# Time gestures:

# Time Conceptualisation in English with Evidence from Gestures in a Multimodal Corpus

Daniel Alcaraz-Carrión

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

Department of Linguistics and English Language

Lancaster University

August 2018

# Statement of ownership

I declare that this thesis is my own work and has not been submitted for the award of a higher degree elsewhere. However, it is important to mention that certain sections of the PhD were elaborated within the framework of an external research group that form the CREATIME Project. The CREATIME research group aims to understand human creativity by studying how humans represent time. CREATIME proposes that temporal constructs are much more complex, creative and flexible than we currently think, and in order to expose these patterns of creativity, we aim to go beyond the lab and the dictionary, comparing big data from speech, film, poetry and gesture. CREATIME is funded by the Seneca Spanish foundation as well as the Spanish Ministry of Economy (Ministerio de Economía y Competitividad), with 10.800€ and 21.500€ respectively. The CREATIME team is formed by several researchers across Europe: Cristobal Pagán-Cánovas and Javier Valenzula-Manzanares from University of Murcia, Inés Olza and Adriana Gordejuela from Universidad de Navarra, Anna Piata from University of Neuchâtel and finally myself, Daniel Alcaraz-Carrión from Lancaster University. Each of the researchers is specialised in certain modes of communication. For instance, Adriana Gordejuela's work focused on temporal patterns in films through the study of flashbacks while Anna Piata's work focused on time and creativity in through language. My role in this research group was to coordinate and lead research on time and co-speech gestures through the NewsScape library. With the assistance of Javier Valenzuela-Manzanares, we collected a large amount of data that served both for the aim of the research group as well as the development of this thesis. The linguistic and gesture data was collected collaboratively with the assistance of a number of undergraduate students. However, I performed the analysis of all the gesture dataset present in this thesis, which became part of CREATIME's temporal gesture database for future research. All the stages that involve the participation of the CREATIME team or other research assistants include a footnote indicating the work that they developed.

# Abstract

This thesis is the first large-scale, corpus-based analysis of time conceptualisation using multimodal 'big' data. It exploits the NewsScape Library, an extensive television database, to investigate co-speech gestures for temporal expressions in English.

This research investigates time conceptualisation by studying gesture patterns that co-occur with temporal linguistic expressions. While much has been written about time conceptualisation based on linguistic evidence, there has been comparatively little based on other modalities. Non-linguistic data has the capacity to help triangulate existing findings if found to be congruent with linguistic data or, conversely, could offer new insights into the way time is conceptualised. Gestures are an invaluable tool for the study of cognition due to their key role in human communication. Gesture is a universal phenomenon, largely unconscious and less monitored than speech and thus it can reveal information not present in language.

In this thesis, I investigate three features of co-speech gestures: axis and direction; language-gesture congruency; and extent or distance. The dataset comprises gestures that co-occur with temporal linguistic expressions which can be further divided into three groups: non-spatial language (e.g. *earlier than*); spatial, directional language (e.g. *back in those days*); and spatial, non-directional language (e.g. *far in the future*).

This research confirms that the lateral axis is frequently employed when gesturing about time, but it also suggests that spatial language uses the sagittal axis more frequently that non-spatial language. Moreover, gestures that co-occur with spatially grounded temporal language tend to be congruent with the canonical direction of time, while this does not occur with non-spatially grounded temporal language. Finally, it seems that temporal distance expressed linguistically is usually analogous to spatial distance expressed through co-speech gestures.

The thesis also reflects on theoretical and methodological issues for gesture studies and some possible steps forward in the study of gesture.

# Acknowledgments

This thesis would not have been possible without the support and assistant of a wide number of people around the world. During this period, I have experienced personal and academic growth that everyone has contributed to and for that I will be forever grateful.

First and foremost, I would like to acknowledge the North Western Consortium Doctoral Training Partnership for their financial and academic support throughout the development of my thesis.

I would like to thank my supervisor, Chris Hart. Your academic guidance and comments during the development of the thesis have been invaluable. Our academic discussions, always humorous and insightful, have helped me to mature and develop both personally and professionally. I cannot think of a better person that could have supervised my thesis.

I would also like to thank Professor Jonathan Culpeper and Doctor Aina Casaponsa for their comments at different stages of the development of my PhD. I want to specially thank Doctor Willem Hollmann for help and comments during my PhD and for giving me the chance to get involved in teaching. Through our conversations I have learned to be a better teacher and researcher.

I am very grateful to my undergraduate supervisor and friend Javier Valenzuela, as well as Cristóbal Pagán. I always have in my memory that meeting in which you warned me about the dangers of becoming a researcher, and I cannot be happier of having made that choice. Thank you for believing in me.

The PhD can sometimes be a lonely journey that requires physical and mental endurance, something difficult to maintain when you are away from home. I was lucky enough to encounter a number of people during my stay in Lancaster that helped me during the process. I want to thank Alicia, for helping me and the rest of the PhD students to keep up with our work when we needed to. To the other Phd students, Javi, Salomé, Kayla, Giulia, Liviana, Roy, Laura, Carolina, Ola and Max. Thank you for the coffee breaks in the department, the beers in the pub and the academic discussions. During my time in Lancaster I also have been able to meet the type of unique people that leave a mark for the rest of your life. I want to thank Emma for being there during

all this journey. From birthdays spent working until later in the department, to trips to London, Murcia or Madrid, and the recent "camp thesis" resort, you have made me feel like "en my home". Thank you for being a true friend. I also want to thank Valeria for all her support during my PhD, even though we are thousands of miles away. Thank you for the conversations and discussions, for the churros at 7 in the morning in Madrid and for always being there. Thank you for being my family in the UK. Lastly, I want to thank Kas. Your support and encouragement have been essential during the last stages of the PhD. I cannot describe your importance in my life, and I will always be grateful for you believing in me.

También hay un número de personas que han tenido un papel crucial en estos tres años a pesar de que no se encontraran en Lancaster. Gracias a mi familia, a mi madre "Flugen" y a mi padre José Pedro por su apoyo y su amor incondicional. Gracias por haberme dado esta oportunidad y gracias por haber confiado en mi en todo momento. Esta tesis es tan vuestra como mía. También quiero dar gracias a mis hermanos Pablo y Raúl, a mis tías Araceli y Marisa. Gracias por confiar en mi y por ayudarme a llegar donde estoy. También me gustaría dedicar esta tesis a mis abuelos. A Saturia, a Luis, a Ginés y a Maruja. Gracias por todos estos años de enseñanzas, siempre os recordaré.

Por último, me gustaría dar las gracias a uno de los pilares fundamentales sobre los que me he apoyado durante estos años, que son mis amigos. Gracias a los "Petazetas"; a Alberto, Mario, Alejandro, Lina, Marina Díaz, Marina Peña, María y Blanca. Gracias por hacerme sentir que no ha pasado el tiempo cuando vuelvo a casa; sois un apoyo esencial. Espero que pronto podamos hacer uno de nuestros reencuentros cerca de un surtidor de gasolina. Gracias a mis amigos del grupo scout "La Milagrosa"; sois demasiados como para poner un listado de nombres. Gracias por las risas, las salidas, los campamentos y las revisiones. Pero, sobre todo, gracias por formar parte de mi vida desde que tengo 8 años. Habéis tenido un papel fundamental en mi desarrollo personal y académico.

# Table of contents

# List of figures

# List of tables

# List of abbreviations

| | |
|---|---|
| **CL** | **Cognitive Linguistics** |
| **CMT** | **Conceptual Metaphor Theory** |
| **T-FoR** | **Temporal frame of Reference** |
| **SPG** | **Source-Path-Goal** |
| **FoR** | **Frame of Reference** |
| **D-Time** | **Deictic Time** |
| **S-Time** | **Sequential Time** |
| **F** | **Figure** |
| **G** | **Ground** |
| **T-Span** | **Temporal Span** |
| **E-Time** | **Extrinsic Time** |
| **S-FoR** | **Spatial Frame of Reference** |
| **CIT** | **Conceptual Integration Theory** |
| **MTL** | **Mental Timeline** |
| **ATOM** | **A Theory of Magnitude** |
| **GS** | **Gesture Studies** |

# Chapter 1

# Introduction

## *1.1 Time and cognition*

Time is perhaps one of the most challenging and complicated concepts that human beings experience. We do not know exactly what it is, or how it came to be, but nevertheless it is an unavoidable part of our existence. I believe we could distinguish three different facets of time we humans deal with. The first one would deal with the nature, reality and objective of time, or to put is simple: what is time? This facet is beyond the scope of this research and is the preserve of other fields in philosophy and metaphysics, addressing the more "objective" characteristics of time and trying to investigate what it is and how it functions. The second one would refer to the experiential basis of time. All human beings have to experience time from birth until death. We all experience ageing, we see how other humans and living beings grow, wither, are born or die. Time is something unavoidable and hence it is a universal phenomenon that is experienced by all of us. The last facet of time, and the one I will be focusing on in this thesis, has to do with how humans perceive time. Even though we all experience time, the way we perceive, understand and conceptualise this concept is subject to construal. This involve a number of processes that are linked to social, cultural and cognitive factors.

One of the most outstanding human cognitive processes is the ability we have to be able to understand and talk about abstract concepts, like time. But in order to do so, we need to talk about these concepts in terms of more concrete domains of experience that share some similarities with the abstract concept. This phenomenon was addressed by Lakoff & Johnson (1980) within the frame of Cognitive Linguistics (CL) through the development of the Conceptual Metaphor Theory (CMT). Conceptual metaphors are a cognitive phenomenon in which a particular semantic area (target domain) uses properties from a different semantic area (source domain). Since time is an abstract concept, it is usually understood in terms of a different domain which is closer to our embodied experience. Traditionally, metaphor has been linked to a more stylistic approach to language; it was thought of as a tool that creative writers employed to elegantly present their work. CMT takes a very different point of view by suggesting

that human reasoning is largely metaphorical, affecting not only language but how we conceptualise the world.

In the case of time, we understand time as it if were a resource, represented in language by expressions such as *running out of time*, *spending time* or *borrowing time*. In these cases, we are talking (and thinking) about time in terms of a resource; time acquires a number of properties that are intrinsic to resources, like the ability to be shared or given, or its quantification.

However, one of the most important domains of experience that language and cognition employ when conceptualising time, and the focus in this research, is the domain of space. Space and time are two interdependent elements that are experientially linked, and this is reflected in language by a large amount of expressions that use the spatial domain to refer to several aspects of time. We can, for example, say that *past times are behind*, or that *the future is yet to come*. There is a lot of research on time conceptualisation based on English linguistic evidence (see Bender et al. 2010; Boroditsky, 2011; Moore 2000, 2006, 2011, 2014; Evans 2004, 2013, amongst others), and it addresses a wide range of linguistic phenomena that support the idea of a strong linguistic (and conceptual) link between space and time. Other authors have addressed time conceptualisation in grammatical phenomena, such as tense and aspect. Here, for example, it is proposed that spatial concepts such as 'boundedness' account for aspectual distinctions, e.g. perfective vs. imperfective (Langacker, 2008). Point of view is also held to feature in the semantics of tense-aspect constructions (ibid).

These theories aim to illuminate conceptual processes through linguistic analysis. However, if these theories are correct, since language reflects general conceptual processes, we should expect to see those same conceptual processes indexed in other modes of communication. While much has been written about time conceptualisation based on linguistic evidence, comparatively little analysis has been made of time conceptualisation based on other communicative modalities. Recent research has emphasised the importance of co-speech gestures for studying metaphorical forms of conceptualisation (Cienki & Müller, 2008; Casasanto, 2009; Marghetis & Bergen, 2014). Specifically, in the case of time, for example, Núñez & Sweetser (2006) studied co-speech gestures in Aymara speakers. They found that the gestures congruent with language placed the past in front of them and the future behind

them. Casasanto and Jasmin (2012) used co-speech gestures to analyse time conceptualisation in English. In this case, however, they found a dissociation between language and gesture. While most of the temporal linguistic expressions in English are related to the sagittal axis, the spontaneous gestures produced while speaking about time were more frequently located in the lateral axis.

Gesture data emerges then as a high-quality data source in order to triangulate existing findings in time conceptualisation if found to be congruent with linguistic data, while, if it is found to differ from the linguistic data, it could provide new insights into time conceptualisation.

### *1.2 Aims and objectives*

The aim of this thesis is to investigate patterns in time-space conceptualisation through the analysis of linguistic data and temporal co-speech gestures. The project exploits the NewsScape Library, an extensive television database constructed as part of the Little Red Hen lab which comprises almost fifteen years of television in English, with more than 250,000 hours of data and more than 2 billion word textual database formed by subtitles. Data is collected from news and talk shows, providing high-quality, contextualised data of individuals that are not aware of contributing to empirical research, and hence it avoids misleading data that could emerge from laboratory conditions.

Through the analysis of large quantities of linguistic and gesture data, this thesis aims to answer the following research questions:

1) Do speakers gesture differently in terms of axis and direction when using spatial and non-spatial linguistic temporal expressions (*e.g. far in the future vs after that*)?
2) When speakers use the spatial domain to talk about time, are there any differences on co-speech gestures in terms of axis and direction when they explicitly state temporal direction vs when it is not stated (*e.g. back in those days vs distant past*?
3) Is there any difference on gesture realisation in terms of axis, direction or congruency in past-related expressions vs future related expressions?
4) Are co-speech gestures congruent with the canonical direction of time represented in language and in other cultural artifacts? Is there any

difference on gesture-language congruency when different temporal linguistic expressions are used?

5) Is temporal distance also reflected on co-speech gesture? Do speakers gesture close to them when referring to temporal proximity and gesture far when referring to temporal distance?

In order to answer these questions, language was chosen as the point of departure in this research. English temporal linguistic expressions were divided in three different groups based on their (non) relation to the spatial domain: non-spatial language (e.g. *earlier than)*; spatial, directional language (*e.g. back in those days*); and spatial, non-directional language (*e.g. far in the future*).

The result was a temporal gesture database that contained around 450 co-speech gestures evenly distributed amongst the three different categories. The analysis of the gesture dataset focused on three key features: 1) axis and direction of gesture; 2) congruency of the gesture with the canonical direction of time; 3) gesture location. Both qualitative and quantitative analysis was performed in order to obtain an accurate picture of the temporal co-speech gesture features.

This research will address a number of issues that so far have not been covered in previous work on temporal gestures in English speakers. Previous literature has addressed temporal co-speech gestures by focusing on temporal language as a whole; the main focus being the perspective that was adopted (Ego-Moving or Time-Moving) or the past-future orientation of the temporal expression (see Section 2.2). This study aims to take the field one step further by focusing on the differences that might arise from the nuisances of the lexical items that are employed to talk about time.

The first research question addresses the possible differences between the use of purely temporal language and spatial temporal metaphors. The absence of a space-time mapping in the first type of language could have an effect on the way these temporal expressions are conceptualise in contrast with the influence of the mapping of the spatial domain in the second type of language. The hypothesis is that we could expect a lesser influence of the spatial domain in temporal language than in space-time metaphorical expressions; spatial properties that are attributed to time due to the space-time mapping (e.g. directionality, orientation) will be less consistent in non-spatial temporal expressions.

This idea is further explored in the second research question, addressing the same differences but in two different types of spatial temporal realisations. Spatial, directional language explicitly states the direction of the temporal period (e.g. *back in those days*), while non-directional language makes no reference (e.g. *distant past*). This research aims to investigate the gesturing pattern is consistent in these two types of gesture realisations. Previous research on time gestures establishes that speakers consistently employ the lateral axis when gesturing about time (see Section 4.3. This analysis will determine whether this is also true when speakers explicitly state temporal direction, located in the sagittal axis or if, on the contrary, gestures mirror language. Besides, this data will be contrasted with the gestures obtained from non-directional language in order to establish if the pattern is consistent across all types of spatial language or if the explicit mention of the temporal location has an effect on the co-speech gesture and time conceptualisation.

The third research question takes a different approach by focusing on the semantics of the temporal expressions rather than on the type of language that is employed. The same types of gesture features are addressed (axis, direction, congruence), but this time the contrast focuses on past linguistic expressions vs future linguistic expressions. The comparison between past and future expressions in the study of temporal co-speech gestures has been a frequent topic (see Section 4.3), although this dimension has never been addressed by using big data. The hypothesis is that past and future co-speech gesture will follow the tendencies that has been observed in previous studies about temporal gestures, that is, when the temporal expressions is referring to the past, the gesture will be located on the left, behind or above them when talking about the past while they will gesture on their right, in-front-of them or below when talking about the future.

The next research question addresses a feature, congruence, that has not been often addressed in previous gesture literature (although it is mentioned on Walker & Cooperrider, 2016). Researchers often assume that temporal co-speech gestures are congruent with the canonical direction of time, due to the influence that the spatial domain has on the temporal domain (see Sections 2.2 and 4.3). However, this assumption is based on usually small quantities of data that were obtained from laboratory experiments. We would expect then to find similar results when analysing

large quantities of data; the co-speech gestures that speakers employ will often agree with the canonical direction of time.

The final research question addresses another gestural feature that so far has not been investigated in any previous research on temporal co-speech gestures: distance. So far, research on time has established that there are features that are mapped from the spatial domain to the spatial domain, such as direction and location, that are present in temporal gesture realisations. (see Section 2.2 and 4.3). However, it is not clear if other spatial features, such as distance, are also represented by co-speech gestures. The analysis of large quantities of data will allow us to establish if this feature is also consistently represented through co-speech gestures.

### *1.3 Thesis overview*

This thesis is arranged in eight different chapters, Chapter 1 being the introductory chapter. Chapters 2, 3 and 4 focus on a literature review of several fields that serve as the basis of the development of this thesis, namely Cognitive Linguistic, multimodal data and gestures; and their relation to time conceptualisation

Chapter 2 focuses on time conceptualisation in language. Here, I briefly describe the main tenets of Cognitive Linguistics, including the different cognitive processes based on Gestalt psychology such as categorisation, frames, embodiment, image-schema or construal operations and I finalise the section by giving a brief account of the Conceptual Metaphor theory. The chapter focuses on the link between language and time-space cognition. Here I describe the experimental proximity between time and space, and later I provide a detailed account of the different types of Temporal Frames of Reference (T-FoR) that previous authors have suggested based on linguistic data. After that, I describe current research on the relation between time and directionality in language. I show, through the use of linguistic data, how several aspects that are only found in space (dimensionality and directedness) are mapped onto temporal concepts through metaphorical linguistic realisations. This sub-section finished by offering a brief account on the topic of language, tense and aspect from a Cognitive Grammar perspective. Finally, the last section of the chapter addresses the importance of several cultural artifacts on language and temporal conceptualisation. Here I describe the role that temporal material anchors, timelines and other cultural artifacts (e.g. calendars) play in time conceptualisation.

Chapter 3 surveys experimental studies of time conceptualisation. This chapter aims to address the importance of using multimodal data in order to investigate conceptual patterns in human cognition. This chapter covers a selection of multimodal and multidisciplinary research that addresses temporal cognition, including behavioural, psychological and neurological research. Some of the topics covered in this chapter tackle the relation between time and spatial experience, the representation of time and space in the brain, the relation between sound, length and duration, or the influence of extra linguistic factors on time conceptualisation. The chapter also includes research on time conceptualisation taking a cross-cultural and cross-linguistic perspective. Here I present research that addresses differences in temporal cognition in languages such as Spanish, English, Mandarin, Swedish or Pormpuraaw, amongst others, and more precisely how these languages employ different spatial frames of reference when conceptualising time, with an effect on the axis and direction.

Chapter 4 focuses on the importance of gesture as a method to study cognition. This chapter offers an overview of the field of gesture studies, indicating the importance of gesture data for the study of cognition and, more precisely, temporal cognition. This chapter offers an overview of the basis and development of gesture studies. It starts by stating a definition of gesture, taking into account perspectives from a wide range of research. Following, I discuss which are the main theories and approaches in gesture studies, giving more emphasis on the theories developed by McNeill (1982). Following, the chapter addresses different methods to approach gesture analysis. Here I offer an overview of the main theories on gesture analysis based on its form as well as its function. Finally, the chapter offers an overview of the current research that has been performed on gesture and temporal cognition. I firstly focus on gesture research performed in English language speakers, and I finish the chapter by referring to some of the research that have investigated temporal gestures in other languages and cultures.

Chapter 5 focuses exclusively in the methodology that has been employed for the development of this thesis. Here I discuss several methodological aspects that involve the database and the data that I have employed, stages in data collection and data analysis as well as an overview of the final dataset. It also provides an overview of the different software and tools that have been employed for the development of the thesis. Here I describe the particularities of the multimodal corpus from which the data has been obtained, the NewsScape library. Next, I describe other tools that have been

employed for the collection and analysis of the data, such as the Edge search engine, CQPWeb and ELAN. Following, the chapter describes the methods employed for the storage and arrangement of the data. It also addresses the data collection process. I distinguish here two different stages: the first stage focuses on the linguistic search criteria that establishes three different groups of temporal linguistic expressions, while the second one focuses on the clip selection criteria for gesture analysis. The next section moves to the next methodological stage by explaining the data analysis process. Here I describe both the qualitative as well as quantitative analysis that has been performed. The chapter finishes by offering a description and overview of the dataset.

The following three chapters, chapter 6, 7 and 8 focus on the analysis of the dataset as well as the different results that have been obtained.

Chapter 6 addresses the analysis of the axis as well as the directionality and congruency of the gesture dataset. Firstly, it offers a general overview of the gesture dataset, indicating the distribution of the gestures in relation to these two features. The chapter further presents the gesture axis data as well as discussing the patterns and tendencies in gesture axial location that have been observed in the three different groups of temporal linguistic expressions. I discuss here gestures that have been performed in the lateral, vertical and sagittal axis as well as "punctual" gestures. Next, the chapter addresses the topic of gesture congruency by comparing the canonical direction of time in a linguistic expression (past-left/future-right) with the direction of the gesture. More specifically, I analyse the level of congruence of these gestures in the three different types of linguistic expressions. Following, the chapter also addresses congruency, but this time I analyse congruency levels between past-related and future-related expressions. The chapter finalises by offering a summary of the results that have been obtained.

Chapter 7 presents an analysis of the gesture location and its distance in relation to the temporal distance that is indicated by the linguistic expression. The chapter starts by offering an overview of the data and the aim of the analysis. Next, it looks at the data from a quantitative perspective, looking at gesture frequency in both distance-related and proximity related language. The chapter analyses in more detail distance-related and proximity-related language, presenting a number of examples that indicate that the gesture distance is analogical to temporal distance. Following, the chapter addresses

gesture location by comparing past-related and future-related language, with not significant results. The next section of the chapter offers a number of examples in which speakers use both a proximity and a distance-related expression together with different co-speech gestures. The last section indicates the possible link between certain gesture shapes and temporal distance.

Chapter 8 presents a discussion focused on several aspects related to time conceptualisation and gesture studies from a methodological point of view, as well as new directions for future research in the field. The chapter focuses on a number of methodological changes and improvements that I suggest based on the analysis that was performed in the dataset, including a more precise way of analysing the form of sagittal gestures. This section also covers a number of improvements in the field of gesture studies and machine learning, present several tools that have taken my dataset as a starting point. Following, the chapter presents a number of ideas that could be taken into account for future research on time conceptualisation and co-speech gestures. I firstly focus on the topic of scriptedness, and how the NewsScape Library and the television setting that it offers can be of great benefit for the study of gesture and cognition. Then I move on to discussing a number of topics relating to time conceptualisation that could be investigated using the NewsScape Library, such as the study of different temporal metaphors or cross-linguistic studies. I also address a number of ideas on future directions for gesture research by using the NewsScape library, like the possible existence of multimodal discursive constructions or the reasons that may make a speaker not to gesture. The chapter concludes by showing some examples of some interesting gesture realisations that I encountered during the initial analysis of the dataset. This includes speakers that employ hand gestures to interact with different items that are located in their surroundings, incorporating them into the temporal construct.

The final chapter of this thesis is chapter 9, which finalises this thesis by offering an overview of the contents of the thesis, as well as indicating the contribution of the same, as well as its limitations and directions for future research.

# Chapter 2
# Time in Cognitive Linguistics

### *2.0 Introduction*

The aim of this chapter is to provide an overview of the main theoretical aspects of Cognitive Linguistics which are required to address temporal metaphor and time conceptualisation. Section 2.1 is devoted to briefly introduce the main psychological theories which serve as the foundation of Cognitive Linguistics. Key concepts such as embodiment, framing, construal operations, categorisation or the encyclopaedic nature of meaning are briefly explained. In addition, a brief explanation of Conceptual Metaphor Theory (CMT) is offered, since this theory stands as the basis for metaphorical and conceptual processes, among which the time-space mapping is included.

Section 2.2 addresses time conceptualisation in language in more detail. Firstly, evidence that supports the experiential basis of the time-space metaphor is provided. Following, a brief overview of different Temporal Frames of Reference (T-FoR) which have been so far suggested is presented. Note that to date there is not a clear taxonomy for classifying T-FoR; each taxonomy addresses certain issues but not others. In this thesis concepts that belong to different taxonomies are used. One key aspect of time conceptualisation which is widely addressed in language is the presence or absence of directionality.

Finally, Section 2.3 addresses a further factor that plays an important role in time conceptualisation; namely the frequent use of a physical element as a material anchor for the conceptualisation. There are several physical structures that serve as material anchors in time conceptualisation. However, here I focus on a structure that is frequently employed not only in cultural practice but also in co-speech gestures: the timeline.

### *2.1 Main tenets of Cognitive Linguistics*

Cognitive Linguistics (CL) is a movement in linguistics which applies the main principles of Cognitive Psychology, and more precisely Gestalt psychology (see Smith, 1988) to language, suggesting that cognitive processes underpinning perception also underpin language. This in in contrast to several key assumptions made in generative approaches.

Croft and Cruse (2004) summarise the main tenets of CL in three main principles:

- Language is not an autonomous cognitive faculty.
- Grammar is conceptualisation.
- Linguistic knowledge emerges from language use.

Language is considered a skill which is integrated within our general cognition, dependent on domain-general cognitive processes such as attention, categorisation, perception or memory. Further, language is not regarded as a componential system in itself where there are clear boundaries between the semantic, grammatical, syntactic and pragmatic levels, but rather these different levels all work together in constructing meaning to language. Nevertheless, research in CL has been traditionally separated in two distinct categories which deal with different dimensions of language: Cognitive Semantics and Cognitive Grammar (Evans & Green, 2006).

There are different cognitive mechanisms which are shared between our linguistic system and other cognitive systems which are involved in the construction of meaning. These mechanisms are also the basis of the development of different CL theories, like Frame Semantics of the Conceptual Metaphor Theory (CMT). Here I will provide a brief overview of the fundamental concepts and theories in CL.

*2.1.1 Categorisation*

Categorisation is considered as one of the most basic human cognitive abilities (Croft and Cruse 2004:74) and it has an important function in the organisation of knowledge and conceptual information, as well as the lexical items which are associated with these concepts. Evans & Green (2006:248) define categorisation as "our ability to identify perceived similarities […] between entities and group them together".

There have been several theories of categorisation, but here I will refer to the modern model, which was suggested by Rosh (1973, 1978), and later developed in CL by Lakoff (1987). The modern model is based on the law of similarity in Gestalt psychology and establishes that categorisation involves constant comparison between different members of the category to some ideal. This model proposes, then, that categories are organised radially around a central prototype, which is the default abstract representation of a category created by the conceptualiser (Rosh & Mervis, 1975). Prototypes have been found to be recognised and mentioned faster as well as being the

first concepts learnt by infants. It is thought that their status as prototypes arises due to their typicality and representativeness (Lakoff 1987). There is not a fixed set of necessary conditions to belong to a category; an item may or may not belong to a category depending on how similar/different to the prototype the conceptualiser perceives it to be.

The different members of the category are, then, gradually arranged around to the central prototype whereby the more characteristics they share with the prototype, the more prototypical these examples are, and thus they are more central. Other members can be considered more marginal or peripheral; they are not the best examples of the category, but they share some characteristics with the prototype.

Figure 2.1 shows how the category BIRD might be arranged for a UK English speaker. A ROBIN or a RAVEN could be considered as good examples of the category while an OSTRICH or a PENGUIN would be marginal members. However, categories do not accurately reflect reality, but they rather represent the *perceived* reality of the conceptualiser. Categories and prototypes may vary across (and within) cultures, and two different speakers might categorise the same concept according to different criteria. If an ostrich is a recurrent and common bird for a conceptualiser, it will be more likely to become a more central part of their category for BIRD.

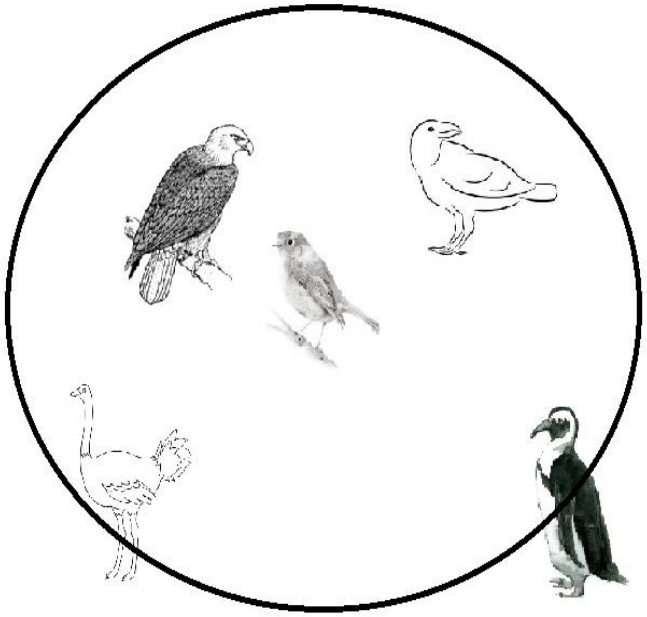


*Figure 2.1 Categorisation of Bird by English speakers*

*2.1.2 Encyclopaedic meaning and frames*

Traditional views of word meaning established a clear distinction between the lexical meaning of a word (denotation) and the extra (usually cultural) information that a particular word evokes (connotation). These two 'types' of meaning were thought to be stored in two completely different systems, making a division between lexical meaning and non-linguistic meaning. Lexical meaning could be decomposed into *necessary and sufficient conditions*; or *semantic features,* (Boy= +MALE, -ADULT) which are binary and complementary (Katz, 1972; Lehrer, 1974). However, CL claims that all the knowledge that we associate with a particular lexical item is part of the meaning of that lexical item. Boundaries between semantics and pragmatics are therefore not clear-cut and a categorical distinction between lexical meaning and world knowledge cannot be made. The term *boy* usually refers to +MALE, -ADULT, but if uttered in the context of a male adult *going out with the boys*, its semantic features change. CL advocates for an encyclopaedic view of meaning, in which everything we know about a concept can potentially be relevant in the meaning-construction process (Barcelona et al., 2012: 16). A clear example of this is the difference in meaning between *bachelor* and *spinster*. Although the only semantic feature that changes between these two words is GENDER, there is wider range of differences that come from other general or 'cultural' knowledge (Fillmore, 1982:131) which suggests that a *bachelor* is a man who lives freely and is (generally) successful while a *spinster* is a woman is too old to any longer be desirable.

Encyclopaedic meaning also takes into account other factors such as the context in which a particular lexical item is used, since it can also have an influence on its meaning. A *candle* in a domestic context would be understood as a source of light, while the same lexical item (and concept) in a religious context would have a different meaning; it would no longer be understood as a source of light but it would rather have a spiritual function. Several models have been offered to account for this phenomenon (*scripts* in Schank and Abelson, 1977; *frame* in Fillmore, 1982; or *domain* in Langacker, 1987; 1991), although they are sometimes used interchangeably (Croft and Cruse, 2004). Here I will be using Fillmore's (1982) notion of *frame,* defined as a 'system of concepts related in such a way that to understand any of them you have to understand the whole structure in which it fits' (p.111).

Frames are not only limited to lexical items, but this notion can also be applied to sentence structure, in which the arguments (semantic requirements of a verb) are accounted for by frames. For example, *talk* would require three arguments: SPEAKER, LISTENER and TOPIC. However, these arguments are not fixed, and the same event can be linguistically coded in different ways depending on the different roles that we profile in the sentence (e.g. *I talked with your friends* vs *I talked with your friends about politics*) through different cognitive mechanisms such as 'scope or attention'. (Evans & Green, 2006; see section 2.2).

*2.1.3 Embodiment and image-schemas*

Another key commitment of CL is the *embodied* nature of meaning, also known as embodiment. Johnson (1987) suggests that language is motivated by our sensory-motor, social and cultural experience. Conceptual structures are created according to our interaction with our environment; our senses allow us to perceive certain characteristics of the world that we store as concepts thanks to the nature of our embodied experience (Evans, 2007: 66-68). Even abstract language can be embodied. For example, in Spanish there is a difference between *ya vi lo que querían decir* (*I saw what they wanted to say)* and *ya me olí lo que querían decir* (*I smelled what they wanted to say*). In the first case, *vi* (*saw*) could be replaced by *understand*¸ while in the second example *olí* (smelled) implies an intuition or assumption. This difference in meaning might be related to the fact that humans (at least in western cultures) rely more on sight for obtaining information than they do by the sense of smell (Barcelona et al., 2012).

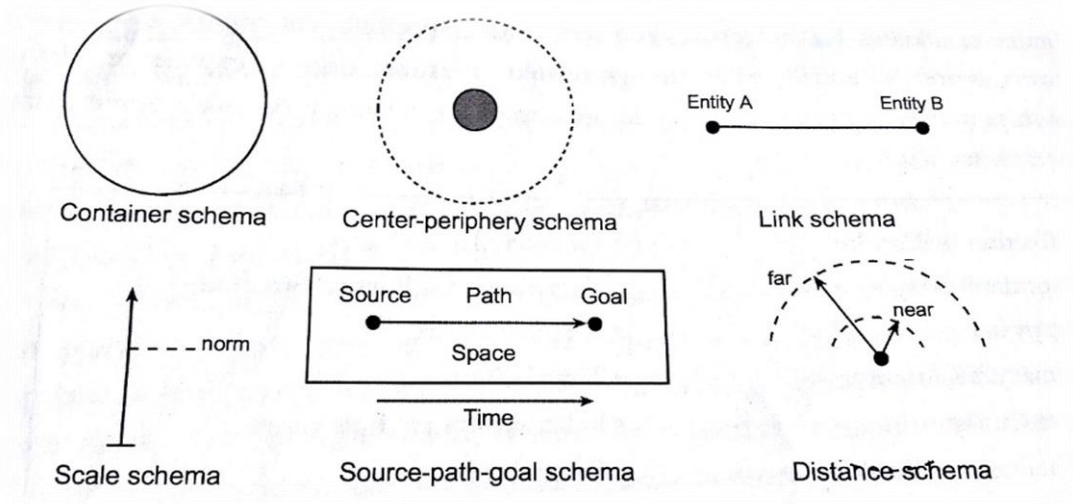



*Figure 2.2 Examples of Image-schemas*

Meaning is then created through our bodily experience and our interaction with the world. Thanks to this, we are able to create basic abstract structures in our minds representing actions which are recurring in the real world. Johnson (1987: 14) refers to this as *image schema* which he defines them as a 'recurring, dynamic pattern of our perceptual interactions and motor programs that gives coherence and structure to our experience'. For example, the recurrent experience of an object moving from one point to another gives rise to the *Source-Path-Goal* (SPG) schema, although there are other recurrent image-schemas including the *Container Schema, Centre-Periphery schema, Link-Schema, Scale-Schema* or *Distance-Schema*. (see Figure 2.2 for models of these schemas).

Image schemas are essential for the conceptualisation of complex concepts, and they are the basis of several cognitive theories both in Cognitive Semantics (see Lakoff & Johnson, 1980; Fauconnier & Turner, 2002) and Cognitive Grammar (Langacker, 1987; Talmy, 2000), although I will only address here the role of image-schema in time conceptualisation.

*2.1.4 Construal and perception*

The conceptual structures and processes described above enable a key function of language: construal realised through 'construal operations'. Construal is defined by Evans (2007: 40-41) as "the way a language user chooses to 'package' and 'present' a conceptual representation as encoded in language, which in turn has consequences for the conceptual representation that the utterance evokes in the mind of the hearer". The same situation can be constructed in different ways depending on the different construal operations that are applied. Alternative ways of encoding a situation reflect alternative conceptualisations of the situation (Lee, 2001: 2). Cognitive mechanisms such as attention, judgment or perspectivisation, which are based on the Gestalt laws of perceptual organisation, are involved in directing construals as they are encoded in the syntactic and semantic structures of the sentence (Croft and Cruse, 2004: 3).

Each of the aforementioned cognitive abilities contribute different dimensions of meaning. Since we are not able to process all of the perceptual information that we are constantly obtaining, we use our attention to select the most relevant information for a particular situation (Barcelona et al., 2012). At the same time, the focus of our attention is surrounded by a scope, which provides us with peripheral elements relevant

to the point of attention (Croft and Cruse, 2004). For example, in the expression *the tip of the finger*, our attention is focused on the area covered by the last phalange of our finger, but other elements like the rest of the finger, the hand or the arm are also relevant to the concept. Even though these elements are not essential, they provide the conceptualiser with useful contextual information about the focus of attention. Attention also allows us to address different entities according to their magnitude (e.g. mass nouns vs countable nouns) or, according to Talmy (2000), add dynamicity to a situation, like in fictive motion (e.g. the *road* goes *up the mountain*).

The second cognitive mechanism is judgment, which is our ability to compare and contrast things. Judgment implies at the same time selection, since we profile a particular element or aspect of a situation at the expenses of others. One of the best examples of judgment is the distinction which is made between Figure (trajector) and Ground (landmark). According to Talmy (2000), the Figure is a moving or conceptually moving entity whose site, direction or orientation is variable, while the Ground is a reference entity which is stationary with respects to the Figure. For example, in the sentence *the laptop is on the desk*, we are contrasting two different elements that have a spatial relation. Here we relate the position of one object, the laptop (figure) to a second element, the desk (ground).

Lastly, perspective also has an effect on language in terms of viewpoint and deixis (Langacker, 2008). The same event can be conceptualised from different viewpoints, which is encoded in language by prepositions or reverse and converse antonyms, among other things. For example, the viewpoint of the conceptualiser changes in:

1.a) I bought a car.

1.b) The assistant sold me a car.

The perspective of the speaker can also be reflected in the coding of spatial and temporal deixis, which involves a coding of place, movement and time relative to the speaker's position. This distinction is easily seen in the following example:

2.a) Today I found my sister right here.

2.b) Yesterday I found my sister over there.

The application of different construal operations allows conceptualisers to adopt different means of locating Figure and Ground. Axial properties such as direction, proximity or sequence between the Figure(s) and the Ground or the position and perspective of the speaker have an influence on how the conceptualiser represents a particular situation. The combination of these contextual or situational elements create a conceptual reference frame, traditionally known as Frame of Reference (FoR). Perhaps the most well-known classification of FoR was proposed by Levinson (1996, 2003), in which he suggested to distinguish among three different spatial FoR: the absolute, the relative and the intrinsic.

*2.1.5 Conceptual Metaphor Theory (CTM)*

One of the peculiarities of language is that it allows us to speak about certain elements by using vocabulary or expressions that belong to different semantic areas. For example, we are able to say *time flies*, even though time is not literally flying, or *I see your point*, even though little can we see with our eyes. We are able to do this because we employ metaphors, that is, we talk about a particular semantic area (target domain) by using properties and characteristics taken from a different semantic area (source domain). Lakoff and Johnson (1980, 1999) developed the Conceptual Metaphor Theory (CMT) in order to address this phenomenon. Metaphor in this theory is defined as 'understanding one conceptual domain in terms of another conceptual domain' (Kövecses, 2010) or 'talk[ing] and, potentially, think[ing], of something in terms of something else' (Semino, 2008:1). A distinction must be made between conceptual metaphor and a metaphorical linguistic expression; the former is the understanding of a conceptual domain in terms of another, while the latter refers to a linguistic realisation of this mapping.

Metaphors often use image-schemas as templates taken from the source domain. For example, in the metaphor TIME IS SPACE image-schemas such as the Source Path Goal schema (*time goes by, time speeds up, time slows down*) or the Container schema (*keep time, save time*) are activated in conceptualising time. The areas of experience that serve as source domains in metaphor are typically ubiquitous and are grounded in physical, biological and socio-cultural experience (Kövecses, 2003).

Moreover, the mapping is selective, that is, some of the aspects of the target domain are highlighted while other aspects are hidden as a consequence of the elements

chosen from the source domain (Kövecses, 2010). Different metaphors are employed to highlight different aspects of the same target domain by using different source domains. For example, we find the metaphors ARGUMENT IS A CONTAINER, ARGUMENT IS A JOURNEY and ARGUMENT IS WAR, which respectively profile the content, the progress and the control of the argument.

One of the assertions that this thesis will be following is that metaphor is conceptual in nature. If this is true, metaphorical conceptualisation should not only be encoded in language, but also be represented in other modes of human communication, such as co-speech gestures.

Lakoff and Johnson (1999), based on Grady (1997), suggest that metaphors can be divided into two different categories depending on which experience they are based on: primary metaphors and secondary (complex) metaphors. Primary metaphors are thought to have a biological root, that is, they are created by a recurrent event which is biologically related, for example the metaphor BODY IS A CONTAINER, which is based in our experiences of substances getting in and outside our body[1]. Primary metaphors are thought to be universal, since they are based on the human experience that we acquire during the early years and they cannot be divided in smaller, more basic metaphors. Languages such as English, Hungarian, Zulu or Chinese have very similar metaphors related to ANGER IS HEAT, since the similar physiological reactions we feel when being angry provide a cognitive motivation to create this metaphor. On the other hand, secondary metaphors are more cultural-oriented and are made up of several primary metaphors. This is illustrated by the metaphor ARGUMENT IS WAR, which is culturally dependent. However, it seems that the distinction between these two types of metaphor would need to be further specified since there is some research in primary metaphors which involve music (see Dolscheid et alia, 2012) or time (see Casasanto & Bottini, 2012; Ouellet et alia, 2010a, 2010b; Boroditsky, 2011; Casasanto & Jasmin, 2012) that suggest that they can also be shaped by culture since they can be 'elaborated' in culturally specific instances as in ANGER IS HEAT INSIDE A COOKING POT.

[1] This phenomenon is directly related with the Embodied Cognition hypothesis and the embodiment of meaning.

## *2.2 Cognitive semantics and time: CMT and conceptual frames of reference.*

### *2.2.1 The experiential basis of time*

Temporal metaphor has been a key topic of research since the inception of CMT, as it exemplifies an abstract universal phenomenon which is metaphorically conceptualised in terms of more concrete domains (Lakoff, 1987; Lakoff & Johnson, 1999). The Master Metaphor List (Lakoff, Espenson & Schwartz, 1991) offers an overview of the different mappings that are recurring in time conceptualisation, grouping them into four main categories: TIME IS SPACE IN MOTION, TIME IS A RESOURCE (MONEY), TIME IS A CONTAINER and TIME IS A CHANGER (HEALER). Although these metaphors provide us with different conceptualisations of time, literature has traditionally focused on the different linguistic realisations of the most prolific one: TIME IS SPACE IN MOTION, which is based on image-schemas such as the Source Path Goal. Nevertheless, further research on the other temporal metaphorical mappings ought to be conducted, since it will undoubtedly provide us with essential information about the different source domains in time conceptualisation [2].

TIME IS SPACE IN MOTION constitutes a metaphor which is thought to be experientially motivated since it is based on a recurrent space-time relation when interacting with the world. Moore (2014) states that concepts such as spatial expectation, arrival and deixis are strongly linked with time. These spatial aspects are correlated to time-related experiences, such as the perception of approach and expectation of collision (Gibson, 1966:143) or the association of an approaching entity with the expected future. The parallels between the spatio-temporal domains and their co-activation provide a motivation for the grounding of the mapping.

