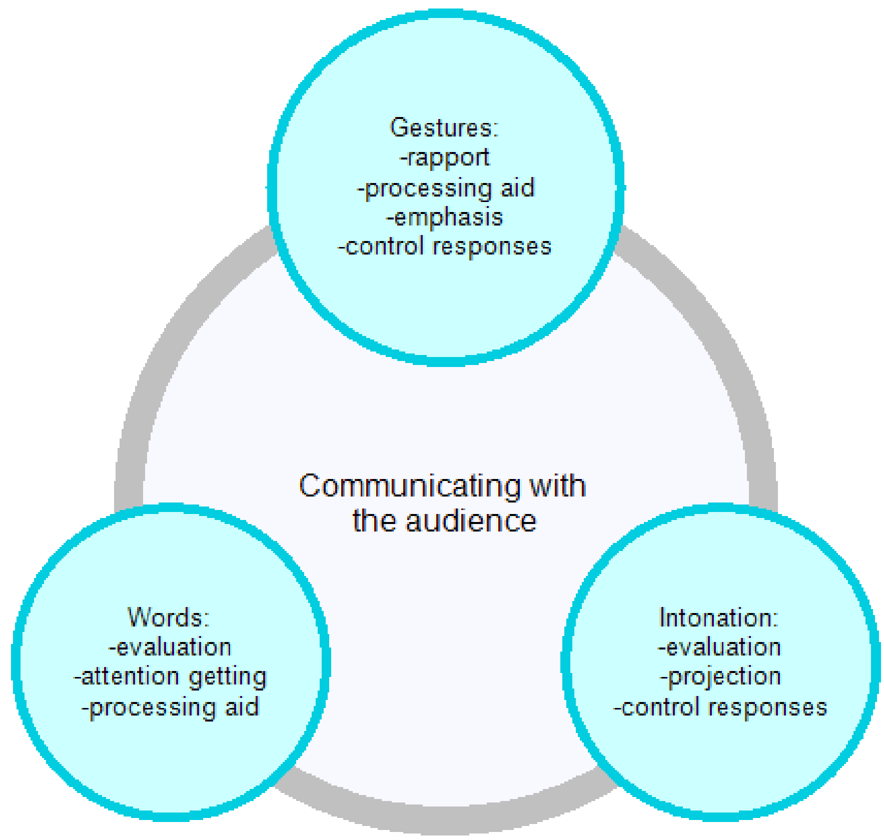
REAlly FAST

you are NOT going to use and ORdinary oven

you are GOING to use the MIcrowave

Multimodal ensemble



Analysis of words:

The speaker directly addresses the audience through two rhetorical questions and an imperative ("think"), getting them engaged. He also implicitly evaluates a situation (slow is bad, fast is better) using intense language ("personal crisis", "realy fast"). In addition, he uses parallelisms with everyday experiences as examples that help understanding and bring his topic close to the audience.

Analysis of intonation:

In units 3 and 20 the speaker uses a low key (80 and 38 Hertz decrease respectively) that has an equating effect and makes his statement sound obvious or taken for granted. This prevents any possible discussion as to whether chemistry and cooking are really similar or whether a microwaves oven is faster than an ordinary one. During the interview the speaker confirmed that this was actually the effect intended in the case of unit 20, but not in unit 3 (he was not consciously presenting the similarity between chemistry and cooking as taken for granted).

He also uses a high key in unit 10 and unit 15 (20 and 47 Hertz increase respectively). In unit 10 this has the effect of presenting this need to wait as something surprising and not what should be expected. As a consequence this waiting time is interpreted as negative, something that should not be there. In unit 15 the high key also adds a nuance of surprise, but in this case at a potential response to this question (=I don't like fast). This is a way of implicitly evaluating fast as good and slow as bad (nobody likes waiting).

In units 3, 6 and 11 there is a high termination (29, 58 and 43 Hertz increase respectively) In units 3 and 6 it is probably anticipating a potential surprise at the parallelism between chemistry and cooking (which is on the other hand counteracted by the low key that presents it as obvious). In unit 11 it prompts an active role from the audience requesting a judgment regarding the need to wait for three days (which the speaker expects to be evaluated as negative, as reported in the interview).

Analysis of gestures and head movements:

“Chemistry and cooking” are synchronous with alternating beats to both sides (OHS position) that visually represents the parallelism. The beat is repeated twice, one for each prominence (chemistry/cooking) in tone unit 4. It is therefore phonologically synchronous and fulfils a cohesive function, bringing together the two elements of a comparison. The speaker actually confirms this was a way of “giving structure" for the audience.

Later, hands moving in circles represent the action of “mixing”. It is a metaphoric, referential gesture in semantic and phonological synchrony with “mix” that works as a processing aid.

Then an open hand supine gesture, as if yielding the turn and inviting an answer from the audience, is performed in synchrony with the rhetorical question “Who doesn’t like fast”. This is a metaphoric gesture that fulfils an interpersonal function. In this case the response of the audience is not verbalised, but is non the less prompted and likely to happen as a mental process. The speaker reported that this OHS gesture was planned and intended to show honesty and create rapport, and agrees with the interpersonal function of inviting an answer.

A gesture with parallel open hands starting at his head and moving outwards represent the action of thinking in unit 16. This is a metaphorical, interpersonal gesture that again prompts a response (in this case mental) from the audience, and is phonologically and semantically synchronous with the imperative “think”.

A closed fist beat intensifies the extreme evaluation in “really fast” (unit 18). It is synchronous with “really" and fulfils a pragmatic function. The speaker also interprets this gesture as a way of showing emotional involvement, which supports his contrast fast= good vs. slow =bad. The same gesture occurs previously emphasizing the prominences in "I really can't wait" (unit 13).

As in the first example (chemistry and cooking), open hand supine gestures alternating to both sides visually represent the contrast ordinary oven vs. microwaves in units 9 and 10. They are semantically and phonologically synchronous with “ordinary” and “microwaves”, and fulfil a cohesive function, since they bring together the two items being contrasted.

Analysis of persuasive activity:

The speaker uses a wide variety of persuasive strategies realised through words, intonation and gestures:

-Evaluates through words and intonation (of having to wait as something negative).

-Projects a particular context of interaction through intonation presenting statements as obvious (everybody prefers fast) or surprising.

-Directs a response towards a shared positive evaluation (fast is better), using intonation and gestures to prompt the audience to react in a particular way.

-Uses processing aids through cohesive gestures, visual representations and parallelisms with common experiences.

-Emphasizes through gestures that intensify evaluations.

-Builds rapport through gestures that suggest honesty.

-Uses attention getting techniques that get the audience engaged (words).

With these strategies the speaker tries to engage the audience, making his text relevant for them and also easy to understand. He also makes his text sound not imposing and emotionally appealing. This behaviour is consistent with what the speaker reported in the interview, especially in what concerns his attempts to keep the audience engaged and his efforts to make his text accessible for a non-specialised audience.

PPKE1

In this rich point the speaker is opening the presentation. He presents a gap and his product as the solution for this potential need. He also highlights what differentiates it and explains what's in it for the users.

Orthographic transcription
Wigoh is the solution. Wigoh stands for "what is going on here?" Wigoh is a free mobile app for local event discovery. So, what distinguishes Wigoh from all the other apps? Wigoh includes it all. Wigoh is a location-based information, but also gives event based information and social-based information. So what's in it for the users? We have a What's hot list, where all the hot events are …of the neighbourhood are [INAUDIBLE], then the social aspect of it, where you can find your friends on, you can find your favourite organisers and you can all bookmark them, and you just stay tuned by a calendar in the app.
MICASE transcription
wigoh is the solution. wigoh stands for what is going on here. wigoh is a free mobile app for local event discovery. so, what distinguishes wigoh from all the other apps, wigoh includes it all wigoh is a location-based information, but also, gives event based information and, social-based information. so what's in it for the users. we have a what's hot list, where all the hot events are_of the neighbourhood are (xx), then the social en_aspect of it, where you can find your friends on, you can find your favourite organisers and you can all bookmark them, and you just stay tuned, by a calendar in the app
DI transcription
1. WIgoh is THE solution 2. WIGOH STANDS FOR WHAT IS GOING ON HERE 3. wigoh is a FREE MObile APP 4. for LOcal eVENT disCOvery 5. so WHAT distinguishes WIgoh from ALL the Other apps 6. WIgoh includes it ALL 7. wigoh is a location BASED inforMAtion 8. but ALso gives eVENT based inforMAtion 9. AND 10. SOcial BAsed information 11. so WHAT'S in it for the Users 12. we have a what's HOT list 13. WHERE all the hot events are of the neighbourhood ARE [inaudible] 14. THEN the SOcial en ASpect of it 15. WHERE you can find your FRIENDS ON 16. you can FIND your favourite ORganisers 17. and you can all bookMARK THEM 18. AND you JUST STAY TUNED 19. BY a CAlendar in the app