Lakoff (1999) states that TIME IS SPACE is a primary, universal metaphor due to its experiential basis, although research suggests that time conceptualisation might vary cross-culturally and cross-linguistically (Ouellet, Santiago, Israeli, & Gabay, 2010; Casasanto & Botini, 2010). However, here I will be focusing exclusively on time conceptualisation in English speakers, although cross-cultural and cross-linguistic research is crucial to have a better understanding of conceptualisation processes in human cognition.

[2] Pagán Cánovas and Teuscher (2012) offer an interesting application of conceptual integration into Michael Ende's *Momo* and the materialisation of time.

There are certain parallels between space and time, with some correlations in how we experience both. Space and time often co-occur in our experience. Thanks to the space-time experiential proximity, it is possible to use spatial terms to talk about time. Galton (2011) suggests that, even though they are experientially close, time and space differ in three main characteristics: magnitude, dimensionality and directedness. Space is made of matter, it is located in a three-dimension space and it is isotropic, that is, it is uniform in all its directions. However, time cannot be measured per se, it is one-dimensional and it has an asymmetric organisation (theoretically, time only can flow in one direction).

But how can we talk about time in terms of space if they are so different? Spatial properties are assigned to the different components of a time construal by the creation of a conceptual frame of reference (Tenbrick, 2011) in which different roles and relations are attributed to the different components of the conceptualisation. Thanks to the experiential similarities between space and time, we are able to employ a Temporal Frame of Reference (T-FoR) when conceptualising time in terms of space. A T-FoR would be then the metaphorical counterpart of a Spatial-FoR (S-FoR) by which the relation between the Figure and the Ground is expressed by applying a three-dimensional space (Huumo, 2015), allowing us to ascribe magnitude (*Five hours*), dimensionality (hand gestures) and directedness (*The years ahead of us*) to time.

*2.2.2 A (brief) overview of T-FoR*

Little agreement has been achieved in the literature for establishing a clear taxonomy of T-FoR. Here I will attempt to present an overview of the different T-FoR that have been proposed by different authors, based mainly on linguistic evidence. Most of the literature agrees on a clear difference between the A (Deictic) and B (Sequential) series of time, but it has been recently suggested that other T-FoR are also encoded in language. For the sake of clarity, I will adopt a mixed terminology based on Evans (2013) and Núñez and Cooperrider (2013), dividing T-FoR into four different categories: A-Series (Deictic Time), B-Series (Sequential Time), Extrinsic Time (Time as Such) and Temporal Span, although other taxonomies are suggested based on a direct mapping from S-FoR, including the absolute, relative and intrinsic S-FoR (see Levinson, 1996, 2003; Bender et al, 2010; Tenbrick, 2011).

*2.2.2.1 A-Series or Deictic time*

Clark (1973) was the first scholar to make a distinction between the A-Series and the B-Series T-FoR. A-Series, also known as 'tensed' or 'Deictic Time' (here D-Time, following Núñez & Cooperrider, 2013), include temporal metaphors in which the speaker (from now on Ego) becomes the deictic center, that is, the temporal expression is conceptuliased according to when and where the conceptualiser is. This T-FoR consists of a path, an Ego and temporal element(s) which are located according to the Ego's perspective. However, the construal might change depending on the perspective that the speaker takes when conceptualising the metaphor. Consider the following examples:

1) **We are approaching winter** here, our nation's greatests are without a roof over their head. (FOX-News, Fox and Friends, 03-11-2015)
2) [They] are willing to take their chances in boats, even if **the approaching winter** makes it more dangerous. (AlJazeera, Inside Story, 11-03-2015).

In 1 the Ego's perspective is emphasised since it is the subject of the sentece, and it moves through a path (time) until it reaches a point (temporal event). All the points located behind the Ego are temporally located in the past, while everything in front of the Ego is located in the future. However, 2 presents a different perspective within the same metaphor. This time the perspective is from the temporal event, which acquires motion and moves through a path until it reaches the Ego. In this case, all the elements that are in front of the temporal event are the past, while everything behind the temporal event is considered the future. In both scenarios movement seems to take place on the saggital axis, suggesting that time may be conceptualised exclusively on this axis. However, gesture data might point in a different direction, since English speakers tend to use a variety of axes when gesturing about time (See chapter 4, section 4.3).

These two realisations of the D-Time T-FoR are known as the Moving Ego (Table 2.1) and Moving Time[3] (Table 2.2), respectively (Clark 1973; Boroditsky, 2000; Boroditsky & Ramscar 2002; Núñez & Sweetser 2006). Note that the deictic center and

[3] Also known as Target-Event perspective point and Reference-Point perspective point (Evans, 2013) or Ego-Centred Moving Ego and Ego-Centred Moving Time (Moore, 2006; 2011; 2014).

the conceptualiser do not necessarily coincide when either the Moving Time or the Moving Ego metaphor is employed. Consider the following:

3) **Lawmakers in Washington are facing a fast approaching deadline** to avert a Government shut-down (KNBC, Today Show, 08-03-2011).
4) We will see where it goes because obviously you've got **your deadline looming**. (CNN, New Day, 20-11-2015).

In this case, the deictic centre and the conceptualiser are not the same; the conceptualiser uses an external Ego when constructing the mapping. The difference with 1 and 2 resides in the viewpoint of the conceptualiser. Example 3 shows a Moving Ego metaphor in which the conceptualiser is off-stage, and the deictic centre is a different Ego, while in 4 we find a similar case in the Moving Time metaphor. Núñez and Cooperrider (2013) refer to this as "intrinsic" and "extrinsic" points of view in the D-Time T-FoR.

| **Source** | | **Target** |
|---|---|---|
| Moving Ego | => | Now |
| Location in front of the Ego | => | Event in the future |
| Location behind the Ego | => | Event in the Past |

*Table 2.1 Deictic Time (D-Time). Moving Ego*

| **Source** | | **Target** |
|---|---|---|
| Ego | => | Now |
| Object approaching Ego | => | Future temporal event |
| Object moving away from Ego | => | Past temporal event |

*Table 2.2 Deictic Time (D-Time). Moving Time*

Overall, the literature agrees on the importance of the Ego in time conceptualisation, although recently some authors have suggested that the emphasis should not be on the element which is moving, but rather on the reference entity (Bender et al, 2010; Sinha et al. 2014).

*2.2.2.2 B-Series or Sequential Time*

The second type of T-FoR suggested by Clark (1973) is known as "B-Series" or "Tenseless time", which is time-based. In this type, the temporal construal is formed by

a path and a series of events which are located along the timeline; the Ego is not placed in the construal. Temporal events are located in a timeline relative to themselves rather than deictically. Moore states that this FoR "establishes the relation between F (Figure) and G (Ground) in terms of an orienting principle that applies to the F and G equally but does not depend on the ego's perspective" (2014: 65). Moore refers to this B-Series of time as "Field-based FoR", although Núñez & Cooperrider refer to it as Sequential Time (S-Time). Consider the following examples:

5) Literally **the calm before the storm** which should hit here in the next few hours and last through most of the week. (KNBC, KNBC 4 News at 9pm, 04-01-2016).
6) Ukraine expresses optimism in **the peace process after multilateral talks** in Paris.(KCET Deutsche Welle Journal, 02-10-2015).

Examples 5 and 6 exclusively refer to the relation between two events, and thus there is no need to include the Ego in the construal. Moore (2014:67) points out that the source frame of this mapping is 'two (or more) entities which are moving in the same direction such that one is ahead of the other' (Table 2.3). In the case of sequential time, Evans (2013) suggests that a prospective or a retrospective perspective can be applied to the S-Time construal, resulting in two different mappings[4]. Example 5 would be an instance of a prospective perspective point of view; emphasis is added on the *early* event (*calm*) since it appears as the subject of the sentence and *before* establishes an EARLY IN SEQUENCE mapping[5] between F and G. On the other hand, 6 adds emphasis on the *later* event, *peace*, and the LATER IN SEQUENCE relation is established by the used of the preposition *after*. In these two cases, we are dealing with a similar conceptual mapping, whose only difference relies on the perspective which is taken in the conceptualisation of the sequence of the temporal events.

[4] Evans (2014) also mentions a 'multiple sanction S-Time' which refers to the inclusion of a deictic element in a S-Time construal, i.e. *Christmas **arrives** before New year*.
[5] Evans (2014:119-121) argues that the use of *before* as a synonym of *in front of* (i.e. *Kneel before your king*) is no longer productive in present-day English, and it only remains as part of some fossilised idiomatic expressions

| Source | | Target |
|---|---|---|
| Entities on a path | => | Temporal events |
| Entity ahead of another entity | => | An event is earlier than another |
| Entity behind another entity | => | An event is later than another |

*Table 2.3 Sequential Time ( S-Time)*

Lastly, Núñez & Cooperrider (2013:223) suggest that S-Time is often conceptualised from an external perspective, that is, the conceptualiser is an external observer of the temporal sequence as if they were watching a train moving from the distance. However, it is also pointed out that some experimental evidence suggests that an internal S-Time viewpoint can be enacted by acoustic stimuli and non-verbal responses, resulting in the speakers placing sequential events along the sagittal axis (Walker et al., 2013).

*2.2.2.3 Temporal Span*

Even though most of the literature agrees on the distinction between the A and B series, Núñez & Cooperrider (2013) purposed the inclusion of a third type of T-FoR, which they named Time Span (T-Span). T-Span concepts are those which refer to a "perceivable or measurable temporal magnitude, such as 'five minutes', 'the whole morning' or 'several months' "(ibid). Through this construal, temporal events acquire characteristics of spatial entities, enabling us to measure their length or quantity (Table 2.4). Within T-Span, it is important to make a distinction between those metaphors that address the duration of the temporal event and those that employ measures for temporal magnitude, even though both can be constructed in terms of spatial magnitude. Consider the following examples:

7) …the two of us thinking this is going to be a **very long film.** (KCBS, Late Show with Dave Letterman, 22-05-2012).
8) But you only can have one, and **through the month of February**, you can only have one election at a time (FOX-News, The Kelly File, 16-01-2016).

Example 7 shows a temporal event, FILM, which is described in terms of spatial magnitude to indicate the extension of the temporal event. The construal appeals to

space to establish the temporal duration of an event. Note that English usually employs terminology related to spatial extension, like *long* or *short*, while other languages like Greek more often use terms like *big* or *small* when talking about time. On the other hand, example 8 offers a slightly different case. This time, the linguistic expression does not refer to a temporal event, but rather a socially constructed period that has been attributed spatial properties. These measures of time are culturally dependent and they rely on 'symbolic cognitive artefacts' which can be based on phenomenological experiences (lunar/solar cycle) but they are not determined by them. The existence of these artefacts makes possible some conceptual structures, like the wide-spread Gregorian Calendar, although different cultures may adopt different cognitive artefacts (Sinha, 2014; see also section 2.4).

| **Source** | | **Target** |
|---|---|---|
| Entity in space | => | Temporal event |
| Spatial magnitude of the entity | => | Temporal magnitude |

*Table 2.4. Time Span (T-Span)*

The main difference between T-Span and S-time T-FoR is that T-Span addresses a particular temporal event or socially constructed time span which is construed in spatial magnitude terms, while S-time establishes a spatialised relation between two (or more) temporal events or socially constructed time spans. However, it is not clear that T-Span temporal expressions construct a T-FoR by themselves; it has been pointed out by different authors (Evans, 2004, 2013; Sinha et al., 2014) that this category might be part of a larger and more general T-FoR: Extrinsic Time.

*2.2.2.4 Extrinsic Time or Time as Such*

The last category that has been suggested for classifying T-FoR was proposed by Evans (2013) and it addresses several issues which had not been previously covered. Evans draws our attention to one feature which is inherent in time but it is absent in space; the notion of transience. This term refers to the unique property of time of not being able to be repeated, that is, an experience or event takes place at a particular time in a particular place but the conditions so that the same event occurs at the same time and at the same place cannot be fulfilled (Galton, 2011). Another characteristic in which time differs from space is that it is (apparently) something eternal, as if it were an

unending event in which everything takes place with no beginning or end. In the words of Evans (2013) this notion of time can be defined as a "time as matrix".

These unique properties of time suggest that there might be another type of temporal relation in which time constitutes the event itself in which other events take place: Extrinsic Time (E-Time). This matrix relation is not grounded in phenomenologically, real experience since humans are not able to experience it, but it is based on the concept of time being something eternal. Consider the following examples:

9) Glad to have you. They say **time flies** when you are having fun. Does it seem like 25 years? (KOCE, Tavis Smiley, 22-12-2015).
10) That's after your local news. Good morning. T**he time right now is approaching 8:56 exactly**. (KNBC, Today Show, 21-01-2011).

In these two cases, time is conceptualised as an eternal matrix which moves towards a particular place. Example 9 makes a reference to the speed by which this temporal matrix advances during a particular event (which also raises the question of time subjectivity, which will be later addressed), while in example 10, time is moving towards a particular 'spatial' point in the time matrix (Table 2.5).

| **Source** | | **Target** |
|---|---|---|
| Entity on space | => | Temporal matrix |
| Entity approaching an object | => | Temporal matrix approaching an event |
| Motion of the entity | => | (Subjective) Speed of the passage of time |

*Table 2.5. Extrinsic Time (E-Time)*

E-Time fixes the event in an absolute way, with no reference to an observer. The reference strategy is based on periodicity, rather than being egocentric or event-based as happens with D-Time and S-Time. According to Evans (2013), T-Span temporal expressions are included within the E-Time T-FoR, since they refer to a particular 'chunk' of the temporal matrix through a symbolic cultural artefact, such as calendars or clocks.

However, it seems E-time is not a universal T-FoR. Sinha et al. (2014) point out that Nuer, a Nilo-Saharan language, does not have a word to refer to "Time as Such", the temporal matrix; temporal reference points in Nuer are constituted by social activities and the relation between social and environmental events.

*2.2.2.5 Alternative taxonomies for T-FoR*

Even though the aforementioned classification of T-FoR are the most widespread, some alternative systems which emphasise the similarities between the spatial and temporal domains have emerged. These systems have attempted to mirror the classical spatial frames of reference (S-FoR) system onto time offering a different approach in the understanding of temporal metaphor.

Bender et alia (2005) suggest a system for T-FoR based on a linguistic analysis of English, German and Tongan which mirrors S-FoR. S-FoR are a set of coordinates employed to construct oriented space within which spatial relations among objects are identified. Here it can be distinguished among three different types: *absolute, intrinsic* and *relative*. An absolute FoR employs a fixed point of reference which is chosen by the conceptualiser. These points are socially agreed upon and culturally accepted, being one of the best examples the cardinal points, although other reference points such as an important mountain, a hill or the sea can also be employed (Bender et al, 2005; Bennardo, 2002). Secondly, an intrinsic FoR is centred on an object and it remains so when the speaker or the object move. The oriented field is independent from the field of the speaker; it depends on the object (i.e. The car is in front of the house). Lastly, a relative FoR is centred on the speaker and it remains so when the speaker moves as, for example, when we say *The car is in front of me* (ibid).

These three different types of S-FoR are then mapped onto time, mirroring this classification for time conceptualisation. Thus, we would obtain an absolute temporal framework in which objects are localised independently from the observer's perspective, like events in a calendar; an intrinsic framework in which temporal relations between events are created according to the observer; and finally a relative framework, which is further developed in three sub-types: basic relative (similar to the aforementioned Moving Time), translation, in which the observer's position is transferred into the point of reference and the perspective is translated, and finally reflection, in which the observer's position is reflected on the reference.

On the other hand, Tenbrink (2011) offers a slightly different perspective on the mapping of S-FoR onto time. She establishes that Clark's B series correspond to an absolute T-FoR, since they do not require the Ego to be present in the conceptualisation, while she makes a further division into different subtypes within the A series, including intrinsic and relative reference frames and several subdivisions of the same (dynamic or static). Nevertheless, neither of these taxonomies offer a description that accounts for E-Time, which includes features of time that cannot be found in space.

Overall, the analysis of the different classification systems for T-FoR leaves us with a wide range of taxonomies, terminology and distinctions that, in many cases, result in overlaps within the existing literature. The following table aims to summarise the T-FoR classification systems that so far have been offered, as well as to indicate the similarities and differences among the different taxonomies.

| AUTHOR | A-SERIES, Tensed time, (Ego-Based) | | | | | | | | B-SERIES, Tenseless time, (Time-Based) | | | E-TIME | T-SPAN |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Clark (1973)** | Moving Ego | | | | Moving Time | | | | Sequence time | | | – | – |
| **Lakoff (1993)** | Time passing is motion over a landscape | | | | Time passing is motion of an object | | | | – | | | – | – |
| **Boroditsky (2000), Boroditsky & Ramscar (2002)** | Ego-Moving | | | | Time-Moving | | | | – | | | – | – |
| **Núñez & Sweetser (2006)** | Moving Ego | | | | Moving Time | | | | Time-Referential point | | | – | – |
| **Bender et alia (2010)** | Ego-Reference Point | | | | | | | | Time-Reference point | | | – | – |
| | Relative Reflection | | | | Relative Translation | | | | Moving Ego | Moving Time | | | |
| **Tenbrink (2011)** | Non-projective dynamic | | | Projective static | Projective dynamic | | | | Absolute temporal FoR | | | – | – |
| | Moving Ego | Moving Time | | | Moving Time | Moving Ego | | | | | | | |
| **Casasanto & Jasmin (2012)** | Moving Attention Perspective | | | | | | | | Moving Attention Perspective | | | – | – |
| **Núñez & Cooperrider (2013)** | Internal D-time | | | | External D-time | | | | Internal Sequential-Time | External Sequential-Time | | – | Temporal Span |
| **Evans (2013)** | Target-event perspective point | | | | Reference-point perspective point | | | | Prospective | Retrospective | Multiple sanction | Extrinsic Time | |
| | Future cluster | Present cluster | Past cluster | Degree of remove cluster | Future cluster | Present cluster | Past cluster | Degree of remove cluster | | | | Event-reckoning | Time-reckoning |
| **Sinha et al (2014)** | Ego-relative temporal motion constructions | | | | | | | | Positional time constructions | | | Time as Such | |
| | | | | | | | | | | | | Event-based | Time-based |
| **Moore (2006, 2011, 2014)** | Ego-Centred Moving Ego | | | | Ego-Centred Moving Time | | | | Sequence is relative position on a path | | | – | – |

*Table 2.6. Summary of the different T-FoR suggested by the literature*

Even though there is agreement in the difference between the A and B series, there is no conformity in how to analyse these two categories, since each author highlights some features while disregarding others. Moreover, E-Time and T-Span have emerged as novel frames for addressing time conceptualisation, although further research on this matter is required if a clearer taxonomy is to be obtained. The inclusion of E-Time raises yet another question: the tight relation between space and time cannot be denied, but it seems there is some limit in the space-time mapping; spatial metaphor does not account for temporal transience. This temporal aspect seems to be inherent to time (Galton, 2011).

Another aspect that remains to be accounted by T-FoR (briefly mentioned in Evans, 2004, 2011) is temporal subjectivity. Consider the following:

11) A tennis match at Wimbledon for the ages because **it lasted ages**. (KTTV-FOX Ten OClock News, 25-06-2010).
12) Just to remind her **how fast 20 years go**. I mean, **blink of an eye**. (KCBS, CBS 2 News at 5, 30-11-2015).

Examples 11 and 12 refer to a temporal event which lasts for a set amount of time. In the first case the tennis match lasted for a long period of time, which is described as *ages* by the speaker, while in 12 20 years are said to last the same amount of time as the *blink of an eye*. Even though each temporal event extends for during a measurable amount of time, each speaker has a subjective perception of how the passage of time is. Evans (2004, 2013) points out that these are cases of *protracted duration* (11) and *temporal compression* (12), that is, the subjectively felt experience of temporal passage. Thus, in example 11 the conceptualiser perceives that the match lasted longer than it actually did, while in example 12 the speakers perceives a long period of time as a short-time event. Different speakers might have experienced the same events in different ways, since the subjective perception of time depends on a wide range of cognitive factors such as mood or emotions like joy or boredom.

### *2.2.3 Time directionality in language*

Extensive literature has been written about how time can be expressed in English language, and it has been pointed out that nearly every aspect of time can be expressed by means of spatial metaphors (i.e. Clark, 1973; Traugott, 1984; Lakoff and Johnson, 1980; Alverson, 1994; Evans, 2004, 2013; Radden, 2004; Moore, 2006, 2014), although

there are occasions in which exclusively temporal language can be employed. Another possibility is to employ different metaphorical realisations of temporal concepts that use other domains of experience, like the TIME IS A RESOURCE metaphor. However, I will just be taking into account two type of linguistic temporal expressions: space-time metaphors (long week, winter is coming, look back in the past) and non-spatial temporal language (the meeting lasted for ages, earlier, later…).

As it has been previously stated, there are a number of features which time and space do not share. Here I will be addressing the relation between space and time concerning their dimensionality and directedness. Galton (2011) states that one of the most distinctive features of time is that it is one-dimensional and anisotropic, while space is arranged in three dimensions and is isotropic. Time is supposed to move across a one-dimensional space and (theoretically) it only has one direction; time is constantly passing and it is not possible to reverse the flow of time. Conversely, space is arranged in three dimensions in which entities are able to move in any direction. One of the consequences of the space-time metaphorical mapping is that these two properties of space are attributed to time, allowing us to go *back* and *forth* in time or to talk about the time *ahead of* us.

Within space-time linguistic metaphors, there is a set of expressions which ascribe time with a certain axiality, while other space-time linguistic realisations do not specify the axis on which time is located. Consider the following examples:

13) **Back in those days**, if you were rich you had 20 children (KOCE, Chalier Rose, 11-03-2013, NewsScape Library).
14) Iran will have now a nuclear weapon and that means **war is upon us** (FOX-News, Shephard Smith reporting, 14-07-2015).
15) Because what they are talking about here is not **the distant past** (MSNBC, The Rachel Maddow Show, 17-04-2014).

In the first two cases, the preposition clearly expresses the directionality of time in relation to the speaker, although the chosen axis is different. Example 13 represents the most common temporal metaphor which is employed in the English language: time is located on the sagittal axis relative to the Ego. The events that took place in the past are located behind the speaker, while future events are situated in front of the speaker. However, English speakers can also conceptualise time vertically, although this type of

metaphorical realisation is much more uncommon. In case 14 time is located along the vertical axis (*upon*) with future and upcoming events above while past events would be located below. These types of linguistic realisations are rare in English language, although they have been reported to be common in other languages such as Chinese (Scott, 1989; Chun, 1997; Boroditsky 2001) as well as other non-linguistic realisations of temporal metaphor (cultural artefacts such as calendars or co-speech gestures). However, this vertical metaphorical realisation is fundamentally different from other type of linguistic temporal metaphors. Firstly, we cannot find an up-down dichotomy as we find in the front-back axis (e.g. the war is below us?=the war is behind us). Secondly, even though further research on the use of the proposition *upon* would be required, the used of the vertical axis could be linked to other non-temporal meanings, such as the negative valence, immediacy or high agency of the event. Lastly, there are no linguistic temporal metaphors that locate time laterally in English language, even though it has been documented that English speakers tend to spatially arrange time laterally, perhaps influenced by their left-to-right writing direction (see Casasanto & Bottini, 2010; Casasanto & Jasmin, 2012; see also chapter 4, section 4.4).

On the other hand, example 15 does not clearly state where the past is axially located, but it indicates how far that temporal point is from the Ego. Both future and past events can be near or far from the Ego in order to indicate how close that temporal event is construed as being. This type of temporal metaphor is also present in other multimodal realisations, for example timelines, in which (occasionally) the distance between two points in the timeline corresponds to the temporal distance between events (Coulson & Cánovas, 2014).

The recurrent front-back linguistic spatialisation of time in English might lead us to assume that time is conceptualised exclusively along the sagittal axis, but incongruences emerge when this data is contrasted with other multimodal data. If metaphorical thinking is part of cognitive conceptualisation processes, we cannot only rely on linguistic evidence, but we need to account for other non-linguistic realisations. Multimodal data, and more precisely gesture analysis, emerges as a crucial element for understanding the embodied conceptual structures of our thought which might not be that obvious in spoken language.

*2.2.4 Tense and aspect: cognitive linguistics perspective*

There are two more elements of language that have a crucial role in structuring in time conceptualisation: tense and aspect. Extensive research has been done on these two aspects of language from a wide range of perspectives but for the sake of concision here I will give just a brief introduction to these two concepts. Hewson (2012) provides a general overview of tense and aspect and, based on Guillaume (1929), suggests that aspect refers to the time which is contained in an event while tense refers to the representation of the time that contains the event. Radden & Dirven (2007) provide a more precise definition by suggesting that tense relates to how the speaker temporally locates a situation from their point of view, while aspect refers to the point of view that the speaker takes when conceptualising a temporal event.

*2.2.4.1 Aspect*

Radden & Dirven (2007:175-197) develop this notion of aspect by pointing out that speakers can conceptualise a situation through two different vieweing frames. Conceptualisers can apply a maximal viewing frame when an event is conceptualised entirely, that is, the temporal event is conceptualised as a bounded event with a clear starting and end point. On the other hand, speakers can also conceptualise a situation by applying a restricted view, through wich events are seen internally and the focus relies on their progression rather than on the starting or the end of the event, presenting an unbounded event with an implicit beginning and end.

When speakers use the non-progressive (perfecive) aspect, they apply a maximal viewing frame and reference is made to its beginning and end-point. These constructions are expressed with simple verb forms in English. On the other hand, the progressive aspect, or imperfect aspect, applies a restricted view and is expressed in English by V-ing constructions (Taylor, 2002; Radden & Dirven, 2007). The crucial difference between perfective and imperfective aspect is that in a perfective process its temporal boundaries (most of the time the end) have to be part of the profiled process; the termination of the process has to be part of the conceptualisation (Taylor, 2002).

Maximal and restricted views can be applied to both events and states, where events are dynamic situations that involve change and states are static concepts that do not involve change. See the following examples:

16) The journalist *interviewed* the president.

17) The journalist *was interviewing* the president.

These two examples show how an event can be conceptualised differently by applying different vieweing frames. Example 16 applies a maximal viewing frame. In this case, the speaker adopts an external view of the event by using the perfective form of the verb. By doing this, the speaker focuses on the event as a whole, highlighting its temporal boundaries. On the other hand, example 17 shows a restricted frame by using the V-ing construction. In this case, the temporal boundaries are out of focus (although they are implicitly there) and the focus is on its progression.

Similarly, these two frames can be applied to states. Considering the following:

18) I *live* in Spain.
19) I *am living* in Spain.

In this case, the concepts we refer to are static; *living* doest not imply change or dynamicity. Example 18 applies a maximal viewing frame to present an infinite view of this enduring state and, most importantly, for an indefinite amount of time. On the other hand, 19 shows a restricted viewing frame, in which the speaker denotes a portion of the ongoing state, making us think that the state has some implicit boundaries, e.g. a time when I began living in Spain, and turning it into a temporal state (Radden & Dirven, 2007).

Another essential notion to take into account when discussing aspect is the boundedness of the events, that is, whether the focus is on the boundaries of the event (bounded) or on the progression (unbounded). Boundedness can be explained in relation to two different criteria: duration and telicity. Firstly, duration relates to the amount of time that an event lasts for. Here we can distinguish between durational events, that are developed for a time span (*I walked home*), or punctual events, that take place in a precise instant (*The gun was shot)*. Telicity refers to how conclusive events are. Telic events have a clear end point, while atelic events do not. An atelic event cannot be finished since it cannot be completed, but it can be stopped. By mixing these two criteria, bounded events can be assorted in for different types: durational telic (*I walked home*), durational atelic (*I cuddle the baby*), punctual telic (*I set the table)* and punctual atelic (*The man fainted*) (Talmy, 2000a; 2000b).

Concerning unbounded events, we find their counterparts on durational telic (*I was walking home*) and durational atelic unbounded events *(I was cuddling the baby*), while punctual telic and punctual telic present a different picture. Since the focus on unbounded events is on the process and punctual events lack a process, the focus here relies on a different aspect of the conceptualisation. In the case of unbounded punctual telic events, the focus relies on the build-up phase that precedes the culmination of the event (*The man is dying*), while in the unbounded punctual atelic case the focus is on a quick succession of punctual acts (I am *sneezing*) ((Radden & Dirven, 2007).

*2.2.4.2 Tense*

Tense relates to how the speakers can temporally locate a situation from their point of view. The instant in which we speak serves as a deictic center (speech time) to anchor speech in a specific spatial and temporal location. This allows the speaker to refer to three different time spheres: present, which refers to events that are happening during the speech time, past, when the events happened before the moment of speech, and future, when the event may take place after the speech is produced. These temporal notions are referred to as deictic times, and although they can be encoded lexically, they are also expressed grammatically through tenses in English, particularly through the use of simple tenses (Radden & Dirven, 2007). Temporal notions can also be expressed with other systems, like an absolute system or a relative system (as seen in sections 2.3.1, 2.3.2 & 3).

The main function of grammatical tense is to ground a situation or action in time, but it also has an influence on the representation of the event as a mental space in our mind. When employing tense, different mental spaces interact in order to temporally conceptualize the event. The main mental space that serves as the basis for any temporal interaction is the base space, which includes all the spatial, temporal and contextual information that the speaker can access during the speech time. However, when speaking about an event, we focus our attention to the temporal period in which the event took place, that is, the event time. Additionally, there are occasions in which the focus might rely on the viewpoint that the speaker takes when talking about the situation. Even though this viewpoint often takes place in the present time (overlapping with the speech time), it can shift to past or future times. The process of viewpoint shifting is known as *reference time*. Thus, when conceptualising a temporal event, three different mental spaces interact (Radden & Dirven, 2007). A base space containing all

the information found during the speech time, a mental space containing all the information connected to the event (event time) and finally another mental space linked to the viewpoint that the speaker takes when talking about the event (reference time).

### *2.3 Material anchor, cultural artefacts and mental timeline*

Time is often associated to a more stable mental structure, space, which makes it easier for us to conceptually manipulate. However, this relation does not only take place at the conceptual level, but it is often manifested on different physical structures, like timelines. Conceptual structures often need to be represented in such a way that they remain stable whilst being manipulated. The stability of these conceptual structures can be increased by associating them to culturally coherent mental schemas (D'Andrade, 1989) or to material structures. Hutchins (2005) states that Conceptual Integration Theory (CIT) offers a framework which can be employed to show how material structures can provide an input space in the creation of blended concepts when combined with other input spaces that increases the stability of the blended structure. Traditionally, CIT points out that the input spaces which are employed when creating a blended space are mental constructs. However, Hutchins suggests that one or more of these input spaces can have a physical form, resulting in what he calls a material anchor.

Hutchins (ibid.) uses the concept of *queue* in order to explain how a material anchor can be employed in a conceptual blend. A queue consists of a number of participants who are arranged in a line and use their body location to establish an order relation. However, the line-like structure is not enough for creating a queue, as we can find other line-like structures which are not queues. It is through conceptual blending that we can attribute an imagined directional trajector to the physical form of the line, providing it with directionality and turning it into a queue. Thus, the concept of queue emerges from the conceptual blend between a physical structure (line) with a mental structure (directional trajectory).

Nevertheless, this might raise some questions concerning the relationship between mental and material structures. Talmy (2001) would say that the relation that is established is not between a mental and a physical structure, but between a mental structure and our conception of the physical structure and its meaning. If we follow this approach, the input spaces in the blending process would be a mental representation of the physical structure (line) and a mental representation of a conceptual element

(trajector). Hutchins (2005) states that a mental representation of the physical space is not required, since this would obscure some important elements of the physical structure. Moreover, in the example of the queue, there is no need to mentally represent a line, because it is already present in the real world. When we count the participants that belong to the queue, we do not use a mental representation, but we refer directly to the physical structure. A physical structure is not a material anchor because it is used in a particular way, but it becomes a material anchor when certain conceptual elements are mapped onto the material object and the material object becomes a representation of the blended concept.

Finally, it is crucial to highlight that humans are able to create imaginary material anchors. Once a material structure becomes extremely familiar, it is possible to imagine it when it is not present in the environment. This allows the conceptualiser to manipulate an imagined material anchor in the physical world. Since we are familiar with the concept of a queue, we can create an imagined queue (for example through hand gestures) so that it is easier for us to understand and manipulate the concept even if it is not present in the physical world.

Concerning time, there are multiple examples of material anchors which are employed for time conceptualisation. Most of them rely on cultural artefacts that have remained for a long period of time. In the case of English speakers, items like clocks, watches, sundials, calendars or timelines are examples of cultural material anchors which are employed for conceptualising different aspects of time. Fauconnier and Turner (2002) offer a detailed explanation of how we create a blended space which addresses a cyclic conception of time, the Cyclic Day Network, by performing a space-time mapping on circular-shaped objects, like watches and sundials.

Recent research shows the importance of cultural artifacts and time conceptualisation. Duffy (2014) performed three different experiments to investigate the role of cultural artifacts and the interpretation of metaphorical expressions about time. In the experiment, participants had to answer the ambiguous time-related question of "Next Wednesday's meeting has been postponed two days. When is the meeting?", which has two possible answers (Friday or Monday) depending on the temporal metaphor that the conceptualiser uses (more on this in chapter 3, section 3.2). Overall, results suggest that participants tend to have a higher amount of *Friday* responses,

probably due to the left-to-right orientation of the calendar, which is also the writing direction in most western languages. In the second experiment, participants performed the same task by using an Arabic calendar. In this case, even though *Friday* was the most frequent response, the right-to-left orientation of the calendar caused some interference, with participants in this condition more likely to say Monday.

There is a wide range of cultural artifacts that play an important role in time conceptualisation., but one of them stands out because of its importance and its salience: timelines. Timelines are culturally constructed artefacts which aim to visually represent a sequence of related events. Even though it seems to be a very simple way of representing time, the first timeline dates from 1753, making it a relatively modern way of mapping time using a physical structure, and it has become an essential component for the reproduction of knowledge and as a tool for thinking (Davis, 2012). The emergence of these cultural artefacts involves a gradual process in which a large number of individuals interact with it. The timeline as a material anchor for time, as it happens with numbers, has developed diachronically so that modern observers perceive it as something completely natural.

Timelines follow two key components in the TIMES IS SPACE metaphor: PROXIMITY IN TIME IS PROXIMITY IN SPACE and TEMPORAL DURATION IS SPATIAL EXTENT, but, according to Coulson & Cánovas (2013) they can also convey a series of meanings that cannot be explained by following the unidirectional mappings proposed by CMT. Thus, they offer CIT as an alternative to explain some of these meanings. For example, figure 2.3 shows a timeline that was created to celebrate the


*Figure 2.3 Timeline celebrating the 50th anniversary of Lego*

50th anniversary of Lego. Generally, timelines use ticks or dots to indicate important events within a particular topic, but here Lego employs pictures of the toys that were created as a metonymical reference to the time at which they were first produced. This timeline is also created by Lego blocks, but they do not have the same metonymical representation that the images do. This timeline goes beyond a TIME IS SPACE metaphor, and it is the result of a blend of different mental, physical and contextual structures (Pagán & Coulson, 2013). The mapping that takes place here is not limited to a space-time mapping, but there are a several domains that contribute to the creation of the metaphorical meaning (colours, images, shapes…).

There are also other examples of timelines in which the relation between time span and space might be different. Sometimes the same amount of space is provided for different events that last different amount of times, while other timeline's spatial distribution might change depending on the time-span of the event. CIT suggests that the conceptual metaphor TIME IS SPACE does not determine the interpretation of the timeline, but it is a constraint that can be flexibly changed depending on the functions and goals of the mappings that take place.

Similar to language, timelines also have a direction and orientation. Western cultures usually follow a left-to-right and top-to-bottom orientation, probably influenced by the writing direction and the layout of written texts. However, this

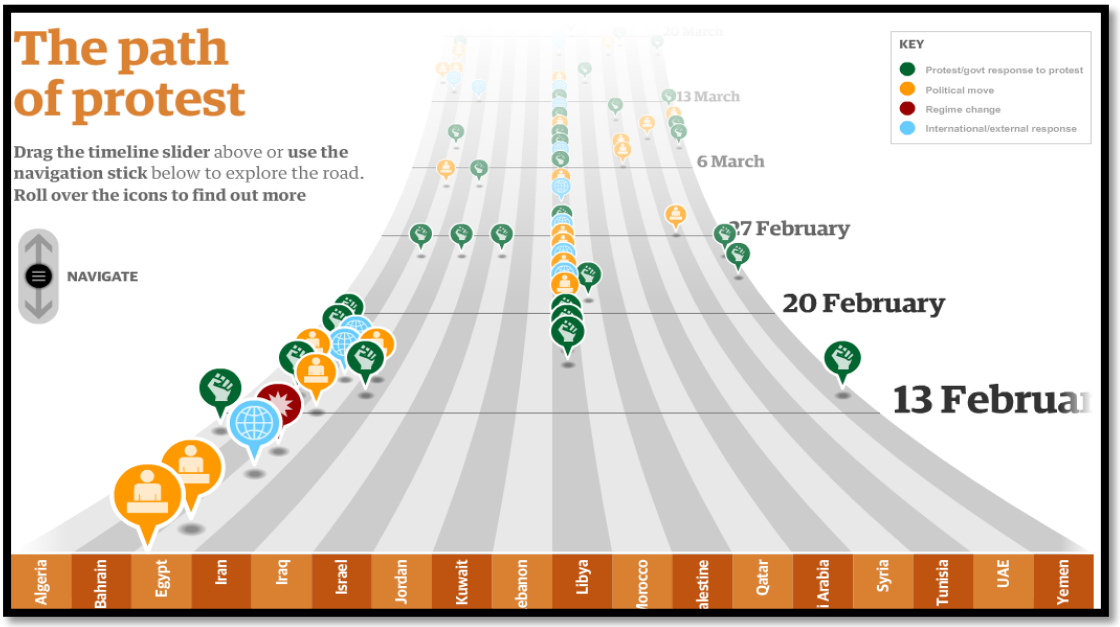



*Figure 2.4 Novel realisation of a timeline, in which future is in front of us and up and past is below and behind us.*

orientation is not congruent with the data we obtain from language, which often follows a front-back orientation. Traditionally, timelines could not be represented in the sagittal axis, so Davis (2012) suggests that the sagittal axis rotated into the lateral or vertical one, with past time on the left and at the top and future time on the right and at the bottom. However, it is possible to create a wide variety of timelines thanks to modern technology, as can be appreciated through 2.4, which shows a sagittal-vertical timeline in which the past is down and behind us and future events are up and in front of us.

Timelines play a very important role in the structure of our cognitive ability; they are not simply space-time mappings but are a linear schema that has been developed in order to meet the needs of time conceptualisation. Thus, timelines show that the time-related conceptual structures are not only a product of language, but they are essential in the structure of our cognitive ability (Pagán & Coulson, 2013). Different authors have proposed that humans may represent time by using a "mental time line" (MTL) (Santiago et al., 2007, 2010; Ishihara et al., 2008; Müller and Schwarz, 2008; Vallesi et al., 2008, 2011; Weger and Pratt, 2008; Di Bono et al., 2012). Bonato et al. (2012) synthesise the main features of the MTL hypothesis as follows:

(1) Time is represented along a spatial continuum similar to a line.
(2) The orientation of the MTL is embodied and culturally influenced; it corresponds with the writing direction.
(3) Time is represented in relative terms; time is coded with respect to a reference point in the timeline.
(4) Spatial attention is involved when using the MTL.

MTL have also be found to interact with other cognitive strategies that are somehow related to space. For instance, research suggests that there are certain parallels with the Mental Number Line (Dehaene et al., 1993; Stoianov et al., 2008; Zorzi et al., 2002, 2012) and that time-space interactions seem to emerge when temporal tasks are characterised by spatial aspects (Torralbo et al., 2006; Santiago et al., 2007; Vicario et al., 2008, 2009; Ouellet et al., 2010a). Furthermore, several studies show that the left side is usually associated with short duration and past events, while right side is associated with long duration and future events (Torralbo et al., 2006; Santiago et al., 2007; Conson et al., 2008; Ishigara et al., 2008). Finally, research also suggests that there is a relation between the MTL and cultural practices like writing direction

(Tversky et al., 1991; Fuhrman and Boroditsky, 2010; Ouellet et al., 2010b) or linguistic factors (Núñez & Sweetser, 2006; Boroditsky and Gaby, 2010; Boroditsky, 2011). It can be inferred from the evidence that there is a link between the processing of time and its relation to space, and even though research supports the usage of a MTL, further study is required to understand how time is mentally represented.

Finally, recent research has suggested that, even though timelines in western cultures follow a predominant left-to-right pattern shaped by the aforementioned characteristics, it is also a very flexible device that can be employed in numerous ways. Pioneer large-scale gesture research has suggested that the imagined timelines which are created by co-speech gestures might be much more flexible than other types of physical timelines (Valenzuela et al., in press), allowing speakers to locate future events on the left and past events on the right or using vertical gestures to talk about time, among others. It will be one purpose of this theses to research further on the timeline-co-speech gesture relation, since it will provide us with very useful data to understand how time is conceptualised in our minds.

# Chapter 3

# Experimental approaches to time

### *3.0 Introduction*

This chapter provides an overview of non-linguistic approaches to time conceptualisation drawing evidence from behavioural, psychological and neurological research, as well as research from cross-linguistic studies.

Section 3.1 addresses non-linguistic research on time-space mappings from a range of non-linguistic disciplines. A general outlook of the current research is provided by showing different pieces of research and theories that use non-linguistic data to investigate time conceptualisation. This section addresses the relation between time and spatial experience, the representation of space and time in the brain through neurological data, the relation between sound, length and duration, the influence of extra linguistic factors on time conceptualisation and finally an overview of time conceptualisation from the 'A Theory of Magnitude' (ATOM) perspective.

Section 3.2 focuses on a range of research that has focused on cross-linguistic and cross-cultural comparison in time conceptualisation, addressing a wide range of languages such as Spanish, English, Mandarin, Swedish or Pormpuraaw, amongst others. More precisely, I focus on the differences that emerge from the use of different spatial frames of reference, changes in axis and directionality and duration. Finally, I also address some cultural differences that emerge from writing direction.

### *3.1 Temporal cognition and multimodal data*

Time is a universal phenomenon that is understood by humans in very different ways. Over the past century, researchers have attempted to investigate this issue by using a wide range of different methods and data that could shed light on time conceptualisation processes and its relation to space. Section 2.3 of this thesis addressed this matter by presenting research performed exclusively with linguistic data that suggested that such a relation is not arbitrary but it is based on our embodied experience, that is, language is reflecting a psychological relation between the spatial and the temporal domains. Here I will present a wide range of research performed in different modalities that reinforces the idea of the conceptual relation between space, time and other domains of experience.

*3.1.1 Psycholinguistic evidence for space-time mapping*

Recent research has suggested that there is a potential role of body motion in temporal cognition, as it is represented in language by expressions such as "the days behind/ahead of me", in which future and past concepts interact with motion. Miles et al. (2010a) performed an experiment in which participants were asked to think how their life was four years ago during fifteen seconds, and later they were asked to think about how their life would be in four years in the future. Results suggest that people tended to lean forward when thinking about the future, while they tended to lean backwards when thinking about the past. Furthermore, another study suggests that apparent forward motion tends to induce more future-related daydreams, while apparent backward motion does so with past-related daydreams (Miles et al, 2010b). Finally, an experiment performed by Hartmann & Mast (2012) located participants over a motion platform which moved forward/backward in the first experiment and upward/downward in the second one. While on the platform, participants had to categorise as quickly as possible future and past-related concepts. Results suggested that future-related verbal stimuli were processed faster during passive whole-body forward motion, although there was no effect in the upward/downward movement, presumably because the participants were English speakers, who do not tend to sue this axis very often in time conceptualisation. The results from these studies serve as empirical evidence of the experiential link between time and space; time is not only represented on the sagittal axis in language, but sagittal motion also has an effect on temporal thinking.