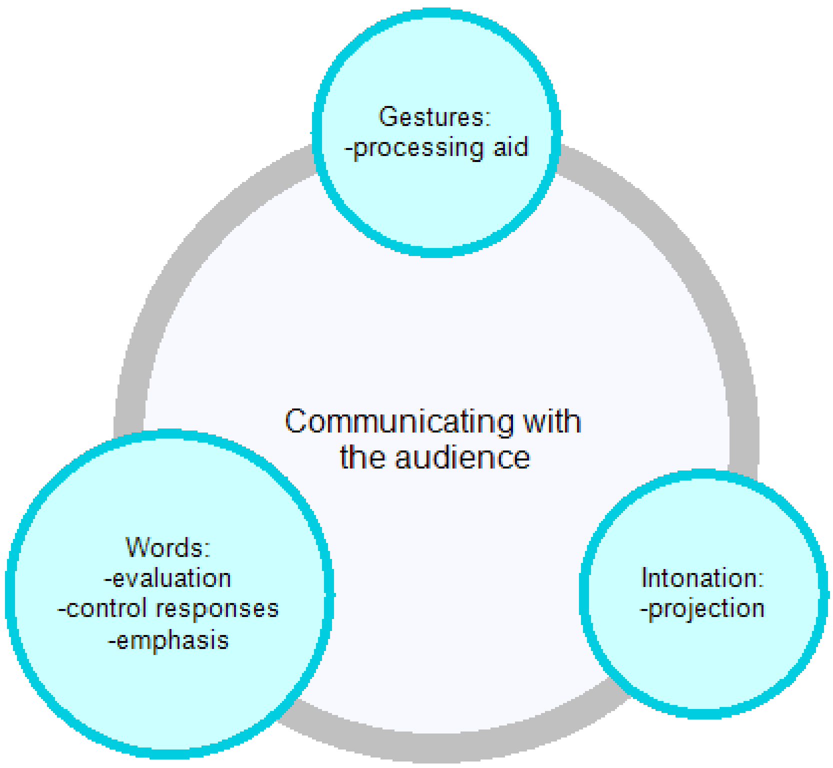
Gestures and head movements

WHERE all the hot events are of the neighbourhood ARE

THEN the SOcial en ASpect of it

and you can all bookMARK THEM

Multimodal ensemble



Analysis of words:

The product is immediately and explicitly presented as the solution to a need, which implicitly evaluates it in positive terms (the very first words in the presentation, unfortunately not captured in the recording, are actually an appeal to the audience to picture themselves in a situation in which they would benefit from using the app). This is something that was explicitly recommended during the training the speaker received.

The product is also explicitly differentiated from competitors ("distinguishes"), and the name is repeated for memorability. In addition, as he describes the type of information the app provides he makes a three part list that also contributes to memorability. Interestingly, towards the end of the excerpt, the functionality of the app is also organised as a three part list with three items: a hot list, the social aspect (embedding yet a third three part list) and a calendar.

The speaker also uses several rhetorical questions to engage the audience and the way the questions are framed prompts a particular response, since they imply that there is indeed something unique and valuable for the users: they do not question "Is there something that distinguishes it?" or "Is there something in it for the users?", but takes this for granted.

Analysis of intonation:

The prominence in "the" as he pronounces "the solution" (unit 1) makes this word selective, and adds the meaning "it is not just one possible solution, it is the *only* solution". The next tone unit presents an unusual concentration of prominent syllables, which can be accounted for by the fact that the speaker is actually explaining the meaning of the acronym that makes up the name of the product. It is necessary to take into account, however, that a remarkable feature of intonation in this extract is the overuse of prominent syllables, i.e. the speaker makes no clear difference between prominent and non-prominent. This provides the speech with a very even pace that sounds slightly unnatural and recited. During the interview the speaker commented that he was actually intentionally slowing down the pace to emphasize the name of the product, but also admits that using a foreign language might also be influencing the overuse of prominence. Also the training he received insisted on memorising the text, which explains the "reciting" intonation.

Unit 5 begins with a high pitch (+61 Hertz) that adds the meaning of discourse initial, and is coherent with the initiation of a new sub-move ("differentiation of the product"). However, the lack of a marked tone contour prevents the catchy phrase (and self-answer to the rhetorical question) "Wigoh includes it all" from standing out in unit 6.

The prominence in "also" in unit 8, and "AND" in unit 9 (which actually stands on its own as a separate tone unit) contributes to create a sense of accumulation as he lists the possibilities of the app and highlights the added value that differentiates it from other apps. The high pitch in "and" (+92 Hertz) also adds a nuance of surprise and innovation. Brief pauses after these two words (see micase transcript) also contribute to emphasize each item in the list.

It is also interesting that the "on" that closes unit 15 is made relatively long. This is a way of indicating that discourse is not finished, and implying that the list of possibilities is endless, which indirectly evaluates the product positively and differentiates it from competitors.

Finally, unit 18 begins with another prominent "and" which has the same cumulative effect as in unit 9. During the interview the speaker confirmed this interpretation of prominences and pauses in the last units, but he also thinks encoding difficulties might have influenced his use of intonation at this point, as a way of winning time while planning the message.

Analysis of gestures:

As the speaker pronounces “neighbourhood” in unit 13 he draws the shape of a semi-sphere with both hands, which represents the delimited space implied by “neighbourhood”. This is a metaphoric, referential gesture. It is remarkable that the gesture is not synchronous with intonation, because "neighbourhood" is actually not prominent.

Next, as he pronounces “social” in unit 14 the fingertips of the index fingers touch, representing the contact implied by “social”. This is again a metaphoric, referential gesture.

Finally, he mimics the action of scrolling down a page on a digital device as a visual representation of what he is saying as he pronounces “bookmark” in unit 17. This is another referential gesture, but this time it is iconic.

Analysis of persuasive activity:

The speaker in this excerpt is using the following persuasive strategies:

-Processing aids (gestures), through gestures that work as visual representations

-Emphasis (words), achieved through stylistic devices such as repetition, three part lists

-Evaluation (words) of the added value of his product as something positive (”the solution”)

-Projection of context of interaction: intonation signals ongoing discourse, which provides a sense of an endless list of advantages that his product offers, while high pitch adds a sense of innovation and pauses contribute to the memorability of each item in this list.

-Control responses (words), prompting to share a positive evaluation in a subtle way through the implicature in the question.

Through these strategies the speaker is making the text easy and memorable, while showing the added value of the product.

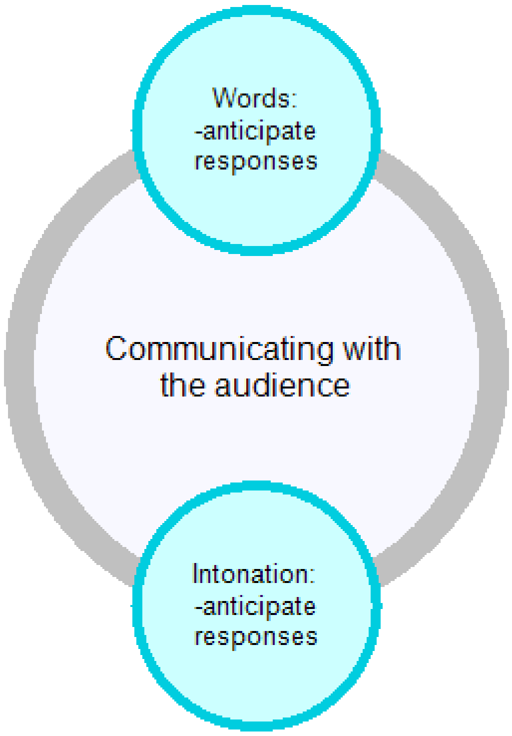
The overuse of prominent syllables makes the text less fluent. This does not contribute to the persuasive effect. The speaker commented during the interview that he believed it was important to project a relaxed attitude and avoid memorizing, something he actually does not achieve fully. He explained that he received intensive coaching to prepare for the presentation during the days preceding the event. Probably this had the effect of taking away spontaneity.

PPKE2

In this rich point the speaker is closing the presentation about his app with an appeal to the audience.

Orthographic transcription
So I have just one thing to say, just download the app and you can see what’s in it. Thank you for your attention.
MICASE transcription
so, i have just one thing to say, just download the app, and you can see what’s in it. thank you for your attention.
DI transcription
1. so i have JUST one thing to SAY → 2. just DOWNload the APP ↘ 3. and you can SEE what’s in it → 4. THANK you for your attention →
Gestures and head movements
 just DOWNload the APP

Multimodal ensemble



Analysis of words:

The speaker makes a direct appeal to audience through an imperative. The audience is prompted to experience the app, hence the use of a perception verb ("see").

Analysis of intonation:

A rather flat intonation makes this catchy close go largely unnoticed. There is, however a significant high termination in unit 2 (+ 425 Hertz) that prompts an active role from the audience and matches the imperative form of the verb. The speaker confirmed this was consciously done for this effect of prompting people to action. The effect is however somehow counteracted by the level tone and mid termination and key in the last unit that closes the speech.

Analysis of gestures and head movements:

Near absence of gestures and head movements reduces the effectiveness of this closing, which is intended to be appealing but is not prominent or memorable enough. The only gesture in this extract (pointed out by the speaker himself during the interview) is a subtle shrugging of shoulders synchronous with "download" that actually gives away a sense of disengagement that contradicts the call to action. The speaker interprets it as a sign of feeling tired, and willing to finish the stressful experience of the presentation.

Analysis of persuasive activity:

Persuasion in most of this presentation is framed as a demand-supply: there is a trendy market niche that we know how to feasibly occupy. It is rational and it works towards highlighting what differentiates the product and the feasibility of the market idea.

The closing presented in this rich point, however is more emotional. The speaker is trying to engage the audience and prompt a response (try the app) by appealing to sensorial experiences ("see what's in it"). The focus here is not on the practical value of the product, but on its emotional value as something simply enjoyable.

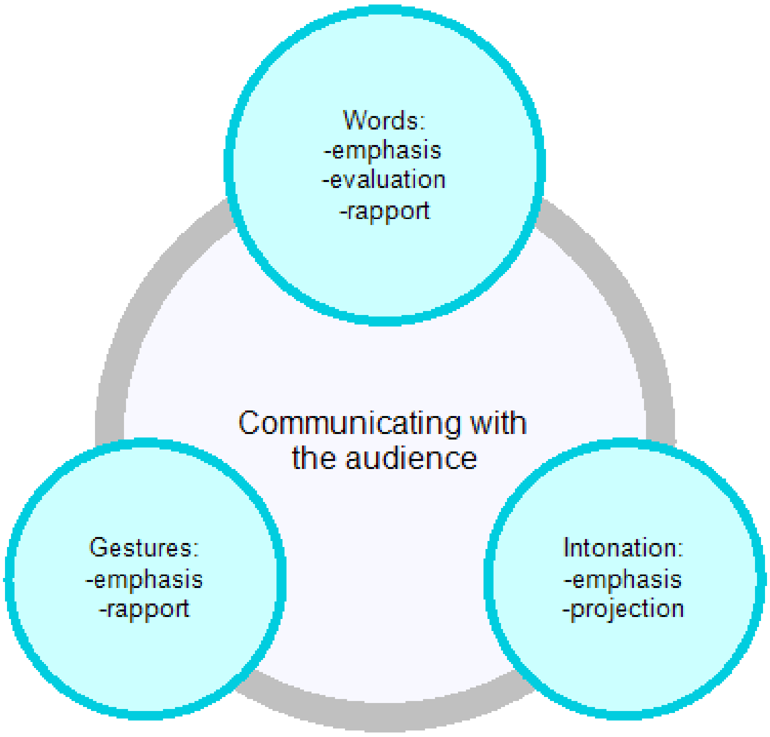
The persuasive effect would then benefit from interpersonal gestures that get the audience involved and prompt them to action, which are, however, lacking in this extract. A more marked intonation would also help make this appeal more prominent and likely to be reacted upon. The speaker agrees with this interpretation and actually mentioned that if he were to repeat this presentation he would actually mimic the action of downloading or even demonstrate it with a phone instead of holding a flicker in his hands.

PPPA1

In this rich point the speaker begins his presentation identifying a need that his product can cover.

Orthographic transcription
You have an app. That app is important to your business, important to your business success. Because of that, you care about how your users experience that app, and you care about how your users feel about that app. But sometimes, you deliver a bad experience. Sometimes, you deliver a real bad experience.
MICASE transcription
you have an app. that app is important to your business, important to your business success. because of that, you care about how your users experience that app, and you care about how, your users feel about that app. but sometimes, you deliver a bad experience. sometimes, you deliver a real, bad experience.
DI transcription

1.	you HAVE an APP	↘
2.	that APP is imPORtant to your business	↘
3.	imPORtant to your business sucCESS	↘
4.	beCAUSE of THAT	↗
5.	you CARE about how your users exPErience that app	↗
6.	and you care about HOW your users FEEL about that app	↘
7.	but SOMEtimes you deliver a BAD experience	↘
8.	SOMEtimes you deliver a REAL bad experience	↘

Gestures and head movements
 imPORtant to your business sucCESS you CARE about how your users exPErience that app but SOMEtimes you deliver a BAD experience SOMEtimes you deliver a REAL bad experience
Multimodal ensemble


Analysis of words:

The speaker directly addresses the audience using the pronoun “you” and emotional language (care, feel). The speaker confirmed this was a way a building rapport with the audience, but also remarked that it was a way of highlighting the USP (unique selling point) of his product: an app that provides information about how people react emotionally when using other apps.

His consistent use of repetition is also quite remarkable, for example in nouns and verbs that carry key concepts in his message, such as “important”, “business”, “care”, “users” and “app”. In particular the repetition of the word “app” directs attention to the product and the market gap it is trying to fill. There are further repetitions that have an intensifying effect and provide the text with a symmetrical structure that makes it more memorable ("But sometimes, you deliver a bad experience. Sometimes, you deliver a real bad experience). The pseudo repetition of "important to your business, important to your business success” is a rephrasing that expands the idea conveyed.

The extract closes with an extreme negative evaluation (“real bad experience”). With this evaluation the speaker is showing a potential danger that can be avoided with this product. During the interview the speaker commented this effect was planned and totally intended. The speaker explained how this initial part was a “rehearsed and scripted presentation” because according to him most people can’t “speak and think at the same time”. This would spoil the effect he intended for his opening, which was to build “emotional resonance” with the audience. In his own words, this was his emotional hook”.

Analysis of intonation:

The speaker uses fall tones in most of the units except in units 4 and 5. This indicates that he is treating the information in this extract as mainly new (the two rise tones most likely indicate that discourse is still not finished, rather than shared information). This use of intonation is a powerful awareness raising tool towards the problem and market gap, but it also contrasts with the way he is lexically framing the information as something that is potentially familiar to the audience (he seems to be asking them to picture a common situation with “you have an app”). The speaker confirmed during the interview that he actually expected the audience to feel identified with the hypothetical situation he was presenting so that a sense of “shared conspiracy” is created, but also agrees that part of the intention is also raising awareness about this potential jeopardy. In unit 3 the prominence on “success” suggests an expansion of the idea initiated in the previous tone unit. The prominence in “real” in unit 8 emphasizes the negative evaluation in “real bad experience". A brief pause immediately following “real” also adds to this effect.

Analysis of gestures and head movements:

Beats with a pointing finger to the audience mark the discourse flow emphasizing key words. They also involve the audience and convey the idea that they are directly affected by the problem presented. Therefore, they have a pragmatic (parsing) function and also an interpersonal function (they try to get the audience involved as directly affected). In particular, beats accompany the underlined words in the following tone units (see the DI transcription for a whole list of tone units):

3. Important to your business <u>success</u>

5. You <u>care</u> about how your <u>users</u> <u>experience</u> that app

7. But sometimes you deliver a <u>bad experience</u>

8. Sometimes you deliver a <u>real</u> bad experience

These beats are generally synchronous with intonation except in the case of “users” in unit 5 and "experience” in unit 7, because these words are not prominent.

During the interview the speaker described these gestures as “scolding” and agrees with the idea that they are meant to prompt people to consider a potential danger and to do something about it.

Analysis of persuasive activity:

The speaker in this extract is using the following persuasive strategies:

-*Emphasis* through beats, prominence, pauses, repetitions and use of emotionally intense language;

-*Projection of context of interaction*: tone choice marks information as new;

-*Rapport*: the addresses to the audience through pronouns (you, your), emotional language (care, feel) and gestures (the beats with an extended finger pointing to the audience) present it as something that is directly affecting them;

-*Evaluation*: the speaker presents a negative evaluation of a potential scenario that his product will help avoid;

With these multimodally realised strategies he tries to:

-*show the value of the product*: presents a market gap that his product can fill;

-make his presentation *memorable*: repetitions and parallel structures make it easy to remember;

-make his presentation *surprising*: in the negative framing that differentiates the message from the general trend in marketing;

-be subtle: words present a hypothetical scenario while intonation and gestures make it directly applicable to the audience in a subtle way.

It is interesting how the speaker directly appeals to the audience creating the impression that each member of the audience is being addressed individually and that his product can solve their particular problems. This can be related to a need to fight general resistance to “ready-made” advertising-like messages that are not spontaneous and honest. This is consistent with some ideas the speaker reported during the interview, mainly that it was important for him to sound “genuine”, and that the key to a good pitch is to think about the “what’s in it for the audience” and making it match their interests. In his own words you have to focus on “why would anyone be interested in this” because “if the why doesn’t resonate with them then they won’t care about any other technical aspect”.

It is also remarkable that the speaker highlights not only what his product offers, but also what the audience may lose by not choosing it: if your app delivers a bad experience you lose customers. This is a trend initiated in this extract and which the speaker maintains throughout the presentation. Research on persuasion (O’Keefe, 2002; Perloff, 2003) suggests this strategy is particularly effective when:

a) the potential negative effect is perceived as real and threatening enough

b) the action required to avoid this negative effect is perceived as feasible

The speaker reported during the interview that he opted for this framing for the sake of originality, to differentiate himself from exaggeratedly positive marketing messages.