However, Boroditsky and Ramscar (2002) suggest that the way people think about time is as much tied to their thinking about spatial motion as it is actual spatial motion. In order to reach this conclusion, they performed several experiments in laboratory conditions and real-life situations. In the first experiment, participants were asked to answer the following ambiguous question: "Next Wednesday's meeting has been moved forward two days. When is the meeting now that it has been rescheduled?" after being primed with ego-moving or time-moving time questions (see chapter 2, section 2.3.2.1 for further information). This question has two possible answers depending on the time of T-FoR that is employed when answering it. If the speaker is using an ego-moving T-FoR, they would conceptualise themselves moving through time, being the past behind them and the future in front of them. In this case, the conceptualiser would respond that the meeting is on Friday, since Wednesday would be

(temporally) in front of the speaker and *two days forward* would be Friday. On the other hand, if the speaker uses a moving time T-FoR, the temporal event is the entity that acquires motion. When the temporal event moves *forward*, it gets closer to the speaker (and thus their deictic centre), ending up on Monday.

Results suggested that participants answered differently to the question depending on the T-FoR that had been primed: moving-ego priming resulted in a higher number of Friday answers while moving-time priming resulted in more Monday answers. Further research was performed in a lunch line of a university café, an airport and a train. Generally, people who had spent some time moving forward (queuing at the lunch line) or who were thinking of doing (people in a moving train) were more likely to take an ego-moving perspective. This research would suggest that not thinking about spatial motion (generally on the sagittal axis) has an effect of the T-FoR that is employed for conceptualising time.

Lastly, some research has been performed to confirm whether or not T-FoR have an influence on how we conceptualise time when using spatial input. Kranjec & McDonoudh (2011) looked at the influence that deictic or sequential T-FoR could have on time conceptualisation by locating participants on a chair with a box in front of them and another one behind them. The researchers then spoke aloud names of different pictures that were located in the boxes, and the participants had to guess in which box they were. The images were related to deictic past/future (World Trade Center or Freedom Tower) or sequential language (butterfly and caterpillar). The same experiment was then conducted by placing the boxes on the left and right of the speakers. Results suggested that more people became aware of the temporal relation of the experiment only during the first study (sagittal axis), giving answers corresponding to the deictic stimulus, but not the sequential ones. However, the second part of the experiment showed different results. When the boxes were placed left and right from the speakers, participants did not become aware of the temporal aim of the experiment, not showing the strategies present in the first part of the experiment.

A different paradigm that has been used to study time mappings involve word processing tasks, in which participants have to categorise temporal words, which are presented in different spatial locations (back/front), as referring to the past or the future (Torralbo, Santiago & Lupiáñez, 2006). Results show that responses were significantly

faster when word location was congruent with the back-past/front-future metaphoric mapping. In a second study, a left-right manual response was included, and responses were again found to be faster when they were congruent to the left-past/right-future mapping. Another experiment which shows similar results was conducted by Weger and Pratt (2008). In this experiment, participants had to decide if a list of actors became famous before or after they were born by manually responding with their left or right index finger. Data showed that names of younger actors were categorised faster and they tended to be responded to faster with the right index, while older actors were categorised faster with the left finger. A second experiment was conducted in which instead of categorising actors, participants had to categorise future and past-related words. Results were the same as in the first experiment, highlighting the importance of the left-right/past-future spatiotemporal mapping.

The results from all these experiments show that there is a strong link between the sagittal spatial experience and time conceptualisation. Spatial motion, thinking about spatial motion and spatial arrangement of items seem to be linked to temporal thinking. It is important to note that most of the results were linked to the sagittal axis or to sagittal spatial movement, which correlates with sagittal temporal language in English. This data presents some interesting questions when contrasted with gesture data and other experimental evidence, which highlights the importance of the lateral axis when manipulating time-related concepts (see section 4.4).

So far, all the previous mentioned research has involved somehow temporal-related language in the experiment which could have had an influence on the time-space conceptual mappings, but a recent body of literature has tried to address this problem by offering auditory cues instead of linguistic ones. Srinivasan & Carey (2010) performed an experiment in which participants saw different lines with different lengths while listening to a tone whose duration also varied. In one of the groups, the duration and the length were positively correlated, while in the other group longer lines were paired with shorter ones. Results in this experiment revealed that adults were better at pairing spatio-temporal information when length and duration was congruent. The experiment was repeated by changing the loudness of the tone instead of the duration, but participants in both groups presented similar results. The same experiment was conducted with infants with similar results. Children in the congruent group noticed the length-duration relation while participants in the incongruent group did not. This last

piece of evidence is remarkably interesting, since it suggests that the functional representation of time-length is not based on its linguistic encoding, but a functional experience overlap. This research also suggests that factors such as the loudness of the sound do not have an impact on time conceptualisation.

Another experiment that appealed to auditory primes analysed the relation between gesture and singing. In this experiment, participants viewed a video in which an actor sang a note while gesturing left or right and then they had to reproduce the duration of the note by pressing a button. The perceived duration of the note was increased by long-distance gesture relative to short-distant ones, although the direction of the gesture (left-to-right or right-to-left) presented no differences (Cai et al., 2013). These results suggest that there is indeed a link between time-space representations and, more precisely, imply that representation of temporal duration is non-directional.

The final experiment that will be reviewed here, which employed auditory information, was performed by Walker et al (2014). In this study, participants were presented with auditory information from speakers who were located on their left, right, back or front. The auditory information consisted of deictic or sequential temporal language and participants had to decide whether the temporal cue was located in the past or in the future. Results showed that participants required more time to make sequential judgements than deictic ones. Moreover, there was an association between sequential judgements and the location of the axis as participants associated earlier judgements with being in front of them and later judgements as behind them. This compatibility effect that took place in the sagittal axis was not found in deictic judgements.

Additionally, a different body of literature has focused on investigating the conceptual relation between spatial length and temporal duration without using linguistic stimuli. For example, Casasanto et al (2009), as well as Bottini and Casasanto (2013), presented data from two similar experiments in which children had to judge spatial information with temporal interference as well as temporal information with spatial interference. In order to do so, children were presented with different videos of animals traveling along parallel paths for different distances or durations. Results showed that children were better at judging distance in the presence of temporal interference than judging temporal duration in the presence of spatial interference,

showing a similar asymmetric time-space relationship which is consistent with the data obtained through similar experiments performed with adults (Boroditsky, 2000; Casasanto & Boroditsky, 2008).

On the other hand, other studies have focused on the relation between spatial distance and temporal duration. Coulson & Pagán (2013) offer some evidence which shows that the use of the timelines' spatial distance in relation to temporal duration might vary across examples. The tenets that PROXIMITY IN TIME IS PROXIMITY IN SPACE and TEMPORAL DURATION IS SPATIAL EXTENT are not maintained the same across timelines.

Another example of the spatial distance-temporal duration paradigm can be observed in a series of experiments performed by Christian et al (2012). They analysed the differences between implicit and explicit interpretation of temporal duration in relation to spatial distance. Participants had to locate birthdays which referred to themselves, a close friend and a stranger in the past, present or future on a timeline whose middle point was labelled as "NOW". Results showed that past events were always placed on the left and future events on the right. Moreover, the amount of space used to represent one year increased as a function of the familiarity between the participants and the target. In the second experiment, participants were exposed to apparent forward-backward self-motion. Participants had to stop the video that was inducing the motion when they thought they had arrived at the temporal target destination. Similar to the results in the first experiment, participants travelled for a longer amount of time when they were referring to events which were more self-relevant.

All in all, the literature suggests that there is a relation between both spatial and temporal length and temporal duration. When length is congruent with temporal duration, participants are faster to recognise it. However, it seems that spatial length has a higher influence on time conceptualisation than temporal length.

### *3.1.2 Time-space in the brain*

Even though there is a strong belief that time is conceptually related to space through common features of their experience (i.e. Evans, 2004, 2013; Radden, 2004; Moore, 2006, 2014, see Chapter 2, 2.3 for further detail), there is still a lack of neural data which supports the claim that temporal concepts are grounded in sensorimotor

representations. Kranjec & Chatterjee (2010) suggest that one way of proving that space domain-specific sensorimotor areas are associated with temporal perception and perspective concepts would be to conduct different brain-lesion studies. In order to cover this gap, Saj et al. (2014) conducted an experiment in which the effects of right-brain damage concerning time-space mappings was analysed. Healthy, right-hemisphere stroke with spatial neglect and right-hemisphere stroke with no spatial neglect participants had to learn about a fictional man and about his likes/dislikes ten years ago and ten years in the future. After that, they had to recall the different items and recognise them as past/future related. Data suggests that patients with spatial neglect recalled fewer items that were associated with the past, although the same pattern was not found in relation to future items. Participants with spatial neglect also mislabelled items belonging the past as future items. Even though much further research requires to be performed using neural data, these results suggest that the ability to represent space is necessary for performing an accurate representation of time.

### *3.1.3 Time and extra-linguistic factors*

One of the most recent bodies of research on temporal cognition that has emerged over the last years looks at the effect of a number of extra-linguistic factors on time conceptualisation. More particularly, Duffy (2014a, 2014b, 2016, 2017) has led a number of experiments that looked at a wide number of factors that may influence time conceptualisation, such as individual differences, personality, power or valence.

Research suggests that factors related to personality or lifestyle might have an influence on the temporal perspectives that we adopt. Hauser et al. (2009) points out that not only the situation in which the speaker is, but also a number of individual differences can have an effect on how time is conceptualised. In order to address individual differences, Duffy and Feist (2014) compared two sectors of the population with two different time-related lifestyles. The first group included university students, who optionally attend university and they are mostly in control of their time, having a higher degree of temporal flexibility. The second group included administrators, who are rather more controlled by time and are more temporally inflexible. Both groups were presented the *Next Wednesday's meeting* question. Results showed that administrators were more likely to respond *Monday* (corresponding to the Moving Time T-FoR) while students were more likely to give a *Friday* answer (corresponding to a Moving Ego T-FoR).

They conducted another experiment in order to investigate whether individual factors such as procrastination, conscientiousness and extroversion had an effect on temporal cognition. Higher levels of extroversion were linked to the Moving Ego perspective, while participants that presented lower levels tended to adopt a Moving Time perspective. This was further corroborated by another experiment performed by Duffy & Evans (2017).

Similarly, results showed that participants that presented a higher level of procrastination were more likely to present a Moving Ego perspective while participants with a higher level of conscientiousness were more likely to adopt a Moving Time perspective. The effect of conscientiousness and procrastination was corroborated by a follow-up study performed by Duffy et al. (2014). Additionally, they showed that people who are on schedule (e.g. for a meeting) tend to adopt a Moving Time perspective, while people who are late are more likely to adopt a Moving Ego perspective.

The last piece of evidence might be related to the amount of power that the conceptualiser has over the temporal event. Duffy & Feist (2016) conducted an experiment to investigate the possible connection between the Ego Moving and Time Moving perspectives and the conceptualiser's degree of power. Results show that participants that maintained a high-power pose before performing the experiment tended to adopt a Moving Ego perspective and hence they thought of themselves moving through time. Finally, Duffy & Evans (2017) showed that there is a relation between the temporal perspective which is taken by the conceptualiser and the valence of the event they are thinking of. Results showed that, when thinking about a positive event, participants were more likely to adopt a Moving Ego perspective, while when referring to a negative valence situation, participants tended to adopt a Moving Time situation.

Overall, results suggest that there is a tendency to adopt a Moving Ego perspective when we are in high-power situations, as well as being linked with positive events. Conversely, the Moving Time perspective tends to be employed in low-power, negative situations. Further research on the effects of extra linguistic factors and individual differences will shed some valuable information on the different cognitive

processes that may have an effect on our cognition and, more precisely, temporal reasoning.

*3.1.4 Time and number: A Theory Of Magnitude vs Mental TimeLine.*

A different approach that has been taken in research about time-space mappings is the analysis of magnitude representation. Some authors have suggested that there could be a shared representational mechanism for time, space and number. This system was firstly proposed in Meck & Church's (1983) Accumulator Model, and it was later extended to A Theory Of Magnitude (ATOM) (Walsh, 2003; Bueti & Walsh, 2009). In contrast to CMT (see Chapter 2, section 2.2.5), which proposes an asymmetrical relation between space and time, ATOM suggests that the domains of time, space and number use the same underlying representational mechanism and are symmetrically interrelated (Dormal & Pesenti, 2013).

However, it remains to be seen which of the theories is the most accurate, since data shows mixed results. On the one hand, several experiments have shown that there is an effect of numerosity on duration processing when employing symbolic and non-symbolic stimuli in children (Droit-Volet, Clément & Fayol, 2003) and adults (Brown, 1997; Dormal, Seron & Presenti, 2006; Oliveri et al., 2008); the reverse influence of duration on numerosity is also present, albeit weaker (Brown, 1997). Crollen et al. (2013) also provide some evidence suggesting a common magnitude processing mechanism. When performing perception and production tasks in number, length and duration, they found that participants tended to underestimate all the magnitudes during perception tasks and overestimate them in production tasks.

On the other hand, Bottini & Casasanto (2013) argue that, even though there are several experiments that suggest there is a symmetry between these three domains, there is a large body of literature that claims that this relation is asymmetric (see section 3.2.3); e.g. length can bias duration interpretation but duration has no effect in length perception.

Last accounts towards the ATOM vs CMT paradigm opt for offering a complimentary view: simple representations of magnitude cannot account for complex time-space-number interactions which are addressed by CMT, while ATOM offers a better explanation when referring to more low-level and language-independent time-space-number associations (Winter, Marghetis & Marlock, 2015)

### *3.2 Cross-cultural differences*

Another area in which a wide range of literature can be found in on the topic of time conceptualisation and the differences that might emerge among different cultures and languages. This includes not only time-related language, but also differences in writing direction. Researchers have focused mainly on comparing the differences of time-spatial properties, with less emphasis on other properties like duration.

#### *3.2.1 Absolute time conceptualisation*

It has already been discussed that languages can use different systems of coordinates to construct oriented space, developing different S-FoR (see Chapter 2, section 2.3). Since the relation between time and space is so close, would there be any differences in time conceptualisation between languages which predominantly use different S-FoR? Boroditsky and Gaby (2010) attempted to shed light on this issue by comparing time conceptualisation between the members of the Pormpuraaw community, an Australian aboriginal community who arrange space in an absolute way (according to cardinal directions) and English speakers, who heavily rely on relative space (according to the ego). After performing two experiments in which they had to arrange a set of cards or draw dots while being seated in different cardinal directions, results showed that English participants always arranged time from left to right, while the Pormpuraaw members arranged it from east to west. Similar results were obtained by Fedden and Boroditsky (2012), whose experiment was replicated with speakers of Mianmin of Papua New Guinea, who also predominantly use the absolute S-FoR.

#### *3.2.2 Axis and directionality*

Two more time-space features that show cross-cultural and cross-linguistic variation are the axis on which time is construed as well as its directionality. This difference can be easily appreciated when comparing two languages like English and Mandarin. Research suggests that, when measuring response time, Mandarin speakers are faster to make temporal judgements after receiving a vertical spatial prime than English speakers are (Boroditsky, 2001). This is because in language, Mandarin encodes temporal relation on the vertical axis. Even though these results were questioned (see January & Kako, 2007; Chen, 2007), they were corroborated later by Miles et al. (2011) who, by performing a similar experiment on Mandarin-English bilinguals, suggested that they employ vertical and horizontal representations of time which are prone to undergo subtle variations due to the cultural context. In ‘Chinese’

contexts, bilinguals were more prone to using vertical representations, while in 'English' contexts they were more likely to use horizontal ones. English and Mandarin speakers also presented some differences in a card-arrangements task. Mandarin speakers arranged a temporal sequence depicted in pictures following a vertical pattern a 30% of the cases, while English speakers never did (Chan & Bergen, 2005). Mandarin-English time-space mappings have also been tested in a 3D environment by Fuhrman et al. (2011). Results were similar to ones obtained by previous studies; English and Mandarin speakers showed a left-to-right representation of time (in accordance with writing direction), but only Mandarin speakers presented a top-to-bottom spatial representation of time, which is congruent with Mandarin linguistic temporal metaphors and writing direction. The last experiment, performed with Mandarin-English bilinguals, suggested that participants who were more proficient in Mandarin and were tested in Mandarin rather than in English were more to arrange time vertically.

In addition to Mandarin-English time mapping studies, some research has also analysed this phenomenon between Spaniards and Moroccan speakers of Arabic (de la Fuente et al., 2014). Even though both languages use future-in-front temporal metaphors, Moroccans have a tendency to place the past in front while gesturing (Román et al, 2014) In this experiment, participants saw a cartoon character with a box in front of him and another one behind him. After that, they read a story in which this character visited a friend who likes plants and in the future, will visit another friend who likes animals. Participants had to place "animal" or "plant" in the different boxes according to their temporal judgement. Results showed that Spaniards tended to place the future in the front box, while Moroccans placed the past in the front box. In a second experiment, participants had to answer a number of questions about the past and the future. This experiment showed that the past-in-front mapping is caused by culture rather than language, since Moroccan participants showed a higher agreement with past statements while Spanish ones showed it with future statements. The last experiment tested if age could be a factor among Spanish speakers for time conceptualisation. Results showed that young Spanish speakers responded according to a future-in-front mapping, while half of the old Spanish speakers used a past-in-front mapping.

### *3.2.3 Duration*

Another of the time related areas in which cross-linguistic research has been performed is cross-linguistic difference in the conceptualisation of duration. One of

these examples can be found in Casasanto et al. (2004), who make a linguistic comparison between four languages in terms of duration. Through the use of linguistic data obtained from Google, they suggest that English and Indonesian languages tend to map duration into linear distance (*long time*), while other languages like Greek or Spanish tend to map duration onto quantity (Spanish *mucho tiempo*, *much time*). Following up on this, Athanasopolous and Bylund (2017) performed a number of psychophysical experiments in Spanish and Swedish monolinguals as well as bilinguals to investigate their conceptual experience of temporal duration. In these experiments they found out that, when talking about duration, Swedish speakers were misled when prompted with length-related stimulus, while Spanish speakers presented the same behaviour when prompted with size and quantity stimulus. Interestingly, linguistic interference disappeared when the linguistic queues were eliminated from the task.

Lai and Boroditsky (2013) performed a linguistic analysis to compare motion in time conceptualisation between English and Mandarin. In this experiment, the cognitive salience of the ego-moving and time-moving metaphors was compared in English and Mandarin monolinguals as well as English and Mandarin bilinguals. Results revealed that Mandarin speakers tended to take a time-moving perspective, consistent with the linguistic analysis which was performed. On the other hand, bilinguals were less likely to adopt this perspective than Mandarin monolinguals, suggesting that there is an influence of the second language on the first one.

### *3.2.4 Writing direction*

Another of the factors that seems to influence how time is spatially conceptualised is the writing direction that is employed by the conceptualiser. People who read or write from left to right arrange time in the same direction, while people who read and write from right to left arrange time accordingly. This reversal in the time-space mapping has been reported by a number of research papers that include picture arrangement, implicit non-linguistic association, auditory tasks or even gesture-related tasks, in languages such as Spanish and English (left-to-right) as well as Hebrew and Arabic (right-to-left) (Fuhrman & Boroditsky, 2010; Tversky et al., 1991; Ouellet et al, 2010). Moreover, writing direction has also been reported to have an influence in blind speaker (Bottini et al., 2015). Reading texts from left to right by using vision or touching seems to lead to a left-to-right horizontal representation of the mental time line (MTL).

Finally, it seems that writing direction not only influences direction, but also axiality. As it was previously reported, vertical time-space mappings in Mandarin seem to be influenced by the writing direction (Fuhrman et al, 2011) and linguistic metaphor. Moreover, Casasanto and Botini (2014) performed an experiment in which Dutch participants were exposed to temporal phrases in different writing directions (left-to-right, right-to-left, top-to-bottom and bottom-to-top). Results showed that the exposure to a new orientation was able to make participants change the direction and the orientation of their MTL, even though this new mapping was contradicting the conceptualiser's usual mapping.

# Chapter 4

# Gesture as a method to study cognition

### *4.0 Introduction*

This section introduces gesture studies and the usefulness of gesture data as a source of information for the study of mental processes and metaphorical thinking. I firstly provide a brief theoretical overview of gesture studies in section 4.1. Here I define the exact meaning of gesture that I take into account by offering a brief overview of some of the definitions that leading gesture researchers have offered. Subsequently, I provide a brief summary of some of the different approaches to gesture studies that have been taken, as well as a more detailed overview of different gesture typologies, giving emphasis to McNeill's proposal for gesture categorisation.

Section 4.2 focuses on two key aspects of gestures: form and function. Firstly, I address Efron's gesture typology, which is considered as one of the first attempts to classify gestures. Then, I describe Bressem's (2013) framework for gesture form annotation, as well as explaining the notion of recurrent gesture. I finish this section by describing a number of gesture typologies that aim to address gesture function, mixed systems that address both form and function and other gesture typologies.

The final section, 4.3, is devoted to justifying the importance of gesture studies in the study of human cognition. Here I refer to a large body of empirical research linked to language learning, metaphorical thinking or human communication which uses gesture as the main source of data. I finish this chapter by providing a thorough description of the research that has been conducted so far in relation to gestures and time. Here I present a wide array of studies which address cross-linguistic gesture information, unconscious co-speech gestures, and temporal gestures in story telling or consciously prompted temporal gestures.

### *4.1 Gestures studies: an overview*

This section provides an overview of gestures studies and, more particularly, the importance of this field for the study of metaphor. First, I cover the different definitions of gesture and state which is the most suitable for the purpose of the thesis. Next, a brief account of the different approaches to gesture studies as well as the most important gesture typologies is given. The section finishes by providing empirical research that

supports the idea that gestures are an invaluable tool for the study of human communication and its relation to metaphorical thinking and, more precisely, temporal metaphor.

*4.1.1 Definition of gesture*

The term *gesture* has been defined several times during the development of Gesture Studies (GS), and it has been attributed slightly different meanings depending on the aspects that different authors wanted to highlight. Rather than relying on a concrete definition of gesture, here I will introduce a number of different previous interpretations in order to establish the most suitable definition for the aim of this thesis.

A good starting point for understanding what we refer to when we talk about gestures can be found in Cienki (2008). He states that, in a very broad sense, gesture can refer to "any wilful bodily movement" (p.6). Although this definition can be too broad for a precise understanding of gesture, it sets a number of features that are crucial to distinguish gestures from other types of human kinetics, like body language.

Kendon (2004) proposes that the best definition of gesture is "visible action when it is used as an utterance or as part of an utterance" (p.7); where an utterance is any ensemble of actions that counts for others as an attempt to give information (Goffman, 1963: 13-14). According to Kendon, (1972, 2004), gestures need to manifest deliberate expressiveness and they need to have at least a certain degree of voluntary control.

A more detailed definition of gesture is provided by McNeill (2005). He suggests a number of features that need to be in place in order for a particular body movement to be considered a gesture. The first characteristic of a gesture is that it is *wilful*; a gesture needs to be manifested with some degree of voluntary control. Body signals which are involuntarily performed, like nervous tics, scratching or other signals like dilatation of pupils do not fall within the definition of gesture, but rather they belong to body language. Secondly, a gesture is not restricted to be performed just with the hands. Gestures can be performed with different body parts like the nose, elbows, feet, or they can even be produced by abstractions from our body, like the gaze. Finally, gestures have a communicative intention. Gestures convey particular messages, either in place of or in conjunction with speech. Here are also included gestures whose aim is the expression of thought or feelings, like the ones produced by facial expressions.

Taking all this into account, Kendon (1980, 2004) suggests that there should be a distinction between *gesture* and *gesticulation*. A gesture includes all the type of body movements that fulfil the aforementioned characteristics, while *gesticulation* refers to the particular movements of the hands when accompanying spoken language. McClave (2000) further extends this notion by suggesting that the head should also be included under this definition, since it can act as a third hand, as well as the legs and feet.

Here I combine some of the characteristics proposed by McNeill (voluntariness and communicative intention) with Kendon's notion of gesticulation. By using the term *gesture* I address any kind of partially voluntary hand movement which appears to have a manifest communicative action. Other types of body movements like head or feet movements, as well as extensions of the gestures (gaze) will not be taken into account. Moreover, gestures co-occurring with speech will be referred to as co-speech gesture. In this case, gestures are expressed contemporaneously with speech and the content part of the gesture is bound with semantically corresponding parts of speech

Gestures can be broken down to different stages in their realisation which constitute a "Gesture Unit" (Kendon, 2004). This term refers to the movement performed by the hand when it starts in a position of rest, performs a particular action and then comes back to rest (generally the same position where it was before). During this action, the gesture-related body part can perform several movements, such as pointing, outlining or sculpting a shap. This stage is known as the "stroke", which McNeill (1992) considers to be the content-bearing part of the gesture. Previous to the stroke, the hands usually need to prepare and organise themselves to perform the action; this phase is known as "preparation". Together, the preparation and the stroke form the "Gesture Phrase", with the stroke being the nucleus of the phrase. A third (optional) stage can be added to the Gesture Phrase, which is known as the "retraction" and it concerns the movement that the gesture makes when returning to a position of rest. Several Gesture Phrases can occur in the same Gesture Unit; the different gesturing body parts can perform several preparations and strokes between the initial position of rest and the final rest (Kendon, 2004).

Kendon (2004) further suggests that gestures (in his sense) are patterned in relation to speech. He observes a relation between tone units, syllables, phrases and patterns in bodily action; speech units are associated with phases of movement of a

shorter duration. The stroke of Gesture Phrases tends to anticipate slightly or to coincide with the tonic centre of the verbal utterance. A semantic coherence is usually established between the gesture produced and the verbal message, what McNeill (1992) named "co-expression". Sometimes the meaning is parallel to the verbal meaning, but it usually complements or adds different nuances to it.

*4.1.2 Approaches to gesture studies*

Gesture studies is a field of interest since Greek and Roman times, with authors like Aristotle or Cicero giving emphasis to the relation between body language and rhetoric. These early scholars defined gesture as the carriage of the body, the posture and actions performed by the hands, head and glance (Kendon, 2004: 17). Gestures also had an impact on research in human communication and the idea of a universal natural language (see Bonifacio, 1616; Bulwer, 1644), and also in the development of sign language teaching methods (see Charles-Michel de l'Epée, 1776). However, it was not until the mid-20$^{th}$ century when interest in the relation between language and thought started to emerge, with authors like Condon & Ogston (1966, 1967) or Kendon (1972) suggesting that gesture and speech were two different dimensions of the same process: communication.

These observations were later developed by several authors, one of the most influential being McNeill (1979, 1985, 1992), whose main concern was the study of the integration and unification of two different representational modes (speech and gesture) in communication processes. It has been argued that co-speech gestures can be employed as a "window to the mind", since they are shaped by the speaker's own meanings. Since gestures are performed in a different mode than speech, they can reveal levels of meaning which are not available through the sole study of verbal communication (McNeill & Duncan, 2000:143). McNeill & Duncan (ibid) proposed a theory to explain the language-gesture-thought integration through the *growth point hypothesis*. According to them, any utterance is the product of the conceptual unpacking of an idea unit; whatever the conceptualiser wishes to express is presented all at once from the beginning. This idea is shaped through imagistic thinking and thinking that involves linguistic categories and it is later conveyed through gesture and speech. This theory contrasts with previous attempts to explain the language-gesture-thought link, like the ones proposed by Levelt (1989), Kita (2000) or De Ruiter (2000), who suggested a linear model to language-gesture integration, in which speech and gesture

are two independent processes which have the same starting point and collaborate to achieve a common goal through different steps in the elaboration of the utterance; language is the main mode of communication, but aspects that cannot be expressed through speech pass to gesture (Kendon, 2004:79).

The idea of gesture as a window to the mind has had an increasing impact in disciplines that aim to investigate the different cognitive processes that are involved in language and communication and, more precisely, Cognitive Linguistics (CL). Since CL takes a conceptual approach to language (see section 2.1), gestures are seen as a way of representing our conceptualisation of the world through physical manual realisations (Glenberg & Robertson, 2000). Language and gestures are employed to refer to our conceptualisation of the world during communicative processes.

A wide range of empirical research supports the idea that gestures can be used as a way to further investigate the relation between communication and cognition. Núñez and Sweetser (2006) support this statement by providing a number of arguments in favour of the inclusion of co-speech gestures in the study of cognition. One of the first arguments is that gestures are a universal phenomenon; co-speech gestures take place cross-culturally and cross-linguistically (see Iverson & Thelen, 1999; Kita & Essegbey, 2001; McNeill, 1992; Núñez & Sweetser, 2001). The second feature is that, as has been already mentioned, they are largely unconscious and they are less monitored than speech. Due to the fact that speakers are not usually aware of the co-speech gestures they perform, they are less monitored than language and thus they can reveal information which is not present there. Co-speech gestures usually tend to be synchronised with speech, and they usually mirror or complement linguistic patterns which are verbally expressed (Núñez & Sweetser, 2006).

Iverson & Goldin-Meadow (1998) carried out an experiment which provides empirical evidence that co-speech gestures are not only aimed at the interlocutor, but the speaker also benefits from gesticulation. Through this experiment, they found out that gestures can be performed without the presence of an interlocutor; for example, people gesture while talking on the telephone or while producing monologues. Research was also conducted on congenitally blind individuals, who also performed co-speech gestures.

Another of the arguments in favour of the use of gestures for the study of cognition is that they are co-processed with speech; speakers whose hand movement is impeded present a slower speech production time and stutterers also stutter in gesture (Mayberry & Jaques, 2000). Gestures can also provide complementary information to verbal communication in order to successfully convey a communicative intention. Kendon (2000) shows that hearers synthesize meaning which is conveyed through both channels, but they are not able to distinguish through which channel meaning was conveyed. Additionally, there is plenty of experimental evidence that supports the link between gesture and speech in language development, further strengthening the hypothesis of the importance of co-speech gestures in verbal language (Bates & Dick, 2002; Goldin-Meadow, 2003; Iverson & Thelen, 1999).

Cognitive Linguistics has found in gestures one of the most valuable sources of information (together with language) to investigate some of the key areas in human cognition such as mental simulation, the representation of mental spaces or construal operations and perspective. However, here I will focus on research performed in two particularly interesting topics, which are metaphorical thinking and the representation of image schemas.

Cienki (1998, 2008, 2013) offers a wide range of research that highlights the importance of studying both oral and gesture communication when addressing metaphorical thinking. He points out that there are different ways in which gestures and language may interact when it comes to the representation of metaphorical thinking. There are occasions in which speakers represent the same metaphorical mapping both in language and in gesture; for example when someone uses the expression *balance all these things* metaphorically while producing at the same time a gesture in which both hands, slightly opened and facing up, bounce up and down (Cienki, 1998), trying to mirror the movement of a balance. Similarly, a speaker can metaphorically talk about prices *rising* or *lowering* while making an upward/downward gesture with their hands. Cienki and Müller (2008) provide further empirical evidence which shows that the metaphor MORAL ACTION IS BOUNDED MOVEMENT is represented both in language and gesture, for example when a speaker performs a fist with the hand which slightly moves forwards while uttering "**push** their moral limits".

Cienki (1998, 2013) also argues that there are some cases in which metaphors can be present in language alone or in gesture alone. Some linguistic metaphorical expressions do not have to be coordinated with a co-speech gesture, or they cannot be represented by one modality. This is the case, for example, for those metaphors that involve colours, since gesture cannot iconically represent this dimension (Cienki, 2013). A different case can be found when we deal with the concept of time. In this case, English language has a large repertoire of metaphorical expressions that refer to the flow of time sagittally (*look back in to the past, look forward into the future*). However, as will be seen in more detail in section 4.4, temporal co-speech gestures tend to be produced along the lateral axis (left-to-right), perhaps due to the relative ease with which gestures can be performed on this axis compared to the sagittal axis. Lateral spatial representation of time is congruent with other cultural artifacts like timelines or writing direction, but this mapping cannot be linguistically encoded, that is, we cannot say that if we move something to the right we moved it forward into the future. The case of the lateral conceptualisation of time can be considered as a case of a metaphor that has different metaphorical realisations in language and gesture.

There is a wide amount of research that supports the importance of gestures as evidence of metaphorical thinking. Some of the research that has been performed involves the role of gesture and metaphorical linguistic expressions in mathematics, the relation between the MORE IS UP or GOOD IS UP metaphors and gesture, valence and spatial gesture arrangement or, as will be indicated in section 4.4, the representation of time thorough co-speech gestures. (Cienki, 1998, 2005; McNeill, 1992; Núñez, 2006; Núñez & Sweetser, 2001; Núñez, 2008, Müller et al, 2013).

Research also offers several pieces of evidence indicating a link between gestures and the reception and production of image-schemas (Cienki, 2013). Cienki (2005) suggests that speakers can employ their previous knowledge on image schemas (see section 2.2) in order to interpret different gestures patterns. Furthermore, research also suggests that image-schemas could motive the production of certain gestures. Some examples of this can be seen in Ladewig (2011), who establishes a link between the CYCLE image-schema and the rotating movement of the hand, or in Williams' (2008) work on the PATH schema and a number of tracing gestures. Cienki (2013) further argues that gestures can be employed to schematically refer to physical objects through the iconic representation of hand gestures, which can be further extended to abstract

metaphorical representations as in timeline gestures. Finally, gestures can also be employed to perform schematic representations of actions, as is exemplified by gestures such as the palm-open-hand to present something (Kendon 2004) or the slicing gesture to indicate cutting (Calbris, 2003). The schematisation of these physical actions can at the same time be employed metaphorically with a more abstract meaning in order to refer to the presentations of arguments, the dismissal of an idea or the end of a process, respectively (Cienki, 2013).

### *4.2 Gesture: form and function*

One of the biggest challenges for gesture studies has been to establish a common typology that could be used to classify the wide variety of gestures that are produced during communication. It is, of course, a difficult task to develop a unique system that accounts for the different interpretations of gestures and hence some of the systems have focused on the form of the gestures while other have put more emphasis on the function.

One of the first attempts to establish the basis for a gesture typology can be attributed to Efron (1972). Even though his aim was not to create a classification system, his work served as the basis for the development of later typologies. Efron (ibid) analysed gestures according to three different properties: (I) spatio-temporal, (II) inter-locutional and (III) linguistic. The first feature, spatio-temporal properties, accounted for the physical and temporal features of the gesture; indicating the type of movement, the shape, the axis, the gesture time-span and the different body parts that were involved in the process. This property of gestures focused on the form and kinetics of hand gestures.

On the other hand, inter-locutional characteristics included all the interactional functions that gestures could have, including prosodic elements or turn-taking. Finally, linguistic properties addressed the semantic features of the gesture in relation to speech or by itself. In this last category, Efron further distinguishes among logical-discursive gestures, which refer to the ideational process of the sentence, and objective gestures, which have an independent meaning regardless of speech. Logical-discursive gestures could be further divided into batons, whose movement emphasises a particular word or phrase in speech, or ideographs, which sketch a particular path or direction of thought. On the other hand, objective gestures could be divided into deictic gestures (pointing), physiographic gestures (depicting the referent or reproducing an action) and emblematic

gestures, which are a standardised representation of meaning. Efron's approach was restricted to gestures that reference the discourse and gestures that referenced the topic, and hence only a small range of representational gestures was covered, basing his analysis on silent motion pictures and the observation of some natural conversation with a small range of topics (Müller, 1998: 93). However, the classification offered by Efron was one of the first steps towards the creation of a gesture typology that could address the different features present in gestures.

Throughout the development of gesture studies, researchers have aimed to create different classification systems to address the complexity of gestures. In the next subsection I will address a number of typologies that have been created to classify either the form of the gesture or its function.

*4.2.1 Gesture form*

Gesture researchers usually focus on four different physical properties of the gesture, based on the system created by Stokoe (1960) aimed at describing the properties of the hand gestures present in sign languages. These four parameters consist of handshape, palm orientation, movement and gesture space (Cienki, 2013). They encapsulate the different physical properties that can be discerned in to gestures. Bressem (2013) proposes a notation system that, by focusing solely on form, tries to address these four physical properties of gesture by establishing a set of categories into which the different parameters can be divided.

The first parameter I will address here is handshape, which refers to the different configurations that the fingers can take when making a particular handshape. Bressem (2013) suggests that the different handshapes can be divided in four basic categories: (1) "fist", (2) "flat hand", (3) "single finger" and (4) "combination of fingers". This description of the hand shape is based on which is the most prominent feature of the hand: the fist, the palm, the configuration of a single finger or the combination of fingers with the palm or other fingers, respectively. No further analysis is required when using categories (1) and (2), but when the hand shape involves the use of fingers, another layer in the categorisation is added which tries to address the finger(s) that are being used and their particular form. The different digits are numbered with the thumb 1 and the little finger 5, allowing for a description of the finger(s) involved in the gesture. Moreover, Bressem distinguishes six different shapes that fingers can take: (1)

"stretched", (2) "bent", (3) "crooked", (4) "flapped down", (5) "connected" or (6) "touching". "Stretched" fingers show no flexing at all, while the categories "bent", "crooked" and "flapped down" indicate different degrees of how flexed the fingers are. When referring to combination of fingers they can appear as "connected", in which the fingers are connected by the tip, or "touching" in which fingers are flapped down and touch each other (Bressem, 2013).

The second parameter, orientation, is based on the distinction made by McNeill (1992), in which he pointed out that it depends on the orientation of the palm as well as the gesture space (see below). Bressem (2013) indicates that the first step to take is to annotate which angle the palm is taking. The four basic angles include (1) "palm up", (2) "palm down", (3) "palm lateral" and (4) "palm vertical". Additionally, Bressem highlights that there are cases in which the palm can also have a diagonal orientation. Once the angle of the palm has been annotated, the second step is to indicate what the orientation of the hand is within to the gesture space. Here four different types are found: (1) "towards the centre", (2) "away from the centre", (3) "towards the body" and (4) "away from the body". If the handshape addresses, for example, fingers or a combination of fingers, it is also possible to indicate their orientation (fingers up or fingers down, for example).

The following parameter addresses the movement of the gesture, and it is probably one of the most complicated features to code. This parameter aims to describe what is considered the most effortful phase of the gesture movement (Cienki, 2013: 668) by taking into account three different aspects: type, direction and quality of movement (Bressem, 2013). Concerning the type of movement that the hand gesture produces, here Bressem distinguishes between movements performed by the gesture as a whole or movements performed by the wrist. Some of the basic movements that have been observed in hand gestures are (1) "straight lines", (2) "arced movement, (3) "circles", (4) "spirals", (5) "zigzag" and (6) "s-line" movements (Bressem, ibid). On the other hand, the wrist can present (1) "bending", (2) "raising" or (3) "rotating" movements. Finally, Bressem also distinguishes amongst different motion patterns of fingers, such as (1) "straight fingers", (2) "arced" or (3) "circle" as well as (4) "beating of fingers", (5) "flapping down", (6) "grabbing movement" or (7) "closing of fingers" for combinations of the digits. Once the type of the movement has been described, it is essential to note the direction of the movement. Hand gestures can be spatially located

in one of the three spatial axes: (1) “horizontal”, (2) “sagittal” and (3) “vertical”. Similarly to hand orientation, the direction of the gesture can also include diagonal movements. When addressing circular or spiral motion, Bressem distinguishes between (1) “clockwise” and (2) “counter clockwise” movements. The last feature concerns gesture movement addresses its quality, which can be analysed in terms of (1) “size” (reduced or enlarged), (2) “speed” (decelerated, accelerated) or (3) flow of movement (accentuated). These three different characteristics can be combined when describing the quality of the movement of a hand gesture.

The fourth and last parameter addresses the location of the gesture in the gesture space or, according to Bressem (2013) the “position” of the gesture. Cienki (2013) points out that this is an especially interesting dimension of gesture, since speakers can exploit the gesture space differently depending on the type of gestures they perform and the communicative context. Gesture space can be divided in four different sectors formed by concentric squares that start in the centre of the speaker and extend outwards (McNeill, 1992). Four different sectors can be distinguished: (1)”centre-centre”, which corresponds to the area around the chest of the speaker, (2) “centre”, corresponding to the space between the chest and the shoulders, (3) “periphery”, which refers to the gesture space slightly further from the shoulders, and (4) “extreme periphery”, which addresses gestures that are performed further from the body. This system is combined with another set of features that concern whether the gesture has been performed up, down, left or right. Additionally, Fricke (2007) offers a more detailed model for categorising gesture space in which four different dimensions can be distinguished: 0 refers to the speaker’s own body, 1 to close distance to the body, 2 middle distance from the body and 3 for distance from the body. These dimensions can be used to indicate both the forward and the backward distance of the gesture according to the speaker’s body by using +/- 1, 2 or 3.

Cienki (2013) further argues that another component that has an effect on the realisation of gestures is effort. He argues that effort is influenced by natural gravity and the constraints of human biology, although other external components such us the location of the interlocutor or the presence of items in the hands might also contribute to it. For each of the parameters aforementioned it is also required to take into account this dimension. Some handshapes, for example, might require a high amount of effort, such as a straight open up palm, or the orientation of the gesture might also be

influenced by effort; performing an open palm gesture facing away from the speaker requires a large amount of effort.

A different approach to the study of gesture form has recently been proposed by Boutet (2010). He proposes a kinesiological system for the study of gesture which takes into account a number of biomechanical and physiological features of the different components that are involved in the production of the gesture. Special attention is paid to the mechanical properties of the hand, the forearm, the arm and how the movement can be extended from one of these areas to the rest. Features such as the flexion, extension, supination, pronation, internal rotation or external rotation are taken into account when analysing gestures from this kinesiological perspective (see Boutet 2010 or Boutet, Morgenstern & Cienki, 2016 for more information on this system).

Form is indeed one of the core features that need to be taken into account when analysing a wide repertoire of gestures. However, as it has been previously stated, gestures are a window to the mind; together with language, gestures have an important role in cognition and conceptualisation, so it is not surprising that another of the core factors taken into account in gestures studies is the study of function. One of the ideas that addresses the pairing between form and function in gesture studies is the notion of *recurrent gestures*. Recurrent gestures are a type of gesture that show a "stable form meaning relation and can be distinguished from singular gestures or iconic and metaphoric gestures due to their conventional character" (Ladewig, 2014). While singular gestures have been described as 'spontaneous gesture realisations' (Müller, 2010), recurrent gestures have a particular form that is shared within the culture and they perform a particular semantic or pragmatic function (Ladewig, 2014). There are a number of gestures that show a recurrent pairing of form and function, for example the "palm up open hand" with a presentational function, (Kendon 2004), "holding away" for denial or negation (Bressem and Müller, 2013) and the "cyclic gesture" for repetition (Ladewig 2010, 2011). Bressem and Müller (2013) offer a repertoire of sixteen different recurrent gestures that have been found in German by describing both the form of the gestures as well as their semantic, pragmatic and illocutionary function.

### *4.2.2 Gesture function*

One of the earliest gesture classification systems whose focus of analysis was the function of the gesture was proposed by Ekman and Friesen (1969). Interestingly,

the authors never talked about gesture, but they referred to it as 'nonverbal behaviour', which was defined as "any movement or position of the face and/or body" (p. 49), although Kendon (2004) points out that they were rather referring to any movement or position which is relevant for communication. Ekman and Friesen (ibid.) focused on developing a typology based on three fundamental features of gestures: origin, usage and coding (see Ekman and Friesen, 1969: 49-63 for a more in-detail explanation of these features), which resulted in a five-category gesture typology which was divided as follows: (1) "emblems"; gestures which have a standardised meaning and usually have a verbal translation ('thumbs up' gesture means 'ok' or 'understood' is most western languages), (2) and "illustrators"; speech-accompanying gestures which aim to illustrate speech. Six different gestures can be found in this second category: (I) "batons", which give emphasis to a particular word or phrase; (II) "ideographs", which sketch the path or direction of thought; (III) "deictic movement", pointing gestures to an object; (IV)"spatial movement", which represents spatial relation; (V) "kinetographs", which depict a bodily action; and (VI) "pictographs", which draw a picture of the referent. The classification system follows with (3) "regulators", which attempt to direct the back-forth interaction during communication; (4) "affect displays", which include facial expressions; and finally (5) "adaptors", which encompass all the touching-of-self and object manipulation gestures (Bohle, 2014). One of the major problems with Ekman and Friesen (1969) typology was that it was not mutually-exclusive, with the analyst being the one to decide in which category some gestures should fit. The authors stated that "illustrators can include the use of an emblem" (p. 68) or that "affect displays can be emblems (...) affect displays and adaptors can serve as regulators" (p. 77). Nevertheless, terms like 'emblem' or 'illustrator' have become widely accepted by the gesture studies community (Kendon, 2004).