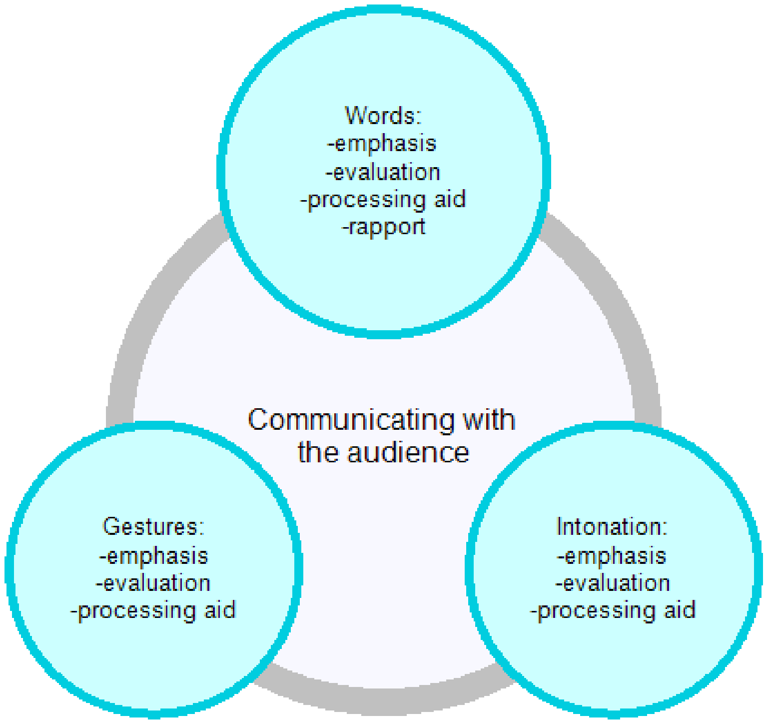
PPPA2

In this rich point the speaker explains what his application can do and what differentiates it from others.

Orthographic transcription
Uxpro is an analytic service that integrates in your app and helps to understand and measure your users' experience and how your users feel about that app. It includes two important concepts. First behaviour and satisfaction tracking in real time. What your users are doing, whether they are failing, whether they are succeeding, how long they need to do things, how they feel about that app. And secondly, really important thing, because the combination of these two things is the real magic. It's, it's collection of users' sentiment in a very pragmatic way with feedback and microsurvey. Letting your users tell you how they feel about how they are doing with your app.
MICASE transcription
uxpro, is an analytic service that integrates in your app and helps to understand and measure your users' experience, and how your users feel about that app. it includes two important concepts. first, behaviour and satisfaction tracking in real time. what your users are doing, whether they are failing, whether they are succeeding how long they need to do things, how they feel about that. and secondly really important thing, because the combination of these two things, is the real magic. it's_it's collection of users sentiment in a very pragmatic way with feedback and microsurvey letting your users tell you how they feel, about how they are doing with your app.
DI transcription
1. UXpro is an Analytic service 2. that INtegrates in your APP 3. and helps to undersTAND and MEAsure 4. your Users exPErience 5. and how your users FEEL about that APP 6. it includes TWO imPORtant concepts 7. FIRST 8. beHAviour and SAtisfaction tracking in real time 9. what your users are DOING 10. whether they are FAILING → 11. whether they are sucCEEding → 12. how LONG they need to do THINGS → 13. how they FEEL about that → 14. and SEcondly 15. REAlly imPORtant thing 16. because the COMbination of these two things 17. is the REAL MAgic 18. it's it's colLEction of users SENtiment 19. in a VEry pragmatic way 20. with feedBACK and microSURvey 21. letting your users tell you HOW they FEEL 22. about HOW they are DOING with your app

Gestures and head movements
 Beats with open hand accompany prominent words that describe how the product works (e.g. synchronous with “understand” or “measure”):  Beats with a finger ring (e.g. synchronous with “behaviour and satisfaction tracking”).  because the COMbination of these two things is the REAL MAgic

Multimodal ensemble



Analysis of words:

The speaker makes use of ordering and listing (“two important concepts”, “first”, “secondly”) to help understanding and make the message coherent and easy to follow.

The intense positive evaluation of the product (“really important”, “real magic”) makes it stand out from competitors.

Finally, the use of pronouns that directly appeal to the audience (“your users”, “tell you”, “your app”) make the message sound relevant to them.

Analysis of intonation:

Prominence in “two” anticipates the listing and “first” and "secondly” are also prominent. These three words also have a high key (+ 75, 80 and 70 Hertz respectively) that is usually associated to discourse initial elements. Both features help to make the structure of discourse clear. Parallel intonation contour (level) through the items in the list of things the product can measure (“whether they are failing, whether they are succeeding…”) adds a sense of accumulation and conveys the usefulness of the application (note that sentence structure is also grammatically parallel). “Feel” is made prominent in units 5, 13 and 21) as a way to emphasize this key concept which is added value of his product (ability to measure how users feel). The prominent syllables in the two last units underline the parallel structure (how-feel/ how-doing) and make the text easy to remember and memorable.

The three brief pauses marked in yellow in the micase transcript give time for the message to sink in and provide the words immediately preceding with some emphasis. The speaker also reported that he consciously tried to slow down for the sake of comprehensibility, especially because he was the only native English speaker.

Analysis of gestures and head movements:

This excerpt is characterised by a frequent use of beats that fulfil two functions:

-Beats with open hand accompany prominent words that describe how the product works (e.g. synchronous with “understand” or “measure”):

-Beats with a finger ring seem to have a more pragmatic function, symbolising the idea of concreteness and specificity and highlighting a unique selling point (e.g. synchronous with “behaviour and satisfaction tracking”). The speaker actually confirmed he usually associates this gesture with “a subtlety”:

Apart from beats, there is also a referential gesture, with fingers coming together to visually convey the idea of “combination”. The gesture is later repeated synchronous with “real magic”.

Analysis of persuasive activity:

The message in this excerpt is transmitted in a very coherent way through the ordering and listing, which makes it easy to follow and likely to be accepted. Frequent beats, prominence and high hey on signposting words contribute to this coherence. The speaker highlights the unique value of the product differentiating it from others, and this is achieved through evaluative lexicon, gestures that convey the idea of specificity (e.g. beats with the shape of a finger ring) as well as prominence on key words (e.g., “feel”) and intonation that adds up an endless list of possibilities of the product. The speaker also appeals directly to the audience through the use of pronouns (e.g. “your users”), while the extreme positive evaluation of the product (e.g. “real magic”) seeks concurrence from them. Interestingly, the speaker mentioned in the interview that he was not happy with this choice of words because “to me I feel the phrase ‘real magic’ is almost a little bit too marketing-ish and I might avoid it if I’m superconscious”

PPPI1

In this rich point the speaker is justifying why using insects for food.

Orthographic transcription

The first question you might ask is why would I bother eating insects? Well, first of all, they are delicious. They taste like nuts, bacon, popcorn; they have all kinds of flavours.

MICASE transcription

the first question you might ask is, why would i bother eating insects. well, first of all, they are delicious. they taste like nuts, bacon, eh popcorn they have all kinds of flavours.

DI transcription

1. the FIRST question you might ASK is
2. WHY would i BOther
3. EAting INsects
4. well FIRST of ALL
5. THEY are deLIcious
6. they TASTE like NUTS
7. BAcon
8. eh POPcorn
9. they have ALL kinds of FLAvours ↘

Gestures and head movements

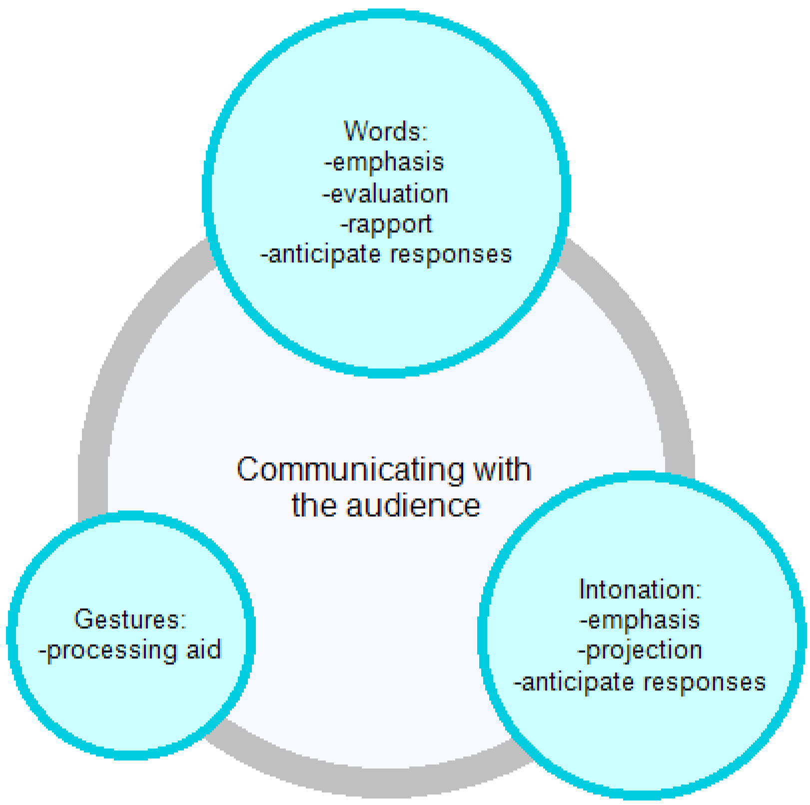
they TASTE like NUTS

BAcon

eh POPcorn

they have ALL kinds of FLAvours

Multimodal ensemble



Analysis of words:

The speaker uses a rhetorical question to anticipate a reaction of the audience, in a way putting himself on their shoes and building rapport. In providing an answer to that question he guides a particular response in favour of eating insects. Interestingly, starting his answer with “First of all” implies there is more than one reason. The extreme evaluation in “delicious” and “all kinds of flavours” also invites to share this positive perception. Finally, the three part list (nuts, bacon, popcorn) provides a sense of completeness and contributes to memorability.

Analysis of intonation:

Units 4 and 5 begin with a high key (+ 69 and + 38 Hertz respectively) that is associated to surprise and denial of expectations. The speaker is presenting this information as potentially shocking for the audience. Unit 5 also ends with a low termination that adds a meaning of "this is discourse final". In doing this the speaker is closing any potential argument, so on the one hand he is acknowledging the information might be shocking, and on the other hand he is leaving no room for debate. Unit 6 begins with a low key (-95 Hertz) that has an equating effect, and matches the purpose of giving examples (delicious= nuts, bacon, etc.) In unit 9 “all” and "flavours" are remarkably prominent. The speaker also uses a high termination (+63 Hertz) that can be interpreted as adding a nuance of "I know this is surprising", together with a fall tone (-121 Hertz) that marks this information as new.

Analysis of gestures and head movements:

As the speaker is listing the different flavours he indexes each item with his fingers. This is a cohesive gesture that is synchronous with the prominences in “nuts”, “bacon” and “popcorn” and brings these different items together, contributing to easy processing and memorability.

As he pronounces “all kinds of flavours” he makes a horizontal open hand supine gesture with both arms moving outwards and downwards in a circular movement that draws a semi-sphere. This metaphorical, pragmatic gesture represents the scope implied by the words. It takes place immediately after the first prominence (“all”) and continues after the prominence in “flavours”.

Analysis of persuasive activity:

The speaker begins anticipating a potential objection from the audience (why insects), and then states the value of the product appealing to feelings (it’s tasty), and immediately following this extract also to ethical or moral values (it’s healthy and it's environmentally friendly). These two strategies boost the interpersonal relationship with the audience. This in turn helps the message to be perceived as not imposed. The self-answered rhetorical question including evaluative language works towards this aim. The use of intonation clearly contributes to this effect as well, presenting information as shocking and at the same time preventing potential argument.

The appeal to feelings, and in particular taste, makes his evaluation more subjective and emotional than rational. This can be related to some of the views about the genre that the speaker reported during the interview: for a good product pitch speakers needs to show their motivation and enthusiasm. His way of personally endorsing the product referring to something as personal as taste (it’s delicious) is a way of showing this personal involvement and motivation.

The stylistic device of the three part list makes the text more memorable, and is supported by the use of counting fingers that make serve as a processing aid and cohesive device. The lexical extreme evaluation is also supported by a pragmatic gesture suggesting its scope ("all kinds of flavours") which seeks alignment with the audience in a non-imposing way.

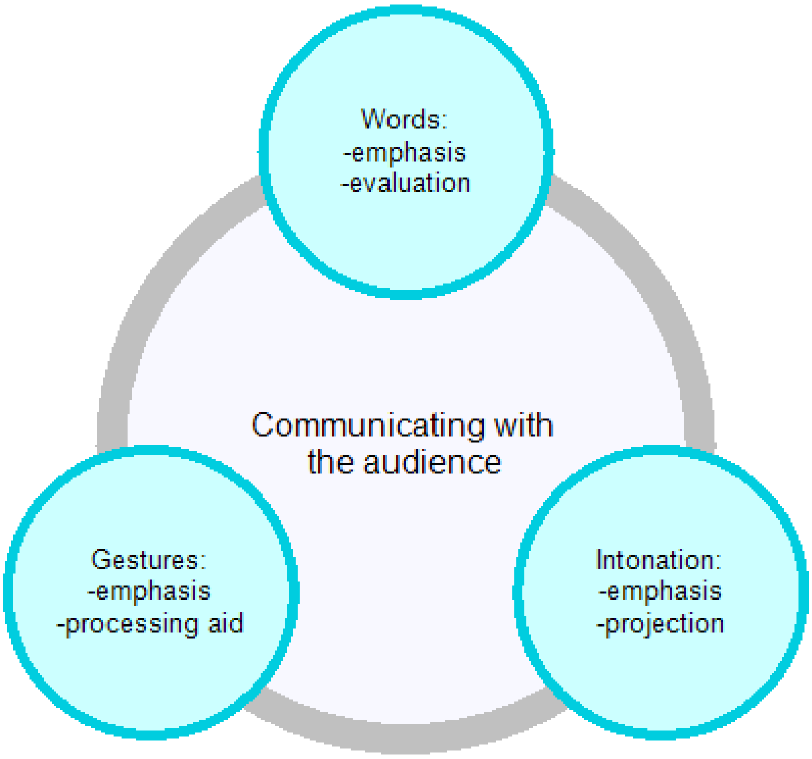
PPPI2

In this rich point the speaker is justifying why the selection of the burger as a format for their product.

Orthographic transcription
We also compared all the different possibilities for insect based food, going from protein shakes, to burgers, to ibars and it turns out that a burger is really the best thing to put on the market right now.
MICASE transcription
eeeh we also compared all the different possibilities for eh insect based food, going from eeeh, eeeh, protein shakes to burgers to ibars. and it turns out that a burger is, really the best, thing to put on the market right now.
DI transcription
1. eeeh we ALso comPARED 2. ALL the different possiBIlities 3. for eh INsect based FOOD 4. going from eeeh eeeh PROtein SHAKES 5. to BURgers 6. to IBARS 7. and it $^{\text{TURNS}}$ out that a BURger is 8. $^{\text{REAL}}$ly 9. the BEST THING 10. to PUT on the MARket right now

Gestures and head movements
 going from eeeh eeeh PROtein SHAKES to BURgers to IBARS  REALly the BEST THING

Multimodal ensemble



Analysis of words:

The speaker uses a three part list to present the options for insect-based food, and uses extreme evaluation with an intensifying adverb and a superlative ("really the best") to justify the selection of one of them: the burger.

Analysis of intonation:

It is remarkable that "burger" (unit 7), "really" (unit 8), "best" "thing" (unit 9) are prominent, because they are the words that carry the positive evaluation of the choice he is referring to. The high key in "turns out" (unit 7, +44 Hertz) and "really" (unit 8, +19 Hertz) adds a nuance of surprise and further emphasizes this choice that the speaker is presenting as preferable.

The fillers before "protein shakes" reveal encoding difficulties. The two brief pauses before "really" (unit 8) and "thing" (unit 9) seem to prepare the audience for the extreme evaluation and add emphasis to it.

Analysis of gestures and head movements:

This extract presents a widespread use of beats throughout, but I have focused on the ones that add extra meaning that can be related to persuasion. The first are beats indexing the list. As he lists the options for insect-based food he reinforces each item with OHS beats, the first one with his left hand, and the rest with his right hand. This is a referential gesture that also acquires a cohesive function in bringing the items of the list together. There are more beats in synchrony with the words “really”, “best” and “thing”. Interestingly these words carry the lexical evaluation, and the beats, with the fingertips of both hands approaching, are used to further emphasize the appropriateness of the choice of the burger, with a pragmatic function.

Analysis of persuasive activity:

The speaker uses a stylistic device (a three part list) to make his text easier to remember. The beats that index each item in the list also provide cohesion and make the text easier to understand. Evaluation in this extract is realized strictly speaking through words, but the other persuasive strategies realised through other modes seem to support it:

-the intensifying adverb "really"

-the beats that emphasize key words

-intonation: prominence in the key words (“really”, “best” and “thing”) and a higher pitch in “really”

-pauses that prepare the audience for this extreme evaluation and add emphasis to it

By making this evaluation so salient the speaker is implicitly seeking concurrence from the audience.

PPSE1

The speaker is differentiating the product (i.e. a waxing device for ski equipment) and comparing it to competitors.