Another key gesture typology was proposed by McNeill's (1992), which has been widely adopted to account for the relation between language, gesture and thought. McNeill accounted for spontaneous gestures produced by speakers, since he suggests that this type of gesture displays the inner thoughts and ways of understanding different concepts. His classification refers to the symbolical interpretation of gestures in relation to the speech they co-occur with (Kendon, 2004: 99). It is important to highlight that, even though the emphasis is on function, this classification also includes form to an extent, for example, when talking about pointing gestures or beat gestures.

McNeill's approach (ibid) suggests a crucial distinction between two large sets of gestures: "imagistic" and "non-imagistic". The former refers to those gestures whose movements aim to reproduce the shape of an object, an action or a pattern of movement, as if they were trying to convey an image through bodily representation, while the latter include "deictic gestures" (pointing gestures) and rhythmic movements that serve as discourse markers (also known as "beat gestures") (Bohle, 2014).

Imagistic gestures can be further subdivided into two different categories: "iconic" and "metaphoric" gestures. An iconic gesture is semantically and pragmatically linked to speech; they complement spoken language rather than being redundant, but they often address physical objects or actions. On the other hand, metaphorical gestures also display an image, but in this case, they present more abstract ideas (Kendon, 2004). Metaphorical gestures also complement spoken language and are not redundant. The key between them is that iconic gestures resemble in some way their referent while in metaphoric gestures, they are not only more abstract but have a more mediated relation (in terms of form) with their referent.

In relation to non-imagistic gestures, McNeill distinguishes between deictic gestures, which imply "pointing movements usually performed by the index finger" (1992: p.80), and beat gestures which are "movements that do not present a discernible meaning" (ibid, p.80). This last assumption, however, has been recently challenged by Casasanto & Yap (2016), who suggested that in fact beat gestures could have sematic value, although further research is required to support this assumption. Non-imagistic gestures can be further subdivided into cohesive gestures, which are employed to join two thematically related but temporally separated segments of the discourse, and butterworths, which are gestures performed when the speaker is trying to recall a word or a verbal expression (McNeill, 1992. p.76-77).

Overall, McNeill's (1992) approach to gesture typology, which was highly influenced by Efron (1972) as well as Ekman and Friesen (1969), seems to be the most adequate for the purpose of the thesis. Additional classification systems have been suggested by other authors, such as Fricke (2007), who points out that the current classification systems do not provide an accurate differentiation of meaning and reference, of reference and signs vs. reference and non-signs as well as gestures and their discursive relation to speech. She takes a semiotic perspective which, even though

solves some of the weaknesses of previous systems by taking into account the speech that accompanies gestures, neglects the interaction or self-regulating functions of the gesture (Bohle, 2013). Additionally, Bavelas (1992) proposes another system which takes a dialogic perspective by distinguishing between emblems and conversational gestures, the latter overlapping with Ekman and Friesen's illustrators. Bavelas further argues that illustrators can be divided depending on whether they refer to the semantic content or if they refer to the addressee, distinguishing between "topic related gestures" and "interactive gestures". Interactive gestures are mainly formed of beat gestures, and they are rarely produced when there is no interlocutor, suggesting that interactive gestures are directed to the interlocutor. Finally, she also points out that a functional approach to gesture studies should be taken, since a single gesture can have several functions at once (Bohle, 2014).

### *4.3 Gesture and embodied experience: time metaphor*

#### *4.3.1 Gesture in English language*

Over the last few years, there has been an increasing interest in the study of gestures and their role in metaphorical thinking. Even though gestures can be employed to represent a wide range of metaphorical mappings, time metaphors have become a particularly interesting topic amongst researchers. Time involves what could be considered a primary metaphor (Grady, 1997); it is something experienced by all human beings regardless of their language or culture. Moreover, temporal conceptual mappings often rely on the spatial domain, which is precisely the domain in which gestures take place. This section reviews some of the most influential literature and empirical research that has investigated the topic of time conceptualisation and co-speech gestures in English language as well as cross-cultural studies.

Even though time conceptualisation is a key topic in metaphorical thinking, there has been little interest on the particular gesturing patterns of English speakers, as Cooperrider and Núñez (2009) point out. Cienki (1998) suggests that there is a tendency for English speakers to project events in imaginary timelines which go from left to right, although there is no mention of the sagittal axis. Cooperrider and Núñez (ibid) addressed this issue by looking at English speaker's temporal gesture patterns, and found that temporal gestures were influenced by culturally shared temporal conceptualisation processes. They also advise that more research on the (likely)

systematicity of gesture patterns with different verbal expressions about time could be highly beneficial to shed further light on this topic.

Another of the most influential experiments on time conceptualisation and co-speech gestures in English was conducted by Casasanto and Jasmin (2012). Here they investigated what the role of co-speech gestures was in temporal language when English speakers talk about time. Firstly, they conducted an experiment in which they asked participants to deliberately gesture about time; here they found out that English speakers spatially located time mirroring language: on the sagittal axis (*look forward to* or *look back*). A second experiment analysed the role of spontaneous co-speech gestures: here results were strikingly different and the transversal axis was used overwhelmingly more frequently.

One of the latest empirical studies that has addressed time conceptualisation in English is attributed to Walker and Cooperrider (2016). They looked at the spatial axis that was employed when speakers performed temporal co-speech gestures. Here they found that people systematically combine left-to-right and back-front spatial axes when gesturing about time; gestures tended to be doubly congruent with the language that speakers were using, since they located the past behind and on the left from the speaker and the future in front and on their right.

*4.3.2 Cross-cultural research in temporal gestures*

Cross-cultural research on time conceptualisation and gestures has been the field in which researchers have focused most of their efforts. One of the most well-known empirical studies was conducted by Núñez and Sweetser (2006), who looked at linguistic and gesture patterns in Aymara, an Amerindian language spoken in the Andean highlands of western Bolivia. They discovered that speakers of this language show a consistent linguistic and gesture pattern in which the future is located behind the ego, while the past is in front. These results are strikingly different from the time-space mapping found in western cultures which tend to locate the past behind the ego and the future in front of it.

Further research has been conducted in a wide range of languages, like Mandarin (Boroditsky, 2001), Yucatec Maya, (Le Guen. 2011), the Yupno of Papua New Guinea (Núñez et al, 2012), Arabic (de la Fuente et al, 2014) or Nheengatú (Floyd, 2016), suggesting the strong effect that the use of different spatial frames of reference as well

as cultural attitudes have on how speakers talk about gesture and conceptualise time. For example, Mandarin speakers are more likely to gesture about time using the vertical axis, possibly due to the high frequency of vertical space-time metaphors present in the language (Boroditsky, 2001). Le Guen (2011) point out that Yucatec Maya has a large linguistic repertoire to express deictic time (D-Time), but when trying to express sequential time (S-Time) the language is more limited. Moreover, speakers of Yucatec Maya do not represent a metaphorical timeline when gesturing about time, but only an opposition between the present or "current time" (located "here") and the "remote time", located at a certain distance from the speaker, without making a contrast between past and future. Furthermore, data from gesture and language suggest that time is not metaphorically located along an oriented line, but rather conceptualised as a succession of completed events with no spatial organisation. The authors suggest that the way time is expressed both linguistically and in gesture is closer to a non-liner, non-directional conceptualisation of time, perhaps closer to a cyclic conception of time.

Another example of differences in time conceptualisation across languages can be found in Yupno of Papua New Guinea. In this case, Núñez et al. (2012) performed an empirical study that suggested that Yupno speakers conceptualised time topographically, which mirrored the way they talked about space. Taking as a reference point a hill which plays an important role in Yupno culture, time was construed either downhill when talking about the past, uphill when talking about the future or it was placed at the location of the speaker when referring to the present. Gesture data indicated that, regardless of the direction speakers were facing, speakers' gestures would be congruent with this spatial temporal arrangement. This resulted in gestures performed in a wide range of locations (backwards, left, right) depending on the location of the hill.

Arabic presents another example of cross-cultural and cross-linguistic differences in gesture in time conceptualisation. In this case, de la Fuente et al (2014) compared temporal gestures performed by Spanish and Moroccan speakers. Both Spanish and Arabic employ sagittal oriented spatial temporal metaphors in language, but gesture evidence shows that, while Spanish speakers' gestures follow this pattern, Arabic speakers located the past ahead of them and the future behind them. The authors argue that, since both languages employ the same type of linguistic temporal metaphors, the difference in temporal thinking is attributed to the temporal focus of the speaker.

Spaniards tend to be more future focused, and hence they locate the future in front of them, while Moroccans are more past focused, giving more importance to older generation and traditional practices, and hence they located the past in front of them when gesturing about time.

The final example that addresses cross-cultural studies and time-related gestures analysis the role of celestial pointing of time in the Brazilian indigenous language Nheenganty. In his study, Floyd (2016) points out that speakers of this language tend to point at different positions of the sky along the east-west axis that corresponds with the location of the Sun in the sky in order to indicate different times of the day.

Lastly, it is important to mention that research is being conducted in different types of sign languages, addressing topics like iconicity in time-space mappings (Wilcox, 2002), the effect of gesture spatial paths on time representation (Selvik, 2006) or the differences between native and non-native speakers of sign language (Nilsson, 2016). Sign language is indeed a fascinating area that will certainly offer crucial information for the understanding of time-space mappings and their relation to the embodied spacialisation of time.

# Chapter 5

# Methodology

### *5.0 Introduction*

This section deals with the methodological aspects of the research. I start by briefly describing the different databases and tools that have been employed during the development of the thesis, namely Red Hen and the NewsScape library and software like the Edge search engine, CQPWeb or ELAN. The next section is devoted to a stage-by-stage explanation of the data collection procedure, focusing on the linguistic search criteria as well the clip selection criteria. This chapter finishes with an explanation of the gesture classification system and analysis scheme that I employ, including both quantitative and qualitative analysis of the data.

### *5.1 Database, software and tools*

A number of different tools were employed as part of the methodology. This section describes the different databases, software and other tools that were employed in order to collect the data as well as to analyse it. All the textual and multimodal data was extracted from the NewsScape archive. Additionally, three main tools were employed for the development of the thesis: (I) the Edge search engine for the observation of the data; (II) CQPWeb corpus software for performing the linguistic searches; (III) the multimodal annotator ELAN for conducting the analysis of different hits; and finally (IV) Excel for the storage and analysis of the dataset.

#### *5.1.1 Red Hen and NewsScape*

The main source of data drawn upon in this study is a multimodal corpus of television news and talk shows known as the NewsScape Library of International Television News (https://tvnews.sscnet.ucla.edu/edge/). This multimodal database was created in the University of California Los Angeles (UCLA) by the Distributed Little Red Hen Lab, a laboratory consortium for research into multimodal communication formed by a number of different worldwide universities. Red Hen aims to develop novel statistical, computational and technical methods for the study of human communication (see more on http://www.redhenlab.org/)

The NewsScape library is a digital collection of more than 350,000 television news programs with more than 250,000 hours of television in different languages

including English, Spanish, German, French and Arabic, among other languages. The result is a multimodal database with more than a 2 billion word textual database in the form of subtitled television news programs that have been recorded since 2004 up until the present day. Data is collected from a wide range of channels such as AlJazeera, BBC, CNBC and Fox News and it only includes television news, excluding series, films or other television content that is formed of artificial speech. Hence, the NewsScape Library offers a massive database of spoken television language which allows for large-scale data searches. More precisely, gestures studies may find here an invaluable source of information, since one of the main constraints in gesture research has to do with the limited number of participants and data that can be collected from laboratory research. This database provides a virtually unlimited amount of speakers; it offers a massive amount of information which is suitable for quantitative and/or qualitative analysis of "big data".

However, the most important feature of this database is its multimodal nature. The NewsScape library allows researchers to look for a particular linguistic expression and obtain textual information from the subtitle dataset as well as a video clip of the exact moment at which the expression was uttered. The inclusion of video and audio gives researchers the chance to address a large number of factors that have an influence on the nature of communication, such as intonation, pitch, interaction with the surrounding context or, the focus of this thesis, co-speech gestures.

Another advantage of this database is that the different individuals that contribute to the formation of the corpus are not aware of this data being used for empirical research purposes, avoiding misleading information that could emerge from the observer's paradox in laboratory conditions. Data is collected from news and talk shows, providing high-quality quasi natural data; television settings might have an influence on the behaviour of the speakers, although most of the conversations found in the corpus aim to be informal and natural, being quite close to spontaneous speech. Television data also permits research to be conducted in other non-conventional communication settings, like public speeches or political meetings, which may show some variation from regular communicative situations. The data is protected by the US Copyright laws, which allow the recording and storage of the shows for research purposes. Furthermore, the data is only available to members of the research group, and the materials can only be accessed through and account on the UCLA.

However, one of the disadvantages of working with a television corpus is that there are many occasions in which the speaker is not present on the screen. Even though the desired linguistic expressions are uttered on speech, there are many occasions in which it is in fact a voice over, since an image or video is being shown and the speaker cannot be seen. Another of the problems with this type of data is that there are moments in which, even though the speaker was visible on the screen, the hands cannot be seen due to the angle of the camera or some object hindering the view. Finally, there were cases in which the same clip is found several times since it has been used in different programs and cases where video/audio is not present on the clip. This means that any data retrieved must be 'cleaned' before it can be reliably used.

### *5.1.2 Edge search engine*

All the information provided by the corpus is presented to the user through the Edge search engine, which has customised tools for qualitative and quantitative analysis, as well as allowing the researcher to use their own annotation scheme for language and multimodal information and presenting textual and audio-visual data.

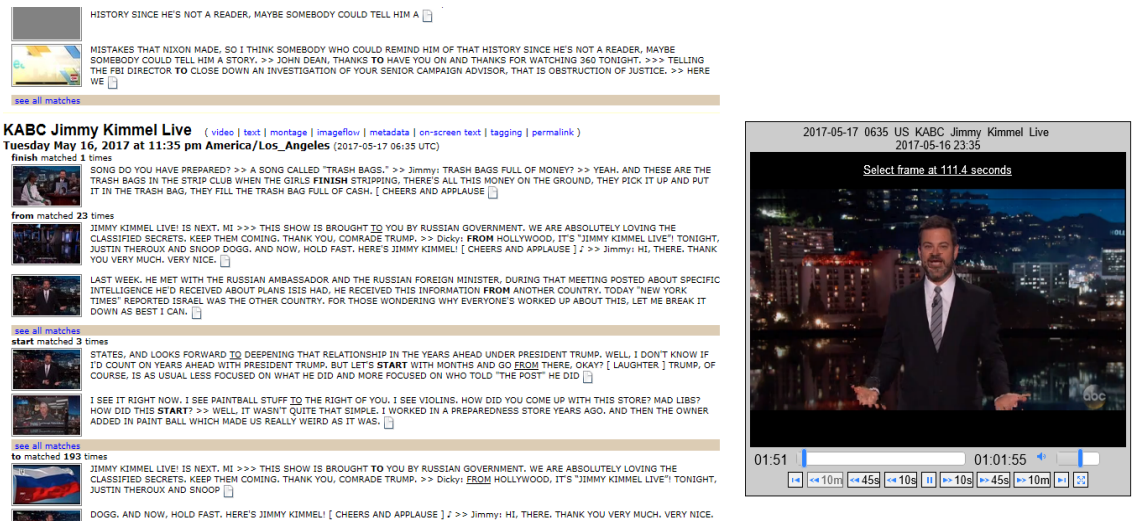



*Figure 5.1 Edge search engine.*

This search engine allows for a quick and easy observation of textual and multimodal information (Figure 5.1). A list of the different hits is displayed on the left, providing the linguistic context of the different search terms, which are highlighted in bold. On the right, the researcher can easily watch a clip of the moment at which the speaker utters the linguistic expression. The reliability of text-video synchronisation is very high thanks to a force-alignment procedure, which involves the automatic transcription of an audio speech segment and the localisation of particular words within

this segment. Force alignment brings the text into synchronisation with the video and the audio.

However, one of the disadvantages of the Edge search engine is that the array of options for performing corpus searches is limited, allowing only for simple very linguistic searches (Figure 5.2). Moreover, the information cannot be automatically recorded; the researcher is required to manually note down features such as day and time, program and linguistic context of each of the hits. In order to solve this problem, the corpus was recently modified so that it could be accessed through CQPWeb, allowing for a number of improvements that will be discussed in the next section.

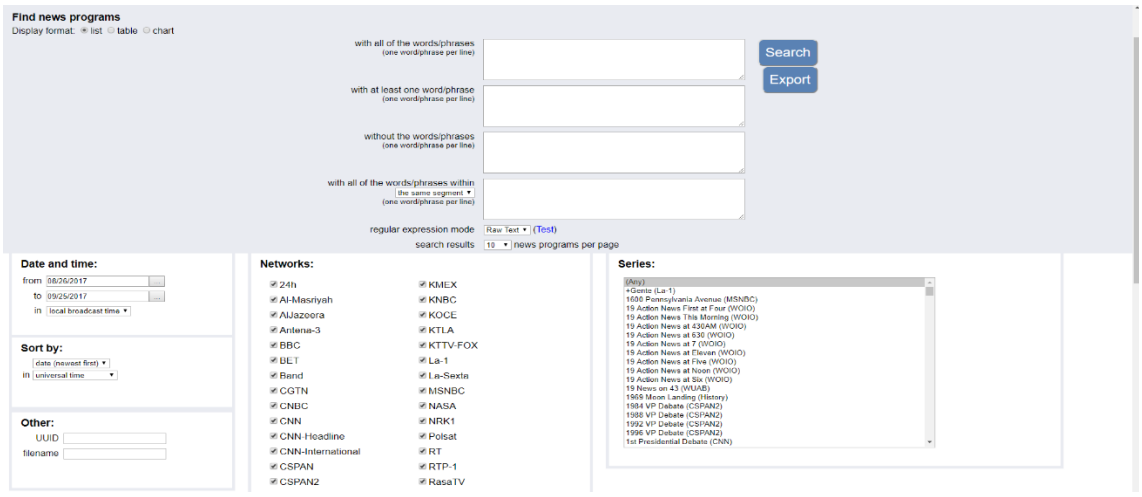



*Figure 5.2 Edge search engine linguistic searches.*

### *5.1.3 CQPWeb*

One of the main benefits of accessing the NewsScape database through CQPWeb is that this software allows for more complex syntax searches as well as corpus parsing. This version of the corpus offers some other additional features not present in the Edge search engine version. Firstly, there is a 95% reduction of 'noise' in the corpus: non-speech related words that appear in the corpus such as indicators of non-speech related sounds ({APPLAUSE}) or terminology that addresses the speaker (*REPORTER, JIMMY KIMMEL)* is omitted from the corpus searches. Secondly, this version allows for a quick and automatic way of downloading metadata such as the date, time or program in which the different linguistic expressions were uttered, something that had to be manually annotated in the Edge search engine. Finally, this access portal also provides one section that includes a link for the observation of the clip (through the Edge engine), which can be automatically downloaded alongside the rest of the aforementioned features (figure 5.3).

|< << >> >| Show Page: 1 Line View Show in random order New query Go!

| No | Filename | Solution 1 to 50 | Page 1 / 2 | |
|---|---|---|---|---|
| 1 | t__7a771c30_4f11_11e2_8f59_001517add6fb | we want to warn you about everything in this hour . | from beginning to end | . do not tr |
| 2 | t__f65fa9f6_4deb_11e2_8b5f_001517add4ae | girl was amazing . and her father too . it hat me | from beginning to end | . what abo |
| 3 | t__a8af53dc_4be3_11e2_9bca_001517add720 | did not stop . took us five months to build this home | from beginning to end | . working |
| 4 | t__c5583820_4696_11e2_8e66_001517add4ae | IT IS -- THAT MOVIE IS BULLET PROOF . it is funny | from beginning to end | and it doe |
| 5 | t__f8b2bcea_4641_11e2_ba62_001517add6fb | it looks like it took maybe 20 or 30 minutes at most | from beginning to end | from the t |
| 6 | t__5ee313e6_3825_11e2_b39e_001517add60e | almost like stifling and so i wrote all these verses about everything | from beginning to end | of this rel |
| 7 | t__0972c8d2_36b5_11e2_8e2a_001517add60e | i could not do anything to help him . the rumba moving | from beginning to end | , spinning |
| 8 | t__35d02e48_35e7_11e2_b3f3_001517add60e | my my boyfriend was really addicted to drugs . the rumba moving | from beginning to end | , spinning |
| 9 | t__78b0087c_3511_11e2_b051_001517add60e | side . we had 16 tiger swan personnel involved in this operation | from beginning to end | . so we ha |
| 10 | t__927b6f5c_353f_11e2_b365_001517add720 | little bit , but i was a grown man before watched it | from beginning to end | , and now |
| 11 | t__927b6f5c_353f_11e2_b365_001517add720 | by myself . i sit and watch alex haley 's epic work | from beginning to end | , every ye |
| 12 | t__0d9eda5e_3548_11e2_b48f_001517add6fb | in those five years . i was not working on casual vacancy | from beginning to end | for five ye |
| 13 | t__7a307076_305a_11e2_8ac1_001517add6fb | york times , " he was there covering the monica lewinsky affair | from beginning to end | and ended |
| 14 | t__f7529de2_2fcc_11e2_afea_001517add6f3 | IT IS -- THAT MOVIE IS BULLET PROOF . it is funny | from beginning to end | and it doe |
| 15 | t__6c224cf4_3017_11e2_a5ad_001517add4ae | lawmakers say they watch this video . they saw the attack | from beginning to end | . and it ev |
| 16 | t__8c483044_2f80_11e2_bf7a_001517add6fb | drone video . a composite put together so they could watch it | from beginning to end | . that actu |

*Figure 5.3 CQPWeb.*

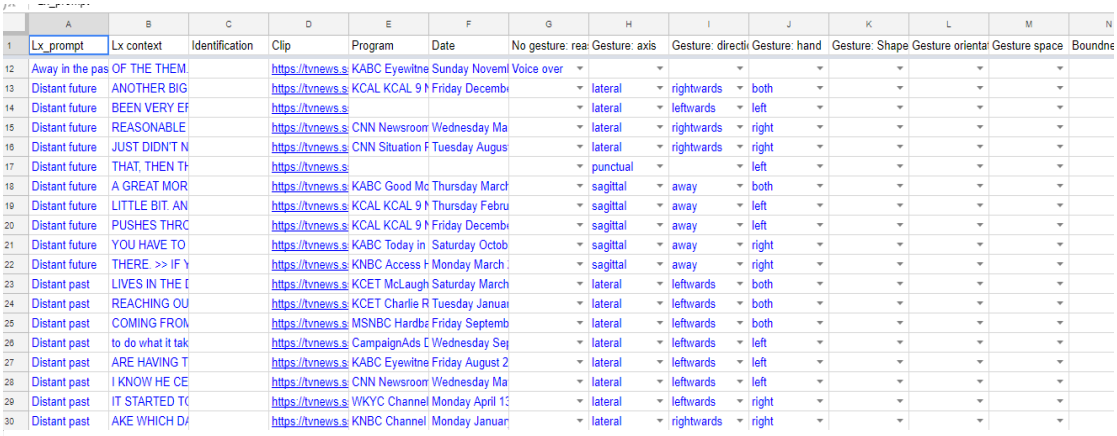
### *5.1.4 ELAN*

ELAN is a specialised software for the analysis of multimodal data that involves audio and video resources. It allows the creation of several annotations on multiple layers of analysis that include both text and multimodal information as well as a more precise manipulation of the different video files.

Video clips that contain a co-speech gesture can be downloaded from the NewsScape library and, together with their linguistic context, added to ELAN in order to perform a more accurate analysis of the gesture features. A layer of analysis can be added for each of the gestural features prior to the observation and annotation of each of the clips.

## *5.2 Data storage*

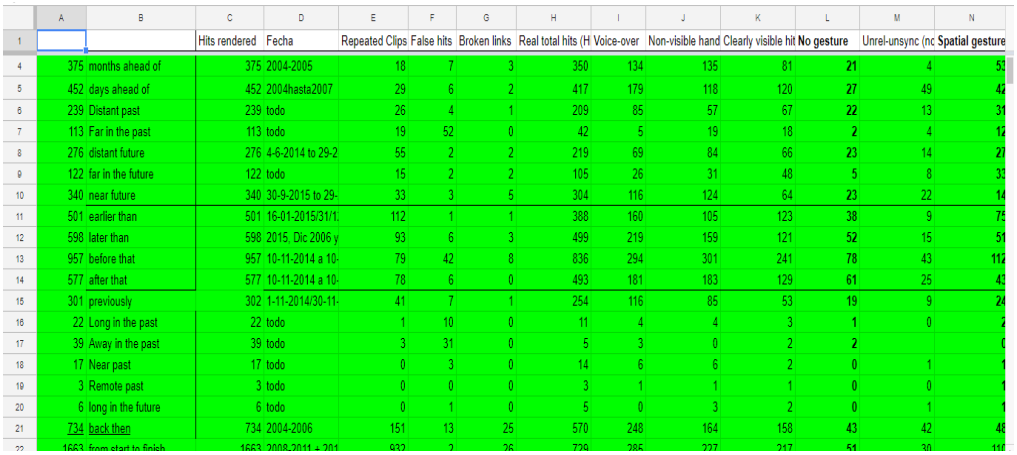
While, ELAN offers a more precise tool for data analysis, is does not offer an optimal way of storing all the data for later analysis. In order to access in a quick and easy way to the data, Excel was employed for the storage and organisation of the data. Two different files were employed for storing the data. The first one (Figure 5.4) contained all the hits that were obtained during the data search, whether they contained a co-speech gesture or not. Data validations tool were used in the tagging of all the gesture features, allowing for a quick and clear organisation and recognition of the different categories.



| | A | B | C | D | E | F | G | H | I | J | K | L | M | N |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Lx_prompt | Lx context | Identification | Clip | Program | Date | No gesture: rea | Gesture: axis | Gesture: directi | Gesture: hand | Gesture: Shape | Gesture orienta | Gesture space | Boundne |
| 12 | Away in the pas | OF THE THEM. | | https://tvnews.s | KABC Eyewitne | Sunday Novem | Voice over | | | | | | | |
| 13 | Distant future | ANOTHER BIG | | https://tvnews.s | KCAL KCAL 9 N | Friday Decembe | | lateral | rightwards | both | | | | |
| 14 | Distant future | BEEN VERY EF | | https://tvnews.s | | | | lateral | leftwards | left | | | | |
| 15 | Distant future | REASONABLE | | https://tvnews.s | CNN Newsroom | Wednesday Ma | | lateral | rightwards | right | | | | |
| 16 | Distant future | JUST DIDN'T N | | https://tvnews.s | CNN Situation F | Tuesday Augus | | lateral | rightwards | right | | | | |
| 17 | Distant future | THAT, THEN TH | | https://tvnews.s | | | | punctual | | left | | | | |
| 18 | Distant future | A GREAT MOR | | https://tvnews.s | KABC Good Mo | Thursday March | | sagittal | away | both | | | | |
| 19 | Distant future | LITTLE BIT. AN | | https://tvnews.s | KCAL KCAL 9 N | Thursday Febru | | sagittal | away | left | | | | |
| 20 | Distant future | PUSHES THRO | | https://tvnews.s | KCAL KCAL 9 N | Friday Decembe | | sagittal | away | left | | | | |
| 21 | Distant future | YOU HAVE TO | | https://tvnews.s | KABC Today in | Saturday Octob | | sagittal | away | right | | | | |
| 22 | Distant future | THERE. >> IF Y | | https://tvnews.s | KNBC Access H | Monday March | | sagittal | away | right | | | | |
| 23 | Distant past | LIVES IN THE D | | https://tvnews.s | KCET McLaugh | Saturday March | | lateral | leftwards | both | | | | |
| 24 | Distant past | REACHING OU | | https://tvnews.s | KCET Charlie R | Tuesday Januar | | lateral | leftwards | both | | | | |
| 25 | Distant past | COMING FRON | | https://tvnews.s | MSNBC Hardba | Friday Septemb | | lateral | leftwards | both | | | | |
| 26 | Distant past | to do what it tak | | https://tvnews.s | CampaignAds D | Wednesday Sep | | lateral | leftwards | left | | | | |
| 27 | Distant past | ARE HAVING T | | https://tvnews.s | KABC Eyewitne | Friday August 2 | | lateral | leftwards | left | | | | |
| 28 | Distant past | I KNOW HE CE | | https://tvnews.s | CNN Newsroom | Wednesday Ma | | lateral | leftwards | left | | | | |
| 29 | Distant past | IT STARTED TO | | https://tvnews.s | WKYC Channel | Monday April 13 | | lateral | leftwards | right | | | | |
| 30 | Distant past | AKE WHICH DA | | https://tvnews.s | KNBC Channel | Monday Januar | | lateral | rightwards | right | | | | |

*Figure 5.4 Storage of qualitative data.*

Additionally, another document was employed to store the data in a quantitative way. Here I stored all the data from the qualitative analysis in a quantitative way. This file contained information about the total number of hits or the number of hits that were categorised as noise of each as well other qualitative features such as axis or direction (Figure 5.5).



| | A | B | C | D | E | F | G | H | I | J | K | L | M | N |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | | Hits rendered | Fecha | Repeated Clips | False hits | Broken links | Real total hits (H | Voice-over | Non-visible hand | Clearly visible hit | **No gesture** | Unrel-unsync (nc | **Spatial gesture** |
| 4 | 375 | months ahead of | 375 | 2004-2005 | 18 | 7 | 3 | 350 | 134 | 135 | 81 | **21** | 4 | 5 |
| 5 | 452 | days ahead of | 452 | 2004hasta2007 | 29 | 6 | 2 | 417 | 179 | 118 | 120 | **27** | 49 | 4 |
| 6 | 239 | Distant past | 239 | todo | 26 | 4 | 1 | 209 | 85 | 57 | 67 | **22** | 13 | 3 |
| 7 | 113 | Far in the past | 113 | todo | 19 | 52 | 0 | 42 | 5 | 19 | 18 | **2** | 4 | 1 |
| 8 | 276 | distant future | 276 | 4-6-2014 to 29-2 | 55 | 2 | 2 | 219 | 69 | 84 | 66 | **23** | 14 | 2 |
| 9 | 122 | far in the future | 122 | todo | 15 | 2 | 2 | 105 | 26 | 31 | 48 | **5** | 8 | 3 |
| 10 | 340 | near future | 340 | 30-9-2015 to 29- | 33 | 3 | 5 | 304 | 116 | 124 | 64 | **23** | 22 | 1 |
| 11 | 501 | earlier than | 501 | 16-01-2015/31/1 | 112 | 1 | 1 | 388 | 160 | 105 | 123 | **38** | 9 | 7 |
| 12 | 598 | later than | 598 | 2015, Dic 2006 y | 93 | 6 | 3 | 499 | 219 | 159 | 121 | **52** | 15 | 5 |
| 13 | 957 | before that | 957 | 10-11-2014 a 10- | 79 | 42 | 8 | 836 | 294 | 301 | 241 | **78** | 43 | 11 |
| 14 | 577 | after that | 577 | 10-11-2014 a 10- | 78 | 6 | 0 | 493 | 181 | 183 | 129 | **61** | 25 | 4 |
| 15 | 301 | previously | 302 | 1-11-2014/30-11- | 41 | 7 | 1 | 254 | 116 | 85 | 53 | **19** | 9 | 2 |
| 16 | 22 | Long in the past | 22 | todo | 1 | 10 | 0 | 11 | 4 | 4 | 3 | **1** | 0 | |
| 17 | 39 | Away in the past | 39 | todo | 3 | 31 | 0 | 5 | 3 | 0 | 2 | **2** | | |
| 18 | 17 | Near past | 17 | todo | 0 | 3 | 0 | 14 | 6 | 6 | 2 | **0** | 1 | |
| 19 | 3 | Remote past | 3 | todo | 0 | 0 | 0 | 3 | 1 | 1 | 1 | **0** | 0 | |
| 20 | 6 | long in the future | 6 | todo | 0 | 1 | 0 | 5 | 0 | 3 | 2 | **0** | 1 | |
| 21 | 734 | back then | 734 | 2004-2006 | 151 | 13 | 25 | 570 | 248 | 164 | 158 | **43** | 42 | 4 |

*Figure 5.5 Storage of quantitative data.*

### *5.3 Data collection*

Data was collected through two different stages that involved the creation of criteria terms and the categorisation of all the resulting hits depending the type of information that was extracted from them. A number of undergraduate students from University of Murcia were personally trained so that they could assist me with the first stage of the data collection (see Statement of Authorship and Responsibility). Students were instructed to detect whether or not the clip showed the speaker who uttered the searched linguistic expressions (see section 5.3.2). Their collaboration helped speed up the process of distinguishing relevant data from noise, since a high amount of the clips collected belonged to the latter.

*5.3.1 Stage one: linguistic search criteria*

The first step in the data collection procedure was to establish a criterion in order to select the different search term items that would be employed for the linguistic search. Since the main aim of this research was to address different aspects of time conceptualisation that have been previously identified in linguistics through the analysis of co-speech gestures, the main interest laid in semantic distinctions. Thus, this research focused on obtaining a reasonable data sample of gestures that co-occurred with temporal linguistic expressions that employed different temporal constructs. The first feature that was taken into account was whether or not the linguistic expression employed a metaphor to refer to time or not. Due to space constrains, I exclusively focus on space-time metaphorical linguistic expressions (*back in the past*) as well as non-metaphorical temporal linguistic expressions (*later than*). Other types of linguistic expression that involve other conceptual domains (TIME IS A RESOURCE, TIME IS A CHANGER) are not accounted for in this research.

This distinction allowed me to distinguish between two types of temporal linguistic expressions: 1) non-spatial temporal expressions that exclusively used the temporal domain to talk about time and 2) spatial linguistic expressions that employed the spatial domain in order to talk about time.

Once this distinction was made, I further subdivided the expressions that employed the spatial domain in two different groups depending on what facets of the spatial domain were being mapped on the expression. The first group was formed by a number of expressions that explicitly stated the direction of time. Spatial directionality was the defining feature of this linguistic expression category.

The second group was composed of spatial expressions that did not explicitly indicate the spatial direction. These expressions made no reference to the directionality of the expression but they gave information about the distance of the temporal period.

Once the distinction between these three different groups was made, I added one more factor, which was the temporal period the expressions were referring to. I intended to include an even number of future and past-related expressions in each of the three different categories.

When elaborating a close list of particular linguistic expressions that dealt with these semantics features, grammar was also taken into consideration. Due to the

different ways of expressing temporality in English language as well as the overwhelming size of the corpus, linguistic searches were restricted to simple grammatical patterns involving temporal adverbs (*previously*), prepositions (*after*) and prepositional phrases (*in the near future, months ahead of)*. This research does not address other types of temporal expressions involving verbal forms, temporal subjects or other more complex grammatical temporal expressions, although further investigation of this topic would certainly be of interest.

Since no systematic study has been done on the relationship between temporal expressions and co-speech gestures, different temporal linguistic expressions that fulfilled the aforementioned semantic and grammatical requirements were collected until a satisfactory amount of gestures was reached in each of the categories. The collection of the co-speech gestures was focused in the three different categories that were aforementioned: non-spatial, spatial direction and spatial non-directional. The gesture threshold that was established was around 150 examples for each of the categories. This amount seemed like a reasonable data size for the purpose of the thesis; it would allow to perform quantitative analysis when looking at patterns in gesture realisation while it would also allow for the qualitative analysis of some of the gestures. Moreover, I tried to maintained a balance on past-related and future related gestures, with a relatively similar number of gestures in each category.

The final list of linguistic expressions comprises the following: (1) non-spatial language: *earlier than, previously, later than*; (2) spatial directional language: *back in those days, months ahead of, time ahead of* and (3) spatial non-directional language *distant past, far in the past, long in the past, near past, remote past, distant future, far in the future, long in the future, near future* (see Table 5.1).

| Category | | Expression | Search time span on NewsScape |
|---|---|---|---|
| Non-spatial | Past | Earlier than | 16/01/2015-31/12/2015 |
| | | Previously | 1/11/2014-30/11/2014 |
| | Future | Later than | 1/01/2015-31/12/2015 |
| Spatial directional | Past | Back in those days. | 01/01/2004-28/02/2016 |
| | Future | Months ahead of | 01/01/2004-01/01/2005 |
| | | Time ahead of | 01/01/2004-28/02/2016 |
| Spatial non-directional | Past | Distant past | 01/01/2004-28/02/2016 |
| | | Far in the past | 01/01/2004-28/02/2016 |
| | | Long in the past | 01/01/2004-28/02/2016 |
| | | Near past | 01/01/2004-28/02/2016 |
| | | Remote past | 01/01/2004-28/02/2016 |
| | Future | Distant future | 04/06/2014-29/02/2016 |
| | | Far in the future | 01/01/2004-28/02/2016 |
| | | Long in the future | 01/01/2004-28/02/2016 |
| | | Near future | 30/09/2015-29/02/2016 |

*Table 5.1 Inventory of linguistic search items*

Some expressions such as *earlier than* or *back in those days* were very common and offered a relatively high amount of co-speech gestures, so the search was limited so some of these expressions and also a limited time frame in the corpus. On the other hand, expressions that formed the non-directional language category rendered a lower amount of hits, so a larger number of expressions as well as a longer search time frame was required in order to obtain a similar co-speech gesture number than in the other categories (see section 5.5 for further details about the number of hits).

This list of temporal expressions does not aim to be an extensive list including all the possible expressions that could be part of these categories. Other expressions could have been added but maintaining a reasonable data sample was key in the correct development of the thesis. In fact, one of the *problems* that can be attributed to the NewsScape library is that it offers a vast amount of data which usually requires to select at some degree the hits to be analysed. The threshold of 150 gesture per category was always taken into account when compiling the database.

*5.3.2 Stage two: clip selection criteria*

Once the linguistic items were chosen, searches were made in the NewsScape library. Each of the hits obtained from this search was individually categorised depending on the type of information that was extracted from them. Firstly, features related to the organisation and identification of the data set were annotated. Each entry was assigned a number, a hyperlink to the NewsScape library to facilitate its retrieval for viewing, the date and the name of the program as well as some textual context (10 words before and after the search word(s)).

Secondly, each clip was individually analysed and classified to distinguish between noise data (categories 1, 2 and 3), data in which the speaker was not fully visible (4 and 5) and clips in which the speaker was clearly visible (6, 7 and 8). Trained undergraduate students assisted me on this task checked and coordinated by me at all times[6]. The categories in which the clips divided correspond to:

(1) *Repeated clip*: cases in which the same clip appeared more than once; this happened quite frequently with expressions that were used in advertisements or videos that were broadcasted several times on television. (2) *False hit*: cases in which the close caption was different from the real speech, that is, the speech did not correspond to the textual data. Additionally, this category included cases in which the search item was part of two different sentences, i.e. *Jenner has not been charged with any crime so* ***far. In the past****, she has pledged…* (KABC Eyewitness News, 4pm, Thursday July 9, 2015 at 4:00pm). (3) *Broken link*: cases in which problems in the recording of the program in the database did not allow its observation. This includes the absence of audio/video, the appearance of interferences in the video (green, black and pink lines) or the loss of signal during the recording. (4) *Voice-over*: there was a large number of hits in which the speaker was not visible; a voice can be heard uttering the relevant expression but the image shows something else, i.e. a video, image, the audience…. (5) *Non-visible hands*: the speaker is partially visible but the camera angle did not allow for a clear observation of their hand(s). There were many cases in which it could be intuited that the speaker was performing a gesture from the arm or shoulder movement, but only cases in which the hands were clearly visible were accepted as valid clips. If the hand(s)

[6] Two researchers from Universidad de Murcia (Javier Valenzuela and Cristobal Págan) as well as a group of four undergraduate students assisted me to classify the dataset between noise data and valid data. Once I obtained the valid gesture data, I individually performed its analysis.

were not visible, the clip was then added to this category. (6) *No gesture*: this category included all the cases in which the hand(s) of the speakers were clearly visible but no co-speech gesture was performed. (7) *Unrelated/out of synchrony gestures*: here all the gestures that were not in synchrony with the linguistic expression as well as the gestures that were considered not to be metaphorical were included. There was a number of cases in which beat gestures (repeated realisations of vertical strokes) were performed while the temporal linguistic expression was being uttered; the speaker performed this gesture frequently before, during and after the linguistic expression. (8) *Relevant co-speech gesture*: this last category included all the temporal gestures that were relevant to the analysis and they were separated into a different database for further analysis.

### *5.4 Data analysis*

All the gestures that fell under the category of *relevant co-speech gesture* underwent a fine-grained analysis both qualitative and quantitative. The qualitative analysis focused on the description of a range of gestural kinetic features of the gestures in order to further classify them. Once the qualitative analysis was done, data was analysed again quantitively by using descriptive statistics in order to discover patterns and tendencies in the data.

#### *5.4.1 Gesture qualitative analysis*

The analysis of the gesture form is based on Bressem's proposal for coverbal gesture annotation (2013), although some modifications were made. This paper contains a very detailed description of a large number of features that can be annotated in gestures, including the shape of the hand, shape of the fingers, levels of stretchiness/bentness of the fingers, different configuration of the interaction amongst the different fingers of the hand. I will only focus on five of them from the wide array of form gesture features that are proposed in this paper. Since the main objective of this research is to analyse general patterns and tendencies in co-verbal gesture forms through big data, just the most salient and informative features will be taken into account. As it has been previously argued, the object of study in this thesis is not the gesture per se, but rather the gesture is employed as a method to study the different cognitive processes and patterns that may be taking place when a speaker makes a time-related gesture.

The categories which have been taken from this proposal of coverbal annotation are direction of movement, separated here in two different features, (1) *axis;* (2)

*gesturing hand*; (3) *direction of gesture*; (4) *shape and (5) orientation;* (6) *gesture space;* Categories (1) and (2) directly address some of the most discussed topics concerning time conceptualisation in language, gesture and other modalities (see section 2.3), although this will be the first time that this type of research is conveyed through the analysis of "big data". Category (3), (4) and 5 address more formal gesture features in order to create a better picture of how the gesture was performed as well as distinguishing them from other types of non-temporal gesture (prosodic-related gestures, presentational gestures…). Category (6) addresses the distance that the co-speech gesture covers when being produced.

One more category, not related to gesture form, was added to the data analysis. Category (7) addresses *congruency*, more particularly the presence or absence of congruency between the temporal linguistic expression and the co-speech gesture in terms of axis and direction of the movement.

Next, I will describe the different tags that were used in each of the categories as well as some additional information about each of them:

Category (1) dealt with the spatial localisation of the gesture in the different spatial axes: *transversal*, *sagittal* or *vertical*. Some of the gestures that were analysed presented a clear axial localisation, while there were other cases in which the gesture could be argued to belong to two different axes (i.e. a diagonal gesture). Gesture whose axis was not easy to detect were labelled as *unclear* and were separated from the database for further discussion. Additionally, a new category labelled as "no axis" was added to the classification in order to account for a set of gestures that were not performed along an axis in particular but rather at one point in space with no clear spatial movement. Even though these gestures presented spatial motion, it was observed that the speaker pointed to a particular point on space rather than performing a gesture to a particular direction.

Category (2) indicated which gesturing hand(s) was used: *left, right* or *both*. This category also accounted for the type of movement that was performed when both hands moved together. More precisely, three different types of two-hand movements were taken into account: *one hand moves and the other stays, both hands move together* or *each hand signals a point*.