Orthographic transcription
But our product is safer, and faster and easier to use than waxing kits, and it offers better quality than achieved at ski shops.
MICASE transcription
but our product is, safer and faster and easier to use than waxing kits, and it, offers better quality, than achieved at ski shops.
DI transcription
OUR product is SAfer →   and FAster →   and EAsier to use than wAxing kits ↗   and it offers BEtter quality than aCHIEVED at ski shops →

Gestures and head movements

OUR product is ^{SA}fer

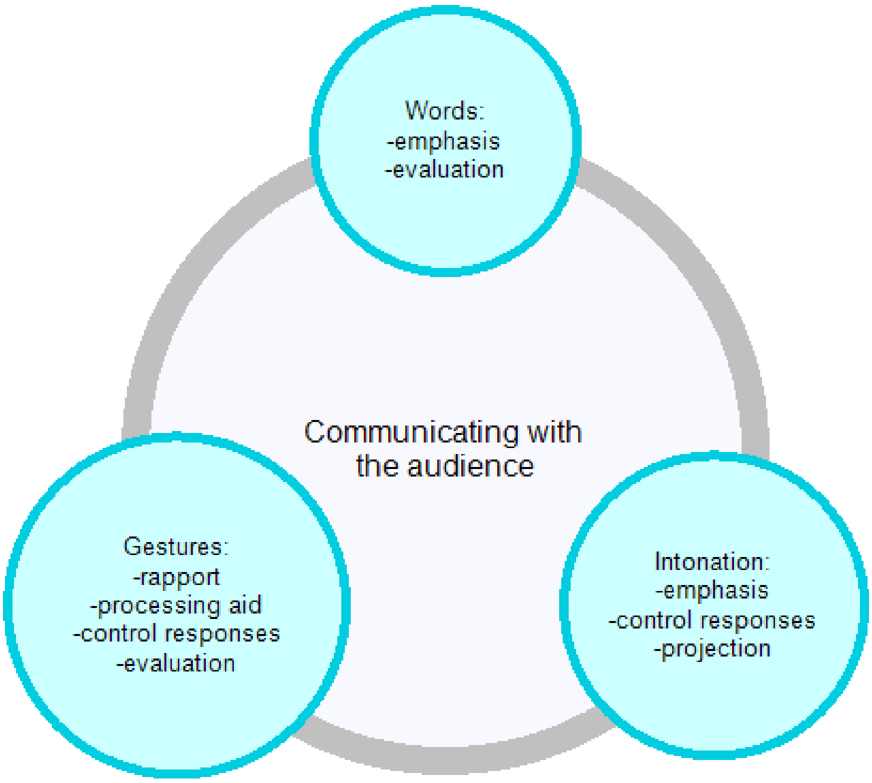
OUR product is ^{SA}fer

and ^{FA}ster

and EAsier to use than $_{WA}$xing kits

and it offers BEtter quality than a$_{CHIEVED}$ at ski shops

Multimodal ensemble



Analysis of words:

The speaker uses positive adjectives in comparative form to evaluate the product and compare it to competitors positioning it as a better option (e.g. safer, faster, and easier). During the interview the speaker confirmed that these adjectives were important because they were the basis of his value proposition and his unique selling point. In addition, he uses a three-part list to enumerate the advantages which provides the text with a predictable grammatical structure. The repetition of "and" also conveys the sense of accumulation. These stylistic devices make the text more memorable.

Analysis of intonation:

The words “our”, “safer”, “faster”, “easier”, “better” are all made prominent. These are important words that identify the product described and qualify it with positive features, and the speaker emphasizes them so that they are remembered.

Though tone is rather flat throughout, the rising tone (+170 Hertz) in “easier to use than waxing kits” is referring, and can be interpreted as presenting information as already agreed upon and not open to discussion. In fact, the speaker recalled during the interview how he had already mentioned this topic before, hence his treating it as shared knowledge.

The high key and termination in "faster" (+33 Hertz) can be interpreted as having a contrasting effect that supports the comparison with competitors. The use of high termination in the two first tone units (48 and 33 Hertz increase respectively) add a nuance of surprise and innovation (what differentiates his product). It also prompts an active role from the audience and invites to judge this evaluation of the product, but interestingly the speaker shifts to a lower termination in the last two tone units (12 and 18 Hertz decrease respectively), which prompts a more passive role and suggests finalization of the topic.

Analysis of gestures and head movements:

The speaker makes a self-deictic gesture as he pronounces "our", drawing attention to the speaker and developer of the product. This gesture in phonological synchrony with words and intonation creates rapport and helps identify the product differentiating it from the rest.

As he starts his list of advantages that his product offers he counts with fingers, visually indexing each item. This cohesive gesture makes this part of the message easy to process. According to the speaker his intention was also to emphasize each item in the list, because they are key in his USP.

The speaker points backwards to the screen as he pronounces “waxing kits” (i.e. competitor products), but keeping his body orientation towards the audience. This is important because by doing this the speaker is making sure the audience does not focus on the image projected but on himself, consequently it seems to suggest that what is being projected does not deserve attention. In addition, the pointing with a thumb suggests that the exact location is not important (Kendon 2004). Therefore, the gesture can be interpreted as disregarding and negatively evaluating what is being pointed at. During the interview the speaker also explained that this gesture is partly a result of his training (he was taught never to face back), but agrees that the way of pointing is disregarding and entails a negative evaluation that he hopes the audience to share.

The final open hand gesture is in phonological synchrony with “better” and metaphorically “offers” this improvement mentioned to the audience. This visual offering metaphor helps to process the message in real time. The speaker added in the interview that this was also a way to show that his product is for everybody and not for the elite. This adds an extra function to the gesture: it is trying to make the audience feel included and build rapport. In this sense, it is also interesting that he makes a brief pause right before “offer” for effectiveness.

Analysis of persuasive activity:

Through the use of a combination of modes the speaker draws attention to bits of the message that he wants the audience to remember. He also makes the text easy to understand and subtly prompts the audience to share an evaluation without being imposing.

According to the speaker himself the strongest point of his presentation was that it was “very logical”, and used a simple problem-solution frame including only basic facts. This helped him remember his script and also helped the audience to follow. This can be related to his use of indexing with fingers and the kinesic metaphor of offering, both aimed at easing understanding. The speaker also mentioned during the interview that as a member of the audience he liked a particular presentation because it was “captivating”. The stylistic devices that he uses in his own presentation (e.g. the list) seem to work towards this same aim (i.e. engage the audience and make them remember).

PPSE2

The speaker closes the presentation with a catchy slogan.

Orthographic transcription

We hope that you can help us with emancipating the winter sporter. Thank you.

MICASE transcription

we hope that, you can help us, with emancipating, the winter sporter. thank you.

DI transcription

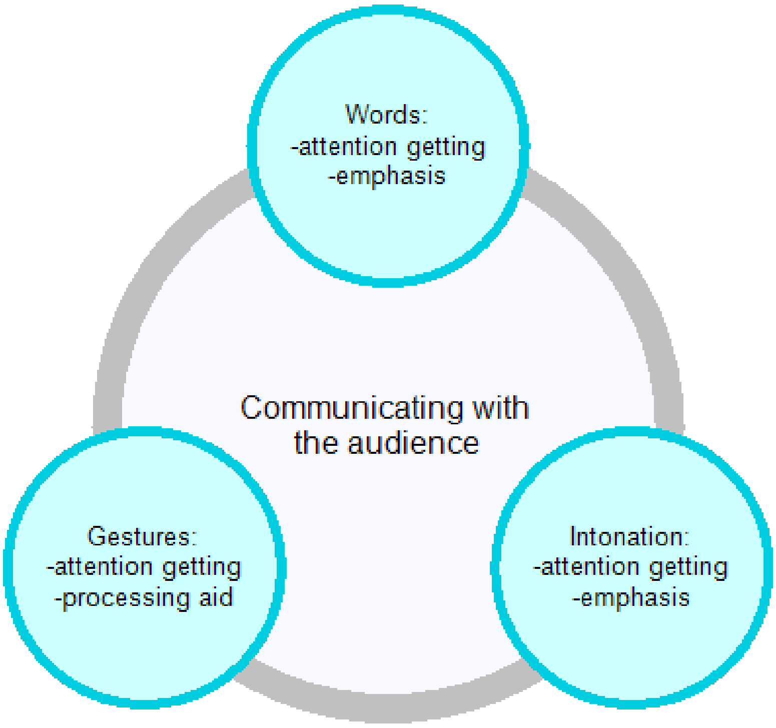
1. we HOPE that
2. YOU can help US
3. with eMANcipating
4. the winter SPORter
5. thank you

Gestures and head movements

YOU can help US

YOU can help US

Multimodal ensemble



Analysis of words:

The speaker uses a slogan which is a direct address to the audience and an implicit request for action. The lexicon used (“help us emancipate”) expresses in a very intense way both the market gap and the solution his product provides:

Need= skiers are forced to hire waxing services or buy expensive equipment that is difficult to use (they are trapped).

Solution= his product frees (“emancipates”) skiers with an easy to use and non-expensive kit.

It is remarkable that the slogan is also used at the beginning of the presentation. A circular structure is thus achieved that contributes to memorability. During the interview the speaker reported that this was consciously rehearsed and enacted during the presentation, precisely to make it memorable.