Category (3) included all the different directions in which the gesture could be performed: transversal gestures could be *rightward*, *leftward*, *inward* (putting hands together) or *outward* (putting hands apart). Vertical gestures were analysed as *upward* or *downward* and finally sagittal gestures could have *away-from-the-body* or *towards-the-body* realisations. Gestures that did not employ an axis were categorised as static. Additionally, there were some cases in which the gestures presented a circular movement, and a clear direction could not be established. In these cases, the gesture direction was labelled as *circular*. Note that when talking about gesture direction the emphasis relies on the stroke phase. The direction of the preparation or retraction stages was not taken into account.

The next category (4) concerned both the shape and the palm orientation. The shape of the hand was categorised as *flat hand*, *fist*, *single finger* or *combination of fingers*. No additional information of the combination of the fingers or their levels of stretchiness/bentness was annotated. Additionally, the orientation of the gesture (5) was annotated; gestures generally presented an *up*, *down*, *front* or *back* orientation. However, there were cases in which the orientation of the gesture was *lateral* with the hand facing right or left (or hands facing each other in the case of gestures performed with both hands).

Category (6) addressed the gestural space which is covered when the speaker produces a co-speech gesture. This included gesture performed in the *center*, that is, when located between both the shoulders, in the *periphery*, when they are performed in the space located slightly further from the shoulders and finally the *extreme periphery*, capturing gestures that are performed further from the *periphery*. However, the analysis of the gesture space when the gesture is not performed in the lateral axis is more complicated. There is no mention of how to address the gestures performed in the sagittal or vertical axes in Bressem's proposal for coverbal annotation, and the body cannot be taken as a reference in order to establish the location of the gesture. I did not take into account gestures that were performed in the vertical axis when looking at gesture space since their number was really low. However, there was a high number of sagittal gestures whose gesture space analysis could not be done following this framework. In order to be as objective as possible, I took into account different anatomical features when analysing the gesture space in the sagittal axis, namely the hand or wrist movement and the level of extension of the arm.

When a sagittal gesture was performed forwards, three different locations could be distinguished. Firstly, the gesture could be located in the centre if the gesture was performed just by using the hands and the wrist, with no motion in the arms, in the surrounding gesture space of the chest. (Figure 5.6).

*Figure 5.6 Sagittal forward gesture located in the centre.*

A sagittal forward gesture could also be located in the external periphery. In this case, I took into account the level of extension and movement of the arm. A gesture would be located in the external periphery if the speaker extended the forearm in order to perform the gesture (Figure 5.7)


*Figure 5.7 Sagittal forward gesture located in the periphery*

A sagittal forward gesture could also be located in the external periphery if the speaker fully extended their arm. This sometimes could be accompanied by the movement of the shoulder or the body (Figure 5.8)


*Figure 5.8 Sagittal forward gesture located in the external periphery.*

Sagittal gestures that were performed backwards turned up to be a problematic case. It is difficult to establish how far from the body a gesture that is performed backwards is, since the backward motion is anatomically limited. Hence, I only established two different gesture locations in backwards gestures: they could be located in the centre or the external periphery.

First, the gesture could be located in the centre if the backward movement was again performed with a small movement of the hands or the wrists in the gesture space surrounding the chest (Figure 5.9).

*Figure 5.9 Sagittal backward gesture located in the centre.*

The rest of the gestures sagittal backward gestures that were not located in the centre presented a similar motion. The speakers generally rise their arm backwards, with the hand at the same level as the shoulder and performed some type of movement with the hand to signal behind them (Figure 5.10).


*Figure 5.10 Backwards sagittal located in the external periphery.*

It is difficult to establish a middle ground as the one offered by gestures located in the periphery in lateral or sagittal forward gestures. Even though this type of gesture is located close from the body, they require a high amount of effort. I suggest then that, even though they are physically close from the body, they might be alluding to an extension of the gesture that locates further back from the speaker.

Category (7) concerned the gesture-time congruency. The canonical direction of time in English language spatially locates past times on the left, behind the speaker or above them, while the future is located on the right, in front of the speaker or below them. English language employs mostly sagittal spatial temporal metaphors (*look back, look forward to*) with vertical language being almost non-existent and lateral metaphors not being used at all. This feature compares the localisation of the gesture to the canonical direction of time and labels the gesture as *congruent* if both coincide or *incongruent* if there is not agreement. Gestures that were performed in the transversal axis with an *inward* or *outward* direction were not considered to be either congruent or incongruent.

The following table (5.2) aims to summarise the different analytical categories that were employed when analysing all the co-speech gestures:

| Analysis features | Axis | Hand | Direction | Shape | Orientation | Gesture Space | Congruency |
|---|---|---|---|---|---|---|---|
| | Lateral | Right<br>Left | Rightwards<br>Leftwards | Flat hand<br>First<br>Single finger<br>Combination of fingers | Up<br>Down<br>Front<br>Back<br>Lateral | Centre<br>Periphery<br>External periphery | Congruent<br>Incongruent |
| | | Both | Rightwards<br>Leftwards<br>Inwards<br>Outwards | | | | |
| | Vertical | Left<br>Right<br>Both | Upwards<br>Downwards | Flat hand<br>First<br>Single finger<br>Combination of fingers | Up<br>Down<br>Front<br>Back<br>Lateral | Centre<br>Periphery<br>External periphery | Congruent<br>Incongruent |
| | Sagittal | Left<br>Right<br>Both | Towards the body<br>Away from the body | Flat hand<br>First<br>Single finger<br>Combination of fingers | Up<br>Down<br>Front<br>Back<br>Lateral | Centre<br>(Periphery)<br>External periphery | Congruent<br>Incongruent |
| | Punctual | Left<br>Right<br>Both | No direction | Flat hand<br>First<br>Single finger<br>Combination of fingers | Up<br>Down<br>Front<br>Back<br>Lateral | Centre<br>Periphery<br>External periphery | Not applicable |

*Table 5.2 Gesture analytical features*

The validity of the gesture annotation scheme was also tasted through the Cohen's kappa inter-coder agreement test. A total of 75 gestures (which conform a 18% of the dataset) were randomly selected and qualitatively analysed through the proposed gesture scheme. The sample was composed of 25 cases of each of the three types of temporal language that were analysed (non-spatial, spatial directional and spatial non-directional). When annotating gestures in terms of axis, the test indicated a substantial agreement between coders ($\kappa$=0.79). Interestingly, most of the disagreement occurred in some gestures that the original coder annotated as lateral, while the secondary coder considered them to be performed in the sagittal axis.

The second feature that was tested was the direction of the gesture. This time a slightly higher score was obtained ($\kappa$=0.81), which again indicated a substantial agreement between coders. Most of the disagreement in this area was also related to the fact that the first coder considered some gesture to be located in the lateral axis, while the second one considered them to be in the sagittal. The disagreement in the axis also produced a disagreement in the direction.

Lastly, the third feature that was tested was gesture distance. This time, only 25 gestures that co-occurred with spatial non-directional language were analysed, since this was the only dataset that was analysed in terms of gesture space. In this case, the test scores slightly lower than in previous cases ($\kappa$=0.78), but it still indicates a substantial agreement between coders. In this case, the cause of the disagreement in some of the cases was to establish the difference between a gesture located in the centre and the periphery as well as a gesture located in the external periphery and the periphery. There were no instances in which the first coder considered the gesture to be located in the centre and the second coder annotated it as extreme periphery. Due to the novelty of the system and the difficulty of establishing a clear boundary in gesture that are located in the periphery, some level of disagreement was expected.

Other features such as hand, shape, orientation and congruence were not tested due to a number of reasons. First, the hand feature does not require the evaluation of two coders, since the level of subjectivity in the analysis of the hands that are part of the gesticulation is low. One coder would be sufficient to establish a clear analysis of this feature. Second, the congruency feature was directly related to the direction of the gesture. Since congruency is indicating if the gesture is congruent with the canonical direction of time in English speakers, it is only required to test the direction of the gesture in order to get a clearer picture of the congruency.

Finally, other features such as direction and orientation were not tested due to the low prominence they had on the analysis. The data obtained was not employed in order to perform any quantitative analysis, and it was only scarcely employed in the qualitative description of some particular cases. If further work in these two areas were to be done if the future, it would certainty require a secondary coder in order to check the validity of the analysis.

*5.4.2 Gesture quantitative analysis*

Once the gesture qualitative analysis was performed, all the data was quantified and stored numerically in Excel in order to perform statistical analysis on the data. Different types of information concerning *noise* data or cases in which speakers did not perform a co-speech gesture could be easily accessed. Similarly, all the information from features 1-6 (discussed in section 6.1) was stored here so that statistical analysis could be performed. The parametric statistical test employed to establish the different relations was a Wald Test, which is employed in cases in which a relationship between the different data items can be established from a sample. This test was employed to investigate the relationships between features 1-7 as well as establishing other parameters such as gesture frequency when using a particular temporal linguistic expression.

***5.5 Database overview***

This section will be devoted to offer an overview of the database that was employed during the development of the thesis, which was compiled after searching in the NewsScape library all the linguistic expressions indicated in section 5.3. 1. Table 5.3 offers a general summary of the distribution of all the hits according to this criterion for each of the linguistic expressions.

The total amount of hits rendered after performing the searches indicated in table 5.1 was 3487 hits. However, there was some variation in the amount of hits after applying the clip selection criteria indicated in 5.3.2. The first stage in the creation of the database was to detect all the cases that could be considered noise. This included clips that were repeated, like news clips or interviews that were repeated in television. The number of clips that were repeated in the corpus was 468, a 13% of the total hits. A small amount of the hits, 122 cases, were considered as false hits, since they do not reflect the original search. Finally, a total of 27 clips could not be analysed since there were some problems with the NewsScape link, the video or the audio file. This left the database with a total of 2897 real total hits, that is, the hits that referred to the linguistic expressions indicated in table 5.1 that were not repeated and excluding the videos that were corrupted.

The second stage of the compilation of the database focused on selecting the clips in which the hands of the speakers were clearly visible. During this stage the size of the hits was greatly reduced since there was a high number of instances in which the hands could not be seen.

A very high number of instances were considered to be voice overs in which there was no speaker to be see, only images, video or other camera shots that did not focus on the speaker. A total of 1073 hits, a 37% of the real hits, fell under this category. Next, there was also a high number of cases in which, even though the camera focused on the speaker(s), it only did on the upper half of their body and the hands were not visible. A total 991 hits were categorised as "no hands" clips, a 34% of the total real hits. This left the database with 1059 (28% of the real hits) cases in which the speaker and their hands were clearly visible and the quality of its observation was optimal for gesture analysis.

Following on the criteria indicated in 5.3.2, the next stage in the creation of the database was to detect in which cases the speakers performed a co-speech gesture and, if they did, if it could be considered to be a temporal co-speech gesture from the 1059 clearly visible hits. A total of 333 (31%) cases showed a clip in which the speaker did not performed a co-speech gesture. A 21% of the data, a total on 222 hits, were considered to be gestures that were not related to time. Finally, the remaining 48% of the clearly visible gestures, 504 hits, were categories as temporal co-speech gestures.

These examples were further subdivided in the axis in which the gesture took place. A 66% of the temporal gestures were analysed as lateral gestures. The remaining gestures were distributed as follows: 19% of sagittal gestures, 2% of vertical gestures, 12% of punctual gestures and 1% of gestures that were unclear and were not included in the final analysis.

| Category | Time period | Ling. expression | Hits rendered | Repeated clips | False hits | Broken link | Real hits | Voice over | No hands | Visible hands | No gesture | No time gesture | Time gesture |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Non-spatial** | Past | earlier than | 501 | 112 | 1 | 1 | 388 | 160 | 105 | 123 | 38 | 9 | 75 |
| | | previously | 301 | 41 | 7 | 1 | 252 | 116 | 85 | 51 | 19 | 9 | 23 |
| | Future | later than | 598 | 93 | 6 | 3 | 499 | 219 | 159 | 121 | 52 | 15 | 51 |
| **Spatial directional** | Past | back in those days | 455 | 48 | 7 | 9 | 401 | 102 | 137 | 162 | 49 | 29 | 74 |
| | Future | months ahead of | 375 | 18 | 7 | 3 | 350 | 134 | 135 | 81 | 21 | 4 | 53 |
| | | time ahead of | 118 | 8 | 17 | 0 | 93 | 30 | 41 | 22 | 6 | 2 | 14 |
| **Spatial non-directional** | Past | Distant past | 239 | 26 | 4 | 1 | 209 | 85 | 57 | 67 | 22 | 13 | 31 |
| | | Far in the past | 113 | 19 | 52 | 0 | 42 | 5 | 19 | 18 | 2 | 4 | 12 |
| | | Long in the past | 22 | 1 | 10 | 0 | 11 | 4 | 4 | 3 | 1 | 0 | 2 |
| | | Near past | 17 | 0 | 3 | 0 | 14 | 6 | 6 | 2 | 0 | 1 | 1 |
| | | Remote past | 3 | 0 | 0 | 0 | 3 | 1 | 1 | 1 | 0 | 0 | 1 |
| | Future | distant future | 276 | 55 | 2 | 2 | 219 | 69 | 84 | 66 | 23 | 14 | 27 |
| | | far in the future | 122 | 15 | 2 | 2 | 105 | 26 | 31 | 48 | 5 | 8 | 33 |
| | | long in the future | 6 | 0 | 1 | 0 | 5 | 0 | 3 | 2 | 0 | 1 | 1 |
| | | near future | 340 | 33 | 3 | 5 | 304 | 116 | 124 | 64 | 23 | 22 | 14 |
| | | TOTAL | 3486 | 469 | 122 | 27 | 2895 | 1073 | 991 | 831 | 261 | 131 | 412 |

*Table 5.3 Summary of the hits obtained through the data collection*

# Chapter 6

# Axis and congruency in temporal co-speech gestures

### *6.0 Introduction*

This chapter is devoted to comparing the gestures that co-occur with non-spatial temporal (*earlier than*), spatial directional (*back in those days*) and spatial non-directional linguistic expressions (*distant past*). The gestures that co-occur with these linguistic expressions are analysed in terms of axis and direction. Furthermore, these two factors are also employed to determine if the gesture is congruent with the canonical direction of time.

I firstly describe in detail the gesture data that I obtained in terms of axis in section 6.1. Here I comment on the axes that the gestures have employed in each of the categories. Gestures performed in the lateral axis are reported to be the most common gesticulation in time related gestures. However, this data shows some interesting trends in the gesture axis realisation in relation to the type of linguistic expression that co-occurs with (non-spatial, spatial directional or non-directional).

In section 6.2 I focus on the directionality of the gesture as well as their congruency with the canonical direction of time. As explained in section 2.4, the canonical direction of time in English located the past on the left, behind or above the speaker, while the future is on the left, in front or below them. This data suggests that the type of temporal linguistic expression may have an effect on the congruency of the co-speech gestures that are produced.

Finally, section 6.3 aims to investigate if the axis in which the gesture is produced (lateral, vertical…) has any effect on the likelihood of the gesture being congruent or incongruent with the canonical direction of time. Furthermore, this notion is investigated in relation to past and future expressions.

### *6.1 Overview*

As indicated in Chapter 5, two main features were employed to arrange the different linguistic expressions. Linguistic expressions were labelled as non-spatial if they did not employ the spatial domain (*later than*) when referring to time by using

temporal language. On the other hand, they could be categorised as spatial if they employed the spatial domain (*time ahead of* or *distant past*).

If the linguist expression employed the spatial domain, it was further analysed in terms of directionality. The linguistic expressions could explicitly indicate the direction in which the temporal period is located (*back in those days*), being a spatial directional temporal expression. Spatial temporal expressions, however, can also omit directionality (*distant past*). In this case, even though the spatial domain is being employed, the linguistic expression is not stating where the temporal event is located but rather how far it is located from the speaker (See Chapter 2, section 2.3).

The different temporal expressions presented in Chapter 5 were re-arranged according to these two features as shown in table 6.1:

| **Category** | | **Expression** |
|---|---|---|
| Non-spatial | Anteriority | Earlier than<br>Previously |
| | Posteriority | Later than |
| Spatial directional | Past | Back in those days. |
| | Future | Months ahead of<br>Time ahead of |
| Spatial non-directional | Past | Distant past<br>Far in the past<br>Long in the past<br>Near past<br>Remote past |
| | Future | Distant future<br>Far in the future<br>Long in the future<br>Near future |

*Table 6.1 Linguistic search teams and linguistic categories*

From the 3,486 hits obtained from the NewsScape library, a total of 412 contained a temporal co-speech gesture (See section 6 for more details on the rest of the dataset). Temporal gestures were evenly distributed amongst the three main categories in the following way: non-spatial linguistic expressions presented 149 gestures, spatial

directional language showed 141 gestures and finally spatial non-directional language contained 122 gestures (Table 6.2).

| Category | Total hits | Visible hits | Temporal gestures |
|---|---|---|---|
| (1) Non-spatial | 1400 | 295 | 149 (50%) |
| (2) Spatial directional | 948 | 265 | 141 (53%) |
| (3) Spatial non-directional | 1138 | 271 | 122 (45%) |

*Table 6.2 Distribution of visible hits and temporal gestures amongst categories*

### *6.2. Axis*

This section will be devoted to analysing the different axes that were employed by the gestures that co-occurred with the temporal expressions mentioned in 7.1. I will firstly comment on the data, giving more specific details of each of the three different linguistic categories. I will later discuss the results, in which I will focus on gestures that were located in three different axes as well as another type of gesture that presented no axial movement.

#### *6.2.1 Axis data*

Data suggests that gestures were most frequently performed the lateral axis, being this the axis with the highest frequency in the three different categories (60%-73%). The second type of gesture with the highest frequency are sagittal gestures, with a range from 12% to 30%. The third category addresses an interesting type of gesture that was found while analysing the data. This gesture has been labelled as "punctual",

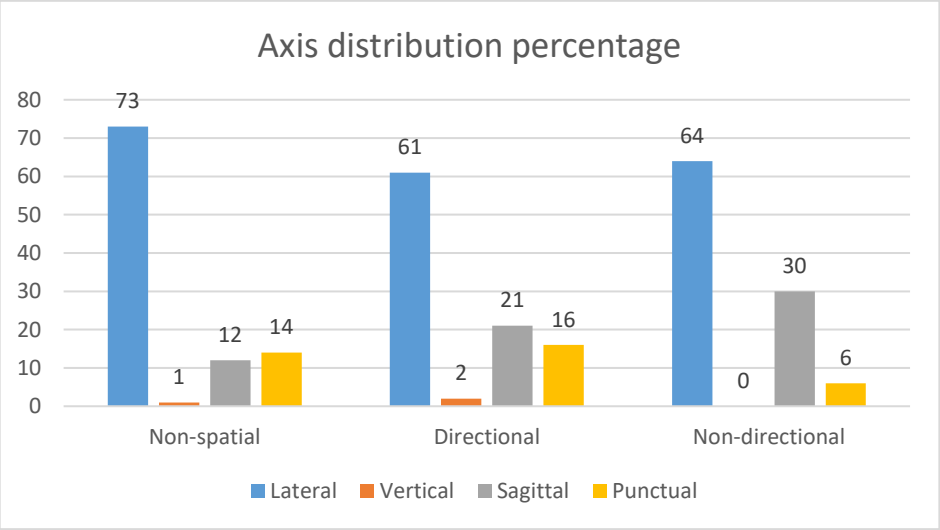



*Figure 6.1 Axis across categories*

since rather than being produced in a particular axis, the gesture pinpoints a point in space (See section 2.2.4 in this chapter for further discussion about punctual gestures). The last interesting result has to do with the presence of vertical gestures. These types of gestures are rare, with only a frequency of a 1%, but nevertheless it indicated that vertical temporal gestures can be produced in English (Figure 6.1).

*6.2.1.1 Non-spatial language*

Non-spatial language included expressions such as *earlier than* or *later than* (see Table 7.1) which used exclusively temporal language to talk about time. Here there was a total of 149 temporal co-speech gestures, 73% in the lateral axis, 14% punctual gestures, 12% in the sagittal axis and finally 1% vertical gestures (Figure 6.1).

Lateral gestures are predominantly recurrent when gesturing about non-spatial language, which is congruent with previous work (e.g., Casasanto & Jasmin, 2012; Walker & Cooperrider, 2016) in time gestures. Even though the rest of the axes are not used that frequently with this type of temporal language, it is worth mentioning that time-related vertical gestures are not usually reported in the literature and this data shows that they can be present in non-spatial language, albeit marginally. The presence of punctual gestures presents an interesting case, since the linguistic items refer to either past or future, but the gestures presents no axial movement. Individual expressions presented the same tendency, with no significant differences in the gesture/axis distribution.

*6.2.1.2 Directional language*

Spatial directional language covered different constructions such as *back in those days* or *months ahead of,* which referred to time in terms of space specifically encoding the direction in which the temporal period was located (front or back, since only the sagittal axis is linguistically referenced in English). A total of 141 directional gestures were obtained from the analysis of the dataset, although 3 of the gestures were discarded since the analysis was not clear. From the remaining 138, 61% were performed in the lateral axis, 2% in the vertical axis, 21% in the sagittal axis and finally 16% gestures were categorised as punctual (Figure 6.1).

Data shows that temporal gestures are commonly located along the lateral axis, although in a lower degree than the other categories. Similarly, there is a reduced number of gestures performed along the vertical axis, reinforcing the hypothesis that

vertical temporal gestures are scarce, yet present in English language. Punctual gestures were present in a relatively high rate, although further discussion is required to determine how to address this category. Finally, there is also a relatively high rate of gestures that are performed along the sagittal axis, presenting a rate which is between non-spatial and non-directional linguistic expressions.

Looking at individual expressions, the construction "months ahead of" seems to rely less on the lateral axis (56%) than the rest of expressions in the categories ("back in those days" with a 61% and "time ahead of" with a 67%), while the remaining gestures are almost equally distributed between sagittal (22%) and punctual (20%).

*6.2.1.3 Non-directional language*

This category addressed expressions such as *distant future* or *near past*, which are space-time linguistic expressions that do not state the exactly spatial location of the temporal period, but rather its distance in relation to the speaker. The total amount of gestures collected in this category was 122, although 1 was discarded since its analysis was unclear. The remaining 121 were analysed in terms of axis as follows: 64% of the gestures were performed in the lateral axis, which constitute the most frequent axis once again. There were no cases of vertical gestures, and there was a marginal number of punctual gestures (6%). The most interesting piece of data was the unexpected increase of the amount sagittal gestures that were performed in this type of linguistic expressions, which was a total of 30% gestures (Figure 6.1).

Most of the results are similar to the information that was collected from the previous categories: high frequency of lateral gestures, some marginal instances of vertical gestures and the use gestures that cannot be located in a particular axis, although the latter appears in a lower frequency than in the rest of the categories. The most interesting piece of information is constituted by the particularly high rate of sagittal gestures in comparison to the other two datasets. A $Chi^2$ test (p-value: .000919) indicated that the high rate of sagittal gestures in non-directional language was statistically significant when compared to the other two categories, rising some interesting questions since it seems there are some factors that make speakers more likely to gesture in the sagittal axis when they use non-directional temporal language.

When looking individually at the different expressions, some interesting results arise. Overall, most of the expressions in this category present a higher number of

*Figure 6.2 Lateral both-hand leftward gesture*

sagittal gestures than the other categories, with expressions like *distant past* (22%) or *near future* (21%). More precisely, there are two expressions that present a remarkable amount of sagittal gesture frequency: *distant future* or *far in the future*, with a frequency of 44% and 34% respectively.

### *6.2.2 Axis discussion*

#### *6.2.2.1 Lateral axis*

Data shows that most of the co-speech gestures are located along the transversal axis, with a very high frequency in all types of language. One of the reason why transversal gestures occur with such a high frequency is because it biologically and visually provides a more useful gestural space (Cienki, 2013). Firstly, the transversal axis corresponds to the biological position of our hands and arms, and thus the physical effort required to manipulate a timeline in this axis in lower than when we are using a sagittal or, more rarely, a vertical gesture. Furthermore, in face-to-face communicative situations, the transversal axis is more visible to both of the speakers, and hence it can be more informative than other spatial axes. The data obtained through the analysis of television news data corroborates previous findings on temporal co-speech gestures, confirming that English speakers tend to use the transversal axis when gesturing about time (Casasanto & Jasmin, 2012; Núñez & Cooperrider, 2013). Consider the following example:

1)… due to the warmth cherry blossoms started to bloom today, **11 days earlier than** normal, and the second earliest on record… (KCET France 24, 22-04-2015).

Example 1 shows one of the many cases of lateral gestures that were found during the analysis of the dataset. In this case, the speaker prepares the gesture when saying *11 days* by extending her finger tips in from of her. From there, she moves both hands leftward, while saying *earlier than*, to finally stop the stroke movement with the same finger position but having moved slightly to the left. This example shows a common right-to-left gesture movement employed to indicate a past event. The future is located on the right of the speaker and the past on the left, triggering a leftward gesture (Figure 6.2).

Even though the lateral axis is the most frequent when gesturing about time, there are some differences when the three categories are compared. When using non-spatial language (*earlier than*) speakers performed a lateral gesture a 73% of the occasions. This trend, although still high, was slightly smaller in other types of temporal language that employed a spatial metaphor, with a 60% of transversal gestures in directional language and a 64% of lateral gestures in non-directional language.

*6.2.2.2 Sagittal axis*

The most interesting piece of information is found when looking at the frequency of sagittal gestures across the different categories. Non-spatial language triggered a low amount of sagittal gestures, with only a 12% of the dataset of these category being a sagittal gesture. This ratio increased to 21% in directional language, and it increased even further in non-directional language, with a 30% of the gestures being performed in the sagittal axis. From looking at the data, it could be hypothesised that the addition of a spatial component might have an effect on the frequency of the spatial axis which is employed when gesturing about time.

When speakers add a spatial component to a time metaphor in directional and non-directional language, the metaphor is always elaborated in relation to the ego. Consider the following examples:

2) He says white people die with us in this movement. Abolitionists were white people **back in those days** (HLN George Zimmerman Trial Live, 19-07-2013).
3) The answer is. Please don't do it. Your retirement is a pot money for **far in the future**. (KCET nightly Business Report, 22-12-2011)

In these two cases, there is a temporal spatial metaphor that is taking the ego as a referent. In example 1, the past event is located behind the speaker, and this is made explicitly by *back*, locating the temporal event in a particular point in space. In example 2, even though the location is not made explicit, the spatial metaphor is also employing the ego as a referent. The event is *far* from the speaker in the temporal sense, rather than being *near*.

Non-spatial language makes no use of a spatial deictic point. This could explain why lateral gestures are more frequent in this type of language. Since there is no reference to the ego, speakers gesture more frequently event in terms of the transversal axis, since it is more anatomically accessible and easily visualised. However, when speakers employ directional or non-directional language, they have to include a spatial deictic component. In this case, the speaker is more likely to perform a sagittal gesture than in non-spatial language, since the temporal expression might be using the ego as a referent.

When looking in more detail at spatial language and, more precisely, to directional and non-directional language (***back*** *in the past* vs ***near*** *future*), there is also a difference in the amount of sagittal gestures that are produced (Figure 6.1). Both cases have a high rate of sagittal gestures when compared to non-spatial language, but also non-directional language presents a significantly higher amount of sagittal gestures than directional language. Going back to the previous hypothesis, when using this type spatial language, a deictic component is required (the ego). However, the main difference between these two types of expressions is that directional language implicitly includes a sagittal component, while this axis is not present in non-directional language. It is hypothesised that the higher frequency in sagittal co-speech gestures in non-directional language is due to the fact that the gesture is trying to cover the information gap which is not included in language.

This can be seen in example 4. This example shows a speaker that employs a sagittal backwards gesture while employing the non-directional temporal construction *distant past*. In this example, speaker prepares the gesture during a brief moment of silence before saying *the past*. She raises her right hand upward with the palm facing left. At the end of the silence, she performs several flaps with the hand in a backward direction, finishing the gesture when saying *the past, the distant past* (Figure 6.3).

*Figure 6.3 Right hand sagittal backward gesture*

4) It's very important. It's sixty - it's more than sixty years. It looks like**… the past, the distant past**, but it is not. But the problem is that it has been used as if it is the only - that the only victims were the Jews. (WWW Democracy Now, 08-06-2009).

Another example of the use of the sagittal axis can be seen in example 5. In this case the speaker is using a spatial directional temporal expression, *months ahead of*, while performing a sagittal gesture. The speaker prepares the gesture while saying *until* by doing a L shape with her left hand, with the palm facing left and the back of the fingers facing her. After that, she moves the left warm away from here, producing a sagittal away-from-body motion while she says *months*. When the hand has finished its motion, it remains still briefly while the speaker says *ahead of*, and briefly after it comes back to a neutral position (Figure 6.4).

5) The pictures are incredible, Jenna. I don't think we're going to really know the scope of the drama **until months ahead of this**. I mean, we can't get in there and assess the damage… (FOX- News Fox Online, 27-06-2011).

When using an expression like *back in the past*, we can gesture backwards, even though the gesture information would overlap with the information included in speech.

However, with expressions like *distant past* or *near future*, the location of the temporal period in relation to the deictic referent cannot be inferred from language. Co-speech gestures would then be employed to cover this information gap which is not found in language.


*Figure 6.4 Left hand sagittal forward gesture*

*6.2.2.3 Vertical axis*

The dataset shows that vertical gestures are very rarely employed when talking about time, resulting into just a 2% of the total gesture dataset. However, they present a very interesting case, since these types of gestures have not been previously reported in the literature. The vertical axis is not very often employed in English language when conceptualising time. This axis is not present in language, although, similarly to the lateral axis, it is present in calendar orientation and writing direction, among others. According to these cultural artifacts, past times would be located above, while future times would be located below (See Chapter 2, section 2.4). However, some of the examples that have been collected show the opposite patter. Considering the following example:

6) The stock market is the leading indicator, looking six **to nine months ahead of time** and jobs, that's a lagging indicator (CNN Newsroom, 03-11-2008).

In example 6, the image on the left shows the speaker preparing the gesture while saying *to nine*. The hand then performs and upward vertical movement, with the palm of the hand orientated vertically and the fingers pointing to one side, doing and inverted L shape. The stroke of the gesture finishes in the second image, while the speaker is uttering *ahead*. Then the speaker goes back to a neutral position, leaving her left hand in front of her body (Figure 6.5). The reasons why the speaker decides to employ a

vertical gesture in this case are not clear, although several hypotheses could be made in order to explain this phenomenon.


*Figure 6.5 Vertical left-hand upward gesture*

Firstly, it could be argued that this gesture, rather than referring to a TIME IS SPACE metaphor is actually referring to a TIME IS A RESOURCE metaphor. Even though time is encoded linguistically in terms of space, the upward gesture could be referring to a MORE IS UP metaphor, indicating the increasing time span in which the stock market looks at. In this case, even though language would be employing a specific metaphor to represent time, gesture would be employed to referrer to a different metaphorical representation of time.

A second hypothesis could be that the speaker is unconsciously modifying her gestures since in order to adapt her gesticulation to the communicative in which she is located. The speaker is in a news program and it is likely that she is aware that part of the screen is occupied by other information. Moreover, she is probably aware of the particular shot of the camera, which is showing her face and part of her left arm. It could be possible that the speaker is modifying her gestures so that they are visible to the audience. The visibility of a lateral gesture would be greatly reduced due to the different information showed onscreen, while the angle of the camera would not clearly convey a sagittal gesture. The speaker employs the gesture space on the top-left side in order to perform the gesture. This could also explain why the gesture is incongruent with the past-up/future-down canonical representation of time. A downward gesture would not be as visible as an upwards gesture.

The second example also presents an interesting vertical gesture. In example 7, the speaker says twice the temporal expression *far in the future*. The first occasion, she

performs a backward gesture with both hands. The second time, she performs the gesture that can be seen in figure 6.6. The speaker prepares the gesture while saying *someday* with a left-hand palm up open hand position. After this, she performs an upward movement, lifting her arm and flexing the forearm. The movement finishes while the speaker says *future*, with the forearm pointing upwards and the palm of the hand slightly bent backwards.

7) Are you done having kids? >> Oh, yes, yes, yes. Absolutely. Some day, far, far in the future, when they're older if I meet somebody. I'm not going to say 100% no. I'm not going to… **someday far in the future** get married… (KABC The View, 31-03-2010).

*Figure 6.6 Vertical left-hand gesture*

This example also presents a particular case that presents no easy explanation. It could be hypothesised that the reason why the speaker performs a vertical gesture is to add more emphasis to the temporal distance of the event she is talking about. The first time she says *far in the future*, she performs a backward gesture. During the second time she uses the expression, she includes the vertical gesture, slightly moving her hand backwards. This gesture could be considered not to be an entirely vertical gesture, but rather a combination of both the vertical and the sagittal axis. If this is the case, the upward movement would be justified, since anatomically it is easier to point backwards when your arm is lifted rather than when moving downwards.

It could also be argued that this gesture is not related to time and it could have other communicative purposes, more precisely indicating dismissal, as if the speaker were throwing the idea. However, the speaker also makes reference to the future some seconds before this gesture realisation (see section 7.6). In this case, the speaker

performs a similar gesture, using both hands and performing several flaps backwards by rising her forearms and moving her hands. As I argue before, this gesture is not an entirely vertical gesture, and the fact that moments before this gesture realisation the speaker performs a similar movement making that co-occurs with "far" reinforces the idea of the temporal nature of the gesture.

*6.2.2.4 Punctual gestures*

Perhaps the most unexpected type of gesture that was registered when compiling this database were a set of gestures that were categorised as punctual. This type of gesture realisation appeared on a relatively low frequency, being a 14% of the gesture dataset in non-spatial gestures, 16% in directional language and 6% in non-directional.

The name "punctual" was attributed to these types of gestures since their main difference with the rest of co-speech gestures is that they are employed to signal just to signal a precise point in space. The rest of the gestures that have been previously analysed contained a particular directionality that was attributed to them thanks to their movement throughout one of the three different axes. In this case, it could be observed how the speaker prepared the gesture so that it made a movement that went from point A to point B. The orientation of the hand or the movement of the wrist also indicated the direction that the gesture had (see all the examples presented in section 6.2 and 6.3, as well as all the examples in Chapter 7).

Punctual gestures present a different case that so far has not been reported in literature that has investigated temporal gestures. In this case, gestures are not performed alongside any particular axis, nor they present a clear direction when they are performed. When using a punctual gesture, the speakers highlight a point in space, usually with a small downward movement with the wrist or the arm. It could be argued that these gestures could resemble beat gestures or have a purpose that is not linked to the temporal expression. However, the different cases that have been examined present speakers that perform a single realisation of this gesture which is perfectly aligned with the temporal expression that they employed. The lack of repetition as well as the absence of this gesture realisation before and after the temporal expression might be an indicator that the gesture is indeed temporal.

The following examples try to show some gesture realisations that have been classified as punctual gestures:

8) Back then, gas prices were through the roof. We were stuck in an unpopular war. Gosh, we've come a long way. **Back in those days**, people thought cocaine was not addictive, that… (KCBS Late Late Show with Craig Ferguson, 23-008-2008)

*Figure 6.7 Punctual right-hand gesture*

Example 8 shows one of the cases that was analysed as a punctual gesture. In this case, the speaker performs a slight downwards gesture with his right hand, synchronised with "**back** in those days". His hand remains slightly open during all the duration, and after he performs this quick gesture the hand retracts to its previous position (Figure 6.7)

9) We are moving forward and we can see a horizon in the **not too distant future** in which so many of the cancers like a few not are merely chronic. (KOCE The PBS Newshour, 30-03-2015).

*Figure 6.8 Punctual left-hand gesture*

Another example of a punctual gesture can be seen in 9. In this case, the speaker shows a different shape in his left hand while performing the gesture, which is acquired while he says “horizon”. His thumb and index are touching each other, with the rest of the fingers touching his hand. During the temporal expression, he performs again a quick downward movement with this hand, which goes back to its original position after the temporal expression (Figure 6.8)

10) It’s already been decided. Find out about raises are made and go in a couple of **months ahead of time**. (KCBS The Early Show, 16-04-2008).


*Figure 6.9 Punctual both-hand gesture*

The last example 10 shows a case of a punctual gesture which is performed with both hands. In this case, we can see a slight variation in the shape of the gesture during its realisation. The gesture commences while the speaker says “a couple”, moment in which both the hands of the speaker are shaped like a fist, although both indexes are slightly extended. She performs a downward stroke during the expression “months ahead of”, while also slightly extending some of the fingers of both hands (Figure 6.9).

As it can be observed, the category of punctual gestures envelopes a wide variety of cases that do not follow the axial and directional pattern that most of the temporal gesture present. At its current state, this category might contain a variety of gestures with different semantic or perhaps even prosodic intention that is still not clear. What they all have in common, though, is that rather than being performed by using one of the spatial axis and taking a particular direction, all of the are quick strokes that are synchronised with the temporal expressions, as if they were signalling a point in space. Punctual gestures show slight axial movement, but the projection of the timeline cannot

be inferred from them. To my knowledge, this is the first time that these types of gestures have been reported in relation to temporal expressions.

Some of these gestures resemble to pointing gestures, but it is not clear which is the object that is being pointed at. A possibility is that the speaker might be using this type of gesture to signal a particular point in the space in which a temporal timeline could be projected. While in some cases speakers performed a leftward or a rightward gesture to talk about a temporal event, some communicative situations might require the speaker to signal a point in the timeline through a punctual gesture. Another of the arguments that could be made is that in fact, these gestures are more related to prosody rather than the semantics of the temporal expression, since they resemble other gesture realisations like beat gestures (See Chapter 4). If this is proven to be the case, this data would nevertheless be very useful; there are occasions in which a speaker decides to perform a prosodic gesture while using a temporal expression instead of performing the more common axial and directional hand movement that is reported in previous literature and experiments.

The reasons or intentions behind the use of punctual gestures is nevertheless still unknown. Further research on other type of temporal expressions as well as a bigger data set would be required to investigate whether or not there is a recurrent pattern in punctual gesture realisation and which is their communicative intention. Analysing a wider range of time expressions to investigate if punctual gestures are produced or performing more precise laboratory experiments in temporal gesture production might shed some light on this topic.

### *6.3 Congruency*

This section will be devoted to discuss whether or not the gestures that co-occur with the temporal expressions indicated in 7.1 are congruent. For this, I will take into account the canonical direction of time in western cultures (See chapter 2, section 2.3), that is, the past being on the left and behind the speaker and the future being on the right and in front of the speaker. I will start by commenting on the data and later on I will comment on the different findings drawn from it.

#### *6.3.1 Congruency data*

When analysing congruency, I compared the direction of the co-speech gestures (left/right, front/back…) to the canonical direction of time (see chapter 2, section 2.4).

Data suggests that gesture congruency levels show some variation depending of the type of linguistic expression they co-occur with. It is important to note that a part of the data set could not be analysed in terms of congruency: gestures that were performed in the lateral axis with both hands and had an inward or outward direction as well as punctual gestures could not be contrasted with the canonical direction of time. The remaining set of gestures was distributed as showed in figure 6.10

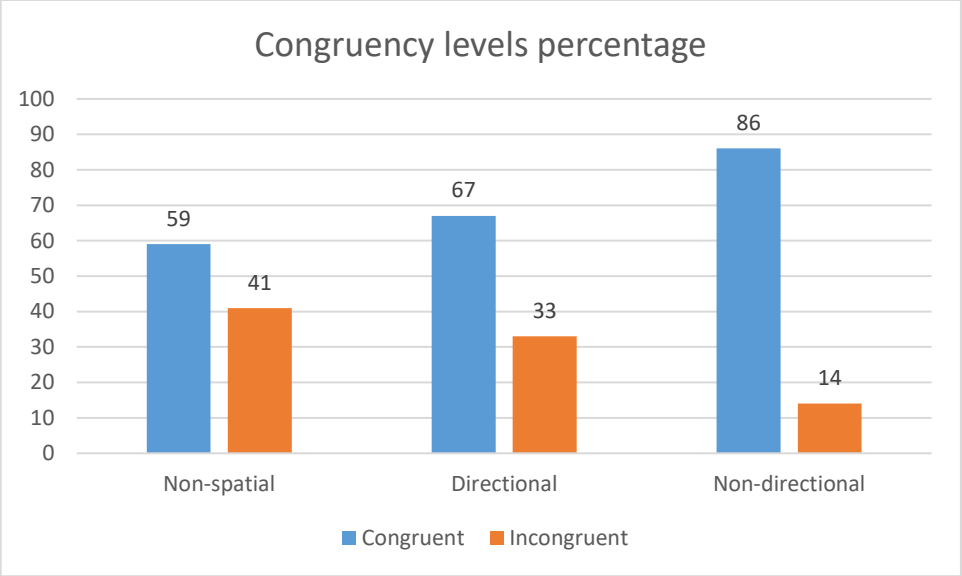



*Figure 6.10 Congruency levels*

*6.3.1.1 Non-spatial*

Data suggest that both congruent and incongruent gestures had a similar frequency rate in non-spatial gesture, as close as 41% for incongruent gestures and 59 % for congruent gestures, suggesting that gestures do not always mirror the canonical direction of time in this particular case. Contrary to expectation, gestures can be incongruent with the canonical direction of time, that is, a past related gesture is performed away from the speaker, or a future-related gesture to the left. No significant results were found in the individual expressions, which presented a ratio similar to the average (Figure 6.10)

*6.3.**1**.2 Directional*

Co-speech gestures tend to be congruent with the temporal expressions they co-occur with (67%), although there are still some cases in which the gesture is incongruent with the canonical direction of time (33%) (Figure 6.10). Some interesting results arise at the individual level; the construction “back in those days” presents a high level of

incongruence, with a 30% of the gestures being incongruent with the canonical direction of the past, while the expression "time ahead of" only presents a 7% of gestures that are incongruent with language.

*6.3.1.3 Non-directional*

Non- directional language presents and interesting case, since an 86% of the gestures were congruent with the canonical direction of time, while just a 14% of the cases were incongruent. (Figure 6.10) This difference shows a statistically significant variation (p-value .03317) of gesture congruency among these three different categories. When looking at the individual level all the expressions keep a low level of incongruent gestures, although expressions such as "far in the future" or "far in the past" present a particularly low amount of incongruent gestures (3% and 8% respectively).

*6.3.2 Gesture congruency discussion*

Non-spatial language presents very similar levels of congruency: 60% of the gestures are congruent with the canonical direction of time, while 40% are incongruent (Figure 6.10). Even though the ratio of congruent gestures is still higher, incongruent gestures are much more frequent than what we previously assumed. Consider the following example:

11) I think our severe weather season this year **will be a little bit later** than normal because of the cold February and the cold March we've had. (WEWS Live on 5, 26-03-2015)


*Figure 6.11 Incongruent future left-hand leftward gesture.*

In this case (example 11), the speaker is using a non-spatial temporal expression, *later than*, while performing a leftward gesture with his left hand. The speaker prepares

the gesture by extending his left arm with the palm facing down while saying *will be*. After that, the speaker performs the stroke of the gesture by moving the arm to the left while saying *a little bit*, creating a leftward horizontal movement. The gesture stops its movement when the speaker says *later*, changing the shape the orientation of the palm, which is facing the right. Even though the linguistic expression is referring to a future time (*later than*), the speaker performs an incongruent gesture, since he moves his hand to the left (Figure 6.11)

One of the reasons that explain why the speaker performs this gesture is that he might be somehow interacting with the outline that is appearing behind him. Timelines generally flow from the left to the right, and so we can see from the screen with the different hours of the day, with the past on the left and the future on the right. A hypothesis could be that the speaker is in fact interacting with that timeline and, he modifies his hand gesture so that the future is congruent with the timeline that is located behind him rather than being congruent with his own mental time line. In other words, the speaker could be aware of the communicative situation in which he is located, and he would take into account the point of view of the viewers in order to modify his hand gestures so that they are congruent with what the viewers can see.