Analysis of intonation:

The slogan is split into several tone units, and “us” coincides with the end of the second one, becoming salient. The brief pauses in between units slow down the pace and let the message sink in, and according to the speaker they are consciously done for memorability. It is also interesting that in unit 2 “You” and “us” are made prominent, making clear who are the participants involved in this call for action. These words are further emphasized through gestures.

Analysis of gestures and head movements:

The speaker makes an OHS gesture pointing to the audience synchronous with "you" in unit 2 (“You can help us”). It is a deictic, referential gesture that visually clarifies who is the addressee in this call for action. Interestingly, as he pronounces “us” in the same tone unit he makes a self-deictic, referential gesture that also clarifies in a visual way who is the other party involved.

Analysis of persuasive activity:

In this extract the speaker is emphasizing important parts of his message and using processing aids. This makes the text easy to understand and likely to be remembered. The processing aids are materialized through gestures, but the emphasis and attention getting is achieved through a combination of words (intense lexicon), intonation (prominence in relevant words such us "you" and "us") and gestures (synchronous with the pronouns).

His use of persuasive strategies is consistent with his concern to make his presentation "logical" and “easy to follow". The use of visual representations and his need-solution framing both serve this purpose. The circular structure that makes the text easy to remember is also coherent with his preference for "captivating" presentations.

PPT01

The speaker gives an example of a potential user for his product in the form of a narrative and explains how the product works.

Orthographic transcription

Now meet my sister. Her name is Ariane, she is a gastroenterologist. She takes a lot of time to take care of her patients and to talk with them. Like the other doctors she draws badly, but her patients understand her. Why? Because she owns Anapad. Anapad is a small plastic tablet, with parts of the body printed on it, and you can write down and erase, very easily.

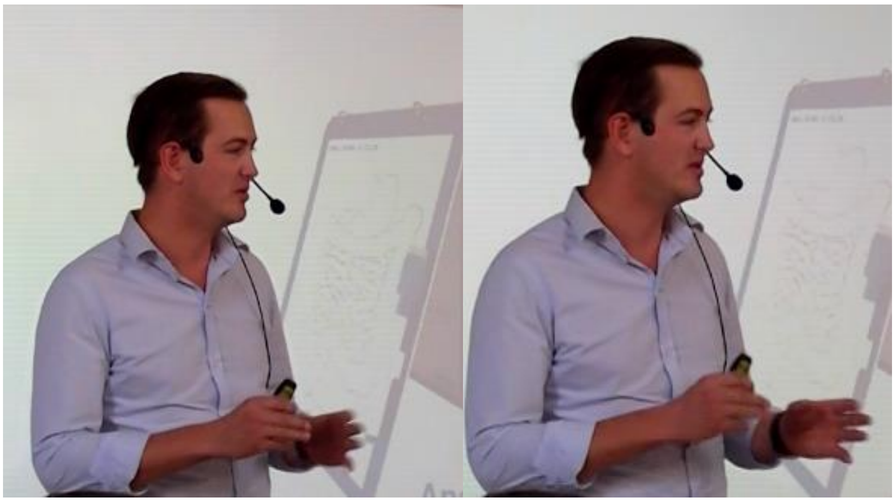
MICASE transcription

now meet my sister… her name is Ariane, she is a gastroenterologist. she takes a lot of time to to take care of her patients and to talk with them. like the other doctors she draws badly, but her patients understand her. why, because she owns anapad. anapad is a small plastic tablet, with, parts of the body printed on it, and you can write down and erase. very easily.

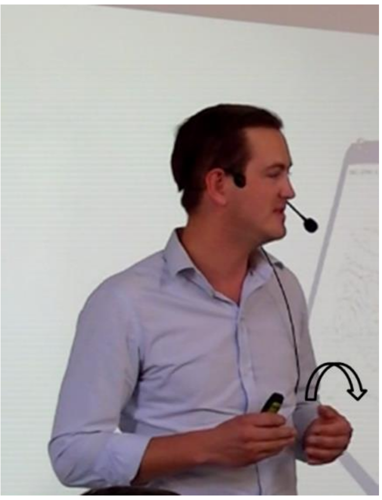
DI transcription

1. NOW meet my SISter →
2. her NAME is ariANE ↗
3. she is a GAStroenteROlogist →
4. she TAKES a lot of TIME →
5. to to take CARE of her PAtients ↗
6. and to TALK with them ↗
7. like the other DOCtors she draws BADly ↘
8. BUT →
9. her PAtients undersTAND her →
10. WHY →
11. beCAUSE she OWNS anapad ↘
12. anapad is a SMALL plastic TAblet →
13. with PARTS of the body PRINted on it ↗
14. and you can write DOWN and Erase ↘
15. VEry easily ↘

Gestures and head movements



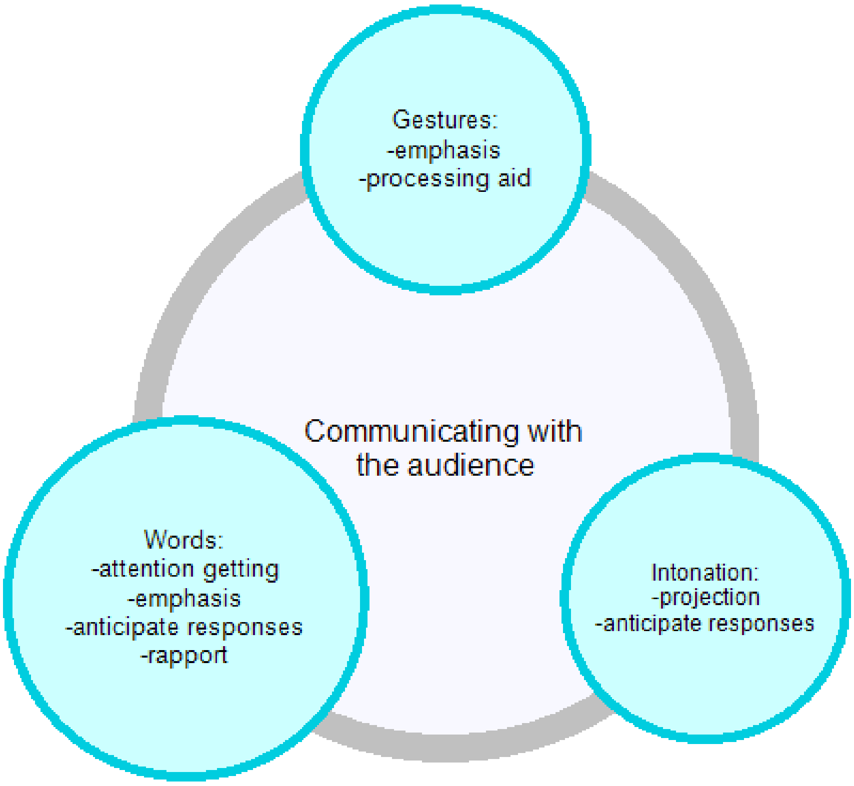
with PARTS of the body PRINted on it



with PARTS of the body PRINted on it

and you can write DOWN and Erase

VEry easily

Multimodal ensemble


Analysis of words:

The speaker tells a personal anecdote that is meant to create rapport, as is the narrative description of product capability (note the use of “you” in “you can write down” to make it personal).

The speaker uses this anecdote to compare the general situation of doctors (“like the other doctors”) and the specific case of her sister, which establishes a contrast between what we have vs. what we need, and therefore highlights the market gap for his product.

The Why? –question anticipates the potential reaction of the audience (surprise, reversal of expectations) and introduces a response to it. The speaker confirmed that his intention with this question was to make people wonder and engage them in reflection. As a matter of fact, he remembers he had planned to make a longer pause right after “why” for effect, but finally missed it because he was nervous during the presentation and spoke too fast.

The name of the product is repeated for memorability. During the interview the speaker reported that the name had been chosen as a marketing technique, precisely because it was easy to remember and would help the audience to perceive this product as real.

Analysis of intonation:

Only half of the tone units carry a tone movement, which makes intonation rather flat throughout. This flat intonation makes the anecdote sound recited and not spontaneous, which in turn makes it less credible and less effective. Still, there are some interesting uses of intonation:

In Units 5 and 6 (to take care of her patients and to talk with them) the rise tone employed (+31 and +62 Hertz respectively) suggests this is treated as shared knowledge: everybody knows that doctors need to communicate effectively with their patients. This sets the ground to establish a need for the product he is trying to sell: a tablet to ease communication. Similarly, in unit 7 (like the other doctors she draws badly) the speaker uses a subtle fall tone (-22 Hertz) that matches the new information he is providing about his sister, but interestingly the mid termination prompts concurrence with this situation and accept it as something widespread. The added meaning is: generally doctors' ability to draw is not enough to communicate with their patients in an effective way. Again, the speaker is reinforcing the need or market gap. The next tone unit, number 8 ("but"), has a subtly higher pitch (+22 Hertz with respect to previous prominence) that adds the meaning increment of “contrary to what you might expect”. The same effect is achieved in unit 10 (why?) (+53 Hertz) which anticipates this reversal of expectations experienced by the audience. Unit 11 (because she owns Anapad) works in a similar way to unit 7: it uses a fall tone (-56 Hertz) that indicates this is new information, but ends with mid termination that prompts concurrence. The speaker uses this to present his product as a feasible solution to the need previously established: the tablet helps doctors communicate effectively.