Even though this is a possible explanation for this case, there are several occasions in which the speakers perform an incongruent gesture even in the absence of an outline or image. Something like that could explain this instance and the high frequency of incongruent gestures is that, since non-spatial language is purely temporal and it is not constrained by the spatial domain, the influence of the different spatial-related cultural artefacts like writing direction or calendars is not that strong. This would allow speakers to gesture more freely without having to adjust to the directionality that space attributes to time. Further research is required on this topic; research on different types of temporal construal or in languages that employ different cultural artefacts from English language (e.g. writing direction in Arab, Hebrew, Mandarin Chinese…) will surely shed some light on this topic. This would also allow the gesture that co-occur with these expressions to be more flexible, and hence the speaker would be able to modify them according to the different communicative situations, as shown in the example.

Directional language presents slightly lower level of incongruence. A total of 70% of the gestures were congruent with language, while the remaining 30% was not congruent. These results are in the line of the main argument that has been suggested: the level of congruency is slightly higher since the speaker is using the spatial domain. However, incongruence is not that rare; one possible explanation for this could be that since spatial information is already found in language, gestures do not need to clarify

*Figure 6.12 Incongruent past both-hand forward gesture*

where the temporal period is located and hence speakers can be more flexible with the gestures. Example 12 is another a case of a gesture which is not congruent with the canonical direction of time:

12) I mean, I feel like the cocktail hour—It's made a resurgence **really, but back in those days**, it was essential. It was part of the culture. (KABC Live with Regis and Kelly, 01-06-2011).

In this case (example 12), the gesture is prepared while the speaker is uttering *really* by positioning both hands in front of her body with the open palms facing her body. After that, while saying *but back in those* she performs a forward gesture with both hands. The orientation of her palms slightly changes, and this time they are position in a more lateral angle. The gesture stops when the speaker says *days*, after producing a sagittal gesture away from the body; this results in a past-related expression in which the speaker produces a gesture whose direction is typically associated with the future (Figure 6.12)

Finally, data related to non-directional language presents some interesting results. In this category of spatial temporal language, congruence levels are much higher

than in the rest of the categories: 86% of the gestures are congruent with the canonical direction of time, while the remaining 14% are incongruent. Similarly, to what it was suggested before, the hypothesis is that non-directional language relies on co-speech gestures to clarify certain information which is not found in non-directional. Congruency levels are higher since co-speech gestures need to reflect the direction of time. In most of the cases, speakers are more likely to perform a temporal gesture that mirrors the canonical flow of time.

Example 13 shows a speaker that performs a leftward gesture while using a past non-directional linguistic expression, *distant past*. The speaker places both her hands, with the palms open and facing each other while saying *did*. Just after, she says *distant past*, while moving both her hands to the left. The lateral motion is also accompanied by a rotation of the hands. They are no longer facing each other laterally, but they are parallel to her body, with the right palm facing her body and the left palm facing the space in front of her (Figure 6.13).

13) Look, I have a lot of respect for what Ralph Nader did in the distant past. I think he makes a good case about thigs… (KCET McLaughing Group, 30-03-2008).


*Image 6.13 Incongruent past both-hand leftward gesture*

The picture that emerges is that non-directional language presents higher levels of sagittality as well as higher levels of congruency. Both pieces of evidence support one of the possible explanations for this phenomenon: co-speech gestures are conveying information which is not found in language but which is important for the correct realisation and comprehension of the temporal blend.

Spatial manipulation of time flexibility is also supported by axis distribution in co-speech gesture realisations. Even though most of the temporal language in English is sagittal, gestures can be easily adapted to maintain a transversal, sagittal or even vertical direction depending on the speaker's needs. It is yet not fully understood what the reasons are that make a speaker prefer one axis amongst the others, but data shows that timelines can be extremely flexible in their realisation.

### *6.4 Congruency across axes: past vs future*

In order to see if there was any relation between the axis and the congruency of the gesture with the temporal period that the speaker was referring to, I also analysed congruency levels across the different axes in relation to past-related and future-related language. Here we will only be addressing the gestures that were performed across the lateral and sagittal axes, since the data from vertical gestures is too scarce.

It is also important to note that the temporal expressions that belong to the non-spatial category (*earlier than, previously and later than*) are not strictly related to the notions of past and future. These expressions indicate anteriority and posteriority of events that, even though most of the times is related to events in the past and in the future, might indicate other temporal periods. Nevertheless, I have decided to include the data in the overall analysis in order to offer a complete picture of the dataset since the results extracted from it do not affect the overall analysis.

#### *6.4.1 Congruency levels across axes: data*

When looking at congruency levels in the different axis, similar trends to the ones presented in 6.2 and 6.3 can be appreciated (Figure 6.14). Non-spatial language presents a 49% of congruent lateral gestures, while a 34% of the lateral gestures were incongruent. Congruency in the sagittal axis is much closer; with a 9% of the gestures being congruent and an 8% incongruent.

Directional language shows a 50% of congruent lateral gestures versus a 20% of incongruent cases. Concerning the sagittal axis, a 17% of the cases were congruent and a 12% incongruent.

Finally, non-directional language shows a 57% of congruent lateral gestures and an 11% of incongruent cases, while the gestures performed in the sagittal axis are mostly congruent, with a 28% of the data being congruent sagittal gestures and a 4% being incongruent.

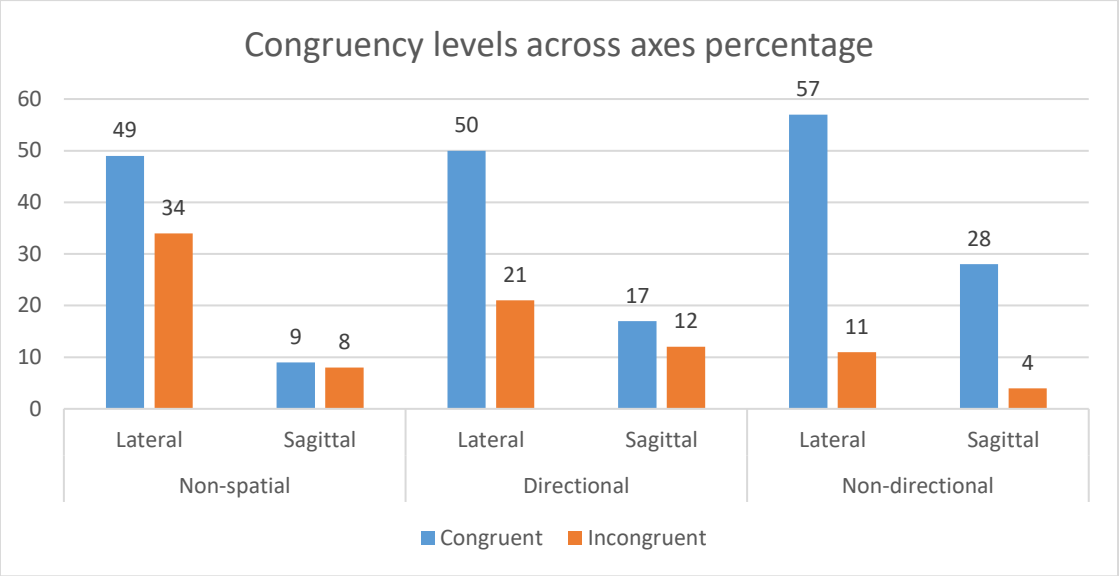



*Figure 6.14 Congruency levels across axes*

*6.4.1.1 Congruency in the lateral axis: past vs future*

Data did not present any relevant findings when looking at the different lateral gesture congruency levels in past and future language (See Figure 6.15). This table shows the ratios of lateral gestures from the overall dataset. Both past and future-related language present a similar trend in all type of expressions. Non-spatial language presents high levels of congruency and relatively high levels of incongruency in both types of language. Past-related language presents a 33% of lateral congruent gestures and a 22% of incongruent ones, while future related gestures presents a 16% of congruent examples and a 12% of incongruent.

Directional language is similar, with similar figures in both past and future language. Past-related language presents a 30% of congruent gestures and a 12% of incongruent ones, while future-related language shows a 20% of congruent gestures and a 9% of incongruent.

Lastly, non-directional language presents a similar trend between past and future expressions. Past-related language shows a 27% of congruent gestures and a 6% of incongruent cases, while future language presents a 30% of congruent gestures and a 5% of incongruent ones.

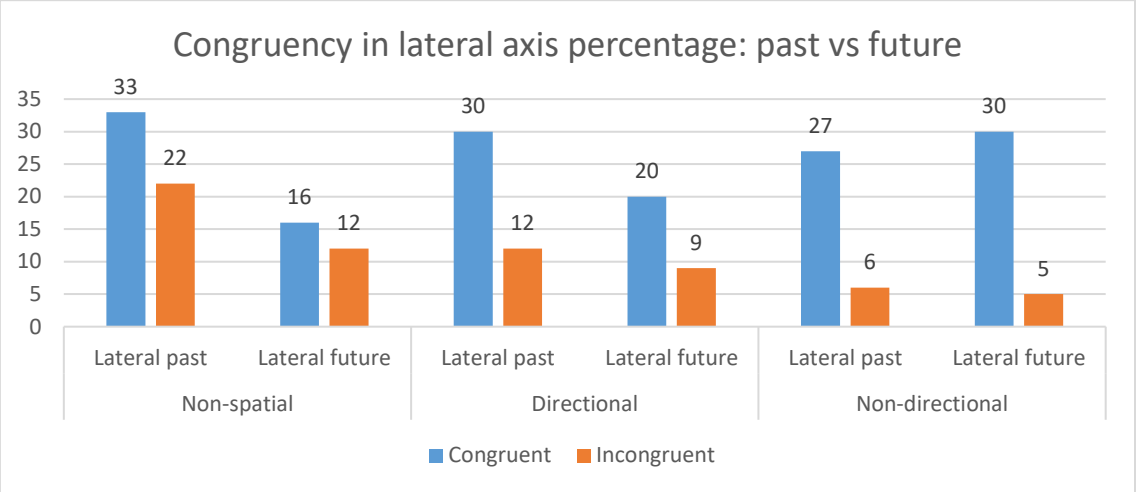



*Figure 6.15 Congruency in lateral axis: past vs future*

*6.4.1.2 Congruency in the sagittal axis: past vs future*

From all the time-related gestures that were performed in non-spatial language, a 3% corresponded to congruent past-related gestures located in the sagittal axis, while an 8% of the data corresponds to incongruent past-related sagittal gestures. When looking at future-related expressions, we obtained a 6% of congruent future expressions, while no incongruent instances were found (Figure 6.16).

Directional language presents a similar trend. From the total dataset, a 4% of the gestures were labelled as congruent gestures that co-occurred with past-related expressions, while an 11% was incongruent. Future language shows the opposite pattern: a 13% of the gestures were congruent with future-related expressions, while there was only a 1% of incongruent gestures.

Finally, the most interesting results are shown by non-directional language. Here we can see that a 6% of the gestures of this category referred to past-related expressions, and they were congruent with the canonical direction of time. A 3% of the dataset contained incongruent past-related gestures in the sagittal axis. Concerning future expressions, a 22% of the data was categorised as congruent gestures with the direction of future, while there was 1% of the gestures that were incongruent.

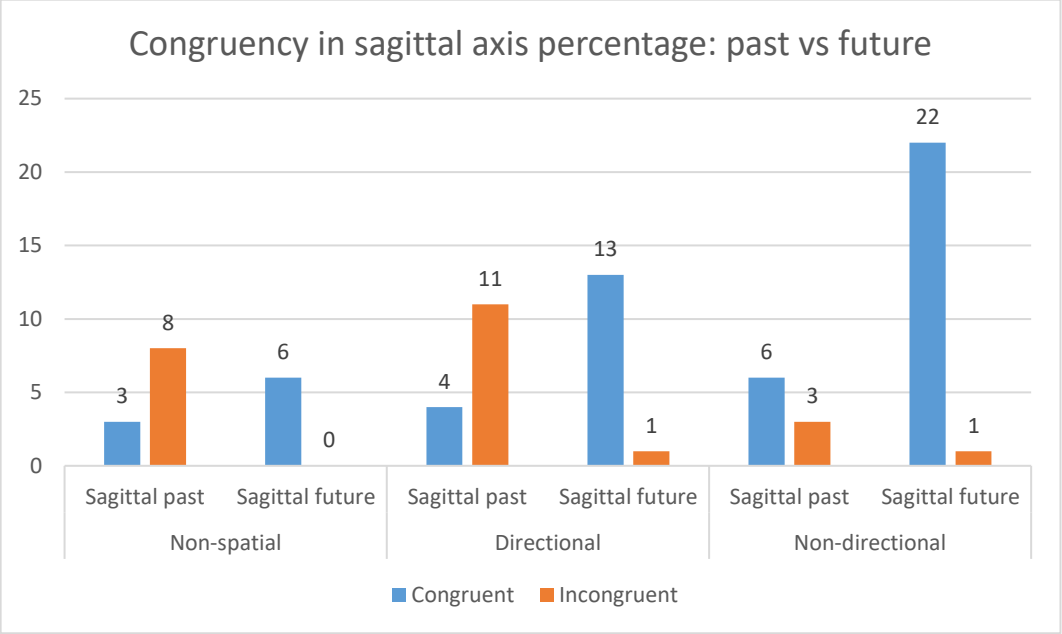



*Figure 6.16 Congruency in sagittal axis: past vs future*

*6.4.2 Congruency levels and axis distribution: discussion*

As it can be seen in Figure 6.16, some interesting findings emerge from the analysis of sagittal gestures in past-related language. Data suggests that sagittal past gestures are more likely to be incongruent in non-spatial and directional language. However, this trend is reversed in non-directional language, in which the amount of congruent sagittal past gestures is higher than the incongruent ones. Interestingly, future language does not present this trend; in fact, congruence levels in sagittal future language are very low.

There is not a definite answer that could explain this phenomenon, although different reasons could be hypothesised. Firstly, this phenomenon could be congruent with argument that was previously suggested. Non-spatial temporal language has no spatial bias, so the speaker would not be that influenced by the spatial constrains in temporal representation. The high rate of incongruent sagittal gestures in past-related language could be explain by the biological constrains of sagittal gesturing.

Anatomically it is easier to gesture in front of us, and the amount of gesture space available is higher. Backward gestures present a more intricated procedure that might complicate its realisation, making the speaker more likely to gesture in front of him/her.

14) I've got a friend who was over in **Iraq previously**, training guys, helping to equip guys. We gave those guys billions of dollars in equipment… (CNN Newsroom, 01-11-2014).


*Figure 6.17 Incongruent past forward gesture*

In example 14 it can be appreciated one of these cases. The speaker performs a sagittal away-from the body gesture that co-occurs with the temporal expression *previously*. The gesture starts with the hand next to the body, probably conditioned by the fact that the speaker was talking about a friend of his, while pointing to himself. He maintains this gesture for part of the utterance, until he says *Iraq*. After this, and at the same time that he says *previously*, he performs a sagittal gesture away from his body, with an open palm facing him. This example shows the difficulty of performing a sagittal gesture backwards; the gesture space available for the speaker is very reduced, and the mechanical effort to perform it is greater than the required one for performing a forward gesture (Figure 6.17)

Something similar might also occur with sagittal gestures in past directional language. In this case, there is a spatial component in language, but the direction in which time is flowing is explicit in language (*back in those days*). Language explicitly represents the temporal event as being behind the speaker, but since this information is already present in language, the gesture might not be required to represent this dimension. The anatomical difficulty for the realisation of the gesture would explain the higher ratio of incongruent gestures.

However, this pattern is inversed in sagittal gestures that co-occur with past-related non-directional language, and some cases in future-related gestures. Language does not explicitly state where the temporal event is located (*distant past*), but the spatial component is present. Since spatial temporal language takes the ego as a reference, the speaker might be more likely to gesture to past events as behind them in order to add the locative dimension, which is not present in language. Even though performing a backward gesture presents a higher level of difficulty, the speaker is more likely to perform a congruent gesture. Future gestures present in this case a very high rate of congruency in the sagittal axis. In this case, both biological and the spatial location of future events (*ahead of*) benefit from an in-front location of the temporal gesture.

The next example, 15, presents an instance in which the speaker performs a future related gesture as well as a past-related gesture, both in the sagittal axis. The first gesture is produced when the speaker says *into the future*. In this case, the speaker performs an away-from-the-body gesture by moving his right hand forwards and extending it, with an open palm facing the left. He keeps the gesture in this position while saying *by tying yourself*. He continues the gesture by performing a towards-the-body gestures with the same hand by rising his forearm. The shape of the hand changes to a fist, with an extended thumb pointing backwards (Figure 6.18).

15) When you are a republican candidate the only way to lead America **into the future** is by tying yourself **to the distant past**.

Another possible reason the explain the high rate of incongruent past gestures in non-spatial and directional language could be that these gestures are in fact not temporal gestures. Some of these gestures could fall within the category of presentational gesture, in which the speaker performs a palm-up-open hand forward gesture, generally to present a topic or idea. In this case, the speakers could be using presentational gesture to present the past event to the rest of the speaker, as if they were presenting the topic to the rest of interlocuters. Nevertheless, this hypothesis would not explain why this tendency is reversed in non-directional language. If we assume that some of these gestures can be classified as presentational gestures, it seems that there are some factors that are making presentational gestures to be more frequent in non-

spatial and directional language, while they are less frequent in non-directional language.

*Figure 6.18 Future forward gesture followed by past backward gesture*

Finally, another reflection that should be taken into account is the possible role of point of view and perspective when speakers perform incongruent co-speech gestures. Most of the speakers that are part of the dataset are accustomed to talk for audiences that are facing them. It could be that they are in fact taking the perspective of the viewers into account, that is, they use the audience's left-to-right or back-front reference points in order to perform the co-speech gesture. If true, viewpoint would also be playing a very important role on the spatial arrangement of temporal co-speech gestures as well the spatial arrangement of temporal information. The information obtained from the analysis of this dataset is not enough in order to establish that viewpoint in fact also a component of temporal co-speech gestures. Further experimental approaches on this issue would certainly shed some light on the role of viewpoint in time conceptualisation.

### *6.5 Conclusion*

This chapter has provided a wide range of information concerning gesture axis, directionality and language-direction congruency. The aim of this chapter was to investigate if there were any differences in gesture realisation when comparing temporal language that employed the spatial domain and linguistic expressions that were purely temporal. The distinction between non-spatial language (*later than*), directional language (*back in those days*) and non-directional language (*distant past*) has resulted to be extremely useful when looking at differences in the gesticulation.

The first finding that the analysis of this data has offered is the confirmation that temporal gestures are often located in the lateral axis. Data shows that in all the different

types of language, lateral temporal gestures are the most common type of gesture. Previous literature already suggested this tendency, and this data has provided the first large scale gesture study that has confirmed this hypothesis.

Another interesting fact suggested in this chapter has to do with a variation in the amount of sagittal gestures that depended on the type of linguistic expression that the speakers were employing. Even though sagittal gestures appear always in a lesser frequency than lateral gestures, their frequency changes depending on the temporal linguistic expression. Non-spatial language presented a low ratio of sagittal gestures. However, directional language presented a higher ratio of sagittal gestures, which was even greater in non-directional language. The hypothesis that has been offered is that certain type of gestures, in this case sagittal, are more likely to appear with certain kind of language, in this case temporal expressions that use the spatial domain. When using the spatial domain to talk about time, speakers are more likely to employ the sagittal axis. One of the possibilities is that, since they take the ego as a reference to talk about temporal concepts, the increase in sagittal gestures would indicate that they also employ the ego as a spatial anchor to gesture about temporal concepts. Moreover, gestures could also be indicating information that is not encoded linguistically, as it is the case of non-directional language. Since this type of language does not indicate where the past or the future are located in relation to the ego, an increase of the use of the sagittal axis when gesturing with this type of temporal language could cover the temporal direction information gap.

The researcher conducted on temporal co-speech gestures also offered some cases of gestures that were located in the vertical axis. These types of gestures are not usually reported on the literature related to temporal co-speech gestures. Vertical gestures appear in a very low frequency, but they are also present in English temporal language. The reason behind the use of vertical gestures is not clear, although one possibility is that other metaphors like TIME IS A RESOURCE or MORE IS UP could be influencing the gesticulation of the speaker in relation to the temporal concept. Other possibilities are related to the environment of the speaker, that might modify the location in which the gesture is located.

Another finding that has be obtained from the analysis of this dataset is the existence of a type of gesture that co-occurs with temporal linguistic expression but that

not show any type of axial movement. Provisionally, these gestures have been categorised as "punctual" gestures. The main characteristic of punctual gestures is that they are not produced along one of the spatial axis, and they do not show a particular direction. Rather than that, speakers signal a point in space, probably a point in a timeline, that might correspond to a moment in time. These gestures appear as single realisations that co-occur with the temporal expression, suggesting that linguistic expression and gesture are indeed related. It is not clear yet which is the intention or the causes for the production of this gesture. Further research of this type of gesture through a larger dataset, other type of experiments or the analysis of different type of linguistic (temporal) expressions would shed some light on this topic.

One of the most interesting findings that this research has shown is perhaps the congruency levels on temporal co-speech gestures. The amount of incongruent gestures is higher than expected, with a high amount of this type of gestures in non-spatial language. One of the reasons that might explain this is that, since speakers are not employing the spatial domain when using this type of language, they do not need to follow the canonical direction of time-space (past on the left and future on the right, for example). Since the language is purely temporal, speakers can situate time in the gesture space more freely. The rate of incongruent gestures, however, slightly diminishes when we look at directional language. Talking about time by using the spatial domain makes speakers more likely to gesture according to the canonical direction of time, but the effect is not too strong since the linguistic expression explicitly states the direction of time (*back in those days*). Gestures would be mirroring information that is also present in the linguistic expression. Finally, non-directional language presents a very interesting case since the amount of incongruent gestures is extremely low. In this case, the gesture would need to indicate the temporal location in space since this information is not encoded linguistically (*distant past*). The gesture could be covering the spatial information gap that is not covered by the linguistic expression.

The last piece of evidence that this analysis has offered has to do with the congruence of the gestures in past and future related expressions. While there is no relevant finding in the past and future gestures performed in the lateral axis, there is some interesting information in the gestures performed along the sagittal axis. Firstly, it seems that past sagittal gestures tend to be incongruent with the canonical direction of time, that is, they are performed forwards. This trend can be observed in non-spatial

and non-directional language, but it is reversed in non-directional language, in which there is a higher number of congruent gestures. The second piece of evidence has to do with the gestures that we employ when we talk about the future. In this case, gestures are congruent in a very high ratio. This means that, when gesturing in the sagittal axis about the future, people also tend to perform a forward gesture.

# Chapter 7

# Temporal distance and gesture location

### *7.0 Introduction*

This chapter is devoted to further investigate the relation between temporal concepts and co-speech gestures through the analysis of temporal distance and spatial distance. Here, I perform a more in-depth analysis of one of the linguistic categories that was described in Chapter 5: non-directional language. Expressions that are included in this category employ the spatial domain when referring to time, but they do not indicate the location of the temporal period. Instead, they indicate how far the temporal period is located from the speaker. Expressions such as *distant past* or *distant future* can be employed to express that the event took place/will take place a long time away, while expressions like *near past* or *near future* indicate temporal proximity.

The aim of this chapter is to investigate if the temporal distance that is expressed linguistically is also represented through co-speech gestures. The hypothesis is that language that expresses temporal proximity will be accompanied by a higher number of gestures that represent spatial proximity, while language that expresses temporal distance will trigger gestures located further from the speaker. Furthermore, it is suggested that gesture shape might have an impact on the semantics of the temporal gesture. Pointing, bounded gestures would indicate that the speaker is conceptualising more precise moment in time, while unbounded repeated flaps with the hands could indicate that the location of the temporal event is more vague and fuzzy.

The first section of this chapter, 7.1, is devoted to describe the dataset that is employed in this chapter, including the range of linguistic expressions language and the gesture features that employ in my analysis. Temporal distance is only implicitly present in non-directional expressions, although it can also be present through other contextual factors in other types of temporal language. Gestures that co-occurred with non-directional expressions are analysed in terms of the location of the gesture by following Bressem's (2013) annotation system (see Chapter 5, section 5.4).

Section 7.2 looks at the data taking into account the frequency of gestures that co-occur with distance-related language and proximity-related language. The most

interesting finding is that, according to this dataset, speakers tend to gesture about temporal distance more often than about temporal proximity.

Section 7.3 includes a detailed description of the gesture data collected through distance-related language and proximity related-language. Here I describe the tendencies and patterns that can be gleaned from the data, as well as presenting a number of examples that are qualitatively analysed. The data shows that when speakers use a linguistic expression to express temporal distance, they are more likely to perform a gesture located further from the body, while when using an expression that indicates temporal proximity they are more likely to perform a gesture that is close to their body.

Section 7.4 further investigates the dimension of gesture location and temporal distance through a comparison of the gestures that are related to the past and the gestures that are related to the future. The main finding in this case is that there is a higher frequency of future-related expressions in the corpus and, as a consequence, there is also a higher number of future-related gestures. Some possibilities for the higher frequency of future-related temporal linguistic expressions are offered in this section.

Section 7.5 includes a number of qualitative examples that were particularly interesting, since speakers made allusion to both close and distant temporal periods. In these examples speakers contrast an event that took place, for example, in the *distant past* with an event that took place in the *near past*, thus offering events that are located at a different temporal distance. In the cases that are presented in this section, speakers also present co-speech gestures that signal two different points in space; one that indicates the close temporal period, and another one that indicates the distant temporal period. A qualitative analysis of this examples is given, and some insights on the gesture and language interaction are offered.

The final section, 7.6, focuses on the possible relation between gesture shape and temporal distance linguistic expressions. Here I offer several examples that indicate that, apart from a frequent open palm gesture in which the hand may face several directions, there are two distinct types of gesture that might be of interest: pointing gestures and palm flaps. I hypothesise that pointing gestures indicate that the speaker is temporally locating the event more concretely, while hand flaps might indicate that the speaker is locating the temporal event in an unspecified timeframe.

### *7.1 Aims and Data*

The aim of this chapter is to investigate if there is a language-gesture relation between temporal distance, which is expressed linguistically, and spatial distance, expressed by co-speech gestures. In order to do so, I take a closer look to a subset of all the dataset presented in Chapter 5. More precisely, I only take into account gesture data obtained from linguistic expressions that fall under the category of *non-directional*. Non-directional language includes time-related language that employs the spatial domain to express temporal concepts. However, this type of language does not include information about the spatial location or direction of the time period, but it makes reference to how far (in spatial terms) the temporal event is located from the speaker. This category includes expressions such as *distant past*, *far in the future* or *near past*.

It could be argued that other types of spatial temporal expressions also make reference to the temporal distance of an event. However, this information is not included in the temporal expression itself, but rather in discourse or contextual factors. Furthermore, temporal distance is a property that depends on the perception of the speaker, being highly subjective. One person can, for example, talk about a memory of their childhood as if it were something that happened yesterday, or about the same event as something that happened ages ago. Non-directional language offers a more objective way of knowing whether the speaker is conceptualising the event as something temporally close or far, since this information is explicitly encoded in language.

The expressions that were employed to perform this analysis were divided depending whether they indicated temporal proximity or temporal distance. The result is a total of 13 different expressions, 7 referring to distance and 6 referring to proximity (Table 7.1)

| Linguistic Expression | Temporal distance | Temporal period |
|---|---|---|
| Distant past | Distance | Past |
| Distant future | Distance | Future |
| Far in the past | Distance | Past |
| Far in the future | Distance | Future |
| Long in the past | Distance | Past |
| Long in the future | Distance | Future |
| Remote past | Distance | Past |
| Near past | Proximity | Past |
| Near future | Proximity | Future |
| Not so distant past | Proximity | Past |
| Not so distant future | Proximity | Future |
| Not long/that long in the past | Proximity | Past |
| Not long/that long in the future | Proximity | Future |

*Table 7.1 Linguistic expressions and temporal distance*

It is important to highlight that the English language presents a reduced number of expressions to indicate temporal proximity. *Near past* and *near future* were the only expression that exclusively referred to temporal proximity that were found in the corpus. The rest of the set of expressions that express temporal proximity are formed by the negation of temporal distance expressions (e.g. *Not so distant future*). The opposite pattern, negated temporal proximity expression to indicate temporal distance, was not very common, although some marginal examples were found. In general terms, the amount of temporal expressions that make exclusive reference to the proximity of a temporal event is quite reduced in English, with *near* + N being the most common construction.

The initial searches of these key terms in the NewsScape Library provided 517 hits of distance-related language and 621 hits of proximity-related language. After eliminating the hits that were considered noise and cases in which the hands were not visible (See Chapter 5, section 5.3), there were a total of 128 hits in distance-related language and 143 hits in proximity-related language in which the hands of the speakers were clearly seen.

The analysis of the final set of visible hits resulted in a total of 122 semantic co-speech gestures. Distance-related expressions presented a higher number of gestures, with a total of 76 cases in this category. On the other hand, proximity-related expressions presented a lower number of gestures with 42 instances. The remaining 4 gestures were not included in either of these categories, since these were cases in which speakers referred both to a distant and a near temporal event (see section 7.5).

These gestures were further analysed in terms of gesture distance, following Bressem's (2013) annotation proposal. Gestures were divided into three main groups depending on their spatial position with respect to the body: *centre* if the hands were located around the chest, *periphery* if the hands were located between the chest and the shoulders and *external periphery* if they were located further from the shoulders. Even though this analysis procedure is relatively clear when analysing gestures performed along the lateral axis, the analysis of this feature becomes more challenging when analysing a sagittal gesture, since the body cannot be used as such a clear reference point.

In this case, I take into account the level of flexion and extension of the forearm. If the speaker showed a fully flexed forearm in a lateral position, with the hand close to the chest, the gesture was analysed as central. Gestures that were performed with a semi-flexed forearm, with some distance from the chest of the speaker were categorised as located in the periphery. Finally, a gesture performed with a fully extended forearm that presented a considerable distance from the speaker was labelled as external periphery. Additionally, there were some instances in which speakers performed vertical flexions of the forearm in order to gesture behind them. Even though this gesture is located close from the body, these instances were categorised as external periphery, since our anatomy does not allow us to signal further back (see further details in Chapter 5, section 5.4).

### *7.2 Gesture frequency: distance vs proximity*

#### *7.2.1 Data*

The first interesting piece of information that I obtained from this data was related to the amount of semantic gestures that were obtained from each of the two categories. Although both distance and proximity related categories counted with a very

similar number of hits in which the speaker (and their hands) were clearly visible (128 and 143 respectively), the amount of co-speech gestures obtained greatly varies.

From the total 128 hits, distance-related language presented 31 cases in which the speakers did not perform a co-speech gesture and 21 cases that were classified as gestures that did not have a clear semantic value (see Chapter 5). The remaining 76 gestures were analysed as temporal gestures. These values drastically change when we look at the data obtained from proximity-related language. From the total 143 hits, there were 53 cases in which the speaker did not perform a co-speech gesture, while there were 48 instances in which the speakers performed a non-semantic co-speech gesture. The remaining 42 cases were temporal co-speech gestures.

The differences in these two categories can be appreciated in Figure 7.1. Proximity-related language presents an evenly distributed trend in the data, each category occupying around 30% of the dataset. This trend, however, changes in distance-related language. The number of non-semantic gestures drops to 16%, while the amount of cases in which there was no gesture also decreases to 24%. The most interesting information emerges from the amount of temporal co-speech gestures obtained from distance-related language, which comprises 60% of the distance hits.

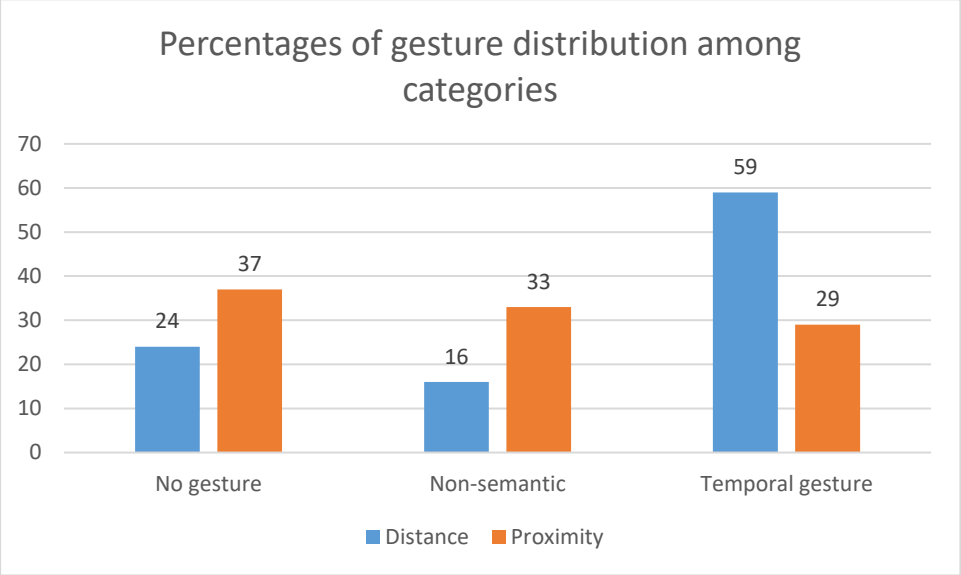



*Figure 7.1 Gesture distribution among categories*

### *7.2.2 Discussion*

The previous data suggests that there are several differences when we look at the amount of temporal gesture that co-occur with distance and proximity related language. Speakers only produce a temporal co-speech gesture when using proximity-

related language 30% of all occasions. On the contrary, speakers perform a temporal co-speech gesture with distance language more often; in 60% of the occasions. A $Chi^2$ test (p-value: .00001) suggests that there is a significant increase of the chances of performing a time-related co-speech gesture when we talk about temporal distance than when we talk about temporal proximity.

It is difficult to determine why this difference in frequency might be present, although one of the possible explanations might be at the conceptual level. Proximity and distance are both concepts that establish a particular spatial distance between the speaker and another entity (in this case, a temporal event). When we think of proximity, we think of it as the immediate space that is around us. Speakers would generally agree that an item is located in the proximity of the speaker if it is situated in the space between the speaker and another, more subjective boundary established by the ego. However, expressing spatial (and temporal) distance is a more complicated topic. The boundary between something being in the proximity and the distance is very subjective, and greatly depends on the individual. Moreover, the amount of space covered by the concept of distance is virtually unlimited. Anything that is located beyond this boundary established by the speaker could be categorised as distant.

One of the possibilities to explain the increase in the distance-related co-speech gestures could be related to the complexity of the concept of distance. Since this concept is covering such a wide amount of spatial configurations, speakers may be more likely to perform more gestures when they are talking about temporal distance in order to express more clearly where the concept is located. This can be appreciated when comparing English demonstratives. In English language, we find one demonstrative to refer to proximity (*this*) and another one to refer to distance (*that*). The amount of items that we can refer to with *this* are reduced to the ones that are located in the immediate space of the speaker, while *that* comprises the rest of the spatial configurations. Similarly, temporal proximity refers to the immediate space; there is no need to specify further the location of the temporal event, since it can be located close from the ego. The familiarity of the speakers with their immediate space would require a lesser level of specification, while distance, as it is such a wide concept, requires further specification, which is provided by co-speech gesture. The gesture would then be employed to provide a better localisation of the temporal period when speakers are talking about temporal distance.

Following this hypothesis, a gesture could potentially indicate different degrees of proximity or distance that are not encoded linguistically, although further analysis as well as a bigger sample of data, together with other experiments would be required to investigate this hypothesis. It would also be extremely useful to analyse if other languages that have different pronouns and determiners that indicate different degrees of spatial distance (*este, ese, aquel* for close, medium and far distance in Spanish, for example) present also a different gesture configuration when talking about temporal distance.

### *7.3 Distance-related language and gestures*

#### *7.3.1 Data*

There was a total of 76 gestures that co-occurred with linguistic constructions that expressed temporal distance, like *long in the past* or *distant future*. 18% of these gestures were performed in the centre, while 32% were located in the periphery. The most interesting piece of information comes from gestures that were located in the extreme periphery, which comprise 50% of the distance-related gestures (Figure 7.1).

Proximity-related language, like *not so distant past* or *near future* presents the opposite trend, even though the data sample is slightly smaller, with a total of 42 co-speech gestures. 14% of these 42 gestures were analysed as extreme periphery gestures, while 38% of the gestures were located in the periphery. Finally, the remaining 48% was formed by gestures that were located in the centre.

Data from both distance and proximity-related temporal language present tendencies that complement the semantics of the expressions they co-occur with. When we talk about events that we perceive as being temporally far from us, we tend to gesture further from our body, while when we talk about events that we perceive as temporally close to us, the gestures are performed closer to our body (Figure 7.2).

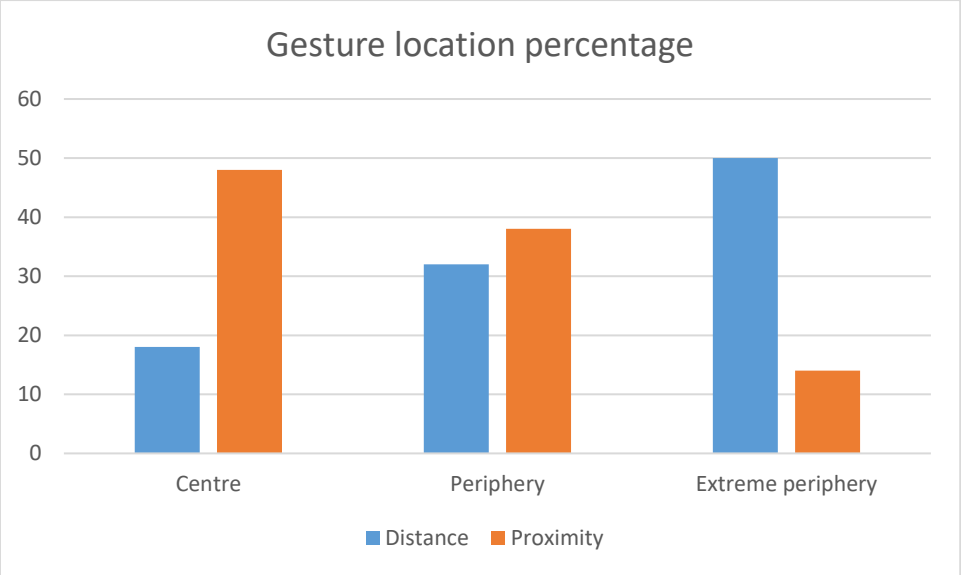



*Figure 7.2 Gesture location*

*7.3.2 Discussion*

This data confirms the initial hypothesis which suggested that when we talk about temporal events that we conceptualise as close in time, we gesture closer to our body, and when we talk about temporal events that we conceptualise as far in time, we gesture far from our body. In most of the cases that have been registered in this dataset, the gesture is representing in space the concept that is expressed through a linguistic expression. The space-time mapping does not take place only at conceptual and linguistic levels, but it is also present in the majority of cases at the gestural level too. Consider the following examples:

1) Here future cast gives us an idea, **high resolution, a better look into the distant future**. You can see here on Saturday… (KCAL 9 News, 26-2-2015).


*Figure 7.3 Both-hand sagittal gesture located in the external periphery.*

Example 1 shows a distance-related gesture that is performed in the extreme periphery. The preparation of the gesture starts when the speaker says “high resolution” (left). Here he positions both hands laterally to the left, with the left palm facing forwards and the right palm backwards. The gesture starts when the speaker says “better look”; here he moves his left hand forwards, until he performs a total extension of his left arm (right). Additionally, the gesture is also accompained by the motion of the whole body. The speaker moves his whole body forward, as if he were pushing something, while perfoming the extension of the left arm (Figure 7.3).

2) Hey folks, I went out and saw the number one movie in America this weekend, “The Martian”, It’s the sci-fi thriller that **takes place in the near future** where… (KCBS Late Show with Stephen Colbert, 6-10-2015).

*Figure 7.4 Both-hands sagittal gesture located in the centre.*

Example 2 presents a case of a gesture that was performed in the centre and co-occurred with a proximity-related linguistic expression. Here the speaker prepares the gesture while saying “takes place”. He positions both hands in front of him, with the palms facing downwards and performing an extension on some of the fingers. Then, when he says “near future”, he performs a very subtle forward movement with both hands. He slightly rises his left hands and rotates it in order to move it forward, but the final position of the gesture is very close to the initial one (Figure 7.4).

According to this data, it seems space-time conceptual patterns that are present in language are often also present (and mirrored) in gesture, at least when it comes to the relation between temporal distance and spatial gesture distance. Distance in time corresponds with spatial gesture distance, while proximity in time corresponds with spatial gesture proximity. However, there are some cases in which this pattern is not

that clear, or it can be reversed. There are some examples that present the opposite pattern: gestures performed in the centre when talking about temporal distance and gestures performed in the external periphery when talking about temporal proximity:

3) … and you know, other thoughts started to come in. But**, unlike in the past, in the distant past**, say a year ago… (WKYC Channel 3 News at 11, 13-04-2015).


*Figure 7.5 Right-hand lateral gesture located in the centre.*

Example 3 presents an instance in which the speaker employs a gesture located in the centre while employing a temporal expression that was related to temporal distance. He holds his right hand at the same level as his shoulder with the open palm facing leftward while saying *"*unlike in the past". After that, he says "in the distant past" while performing a short leftward motion with his right hand, without any change in the shape of the hand (Figure 7.5).

Another example of a gesture that does not correspond to the distance that is indicates by the linguistic expression is 4:

4) …during the war and after that war is a mistake that the country I think now is fairly unified in trying not to repeat. But Vietnam **is not so far in the past**, though, that you can just fix that problem… (MSNBC The Rachel Maddow Show, 8-01-2013).


*Figure 7.6 Left-hand lateral gesture located in the external periphery*

In this occasion, the speaker is talking about a temporal event that, according to her, is temporally close to that moment. However, she performs a hand gesture in the external periphery. Both the speakers' hands are on the table in resting position with both palms facing downwards. When she says "not so far in the distant past" she moves her left hand leftwards, extending her forearm and opening the palm, which is facing to the front (Figure 7.6).

These two examples are cases in which the speakers do not follow the gesture pattern that has been previously observed; the location of their gesture does not correspond to the temporal distance that they are indicating in the linguistic expression. Further research in these particular cases is required in order to further understand what are the circumstances that make speakers perform a gesture that does not correspond to the linguistic expression they are using. In example 3, for instance, the speaker is located in a very specific communicative setting (a press conference) that may have an impact on the type of gestures that he produces. The fact that the speaker has to address a large number of people in a particular setting may have an impact on the type of co-speech gestures that are produced. Other discursive, prosodic and contextual factors may also have an impact on the type of co-speech gestures that are produced, although further research is required to investigate which factors have an effect and to what extent.

Similarly to the information discussed in Chapter 6 (see section 6.3.1), gestures appear to be an extremely flexible tool that can adapt to the speaker's needs when it comes to the expression of temporal distance. The gesture mirrors the information that is encoded in language most of the time, but there are a number of occasions when the gesture can perform something completely different. Even though the reasons behind the modification of the gesture are still unknown, it can be said that temporal gestures

are indeed extremely flexible. Even though in most of the occasions they mirror the temporal distance with spatial gesture distance, speakers are also able to decide not to perform this mapping and modify their gesture according to the situation.

### *7.4 Past vs future language and gesture location*

#### *7.4.1 Data*

This section aims to investigate if there is a relation between the temporal distance that is covered by the expression (proximity or distance) and the temporal period the expression is referring to (past or future). The first piece of information can be appreciated when looking at the total amount of hits for both future and past language in the whole corpus. Future language had a much higher frequency in the initial searches, with 744 hits, while past expressions appeared in a smaller ratio, with only 394 hits. It seems that, when looking at the whole corpus, it is more normal to talk about temporal future distance than temporal past distance.

The results show a similar trend when looking at the examples in which there was a co-speech gesture. In general terms, the dataset shows a higher ratio of future related gestures in both proximity related language and distance related language (Figure 8.7. However, a $Chi^2$ test (p-value: .183623) suggest that this difference is not significant. The data suggest that when we gesture about temporal distance, we gesture more about the future than about the past.

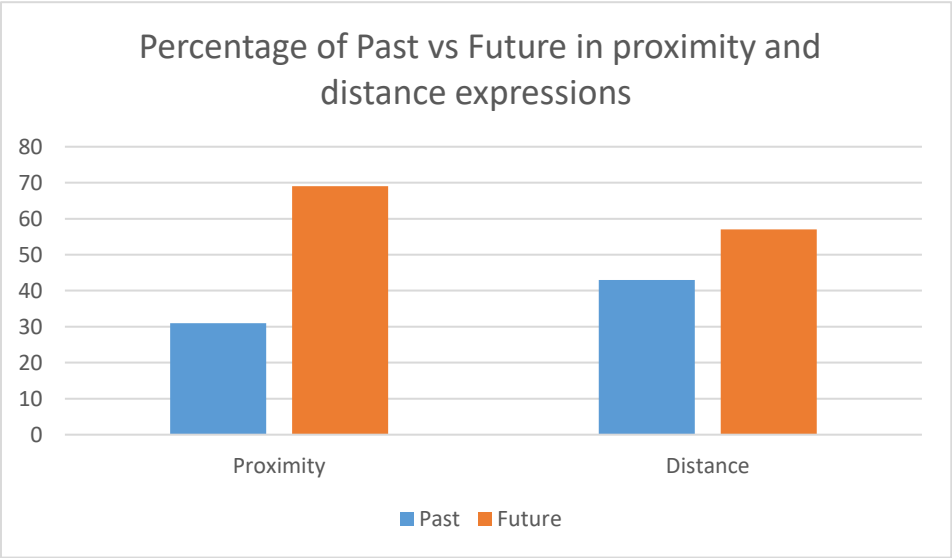



*Figure 7.7 Past vs Future in proximity and distance expressions*

Other dimensions of the dataset were also investigated with no interesting results. When looking at the relation between the location of the gesture (centre, periphery and external periphery) and the temporal period the expression was referring to, results mirrored the information explained in section 7.1

*7.4.2 Discussion*

The data indicates that, when talking about temporal distance, there is no significant differences between past and future expressions. Even though the results are not significant, it seems that there is a slight trend towards the future gestures.

Recent research has suggested that cultural factors, and more precisely how forward thinking or traditional a culture is, might have an impact on the conceptualisation of the past and the future and the direction in which speakers gesture (see de la Fuente et al., 2014 and Chapter 3, section 3.3.2). It could be hypothesised that this might also have an impact on the frequency that speakers talk about the past or the future. More forward oriented cultures will be more focused on the future, and thus they would be more likely to talk about how far in the future certain events are.

Even though this data is interesting, much further analysis is required to establish a robust hypothesis that might explain the increase of future expression when talking about distance. Broadening the analysis to a wider range of future and past temporal expressions or conducting further corpus study on the frequency of future and past related terms might shed some light on this issue.

***7.5 When both distances are present***

Another phenomenon that is recorded in this dataset has to do with the presence of both proximity and distance related language in the same sentence, together with two different hand gestures that correspond to the two different temporal distances. From all the dataset, there were only 4 examples that were categorised as containing both proximity and distance-related language in which speakers performed two co-speech gestures.

Example 5 shows the first case in which this type of language and gesture interaction took place:

5) …republicans have made their choice to stand with the special interests and stand up for failed policies **of the recent past or the distant past** on election day… (CampaignAds Democracy Party, 8-8.2010).

In this example, the speaker performs two hand gestures, one with each hand, that co-occur with the expressions *recent past* and *distant past*, one after another. Firstly, the speaker performs a leftward gesture with his right hand towards the centre of his body. The palm has a vertical position at the beginning of the gesture, but it slowly tilts until it faces downwards. This gesture is produced at the same time that the speaker says "of the recent past" (Figure 7.8.1).

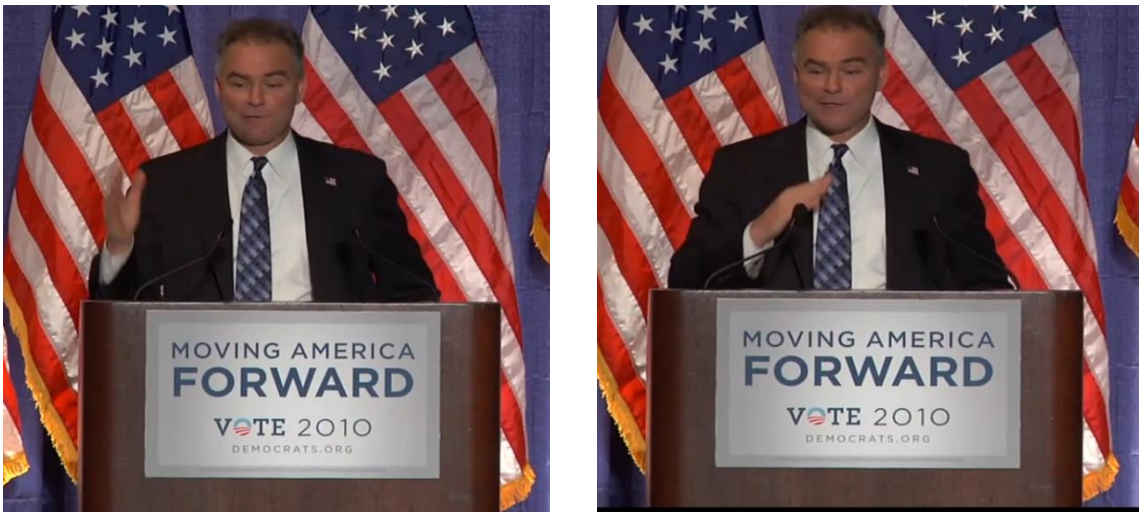



*Figure 7.8.1 First gesture. Right-hand lateral gesture located in the centre.*

After that, the speaker lowers his right hand until it is covered by the stand. Immediately, while he says "or", he starts to raise his left arm. He produces then an arc-shaped gesture starting on the stand and finishing with his left arm completely extended to the left. All the time the palm remains facing downwards (Figure 7.8.2).


*Figure 7.8.2 Second gesture. Left hand gesture located in the external periphery.*

This example shows the relation that conceptual temporal distance has with gesture spatial location. The speaker is assigning a determined amount of space with his gestures to the two different temporal periods he is referring to. The recent past is indexed by a small right-handed gesture located between his arms, while when mentioning the distant past, he performs a large gesture with his extended left arm that covers as much space as possible. Temporal distance is not only represented linguistically, but gesture makes a direct reference to the spatial component of the temporal metaphor.

It could be argued that these two gestures, rather than referring to a spatio-temporal mapping, could be employed to present the two different events that the speaker is talking about ("the recent past **or** the distant past"). Even though this could be possible, the contrast in the amount of space covered by the two different gestures might suggest that the intention of both gestures could be more than just presenting those two ideas.

Another case in which two gestures co-occur with proximity and distance-related language is 6.

6) Two years ago, they were gloating over their debt ceiling hostage strategy. No president **in the near future, maybe in the distant future**, will be able to… (US MSNBC The Ed Show, 19-01-2013).

In this case, the speaker performs both gestures with his right hand. In the first image, he prepares his right hand to perform the first gesture by raising the hand, with the palm facing forwards (Figure 7.9.1). While saying "in the near future", he performs a small stroke to the right, reaching the position that is appreciated in the second image (Figure 7.9.2). After that, he maintains the same position with the gesture until he says "in the distant future". At this moment, he performs another stroke to the right, going slightly further than with the previous gesture (Figure 7.9.3).

This example offers a subtler case in which both temporal proximity and temporal distance are present and accompanied by two co-speech gestures. The amount of gesture space covered here is not as large as the previous example, but both strokes are easily detected. This example helps to reinforce the hypothesis that time gestures are extremely flexible, and speakers can adapt them to the discursive situation.


*Figure 7.9.1 Preparation*


*Figure 7.9.2 First stroke*


*Figure 7.9.3 Second stroke*

***7.6 Gesture shape and temporal distance***

The last topic that is addressed in this chapter will be the possible relation between the shape of the gesture and the temporal expression it co-occurs with. When analysing gesture shape in this dataset, it could be observed that there are certain patterns in the shape of the hands when speakers perform a co-speech gesture.

Probably the most common shape was a gesture produced with an open palm, facing rightwards or leftwards if the gesture was lateral or downwards or upwards if the gesture was sagittal. Images 7.3-7.6 show this type of shape, in which the speaker opens their palm while performing the gesture motion.

However, there is a set of cases that, even though they are less frequent, they present very different gesture shapes that might have an influence on the semantics of the gesture. These gestures could be categorised in two different groups, according to the different shape that the hand takes. The first group includes a number of gestures that could be considered to be pointing gestures, while the second group includes cases

in which the hand acquired a more relaxed position while, in some occasions, performing several flaps.

There were several occasions in which the speakers employed a pointing gesture when referring to the temporal distance of an event. Example 7 shows how the speaker performs a backwards gesture with her right hand while pointing with her thumb backwards (Figure 7.10).

7) I think by tonight; the storm system should be **long in the past**. (KTTV-FOX Morning News at 6AM, 14-9-2009).


*Figure 7.10 Backwards pointing gesture*

Example 8 offers another similar case in which the speaker performs a backwards gesture with his right hand while pointing backwards. Even though it is difficult to appreciate in the image, he points backwards with his thumb, while keeping the rest of his hand as a fist (Figure 7.11).

8) …because when you are a republican candidate the only to lead America into the future is by tying yourself **to the distant past.** (Comedy Central Colbert report, 10-01-2012).


*Figure 7.11 Backwards pointing gesture*

The change from the highly frequent open-palm gesture to these two specific types of gesticulation might indicate that they are indeed employed for a specific purpose. Even though further research is required, it could be hypothesised that pointing gestures which co-occur with distance related language indicate that the speaker is conceptualising a more concrete or specific moment in time. In example 7, the conceptualiser is talking about a storm that is probably happening at that specific moment. She says that by that day at night time, that storm should be "long in the past", while pointing backwards with her thumb. The moment in time she is referring to seems to be quite specific according to the information we get from the weather forecast.

Example 8 offers a very interesting case, in which the speaker is referring to a specific moment which is represented by the image that appears next to him. In this case, he says that republicans tie themselves to the "distant past". At that moment, an image of a mummy or other type of aged body in a background with cave paintings appears. Even though the moment in time he is referring to is not too specific, he is referring to a graphical representation of that moment in time. It could be possible, then, that in this case language, gesture and image are interacting to convey a very precise message, which in this case is humoristic.

The second group of gestures that present a specific shape includes some examples in which the speakers perform several flaps with their hands. In this case, all the examples included language that referred to distant periods of time, either past of future. There were no instances of this type of gesture that made reference to an event located in the temporal proximity. Example 9 shows how the speaker performs several

backward flaps with both hands while repeating several times “far” and finishing the gesture with “in the future” (Figure 7.12).

9) Is this it for you done having kids? >> Oh, yes, yes, yes. Absolutely. Someday, **far, far, far in the future**, when they’re older if I meet somebody… (KABC The View, 31-03-2010).

*Figure 7.12 Repeated backward hand flap.*

Another example that shows a similar tendency is 10. In this case, the speaker performs several backward flaps with her right hand while referring to “the past, the distant past” (Figure 7.13).

10) It's very important. It's sixty - it's more than sixty years. It looks like**… the past, the distant past**, but it is not. But the problem is that it has been used as if it is the only - that the only victims were the Jews. (WWW Democracy Now, 08-06-2009).

*Figure 7.13 Repeated backward hand flap.*

The second set of examples offers a very different type of gesture, and I hypothesise that it could be employed to suggest a very different type of meaning. In contrast with the more concrete pointing gesture that was employed in the previous examples, these two cases offer gestures in which the hand is relaxed and performs a flap several times. In this case, speakers would not be talking about a concrete moment in time, but rather the opposite: a moment in the past or in the future that is not very specific for them, or that is located at a very long temporal distance. In the first case (9), the speaker says that she might have kids again "far, far, far in the future". The flaps she performs with the hand co-occur with the repetition of the word "far". This contributes to the idea that the speaker is not thinking about a precise moment in time, but rather to an event that might take place at an unspecified moment.

The following example (10) presents a similar case, in which the speaker is talking about an event that took place a long time ago. She does not quite know the precise moment, since she says "more than sixty years ago". After that, she says that that event looks like "the past, the distant past" while performing several flaps with her right hand backwards. The speaker would be referring to an event that happened in the past, but at an unspecified time. She has a vague idea of the temporal location of the event, but both the language and the gesture might indicate that the precise moment is not clearly defined.

Much further research is required to find out if indeed gesture shape can have an effect on the semantics of temporal gestures. However, this data could be connected with some previous research on gesture boundedness and verb aspect (see Cienki et al., 2016) in which a link is established between bounded gestures and bounded events and unbounded gestures and unbounded events, at least in French. The data obtained from this analysis may indicate that this phenomenon takes place also in English language and it is not restricted to verbal aspect. When speakers talk about temporal events, the gesture might be an indicator of how the aspect of the event is being conceptualised. Pointing, bounded gestures would then be used to indicate more clear and precise moments in time, while flaps with a relaxed hand could indicate temporal events that happened in a less defined and blurry moment in time. Further research on this topic would be highly beneficial to the study of temporal gestures; performing a more in-depth study of gesture shape or performing more controlled experiments in a laboratory could shed some vital light on this topic.

### *7.7 Conclusion*

This chapter has been devoted to the analysis of the location of gestures that co-occurred with temporal distance linguistic expressions. Language which indicated temporal proximity as well as temporal distance was taken into account. A number of findings can be drawn from the analysis of the co-speech gestures that are employed with these expressions.

Firstly, even though the initial number of hits was the same, the data indicates that it is more normal to perform a co-speech gesture when we talk about temporal distance than when we talk about temporal proximity. 30% of the data in proximity-related language contained a temporal co-speech gesture, while this ratio was double, 60%, in distance-related language. The hypothesis is that distance is a very complex and wide concept that includes multiple special configurations. Something that is located *far* from the speaker can be located anywhere between the boundary that the speaker chooses and the infinite. On the other hand, language that refers to the proximity of the temporal events heavily relies on the space around the ego, which makes a clearer and more familiar concept. Since distance is such a complex concept, co-speech gestures could be employed to be more specific about the temporal location of the event.

The second finding is that, when gesturing about temporal distance, gestures that co-occur with proximity-related language tend to be located close to the speaker, in the centre of the gesture space. Conversely, gestures that are employed with distance-related language are more likely to be located far from the speaker, in the extreme periphery of their gesture space. This data suggests that the role of temporal co-speech gestures is not limited to dimensions such as direction (see Chapter 6), but it can also represent temporal distance through spatial distance. This contrast between gestures located in the centre with a smaller amount of gesture space indicating temporal proximity and gestures located in the external periphery when indicating temporal distance is also observed in cases in which the speaker talks about both types of temporal location. Finally, even though speakers tend to follow this pattern, there is some degree of flexibility in gesture representation when indicating temporal distance. Speakers are also able to employ gestures that do not mirror the temporal distance indicated by the linguistic expression, although this happens in a smaller frequency. More research is required to investigate what are the causes that make speakers perform these types of gestures that do not follow the general pattern.

The next point that was investigated in this chapter was the possible relation between past and future temporal expressions and gesture distance and location. There were not significant results in this area, so it seems that there is no relation between gesture distance and the temporal period the expression in referring to. However, data suggests that there is a very high number of instances of expressions that refer to future events; almost twice as many as past expressions. Although this has been a marginal result it could be interesting to investigate the reasons for the high frequency of future related temporal distance expressions.

The last finding that was suggested in this chapter has to do with the possible relation of the gesture shape and its semantics when it comes to temporal distance. The most common type of gesture shape in this dataset was the open palm with different types of orientation depending on the axis in which the hand was employed (lateral, down, up…). However, the most interesting cases were two different groups of gestures that presented some particular shapes. The first group contained a number of examples in which the speakers used a pointing gesture, usually using their thumb to point backwards. It is suggested that the gesture shape and, in this case, the pointing gesture might be employed by the speaker to indicate a more precise moment in time through a gesture realisation. The second group of gestures presented a relaxed hand that performed several flaps. In this case, the gesture could be employ to indicate that the speaker is not sure about the exact temporal distance at which the event took place, since rather than indicating a precise spot in space as was done with the pointing gesture, this gesticulation is more vague. Further research is required in this area, but it is possible that these results could be linked to previous research that was performed on the topic of boundedness of gesture and boundedness of events (perfective vs imperfective).

# Chapter 8

# Gesture studies and time conceptualisation: thought and directions

### *8.0 Introduction*

The following chapter does not focus on particular data aspects, but rather I comment on methodological and theoretical issues that should be taken into consideration in further research involving the study of gesture and/or (time) conceptualisation.

The first section focuses on different methodological issues that emerged during the development of this thesis. Here I discuss some problems and possible improvements to the gesture analysis frameworks I have used. Traditionally, gesture analysis frameworks have focused either on the form or the function of the gesture, although I suggest that there are occasions in which other properties, like physiological, can be employed when analysing features such as gesture location. Additionally, I discuss a type of gesture that seems not to present a clear axial movement: punctual gestures.

Another important methodological issue that is covered has to do with the progress that has been made on the development of machine-assisted gesture analysis thanks to the data obtained from this thesis. I briefly comment on an automated annotation system that, even though I have not directly programmed, has used the temporal gesture dataset created in this thesis as well as the gesture annotation for its development.

The second section will focus on the possibilities and advantages that the use of big data can offer to the study of gesture and cognition. By using as a starting point the data and results obtained in this thesis, I will suggest here a number of ways to continue research on temporal gestures, as well as other aspects of interest for gesture and cognitive linguistic researchers.

Here, I focus on the contribution that the Red Hen Lab and the NewsScape library can make to gesture studies and more precisely to other types of temporal gesture studies that could be conveyed through this dataset, including a different set of

expressions, different type of metaphors or different languages. I also cover a number of factors that could have an effect on the different co-speech gestures that speakers are producing. Some factors, such as number and position of speakers, have already been found to have an effect, but this can be studied quantitatively through this dataset. Other factors that can be studied here have to do with the reasons why a speaker does not perform a gesture, the communicative genre and situation and its effect on gesture or scriptedness and co-speech gesture, amongst other.

I finish the chapter by presenting a number of examples that might show how speakers are able to interact with external or imaginary physical objects through gestures in order to convey a more precise temporal concept.

### *8.1 Gesture studies: methodological developments*

This section is devoted to discussing a number of methodological advances that could be implemented in current gesture analysis frameworks. The first section makes some suggestions to help achieve a more accurate description of the location of co-speech gestures. Traditionally, the body has been used as a referent to establish the relative position of the co-speech gestures, although this cannot be achieved when describing sagittal gestures. Hence, I suggest that a number of physiological features can be used when describing the gesture location of sagittal gestures.

The second point I discuss here has to do with certain gestures that present no axial movement (described in Chapter 6). Even though further research is required, perhaps the current gesture analysis framework could be expanded to include these type of gesture realisations.

The last section describes the advancements in research software that has been achieved thanks to the time gesture dataset created in this thesis. More precisely, I describe a software that is currently able to assist gesture researchers to perform (to some degree) large scale gesture analysis and which, in the future, could evolve into an automatic gesture recognition software.

#### *8.1.1 Gesture framework. Improvements over the current ones*

Gesture analysis frameworks tend to focus either on the form or the function of gesture realisations in order to provide an accurate description of them depending on the researcher's needs. The main aspects of interest that were taken into account in this thesis had to do with form features that included axis or direction, for example. In order

to address the most important gesture form features, I employed Bressem's (2013) gesture annotation framework as the basis for analysing gesture forms. In this framework, Bressem addresses a large number of formal gesture features that have been discussed over the years and aims to establish a common gesture form framework for scholars. This annotation framework has been found to be extremely useful and it covers some of the main formal gesture aspects that can be identified in co-speech gestures. However, during the development of this thesis, it became apparent that there are some areas that could be expanded and improved in order to offer a more complete framework; a more precise set of features to locate gestures in the gesture space and a more completed system for gesture axis localisation.

#### 8.1.1.1 Gesture location in the gesture space

The first formal gesture aspect that would require some discussion has to do with the gesture location in the gesture space, and the parameters employed to locate a gesture in relation to the body. Bressem (2013) employs McNeill's notational system for gesture space (1992), which divides gestures into four different areas according to the body, including the area between the shoulders (this one having two different sections), the area around the shoulders and the area beyond the shoulders (see chapter 4, section 4.2.1). Additionally, she includes Fricke's three-dimensional model (2005, 2007), which includes four different dimensions to the gesture space: from 0 referring to the speaker's own body to 3 referring to far distance from the body. Even though this system is useful for establishing a general gesture location, it is highly dependent on the analyst's judgement. There are not any clear and objective parameters that establish the differences between a gesture performed in 1 (close distance from the body) and 2 (medium distance from the body); all rely on the analyst's judgement.

What factor(s) are taken into account when addressing gesture location? Both systems use the body as a referent to establish an objective location for the gesture. Fricke's model, even though useful for some situations, does not provide accurate measures for establishing gesture location for all the different gesture configurations that can be obtained, and thus it is a bit complicated to accurately locate gestures in a three-dimensional space. On the other hand, McNeill's (1992) system employs the body as a referent to locate gestures, although this system exclusively works when gestures are performed on the lateral and vertical axis. By using the body as a referent, for example, a gesture can be located in the external periphery if it is located far from the

shoulders. However, there are some gesture realisations in which we cannot refer to the body in order to measure distance; gestures that are performed in the sagittal axis present some problems when using this system, since the body cannot be easily taken as a referenc..

When analysing gestures in the sagittal axis, we cannot use as a referent any part of the body; the gesture is perpendicular to our body and a clear reference cannot be established. This becomes more problematic when we look at sagittal gestures that are performed backwards. Even though Fricke's system addresses this type of movement (-1, -2 and -3) backward gestures present a very different case from gestures performed in other axes and directions. Gestures performed in the lateral and vertical axis, as well as gestures performed in front of us do not present many constraints in terms of the gesture space available for the gesturer. However, backward gestures present a completely different case; the amount of gesture space in this direction is very limited, since our anatomy does not allow us to perform gestures much further back from our shoulder. If we try to apply, for example, Fricke's dimensional framework in this case, the gesture would very frequently be located close to the body, since it is anatomically impossible for a backward gesture to be located far from the body.

Thus, I suggest that a different system should be used when measuring sagittal gesture location. Instead of a system that exclusively looks at the relative distance between the gesture and the body, paying attention to physiological factors, such as the degree of extension or flexion of the arm, could be of great benefit to locate a gesture performed in the sagittal axis. Speakers tend to modify the distance covered by a sagittal gesture by modifying the degree of flexion or extension of the forearm, or by rotating the arm in different ways. Looking in more detail at the degree of extension or flexion of the forearm could serve as an indicator to establish whether a gesture is located at a close, medium or far distance the body, for example.

The different types of configurations that can be obtained are explained in section 5, but here I offer a more detailed account of the anatomical features that we could take into account. Let's start by looking at sagittal gestures that are performed away from the body. When looking at the possible configurations that locate the gesture at the immediate gesture space, the forearm can be positioned laterally or vertically, and it is usually presented with a semi flexion and, more rarely, a full flexion. This position

could be considered to be the central location of the gesture; here the speaker manipulates the gesture space that is located in front on them. The hands are positioned laterally or vertically in front of the speaker, who is able to manipulate the space directly in front of them (Figure 8.1).

*Figure 8.1 Central sagittal gesture.*

The next gesture location would correspond to a medium distance. In this case, the speaker would perform a slight movement away from their body that would make the forearm to be into a semi-extended position together with a semi-extension of the upper arm. This position can also be accompanied with a rotation of the elbow, adding a slight upward movement if the arm was positioned laterally. If the arm was positioned vertically, the movement will always be accompanied by downward motion, and occasionally upward motion if the arm was completely relaxed (Figure 8.2).


*Figure 8.2 Medium location of away-from-body sagittal gesture.*

Finally, the last position would correspond to a gesture that is located far from the body. In this case, the forearm would be completely extended, as well as the upper arm. The furthers distance from the body is usually associated with a full extension of the arm, which is usually positioned in a completely straight line, although there are occasions in which the extension is not total (Figure 8.3).


*Figure 8.3 Far location of away-from-body sagittal gesture.*

Sagittal gestures that have a backward motion, that is, they are performed towards the body, would have a slightly different classification. While determining a close, medium and far location is relatively easy when analysing away-from-the-body gestures, the amount of gesture space available for the speaker to gesticulate greatly varies in backwards sagittal gestures. For this reason, I suggest that we can only clearly distinguish between backward gestures that are performed close or far from the speaker, with no medium distance.

*Figure 8.4 Central towards-the-body sagittal gesture.*

Gestures performed in the centre would be located in a similar position as away-from-body gesture. The hand(s) would be covering the central space located directly in front of the speaker, with a semi-flexed forearm(s) and afterwards the speaker would flex the forearm(s) even more, or move the hand(s) towards the body. This type of gesture can also be performed with a vertical semi-flexed forearm in which the speaker fully flexes it, but with no movement in the upper arm (Figure 8.4). Both away-from-the-body and towards-the-body gesture would share the same central gesture space, with the only difference being the direction in which the gesture is performed.

As previously explained, I argue that backward sagittal gestures would not have a position that corresponds to medium distance from the body. Due to the anatomical difficulties of performing a backward gesture, it is difficult to find a middle point into the arm extension when performing a backward gesture. Speakers tend to employ the gesture space in front of them when performing a central gesture, but the amount of space they can cover is limited. In order to perform a gesture further from their body, they need to rotate the forearm until it has a vertical position and after that perform a backward rotation of the arm. Moreover, speakers occasionally perform some sort of gesture with the hand in order to signal backwards, such as pointing or performing different backward flaps (Figure 8.5).


*Figure 8.5 Far distance towards-the-body sagittal gesture.*

So far, I have covered the most usual gesture locations in which sagittal gestures can be placed. However, drawing on examples from my temporal gesture database, I have encountered several cases in which the speaker tries to extend the gesture further from their body through different means. I would like to suggest that speakers are also capable of locating a gesture even further away from their anatomical limitations. In

this case, the gesture would not be located at the limits of the gesture space of the speaker, but, through different means, the speaker would be trying to project the gesture even further; some of these cases include body motion and amplitude of gesture (Figure 8.6), repetition of the gesture and the temporal linguistic expression (Figure 8.7) or gaze (Figure 8.8).

Consider example 1 (Figure 8.6):

1) Just coulnd't do it, could you? You're from where? >> Holland>> How does one from…>> We take the boat. **Back in those days**, we too the boat. (KCET Tavis Smiley, 19-07-2008).

*Figure 8.6 Backward body motion and big amplitude backward hand flap.*

This example shows a backward sagittal gesture with the right hand that also involves body motion. The gesture starts with the speaker placing her right hand in front of her, with the palm facing left. Just as she says "back in those days", she performs a big-amplitude-gesture that goes from the gesture space in front of her to slightly behind her. Her arm slightly tilts backwards, as well as the palm and fingers of the speaker. Moreover, she also tilts her body so that it is easier for her to reach a location which is further back from her.

The next example, 2, (Figure 8.7) shows a case in which the linguistic repetition of the term that involves distance is mirrored by a repetition in the gesture.

2) Are you done having kids? >> Oh, yes, yes, yes. Absolutely. **Some day, far, far in the future**, when they're older if I meet somebody (KABC The View, 31-03-2010).

This example was already explained in section 7.7, but I would like to highlight the fact that the speaker repeats several times a spatial term, *far*, and that this repetition

is synchronised with several hand flaps that the speaker performs backwards. With both hands, the speaker performs several upwards gestures, with her palms rotating from a lateral position to a backward position.

*Figure 8.7 Repeated backward hand flap.*

The last example, 3, (Figure 8.8) shows a case in which the hand gesture is heavily accompanied by the gaze of the speaker:

3) And you know it's only going to get worse because rain's just a vague memory around here isn't it? >> That's right vague memory **so far in the future** don't even know Johnny. (KCBS CBS 2 News at 6, 18-07-2007).

*Figure 8.8 Sagittal forward gesture accompanied with gaze*

This last example shows the speaker performing a sagittal forward gesture with her left hand. She first moves her left hand upwards and then she straightens the arm, with her palm facing down. Most interestingly, during all the gesture her gaze is focused on the direction of the hand. The speaker then gestures and looks to a point located slightly above her.

In these cases, the speaker would not only be trying to locate the gesture far from their body, but they would be aiming to extend the gesture out from their gesture space. These types of gesticulations could be classified as *gesture extensions*. However, more research is required on these type of gestures, and a larger sample is required to establish whether or not these gesture extensions happen with other types of linguistic expressions. The examples in figures 8.6 - 8.8 present, in my judgement, cases in which the speaker could be extending the gesture further from their body, but nevertheless I categorised them as gestures located far from the body until further data is collected. These gesture extensions would belong to a grey area in which we would need to take into account different factors that belong both to the form and the function of the gesture. It would be necessary to study these types of gesticulations with further detail in order to detect patterns in other formal aspects of the gesture or in its function

#### 8.1.1.2 Gesture axial location

Another aspect that could be included in frameworks for gesture formal features has to do with the axial location of the gesture. Current gesture frameworks, including the one used initially in this thesis (see Bressem, 2013), state that co-speech gestures can be located in one of the three different axes of dimension: lateral, sagittal or vertical. Most of the frameworks also include information about the axial directionality that the gesture has, providing an accurate picture of where the gesture is produced and to which location it moves. With this information we can picture the movement and trajectory of the gesture in three dimensions; a gesture can, for example, move to the left along the lateral axis or upwards in the vertical axis.

However, during the analysis of some of the dataset of this thesis, a set of gestures that could not easily be located along any of the axes presented themselves. I analysed these gestures as *punctual gestures* since rather than being located along a particular dimensional axis with a direction, these gestures were employed to index a specific point in space (see chapter 6, section 6.2.2 for more information on punctual gestures).

Gestures located in the lateral, sagittal or vertical axis usually present a clear direction, with the hand moving to the left or in front of the speaker, for example. However, punctual gestures do not show this type of movement. When performing a punctual gesture, speakers tend to perform a slight downward or upward movement

with their hand and forearm. The main difference between punctual and vertical gestures is that, even though both present vertical movement, vertical gestures have directionality, and the hand follows a certain motion to a direction, while punctual gestures present a small and quick stroke employed in order to signal a point in space. Since they are just signalling a point in space, we cannot attribute a direction to the gesture.

When a speaker performs a lateral leftward gesture while saying, for example, *back in those days*, we can assume that the gesture is locating that temporal event (the past) on the left of the speaker. However, if instead of a leftward gesture the speaker performs a punctual gesture, we cannot define the directionality of the gesture. Even though the speaker is using a past expression, they are using a hand gesture to pinpoint a point in space, with no clear directionality.

Formal gesture features currently include a wide range of features to spatially locate gestures, but it seems that the framework could be further extended to include gestures that do not present a clear directionality and axial localisation. Hence, a fourth category could be included when analysing gestures in terms of their axial localisations. Gestures that are employed to signal a point in space presenting no concluding direction could be categorised as punctual gestures.

*8.1.2 Gesture studies and machine learning. New insights and possibilities*

The creation of the temporal co-speech gesture database employed for this thesis has been the first large-scale research of temporal gestures through big data. The nature, quality and size of the dataset has given me the opportunity to collaborate with a number of researchers in other disciplines in order to create new software and tools that could exploit the dataset. Perhaps the most important tool that has been developed thanks to the temporal co-speech gesture dataset obtained through the analysis performs in this thesis is a software that allows for machine assisted gesture analysis and, potentially, automatic gesture recognition.

This system allows for a range of automated searches like person and speaker detection; head, hand and shoulder detection and the recognition of timeline gestures. However, this tool is still at a very early stage of development, so constant feedback is required to refine its accuracy. This feedback consists on the human revision of the data that the program provides in order to offer a more accurate gesture analysis.All the

credit for the creation of this tool goes for Sergiy Turchyn and Soumya Ray, from the Department of EECS of Case Western University, amongst other contributors (see Turchyn et al. 2018), although the software required a large amount of gesture input that was obtained mainly through the timeline gesture data presented in this thesis. This program, then, can potentially recognise the different gesture features (axis, direction…) that have been included in this thesis, although it also can offer other information such as head or shoulder movement.

This software uses as a base the multimodal annotation tool ELAN to show the output of an automated tagging through different layers of analysis. Through a combination of supervised algorithms and unsupervised heuristics, the program detects different video fragments that contained relevant information. Each time the program detects one of the features that it can automatically recognise, it adds a time stamp on ELAN that covers the time frame in which that feature appears on the video (Figure 8.9; Turchyn et al 2018). The dataset of clips on temporal language employed in this thesis as well as other NewsScape-based clips containing other type of language and gestures were employed to train the classifier. Thanks to this, the program is not limited to hand gestures, but it also includes other features such as head or shoulder movement.

At its current state, the program can use face recognition to detect both people on screen and speakers. By looking at the motion on the lips, the program is able to

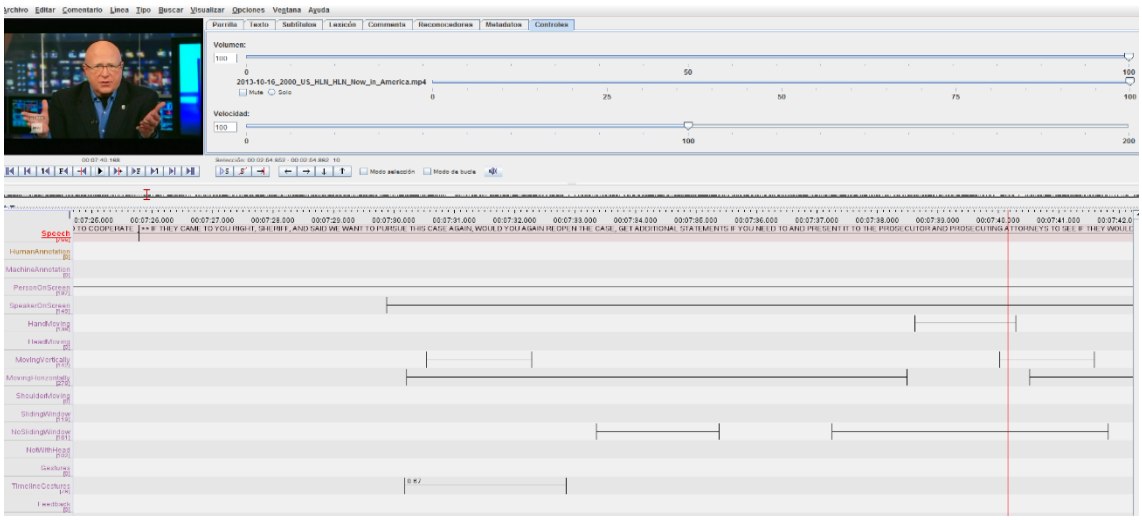


*Figure 8.9 Machine assisted tagging through ELAN*

detect if the speaker on screen is talking or not (Tuchyn et al, 2018). This feature is able to solve one of the biggest drawbacks in using the NewsScape dataset. As indicated in chapter 5, section 5.5, a large number of the clips obtained in the initial searches of

the dataset contained videos in which there was no speaker on screen. These cases had to be detected through a lengthy manual revision of the instances. By allowing the software to detect the presence or absence of the speaker, it would be possible to greatly speed up the data collection process when using the NewsScape library by excluding all these instances, reducing the dataset to cases in which the speaker is present on screen with a precision of 95% (Tuchyn et al, 2018).

The program can be also employed to detect different body parts that could be of interest for gesture scholars. Even though the accuracy is just around 50%, the program can detect head, hand and shoulder motion. The accuracy on the detection of these features is highly dependent on factors such as the colour of the background of the speaker, the angle of the camera or the number of speakers. Even though its accuracy is not very high at present, feeding the program with further input would increase its reliability. Furthermore, the program can be tweaked so that it adapts to the setting of a particular television program. The software would then be able to easily recognise the settings and camera angles that are employed in that television show, offering in exchange an increased accuracy (Tuchyn et al, 2018).

Finally, the program can also detect to some extent hand gestures. This feature is still at a very early state, but some promising results have been obtained through feeding the program with a large number of lateral timeline gestures. By looking at frame-by-frame hand motion as well as hand angles, the program is able to automatically recognise to some extent lateral gestures. Nevertheless, this feature is at an extremely early stage, and further development is required in order to improve it further.

The use of machine learning in gesture studies offers new possibilities for the study of gesture and cognition. Even though much further work on this area is required, the early results have been promising. The software is not able to perform accurate fully automated searches in its current state, but it allows gesture scholars to perform machine-assisted searches with an increase in efficiency and productivity (see Tuchyn et al 2018). In the long term, this tool would allow researchers to look for a particular hand or arm motion across a large amount of data. We would then, for example, be able to look at gesture-language patterns from a completely new perspective; we would no

longer look at the gestures that co-occur with language, but what linguistic patterns are recurrent with a certain type of gesture.

## 8.2 The NewsScape library: future work on gesture and cognition

The use of big data and more particularly the use of the NewsScape library has proven to be an invaluable tool for the study of gesture and (temporal) cognition. In this section I address a number of ideas and areas that could be addressed by using data from the NewsScape library.

I firstly discuss a perhaps controversial area that is related to the scripted nature of the data obtained through the NewsScape library, and its potential influence on gesture realisations. I discuss the validity of the data obtained through the NewsScape library and compare with other sources of linguistic and gesture data that are obtained by other means.

Secondly, I address certain topics related to time conceptualisation and co-speech gesture that, due to space constraints, could not be addressed and included in this thesis. The NewsScape library offers a wide range of possibilities to further study certain areas on time conceptualisation and co-speech gestures that includes the study of different temporal constructions, temporal metaphors that do not rely on the spatial dimension or cross-linguistic studies, for instance.

The last section is devoted to suggest a number of factors that might be of interest for the study of cognition and co-speech gestures. Here I cover a number of additional gesture features that could be taken into account in further research as well as some studies that could be conducted thanks to the NewsScape library, including the reasons that could influence gesture absence, the possible existence of multimodal discourse constructions or the influence of contextual elements in gesture realisation. The last point I cover here presents a number of examples that were extracted in some initial searches of temporal constructions that include cases in which speakers interact with external or imaginary physical objects in order to convey a temporal concept.

### *8.2.1 NewsScape and television setting: scriptedness and co-speech gestures*

The NewsScape library offers a large repository of multimodal data that can be employed for linguistic and multimodal analysis. However, one of the characteristics that makes this type of data different from other sources of data is that all the communicative exchanges that are produced here are performed in a television setting.

One of the criticisms that could be made of the NewsScape television database (argued in Chapter 5) is that, since the main source of data is made of television news shows, the communicative interactions are subjected to certain level of scriptedness, which may greatly vary from one situation to another. While some of the communicative exchanges that take place in television could be argued to have a high level of scriptedness, like a monologue, a political speech or someone reading from an autocue, there are other communicative exchanges that present a smaller degree of scriptedness. An interview between a journalist and an actress or the political discussion of the current state of the government in a talk show will cover a set of questions and topics that have been previously established, but the different communicative exchanges that emerge from them are highly spontaneous, providing high quality linguistic and gesture data.

In fact, it is very complicated to draw a clear line between what can be considered scripted communication and what can be considered spontaneous. The level of scriptedness that is found in certain television settings could be equated, for example, to some of the scenarios that are found in laboratory experiments in linguistic or gesture research. For example, one of the most common methods in gesture research consists of showing participants a film and then asking them to the retell the story to other subjects (McNeill & Levy, 1982). In this case, the communicative exchange that takes place is based on a previously agreed topic (in this case, a film) and the researchers may influence the conversation by asking certain questions (or through a questionnaire) that they find interesting in the same way that a television presenter could shift the conversation to the topic of interest. Other experiments, for example, gather their data by performing interviews about different topics like anecdotes or personal information about the speakers. (McNeill & Levy, 1982). Even though these types of communicative events can be considered to be spontaneous, there is also a certain degree of scriptedness on the communicative exchange that will take place: the topic of the conversation does not usually emerge spontaneously, but it is previously established, and the responses that the participants provide are subjected to a certain level of scriptedness, since they are based on events that they have experienced or previously talked about.

It is not clear either to what extent a scripted or non-scripted communicative exchange influence some multimodal factors such as gesture. The differences are not clear cut, and further research is required to establish what are the factors that condition

the level of scriptedness that a communicative exchange has. Some people may argue that a highly-scripted communicative exchange might present a lower amount of gestures, since the speaker needs to adjust to a strict script. However, there are some occasions like oral presentations or monologues in which, even though communication is highly scripted, the speakers performed a great amount of gestures due to factors that still remain unknown.

Rather than being considered something problematic, the television data that the NewsScape library offers should be considered a data source that shows a great variety of communicative situations that are of interest for research on multimodal communication. One of the possibilities that this dataset offer is, for example, being able to analyse settings that are considered to be highly scripted and analyse the reasons why certain linguistic and gesture realisations take place during those communicative exchanges. Some of the communicative exchanges that I personally find of great interest are political speeches, in which both the language and the gestures are often carefully measured in order to convey a certain message, and humoristic monologues, in which the usage of both language and gesture interacts in order to achieve humoristic aims. The multimodal communication employed in these two communicative genres could be analysed through the use of large amounts of data thanks to the NewsScape library, setting the first body of research on the relation between scriptedness and multimodal communication.

### *8.2.2 Temporal cognition and NewsScape library: future steps*

The multimodal television news database that the NewsScape library offers a whole new range of possibilities to continue research on temporal language as well as temporal expressions. As previously explained in chapter 5 section 5.3.1, the temporal linguistic key terms employed in this thesis were simple grammatical constructions (temporal adverbs, prepositions…). During the development of the thesis, a number of improvements to the CQPWeb search engine for the NewsScape library were made, which allow more refined searches when it comes to temporal linguistic expressions. The database was parsed with Stanford Core NLP and it also includes tagging for temporal expressions. This allows for a more refined search of temporal linguistic expressions that would allow, for example, look for calendric units (*minute, hour, day*) in the subject position as well as the verbs they co-occur with. This is especially interesting when looking, for example, to the type of motion verbs that might occur with

these temporal subjects and more precisely if they convey any type of speed or manner of motion. From the gesture perspective, similar characteristics to the ones investigated in this thesis could be investigated (e.g. axis, direction…) but more interestingly, other gesture features such as shape or gesture movement might show some interesting results when contrasting them with the temporal expression they co-occur with. More advanced features that could be included might deal with, for example, gesture speed, by measuring the frames that the gesture employs during the realisation of the stroke.

Another direction in which research on time conceptualisation could head is the investigation of other type of temporal linguistic metaphors as well as the gestures they co-occur with. In comparison to the TIME IS SPACE metaphor, little research has been conducted from the linguistic and gesture point of view into other temporal metaphors such is TIME IS A RESOURCE or TIME IS A CHANGER. Looking at both the linguistic and gestural realisation of these two metaphors is needed in order to further understand the different ways humans can conceptualise time. If gestures are found to frequently co-occur with these expressions, it would be interesting to see if the different patterns that have been observed to take place in spatial temporal language (see chapters 6 and 7) also take place when using this metaphor. Conversely, if there are differences in co-speech gesture in terms of axis, direction or shape, this might be contributing to the idea that a different conceptualisation is being used.

Even though this thesis has focused exclusively on the English television data found in the NewsScape library, a large amount of data from other languages can also be found in the database. Thanks to the collaboration of several universities across the world, the NewsScape library also includes a large repository of television news in Spanish, French, German, Arabic, Italian, Portuguese or Japanese, amongst other languages. The large amount of data and the variety of languages would allow for performing different cross-linguistic and cross-gestural studies focused on time conceptualisation. The data on gestural patterns obtained in this thesis could be compared to other languages to see what the similarities and differences across languages are. Languages like Japanese or Arabic, which have a different writing direction from the rest of European languages, could present different gestural patterns from the ones observed in English due to the influence of this cultural practice. The high amount of data as well as the easy access make it possible to obtain similar data

samples that could then be compared in order to obtained a clearer picture of the language-gesture relation in time conceptualisation across different languages.