Analysis of gestures and head movements:

As the speaker says "parts of the body" he first brings his hands together and then moves his left hand outwards. The gesture is not totally synchronous with words and intonation; it slightly follows the prominence in "parts". This metaphoric gesture can be interpreted as representing the range of body parts available.

Immediately following the previous gesture, as the speaker pronounces "printed on it" he moves his hands from a resting position close to his body to an open hand vertical position in a semi-circular movement outwards. This metaphoric gesture can be interpreted as symbolising the availability of the "printed" information which visually "comes out".

As the speaker pronounces "write down and erase" a beat is repeated twice (one time at each side) to reinforce the contrast write/erase. The speaker also pointed out during the interview that first time he makes this beat he actually also briefly mimics the action of writing, as if holding a pen with one hand and writing on his open palm.

Finally, as he says "very easily" both hands with palms facing up (OHS) move circularly in a metaphoric gesture that suggests the smooth flow conveyed by "very easily".

All these gestures are referential and make the meaning easier to process for the audience. The beats can also be considered cohesive, since they connect both elements of the contrast write and erase.

Analysis of persuasive activity:

Persuasion in this rich point is built around two key aspects: creating rapport with the audience and showing a market gap. To that aim the use of narratives in this piece is especially appropriate because:

-it creates rapport and shared ground with the audience

-it provides a specific example that is more convincing than abstract figures

It is interesting that during the interview the speaker reported that for him the most important thing in a pitch is to "get the attention of the audience" and mention something “they are used to”. This can be related to his use of narratives to create rapport with the audience, something that he remembers as actually a conscious effort. As he puts it himself, his intention was to “impersonate the problem” and make himself likable.

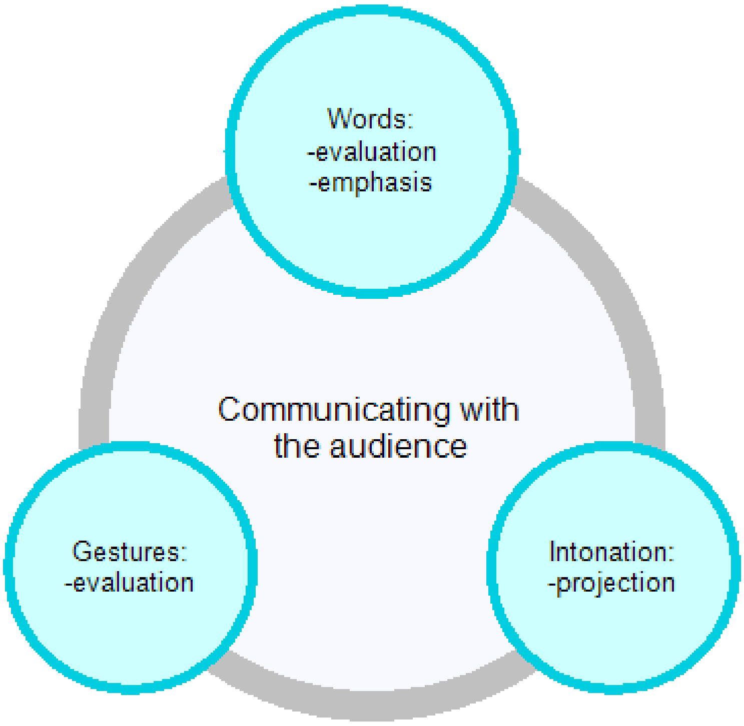
Altogether, this makes the message credible and more memorable. However, the speaker uses a rather flat intonation that makes the text sound recited at points (in Brazil's terms, oblique orientation), especially as he introduces the narrative with “now meet my sister”. This takes away vividness to the narrative because it does not sound spontaneous enough, and in turn makes it less credible. As a matter of fact the speaker admitted during the interview that he “rehearsed a lot”, almost memorising the text.

It is also notable that the few gestures employed are all concentrated in the very last part of the extract, as the speaker describes the product. This is probably because the speaker needs more visual support to convey the way users can interact with the product and show its potential use. The speaker confirms this interpretation, and during the interview he reported having “a visual mind”, and a tendency to explain things trough drawings.

PPT02

The speaker explains the benefits of his product and compares it with competitors.

Orthographic transcription
Pharmaceutical companies always need to, to, to, are always eager to spread their image across the doctors by offering what they can. Sometimes it’s a travel, sometimes it can be a useless gadget. With our product eh, the, with our product the, the advantage is that we offer them something that the patient and the doctor can use in a very positive thinking way and it’s a good image for the company.
MICASE transcription
pharmaceutical companies, always need to to eh to are always eager to spread their image across the doctors, by offering what they can. sometimes it’s a travel, sometimes, it can be, a useless gadget. with our product, eh…the the_with our product the, the advantage is that we offer them something that the patient and the doctor can use in a very positive thinking way, and it’s a good image for the company
DI transcription
1. pharmaceutical COMpanies always need to to eh to → 2. are always Eager → 3. to spread their Image across the DOCtors ↗ 4. by Offering what they CAN → 5. someTIMES it’s a TRAvel ↗ 6. someTIMES it can BE → 7. a USEless GADget → 8. with our proDUCT → 9. eh the the → 10. with our proDUCT the → 11. the adVANtage is that we offer them something ↘ 12. that the PAtient and the DOCtor can use → 13. in a very POsitive thinking way ↘ 14. and it’s a good iMAGE for the company ↗

Gestures and head movements
 a USEless GADget
Multimodal ensemble


Analysis of words:

The speaker uses lexical evaluation to compare his product with competitors (useless vs. positive), also intensified through an adverb (very).

The repetition of “sometimes” provides the text with a parallel structure. There is another repetition of “with our product” which rather than being intentional reveals encoding difficulties (as further suggested by the filler “eh”). The speaker confirmed this difficulties and explained during the interview that he “lost track” at this point. This has a certain negative effect of bringing down the credibility of the message. Also, it makes the comparison between his product vs. competitors less powerful because it makes both elements compared more distanced, and delays the positive evaluation which reduces the impact of the contrast useless vs. positive.

During the interview the speaker pointed out that he had planned to say “pharmaceutical companies always try to” but came up with “are always eager” which is actually more intense language and helps emphasize the need or market gap. He also intended to say “offer a product” and came up with “offer them something”, which is too generic and neutral to be perceived as positive. He also regrets not repeating the name of the product once more for memorability.

Analysis of intonation:

Intonation is again quite flat throughout, but there is an interesting use of rise tone in unit 3 (+28 Hertz), which suggests the speaker is treating this as shared knowledge already agreed upon. Interestingly, this corresponds to his establishment of a market gap (pharmaceutical companies need to spread their image). On the other hand, the fall tones in unit 11 and 13 (-50 Hertz) conveys this is new information. It is remarkable that the semantic content expressed here corresponds to what differentiates his product from other "useless gadgets", and is coherent to treat it as new, since it is the unique selling point of the product and the take-home message.

Analysis of gestures and head movements:

As the speaker says “a useless gadget” he makes an OHS deictic gesture with his right arm moving outward which is interesting because of its unspecificity: the speaker does not point with precision to direct attention, but rather represents a range of useless things that pharmaceutical companies offer and which are at that moment projected on the screen, but asking the audience not to focus on them. Interestingly he does not shift focus either, but keeps focusing on the audience. The gesture can also be interpreted as discarding, emphasizing that his product can be a much better gift for doctors and prompting to share this evaluation. It acquires then a pragmatic function. The gesture is in semantic and phonological synchrony with words and intonation: it starts slightly before the first prominence in unit 7 and continues up to the second prominence.

Analysis of persuasive activity:

The speaker uses the following three persuasive strategies in this extract:

-*Emphasis (words)*: through the repetitions and intense language

-*Projection of context of interaction (*intonation): marks new vs. old information

-*Evaluation (words, gesture)*

Persuasion in this extract is mainly articulated through evaluation, and it is interesting to see that this positioning of his product is enacted multimodally: lexicon positively evaluates his product (very positive), and negatively evaluates competitors (useless), but this lexical evaluation is co-expressive with the discarding gesture that negatively evaluates competitors and with the fall tone that highlights the unique selling point as new information.

In short, the speaker highlights what differentiates his product and makes it unique using intonation, gesture and words. At the same time, he makes his message not imposing. The prompt to share an evaluation through gestures is more subtle than if it were expressed through words. Note also how the comparison at the lexical level is not explicit, rather it has to be inferred (there is no comparative use of adjectives that explicitly positions his product as better). Finally, the speaker makes his message more memorable trough the parallel structure achieved by the repetition of “sometimes”. However, the encoding difficulties revealed through the unintended repetitions and the false start (and confirmed by the speaker) make the message less credible.