Another interesting aspect that could be investigated through the NewsScape library by taking a cross-linguistic approach is the relation between temporal and gesture distance (see chapter 7). While English language only distinguishes between two types of spatial distance through the use of demonstrative pronouns (*this* and *that*), other languages like Spanish distinguish three types of spatial distance through the demonstratives *este, ese* and *aquel*, for close, medium and far distance respectively. As previously stated in chapter 7, English speakers tend to gesture close to them when talking about a "close" temporal event while they gesture far from them when the event is "distant". It would be very interesting to see if Spanish speakers are able to distinguish a third type of temporal distance, a "medium" distance, and, in case they do, if this is also represented by gesture.

Additionally, another topic that could be investigated by using data from the NewsScape library has to do with the possible relation between phonetics and gestures. Even though this idea is based solely on my own observations of the data, it is possible that there are occasions in which co-speech gestures are synchronised with different phonetic elements of speech. Through the observation of the temporal dataset, I have encountered some cases in which speakers may lengthen some vowels in words such as *long*, in order to add more emphasis, while producing a hand gesture. It is possible that the phonetic extension co-occurs with the gesture realisation; the longer the vowel is extended, the longer the gesture is. Even though this is a hypothesis based on general observation of the data, it would be interesting to see if indeed there is a correlation between phonetic extension or narrowing and gesture realisation.

One final topic on time conceptualisation that could be investigated through the NewsScape library has to do with the possible relation between temporal language, co-speech gestures and boundedness. Cienki et al (2016) have previously addressed this topic by looking at languages such as Russian, French and German, with some interesting results. The database that the NewsScape library offers could be employed to replicate this type of study, including a variety of languages and temporal constructions that could be investigated. Most importantly, big data could be employed

to see if there are any recurrent patterns when it comes to linguistic or conceptual boundedness and gesture realisation.

*8.2.3 Gesture and other factors*

The use of big data also allows researchers to investigate other factors in the language-gesture relation that are not easily studied with a smaller data sample. For example, researchers can use the NewsScape library to investigate if a linguistic construction co-occurs in a high or low frequency with a particular gesture realisation. Big data would allow research to investigate if a linguistic expression or a set of linguistic expressions tend to be produced accompanied with a co-speech gesture. An example of this can be found in part of the dataset that has been found in this thesis. The data obtained in this thesis (Chapter 5, 6 and 7) shows that 70% of the temporal expressions were accompanied by a co-speech gesture, while in the remaining 30% of the cases the speakers decided not to perform a co-speech gesture. This suggests that speakers are likely to perform a co-speech gesture when they are talking about temporal concepts.

The study of gesture and language frequency is not limited to temporal cognition. Recent research has investigated if certain entrenched linguistic constructions are more likely to appear with a particular set of co-speech gestures. Valenzuela (in press) looked at the expression "What do I care?" using the NewsScape library. Through the analysis of a wide number of examples, he notes that speakers tend to adapt a similar body posture and perform a similar gesture when they use this expression: shrugging of their shoulders, opening their arms and rising both forearms while their palms are facing upwards; usually accompanied with a small head shake. The author continues his arguments by stating a number of reasons which suggest that this combination of gesture and linguistic construction could be in fact a multimodal discourse construction. The data offered by the NewsScape library can result in an invaluable tool for the study of the potential existence of multimodal discourse constructions. The amount of information on gesture data that can be gathered on certain linguistic constructions through laboratory research can be somehow limited. However, the size of the NewsScape library would allow researchers to investigate if certain linguistic constructions can also be accompanied by particular gestures.

Another of the ideas that could be investigated through the data offered by the NewsScape library has to do with cases in which speakers decided not to perform a co-speech gesture with a linguistic expression, or gesture absence. As explained in chapters 5, 6 and 7, the data suggests that, when using a temporal linguistic expression, speakers are more likely to perform a co-speech gesture. Speakers performed some type of gesticulation 70% of the times, while in the remaining 30% they did not perform a gesture (see chapter 5, section 5.5 for more details on this). The data suggests that the tendency is to perform a co-speech gesture with a temporal expression, but what are the factors that make a speaker not perform a co-speech gesture? Through the analysis of a big set of data in which speakers do not perform a co-speech gesture with an expression that is usually accompanied by one, it could be possible to investigate if there are a set similar features amongst the different cases in which the speakers do not perform a gesture.

This task would indeed be very complicated, since there are several variables that might have an effect on the presence or absence of a co-speech gestures. One of the factors that I have observed from the dataset that might have an impact on gesture realisation is if the hands are free or not. There are several cases in the television in which the speakers employ a temporal expression while holding, for example, a microphone with both hands or holding pen and paper. Even though there are some exceptions, speakers usually do not perform co-speech gestures when they are not able to freely move their hands. Further research is suggested, but it could be possible that speakers avoid using certain hand gestures when their hands are busy with other tasks or holding certain items.

Another feature that might be linked to the presence or absence of the gesture could be related to the television setting; speakers might be aware of factors such as how visible their body is with respect to the angle that the camera is showing or if there are any extra elements on screen (graphs, images, videos) that might be covering part of their body. These factors might condition the realisation of co-speech gestures and perhaps, even cause a speaker not to perform a certain gesture.

The number of factors that could have an effect on the realisation or absence of a co-speech gesture can be very extensive. Features such as the topic the speakers are talking about, the spatial position of the speakers, the gestures that the speaker has

performed before and after that linguistic expression, individual differences or other discursive and prosodic features are some examples of factors that could be affecting the absence of a gesture in some cases. Studying gesture absence is indeed a very complicated task that would require detailed and thorough analysis. Even though it is a challenging task, the NewsScape library offers the data necessary to study this phenomenon in multimodal communication. This type of research does not have to be limited to the study of time conceptualisation and temporal language/gesture, but it could be easily extended to other metaphorical realisations that are often represented both in language and in gesture (physical or abstract metaphorical motion, embodied metaphors…).

The final idea that may be worthwhile investigating through the NewsScape library is a more detailed study of the type of gestures that are produced by some recurrent television personalities. The availability of large amounts of multimodal (and gesture) data of some television personalities like presenters or actors would allow for more detailed study of the type of gestures that each individual performs. We may find that each individual performs a different type of gesture when employing the same linguistic construction, or perhaps we could detect gestural patterns in each individual. Of course, more detailed research questions would be required to perform such study, but the NewsScape library offers the possibility to analyse gesture realisations of an individual through big data. While this is theoretically possible to perform in a laboratory setting, it would be very complicated to achieve due to constraints on time, resources and participants.

*8.2.4 External items, gesture and time conceptualisation*

The last point I worth mentioning in this section has to do with some interesting gesture realisations I found while analysing my data. Some of these gestures were performed with a different set of temporal linguistic expressions that were not included in the final version of the thesis, but nevertheless they present a very interesting case for future research. The following examples present a set of gesture realisations that co-occur with some temporal linguistic expressions. What makes these gestures somehow special is that it seems that the speakers are exploiting a physical object in their surrounding space or a point in space in order to anchor the temporal event.

While in the different co-speech gestures presented in this thesis speakers presented variations in direction or the axis in which the gesture was performed, these examples show how speakers modify their temporal gestures so that they can employ an object or a point in space to anchor the temporal event they are talking about. Consider the following examples:

4) Because they—because Steve Jobs, the head of Apple, is very famous for wanting to control, **from start to finish**, the design of his products—the hardware, the software, the whole thing (KCET Charlie Rose, 30-06-2007).

*Figure 8.10 Speaker using the phone as ending point of a process.*

Example 4 shows a case in which the speaker uses a physical object in his surroundings that may serve as a *material anchor* for the temporal period he is referring to through language. The speaker is talking about how involved Steve Jobs is with the production of his products. During part of the conversation he is holding a phone (probably and iPhone) in his left hand. At this moment, when he says that "Steve Jobs wants to control **from start to finish** the design of his products", he performs the gesture that is appreciated in Figure 8.10. The speaker performs a gesture with his right hand that co-occurs with the expression "from start to finish". The first stroke (from start) is performed to the right of the speaker (which is usually employed for future events) while during the second stroke he performs a leftward movement, with his hand stopping next to the phone that he is holding.

One of the hypotheses that may explain what is happening here is that, as previously explained, the speaker could be using the phone as a material anchor (see Chapter 2, section 2.3) in order to refer to the end of the process. The speaker is talking about the control of the inventor of that type of devices during the production of the same. When he gestures to his right, referring to the past, he might be referring to the

beginning of the process of making a phone; while when he gestures to the left he uses the phone as an anchor to signal which is the end of the process he is referring to. The end, the "finish" of Steve Jobs design of his products is the item that he is showing. The phone one additional element that is located in the timeline, and it has a specific purpose (indicating an end point).

Another example that is very similar to this one can be found in example number 5:

5) I didn't read the book, I don't know what else is in there so I don't know if it begins **from the origin of man until the future**, or if it is **s**pecific to that day. (KOCE Charlie Rose, 25-10-2012).

*Figure 8.11 Speaker using a book as the starting point of an event.*

In this case, the speaker is talking about the contents of a book. While she says that she does not know if the book begins with the "origin of man", she performs a pointing gesture with her right hand in a palm up, open hand position to the book. After that, she refers to "until the future", talking about the contents of the book, while she performs a rightward gesture with her hand, rotating her elbow until the palm of the hand is facing laterally (Figure 8.11). This case offers a similar example to example 1. In this case, the speaker is including a physical object, the book, in order to talk about the temporal development of the same. She uses the book to indicate where "the origin" is located in space, while performing a rightward gesture for indicating the future. Similarly to the previous case, a physical element is included in the temporal construct, although in this case it is employed to locate the beginning of the event.

There are other cases that show a similar tendency on how physical objects might interact with temporal conceptualisation, although they present some differences. Consider the following example:

6) … He tells you when to put the pot on, when to sauté the onions. You don't have to think. **From start to finish**, he holds your hand…

In the case, the woman on the right side of the images in Figures 8.12 performs a gesture with her left hand that is very similar to the previous example. Here, the speakers are talking about the advantages of a new cooking book, "Food 52", which is located at the left bottom of the images. While taking about the book, the speaker on the right performs a vertical top-to-bottom gesture while saying "from start to finish", referring that the book, or rather the person who wrote the book, will hold your hand through the cooking process "from start to finish".

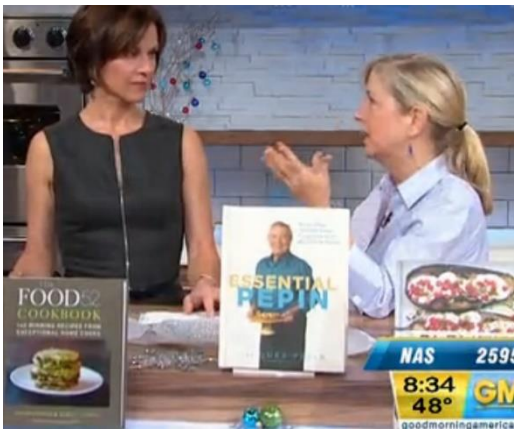



*Figure 8.12.1 Gaze focused on interlocutor before the gesture*


*Figure 8.12.2 The gaze is directed to the book. The gesture is prepared.*


*Figure 8.12.3 The speaker performs a downward gesture with her left hand.*

The first interesting detail about the gesture has to do with the gaze of the speaker. Just before performing this lateral gesture and employ the temporal linguistic construction, her eye gaze moves from her interlocutor to the book she is talking about (figure 8.12.1 and figure 8.12.2). Once her focus has shifted from the interlocutor to this other item in space, she performs this vertical downward gesture, that might mirror top-to-bottom direction when reading, making reference to the cooking process that is described in the different recipes in the book. It could be argued that, similarly to the previous example, the speaker is employing an item that is in her immediate situational context together with a linguistic expression and a co-speech gesture in order to convey a temporal concept; the speaker is using linguistic, gestural and contextual elements in order to express a precise temporal concept. Nevertheless, there are some differences between both examples. In this case, the book is not part of the timeline, and it is not interacting directly with it. However, the eye gaze as well as the vertical gesture referring to the book seems to indicate that the presence of the item is having an effect of the gesture (and, potentially, conceptual) representation of the temporal concept.

The last example presents a case in which, again, the speakers interacts with an object in their surroundings in order to convey a temporal concept. However, this time the material anchor, or rather the *imagined material anchor* is created through the interaction of the gestures the speakers produces and a table at which the conversation is being held:

*Figure 8.13.1 On the left, the speaker performs a downward gesture with his left hand. On the right, he performs a second stroke with his left hand, creating a timeline between both hands.*

7) …I think, well, that's the Jewish story—Exactly, exactly. **From genesis to revelation**. There are people who wrote the word of God that we honor and love under **Egyptian oppression, Syrian oppression, Babylonian oppression, Persian oppression, Greek oppression, Roman oppression**…

In this example, the speaker is talking about the Jewish history from a religious point of view. The first temporal gesture he performs co-occurs with the temporal expression *from genesis to revelation*. At this point, the speaker performs a gesture that consists of two different strokes. First, he moves his hand to his right in order to located *the genesis* (Figure 8.13.1, left), and then he moves his left hand to the right in order to locate *the revelation* (Figure 8.13.1. right). In both cases, his hands are opened and facing inwards. However, the gesture does not conclude there. The speaker holds that position for a few seconds, and after that he employs the timeline he has created with his hands in order to locate a number of temporal events that correspond to that timeline.

*Figure 8.13.2 The speaker signals different points in the timeline using his left hand.*

The hands of the speaker are employed as temporal limits to the sequence of events he refers to. After producing this initial gesture, the speaker performs a number of downward strokes with his left hand, while it remains opened and facing inwards. Meanwhile, his right hand remains stationary, facing to the opposite direction. The speaker performs a number of strokes that correspond to each of the temporal periods he is referring to in his speech: Egyptian, Syrian, Babylonian oppression… Each of the linguistic constructions is accompanied by a rightwards stroke of his left hand, until it reaches the same position of the right hand when he mentions Roman oppression (Figure 8.13.2).

In this case, the speaker is also employing a physical object in order to express a temporal concept, but it is slightly different from the previous cases. While examples 1 and 2 presented cases in which speakers employed props, or physical objects that were located in their surroundings, in this case the speaker is creating his own timeline by using his hands as the limits of the same. After creating this imaginary timeline, he starts manipulating it in different ways in order to convey temporal information with a higher amount of precision than if he had conveyed it just by using language.

The four examples that I have presented here suggest that temporal cognition and more precisely the interaction between gesture and language when communicating about temporal cognition might be indeed much more complicated. Even though further research is required, as well as a higher number of examples, it seems that humans are not limited to use language and gesture to convey temporal messages, but they also integrate contextual elements to the temporal concept they are trying to convey. From the information presented through example 3, it seems that not only are humans able to use contextual elements, but, if they think it is necessary to employ a physical item so that the message is clearer and more precise, they are also able to create them online and manipulate them according to their needs. The picture that emerges is that temporal cognition and timelines may follow certain patterns that have been discussed throughout the study of temporal cognition (direction, axis… see Chapter 2, section 2.2 and 2.3, as well as Chapter 3, section 3.1 and 3.2 and Chapter 4, sections 4.3.2), but when looking at more natural communicative situations we can observe how conveying temporal messages is not as clear and easy as we initially thought.

Even though these examples show how speakers can interact with physical objects when conveying a temporal message, they do so in slightly different ways. It could be possible to suggest a typology based on these. This classification is by no means extensive, but rather serve as a basis for gesture-physical object interaction in time- conceptualisation. The first category would include examples in which the speakers include contextual elements and items in a physical temporal representation. In this case, the different items interact with the timeline in different ways; the speaker can employ certain elements to indicate that they are, for example, the beginning or the end of the event or an element that refers to the past. The second group would include cases in which the physical item is not directly included in the timeline, that is, it does not participate in the physical realisation, but it is present in the surroundings and it has an influence on how the speaker is conceptualising that particular event. Lastly, the third group would include cases in which the speaker has no direct access to a specific physical object, and thus he or she makes an imaginary item by using hand gestures. This would include, for example, the online creation of a timeline in order to designate different temporal periods in the same.

### *8.3 Conclusion*

This chapter has aimed to address a number of issues relating to methodological aspects of gesture research as well as the development of new tools for the study of multimodal communication. Furthermore, I have discussed several possible areas of study both in temporal cognition and gesture studies that could benefit from the amount and the type of data offered by the NewsScape Library.

Concerning methodology, I have made some suggestions that address the current gesture analysis frameworks which are employed to analyse gesture space and gesture axis.

My first suggestion is that there might be some type of gestures, mainly gestures located in the sagittal axis, that cannot be objectively located in the gesture space. Gesture frameworks use the human body as a reference frame to describe the position of the co-speech gestures and, even though this method is valid and highly accurate for lateral and vertical gestures, it is vaguer when analysing sagittal gestures. In order to achieve a more objective analysis of the gesture localisation, I suggest that a set of physiological characteristics, including the levels of flexion and extension of the

forearm and the upper arm as well as the rotation of the elbow and the shoulder. I also suggest that gestures which are performed in the sagittal axis, with a backward direction, can also be located close or far from the body; the reduced amount of gesture space available for performing a gesture does not permit speakers to perform these gestures in a middle distance.

I finally propose that a fourth type of gesture distance different from the classic *close, medium, far* configuration can occasionally emerge. This fourth level in gesture distance would include cases in which speakers try to extend a gesture further from their gesture space by emphasising the gesture throw the gaze, body motion or the amplitude of the gesture.

The second suggestion for an improvement in current gesture analysis frameworks has to do with the inclusion of gestures that do not present a definite axial movement, or punctual gestures (see Chapter 6, section 6.2 for more information). These types of gesture seem to be more frequent than expected, and the lack of axial movement highly differentiates them from the rest of gestures that are performed along one of the spatial axes.

Another of the methodological improvements that I have discussed in this section is the development of new software for the automated recognition and analysis of gestures. This type of machine assisted gesture analysis will have a very positive impact on future research on gesture studies as well as multimodal communication. This program uses as a base the gesture dataset obtained in this thesis to elaborate a software that is able to recognise a number of bodily and gesture features in multimodal data. The software presents high accuracy when recognising people or speakers or screen and, even though further improvement is required, it is showing promising results when performing automated searches of concrete gesture realisations (e.g. lateral or vertical gestures).

Secondly, this chapter covered a series of possible areas of research that would be of great interest for different areas in time conceptualisation, gesture studies and multimodal communication.

The first topic has to do with the influence of scriptedness in the television setting, and its possible influence in linguistic and gesture communication. Television data includes communicative exchanges that are highly scripted, like monologues and

political speeches, while in other occasions the level of scriptedness is more sublte, like an interview or debates in news shows. It is difficult to establish a clear distinction between a type of communication that is spontaneous and scripted communication; data collected through interviews in a laboratory is often considered to be spontaneous even though it follows a topic and a number of questions that have been pre-established. Additionally, the study of highly scripted television communicative exchanges would be highly interesting for research in gesture studies; even though a television monologue and a news reader follow a very precise script, there is a large difference in the type of gestures they produce.

I also include in this section a number of possible topics that could be researched through the NewsScape library in order to further expand the topic of time conceptualisation. The first study that could be conducted after this thesis could focus on investigating the linguistic and gesture patterns of different type of temporal linguistic constructions in the NewsScape library. This would include, for example, the analysis of temporal subjects (temporal units, event…) and their interaction with different spatial verbs. Some of the particularly interesting areas would include the manner and speed of motion.

Research on type conceptualisation does not need to be limited to the TIME IS SPACE metaphor. The NewsScape library could be also employed to perform linguistic searches of other temporal metaphorical realisations, for instance TIME IS A RESOURCE. Analysing the linguistic constructions as well as the possible gesture realisations that co-occur with them is likely to yield some interesting results on temporal cognition.

Another follow up study that could be conducted could replicate the research questions of this thesis in different languages also present in the NewsScape library. Spanish, for example, presents an interesting case; it could be investigated if, since Spanish makes a distinction amongst three types of spatial distance (close, medium and far), this distinction is also present in gesture and temporal cognition.

The last topic on time conceptualisation that I addressed in this section is inspired by Cienki et al (2016), who investigated gesture and conceptual boundedness in German, French and Russian. Replicating this type of study including a wider variety of linguistic constructions as well as including different languages through the use of

the NewsScape library would be highly beneficial to continue research in this area. By using big data, it would be possible to detect and establish clearer patterns of gesture realisation and temporal boundedness.

Finally, this section also included a number of ideas that could be of interest for researchers on multimodal communication and gesture studies. The NewsScape library can be employed to investigate a wide range of dimensions in multimodal communication, including the possibility of multimodal discourse constructions. Similar to the study conducted by Valenzuela (in press) on the linguistic expression "What do I care?", there might be other types of linguistic construction that tend to be paired with particular co-speech gestures.

Another interesting, albeit very complicated research topic would focus on investigating the reasons that might influence gesture absence. The data on temporal gestures indicates that speakers gesture about time 70% of the occasions. However, what are the reasons why speakers do not gesture 30% of the times? The list of factors that might influence this is indeed very extensive: distribution of the speakers, topic of the conversation, genre, presence or absence of items in the hands… Little research has been performed on gesture absence, but through a careful and precise analysis of the data it would be possible to establish a set of common factors that condition gesture absence.

Another area in which the NewsScape library could be useful has to do with one of the features of the nature of the database. Since it is a television corpus, there are certain television personalities that are frequently featured on screen. This particularity could be exploited to gain access to large amounts of data of a particular speaker, which could show some interesting information on different areas. By analysing data of the same television personality in different areas (interviews, debates, monologues) we could see if there is any type of difference in their gesture realisations depending on the communicative situation they are in. Other features, such as individual differences in gesture realisations could be also analysed.

The last point I have discussed in this chapter tackles the possible inclusion of physical objects when speakers convey a temporal concept. In this section, I have shown three examples in which speakers interact with other items in order to convey a temporal idea. I also proposed that three types of interactions can be given: 1) the physical object

can be present in the context and be used as a part of the temporal construct (e.g. the starting point or the end point); 2) the item can be present in the context but it is not employed in the timeline, although it triggers a particular gesture in the speaker (e.g. performing a vertical gesture when referring to a book); 3) there is no physical item and the speaker creates an imaginary item (e.g. timeline) that they manipulate when conveying a temporal idea.

This chapter shows that there is still much research to be performed on the topic of gesture and temporal cognition. Most of the research on temporal cognition that has been performed so far has focused on the linguistic dimension, and the experiments that have covered the co-speech gestures usually contain a limited amount of data. The NewsScape library is an extremely useful tool that allows research to perform a wide amount of studies by using high quality data. Thanks to this tool, different types of research that were not possible before can now be performed.

# Chapter 9

# Conclusion

## *9.1 Overview of the thesis*

The aim of this thesis has been to investigate patterns in time-space conceptualisation through the analysis of linguistic data and temporal co-speech gestures. More precisely, it has exploited the multimodal television corpus of the NewsScape Library, an extensive archive of television data which includes both textual data formed by subtitles as well as audiovisual data from news and talk shows. The use of this database has enabled a large-scale analysis of temporal co-speech gestures in a contextualised environment based on big data.

The thesis started by offering a general overview of the main tenets of Cognitive Linguistics, as well as the main cognitive processes based on Gestalt psychology. It also covered research on Temporal Frames of reference, time conceptualisation in language and the role of cultural artifacts in temporal cognition. The thesis also covered the importance of multimodal research on time conceptualisation, and more precisely the importance of gesture.

The data obtained in this thesis through the gesture analysis of the NewsScape archive offered a wide amount of gesture information for the study of time conceptualisation. The analysis of the co-speech gestures that were produced when speakers used non-spatial and spatial language resulted to be very fruitful, with a number of interesting trends that can be observed in the data. The first finding confirms that temporal co-speech gestures are often located in the lateral axis, being the most common type of gestures in non-spatial language as well in the two variants of spatial language (directional and non-directional).

However, when looking at gestures that are produced on the sagittal axis, some variation can be appreciated. The data suggests that non-spatial language presents a very low ratio of sagittal gestures, but it increases when looking at spatial language. The data suggests that when speakers use the spatial domain, even though the most employed axis is the lateral one, there is an increase in how often they use the sagittal axis Moreover, there were also differences between non-directional and directional language; linguistic expressions that did link a the past or the future with a particular

direction (e.g. *distant past*) presented a higher number of sagittal gestures that language which explicitly stated the direction of time (e.g. *back in those days*). One of the possibilities that could explain this increase in the sagittal axis is that since these spatial expressions are deictic, they might be using the speaker (ego) as a spatial anchor to talk about temporal concepts. In this case, the gesture would be taking as a referent the ego, which could explain the higher number of sagittal gestures. In the case of non-directional language, which shows the highest ratio of sagittal gestures, the gesture could also be employed to cover an information gap that is not present in language. Since this type of language does not indicate where the past or the future are located in relation to the ego, a gesture located on the sagittal axis could be covering this temporal direction information gap.

An unexpected finding that was obtained through the analysis of the dataset is that there were also a number of cases in which speakers performed gestures in the vertical axis. Even though the number of cases was quite scarce, the use of vertical gesture is not frequently reported in temporal gesture literature. Moreover, another type of gesture that showed no axial movement was also located in the dataset. Provisionally, these gestures have been categorised as "punctual", since rather than being performed alongside an axis with a particular direction, they are employed to signal a point in time.

Another aim of this thesis was to investigate whether or not co-speech gestures are congruent with the canonical direction of time, which in Western cultures locates the past behind, on the left and above the speaker, while the future is in front of, on the right and below them. The first finding is that gestures tend to be congruent with the canonical direction of time but this congruency level varies across non-spatial, spatial directional and spatial non-directional language. Non-spatial language presents a very similar rate of congruent and incongruent gestures. One of the reasons that might explain this is that since this type of language is not using the spatial domain, gestures do not need to reflect the directionality that time is attributed through its mapping with space. By exclusively using the temporal domain, speakers would not need to be consistent with the direction that is attributed to time through the use of the spatial domain (e.g. left-right=past-future). It would be interesting if this effect could be mirrored through laboratory-controlled conditions by priming English speakers through the use of purely temporal language and priming a different group with space-time

temporal linguistic expressions; and then observing if there are any differences on where they locate temporal events in, for example, a timeline.

When looking at spatial language, there is also a difference between directional and non-directional expressions. Directional language presents a slightly higher ratio of congruent gestures than non-spatial language. When speakers use the spatial domain to talk about time, they are more likely to gesture according to the canonical direction of time, but the effect is not too strong compared to non-directional language since, in this type of language, directionality is already explicitly stated in the linguistic expression. The gesture would be then mirroring information that is present in language. In the case of non-directional language, the amount of incongruent gesture greatly diminishes. In this case, gestures are congruent with the canonical direction of time in most of the cases, perhaps because the gesture is employed to indicate where the temporal period that is being addressed linguistically (e.g. *far in the future*) is located, since this information is not encoded linguistically. Differently from non-spatial expressions, non-directional language employs the spatial domain. This link to the spatial domain, and more precisely to the directionality that is attributed to time through its mapping with space, would be represented through gesture since it is not linguistically encoded.

The data also shows some differences between expressions that refer to past and future events. No significant differences were observed when analysing the direction or axis of past and future-related language. However, some differences were observed when looking at the congruency levels. It seems that spatial non-spatial past expressions' sagittal co-speech gestures tend to be incongruent with the canonical direction of time. In this case, speakers would gesture away from them when referring to past events. This trend is reversed in non-directional language, presenting a much higher number of congruent past co-speech gestures. In the case of future-related expressions, both types of spatial language present a high congruency ratio, that is, most of the gestures are performed away from the speaker. Even though there is not a definite answer that could explain this phenomenon, it could be hypothesised that since non-spatial temporal language has no spatial bias, the speaker will be less likely to perform gestures that follow the past-back and future-front pattern. This high rate of incongruent sagittal gestures in non-spatial past language could be related to the anatomical constrains of sagittal gestures; performing a backward gesture is anatomically more difficult, with a lower amount of gestural space and more demanding movements.

Another explanation could be linked with the idea of these forward gestures falling within the category of presentational gestures. Nevertheless, this hypothesis would not explain why the tendency is reversed when speakers used non-directional language.

The final question that this thesis addressed was if linguistic temporal distance was analogous to co-speech gesture spatial distance. The first finding is that we are more likely to perform a co-speech gesture when we talk about temporal distance than when we talk about temporal proximity. One possible explanation for this has to do with the complexity and subjectivity of the concepts of proximity and distance. Proximity generally refers to a relatively controlled number of spatial configurations that fall within a subjective barrier we establish and the ego. On the other hand, distance refers to all the remaining spatial configuration that are beyond that subjective barrier that we establish. The picture that emerges is that, while proximity refers to a controlled number of spatial configurations, distance refers to an unlimited amount. The higher frequency in gestures that are related to temporal distance could be related to the higher complexity of the concept of distance, while temporal events that are situated in the proximity would be easier to located. From the proximity-related language dataset, only a 30% of the data contained a co-speech gesture, while this ratio was double (60%) in the distance-related language dataset. Another finding is that it seems that the distance covered by the co-speech tends to be analogous to the temporal distance addressed by the linguistic expression. When speakers use temporal proximity-related language they tend to gesture close to their bodies, while the gesture far from their bodies when using temporal distance-related language. The picture that emerges is that temporal co-speech gestures are not limited to the representation of dimensions such as direction or axis, but also other spatial properties such as distance.

The analysis of this data suggests that gestures are mirroring the information concerning the distance of the temporal event that is present in language. It could be argued then that the gesture is actually not related to the conceptualisation processes of time, but rather it is used in order to mirror the relevant lexical items that are being employed. Even though it may be possible that the gesture is representing the lexical items rather than representing the temporal concept, this idea in in conflict with some of the findings that were obtained when analysing the direction and axis of the gesture. As it was previously argued, speakers only employ the sagittal axis a 21% of the occasions when using spatial language that explicitly states the direction of the temporal

period (e.g. *back in those days*). If the main purpose of the temporal gesture were to mirror language, we would expect a much higher frequency of sagittal gestures in this type of language. In the same way, the high number of lateral gestures that is employed when talking about time is not mirroring any components that are present in language. We could assume then that since temporal gestures do not consistently mirror the spatial dimension of axis that is present in language, the main objective of the gesture is not to mirror language in these cases.

Of course, it could also be possible that, in cases in which speakers mention temporal distance, gestures are in fact mirroring the lexical item, but this would mean that speakers would tend to mirror lexical aspects concerning distance through gesture while they do not do so when using lexical language concerning direction. Future research could expand this notion to the use of calendric units or other temporal events. Through the analysis of this type of data, it could be assessed whether or not the gestures that co-occur with these expressions are also representing the dimension of distance. Nevertheless, it should be noted that calendric units usually have a very strong subjective bias; we can talk about the birth of our first child as if it were yesterday, even though it happened 30 years ago, while we could consider a fight between us to be *far in the past*, even though it happened a month ago. Through a more detailed qualitative analysis of the context of the different calendric units, it could be establish if the speaker is considering the event as being something located in the temporal proximity or is temporally distant. This could be then contrasted with the gesture in order to establish if gesture distance is indeed related to temporal distance.

Finally, the data also suggests that gesture shape may contribute semantic information relating to temporal distance. The data shows some cases in which speakers used pointing gestures, usually with their thumb pointing backwards, in order to indicate a more precise moment in time through gesture shape. Conversely, there were occasions in which the gesture shape presented a more relaxed hand, usually performing several flaps. In this case, the gesture shape could be suggesting that the speaker is not sure about the exact temporal distance of the temporal event.

### *9.2 Contribution to the field*

This thesis has contributed in different ways to the current research on time conceptualisation and co-speech gestures by performing a large scale study of temporal

gesture in a multimodal corpus. The use of big data highly benefits the study of cognition, conceptualisation and multimodality, since it allows researchers to look at patterns that are repeated across a large number of instances. Moreover, the multimodal nature of the dataset makes it an extremely useful tool to explore extra-linguistic features, such as gesture. Gesture studies have typically relied on data obtained from laboratory experiments but through the use of multimodal corpora such as the NewsScape, the availability as well as the quantity of the data could increase. The use of the NewsScape library makes it possible to test hypothesis on a larger, more authentic dataset. Moreover, the resulting temporal gesture database that has been created is one of the largest temporal gesture repositories that have been created so far, and its data could be employed to perform further research on time, cognition and gesture but also in other fields concerning, for example, emotion or multimodal constructions

The analysis of the dataset has confirmed the hypothesis that gestures tend to rely on the lateral axis, although there is also evidence that suggests that gestures can be also performed in the sagittal or vertical axis as well as present no axial movement. Moreover, it also has suggested new hypotheses concerning language-gesture relations, more particularly the possible relation between axis, direction and congruency, and the type of (spatial) temporal language that speakers use. This thesis also investigated the possible relation between temporal duration and gesture location, achieving very promising results. Finally, it has been suggested that gesture shape could also be semantically significant when observed in gestures that aim to indicate temporal duration. Following the steps of Boutet et al. (2016) it seems that linking this phenomenon to linguistic aspect and gesture boundedness might be promising.

Finally, this research has also made a number of methodological contributions. Firstly, the gesture dataset that was compiled for the completion of this project has already been shared with other researchers in order to elaborate and improve software and tools for multimodal annotation. Some of these examples are the automatic gesture recognition software created by Tuchyn et al. (2018) and improvements on CQPWeb for NewsScape partially based on my feedback or the elaboration of a rapid gesture annotator.

Lastly, this thesis has also suggested a number of methodological changes in the current gesture frameworks that could solve some of the problems that were

encountered during the development of the thesis. The changes include the expansion of axial gesture localisation so that it includes gestures that have no axial movement as well as a physiological approach to gesture location; more precisely the location of sagittal gestures.

### *9.3 Limitations and implications for future research*

In spite of the number of contributions of this study to the field, there are a number of limitations that also need to be acknowledged. The first one is that there are a large number of temporal linguistic structures that could not be addressed in the analysis. The repertoire of temporal linguistics expressions in English is extremely rich, and here I have only addressed a small fraction of the different temporal constructions that can be employed in English. Further research would highly benefit from performing both linguistic and gesture analysis on different types of temporal expressions, for instance looking at the different linguistic and gesture patterns of calendric units (*e.g. year, month, day*), looking at constructions that employ an Ego-Moving as well as a Time-moving perspective or constructions that provide temporal concepts with speed or manner of motion (*e.g. time flies, the match lasted an eternity*).

Another limitation is that the clarity of some of the gestures as well as the accuracy of part of its analysis (*i.e. gesture location*) was not optimal due to the different angles and positions that the camera focuses on in television setting. More refined equipment, such as motion detection, would be an invaluable tool for studying the exact distance that a gesture has covered, but this would have been in detriment of the large amounts of data that were obtained.

There are also several topics that, due to space limitations, I have not been able to discuss in depth. Further research should address more thoroughly "punctual" gestures. It is highly possible that within this broader category of gesture different subgroups can be made, particularly if further analysis is performed on features such as gesture shape. Another topic that I have not been able to cover in depth has to do with the interaction between language, gesture and extra-linguistic items when conceptualising time. This type of information is unique from this multimodal corpus, since during laboratory experiments participants are often located in a neutral space in which there are not many props they can interact with. It seems that time

conceptualisation might go beyond language and gesture, with some temporal concepts mapped onto other contextual elements.

As a final reflection I would like to suggest that we are only scratching the surface of time conceptualisation and the language-gesture interface. Most of the research that has been conducted on this topic has focused on either isolating language or performing experiments with, sometimes, restrictive environments. Conceptual knowledge and, of course, temporal concepts are conveyed through language and through gesture, but these are only two of the components of a much larger set of factors that co-occur in communication. Studying the different components separately through laboratory experiments is of course useful and necessary, but we also need to study how these elements interact in real, contextualised communicative situations.

# References


Barcelona, A, Ibarretxe-Antuñano, I. & Valenzuela, J. (2012). *Lingüística Cognitiva.* Barcelona: Anthropos.

Bates, E. & Dick, F. (2002). Language, gesture, and the developing brain. *Developmental Psychobiology, 40*, 293-310.

Bavelas, J., Chovil, N., Lawrie, D. & Wade, A. (1992). Interactive Gestures. *Discourse Processes, 15*(4), 469-489.

Bender, A., Beller, S. & Bennardo, G. (2005). Spatial frames of reference for temporal relations: A conceptual analysis in English, German, and Tongan. In L. B. B. Bara, *Proceedings of the Twenty-Seventh Annual Conference of the Cognitive Science Society* (pp. 220-225). Mahwah, NJ: Lawrence Erlbaum.

Bender, A., Beller, S. & Bennardo, G. (2010). Temporal frames of reference: Conceptual analysis and empirical evidence from German, English, Mandarin Chinese and Tongan. *Journal of Cognition and Culture*, 283-307.

Bennardo, G. (2002). Mental images of the familiar: Cultural strategies of spatial representations in Tonga. In G.Bennardo, *Representing space in Oceania: Culture in language and mind.* Canberra: Pacific Linguistics.

Bohle, U. (2014). Contemporary classification systems. In A. C. Cornelia Müller, *Body-Language-Communication: An International Handbook on Multimodality in Human Interaction* (pp. 1453-1461). Berlin/New York: de Gruyter.

Bonifacio, G. (1616). *L'Arte de cenni.* Vicenza.

Boroditsky, L. & Gaby, L. (2010). Remembrances of Times East: Absolute Spatial Representations of Time in an Australian Aboriginal Community. *Psychological Science, 21*(11), 1635-1639.

Boroditsky, L. & Ramscar, M. (2002). The roles of body and mind in abstract thought. *Psychological Science*, 185-189.

Boroditsky, L. (2000). Metaphoric structuring: Understanding time though spatial metaphors. *Cognition, 75*, 1-28.

Boroditsky, L. (2001). Does Language Shape Thought?: English and Mandarin Speaker's Conceptions of Time. *Cognitive Psychology, 43*, 1-22.

Boroditsky, L. (2011). How Languages Construct Time. *Space, Time and Number in the Brain*, 333-341.

Bottini, R., Crepaldi, D., Casasanto, D., Crollen, V. & Collignon, O. (2015). Space and time in the sighted and blind. *Cognition, 141*, 67-72.

Boutet, D. (2010). Structuration physiologique de la gestuelle: Modéle et tests. *Lidil*, 77-96.

Boutet, D., Morgenstern, A. & Cienki, A. (2016). Grammatical aspect and gesture in French: A kinesiological approach. *Russian Journal of Linguistics*, 132-151.

Bressem, J. & Müller, C. (2013). A repertoire of recurrent gestures in German. In A. C. C. Müller, *Body-Language-Communication. An International Handbook on Multimodality in Human Interaction* (pp. 1575-1592). Berlin/Boston: De Gruyter Mouton.

Bressenm, J. (2013). A linguistic perspective on the notation of form features in gestures. In A. C. C. Müller, *Body-Language-Communication: An International Handbook on Multimodality in human Interaction* (pp. 1079-1098). Berlin, Boston: De Gryter: Mouton. Retrieved 5 24, 2015, from http://www.janabressem.de/Downloads/Bressem_notating%20gestures-rev.pdf

Brown, S. (1997). Attentional resources in timing: interferences effects in concurrent temporal and nontemporal working memory tasks. *Perception & Psychophysics, 59*(7), 1118-1140.

Bueti, D. & Walsh, V. (2009). The parietal cortex and the representation of time, space, number and other magnitudes. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*, 1831-1840.

Bulwer, J. (1644). *Chironomia: or, the art of manuall rhetorique with the canons, lawes, rites, ordinances, and institutes of rhetoricians, both ancient and moderne, touching the artificiall mamaging of the hand in speaking.* London: Harper.

Bylund, E. & Athanasopoulos, P. (2017). The Whorfian Time Warp: Representing Duration Through the Language Hourglass. *Journal of Experimental Psychology: General*, 911-916.

Cai, Z., Connell, L. & Holler, J. (2013). Time does not flow without language: Spatial distance affects temporal duration regardless of movement of direction. *Psychon Bull Rev*, 973-980.

Calbris, G. (1990). *The Semiotics of French Gestures.* Bloomington: Indiana University Press.

Calbris, G. (2003). From cutting an object to a clear cut analysis: Gesture as the representation of a preconceptual schema linking concrete actions to abstract notions. *Gesture*, 16-46.

Casasanto, D. & Boroditsky, L. (2008). Time in the mind: Using space to think about time. *Cognition*(106), 579-593.

Casasanto, D. & Bottini, R. (2010). Can mirror-reading reverse the flow of time? In T. S. C. Hölscher, *Spatial cognition VII* (pp. 1342-1347). Berlin Heidelberg: Springer. Retrieved March 19, 2014, from http://mindmodeling.org/cogsci2010/papers/0371/paper0371.pdf

Casasanto, D. & Jasmin, K. (2012). The Hands of TIme: Temporal gestures in English speakers. *Cognitive Linguistics*, 643-674.

Casasanto, D. & Yap, D. (2016). The semantics of beat gestures. *Conference presentation: ISGS7.* Paris.

Casasanto, D. (2009). Embodiment of Abstract Concepts: Good and Bad in Right- and Left-Handers. *Journal of Experimental Psychology: General, 138*(3), 351-367.

Casasanto, D., Boroditsky, L., Philips, W., Greene, J., Goswami, S., Bocanegra-Thiel, I. et al. (2004). How deep are effects of language on thought? Time estimation in speakers of English, Indonesian, Greek and Spanish. In D. G. K. Forbus, *Proceedings of the 26th Annual Conference Cognitive Science Society* (pp. 575-580). Hillsdale: Lawrence Erlbaum Associates.

Chan, T. & Bergen, B. (2001). Writing direction influence spatial cognition. *Proceedings of the Twenty-Seventh Annual Conference of the Cognitive Science Society.*

Chen, J. (2007). Do Chinese and English speakers think about time differently? Failure of replicating Boroditsky (2001). *Cognition, 104*(2), 427-436.

Christian, B., Miles, L. & Macrae, C. (2012). Your Space or Mine? Mapping Self in Time. *PLOS, 7*(11), 1-7.

Chun, L. (1997). Conceptualizing the world through spatial metaphors. An analysis of UP/DOWN vs. SHANG/XIA metaphors. *Proceedings of the 19th Annual Meeting of the Cognitive Science Society.* Mahwah, NJ: Erlbaum.

Cienki, A. & Müller, C. (2008). *Metaphor and Gesture.* Amsterdam; Philadelphia: John Benjamins Pub. Co.

Cienki, A. (1998). Metaphoric gestures and some of their relations to verbal metaphoric expressions. In J.-P. Koenig, *Discourse and Cognition: Bridging the Gap* (pp. 189-204). Stanford, CA: Center for the Study of Language and Information.

Cienki, A. (2005). Gesture and the question of literal versus non-literal reference. In B. &. Lewandowska-Tomaszczyk, *The Literal and Nonliteral in Language and Thought* (pp. 281-298). Frankfurt am Main: Peter Lang.

Cienki, A. J. (2008). Why study metaphor and gesture? In A. &. Cienki, *Metaphor and gesture* (pp. 5-25). Cambridge: Cambridge University Press.

Cienki, A. J. (2013). Cognitive Linguistics: Language and gestures as expressions of conceptualisation. In A. C. C. Müller, *Body-Language-Communication: An International Handbook on Multimodality in Human Interaction* (pp. 182-201). Berlin: De Gruyter Mouton.

Cienki, A., Morgenstern, A., Müller, C. & Boutet, D. (2016). Linguistic aspect, tense and gestural movement quality in French, German, and Russian utterances. *International Society of Gestures Studies 7.* Paris.

Clark, H. (1973). Space, time, semantics, and the child. In T. M. (Ed), *Cognitive development and the acquisition of language* (pp. 27-63). New York: Academic Press.

Condon, W. & Ogston, R. (1966). Sound film analysis of normal and pathological behaviour patterns. *Journal of Nervous and Mental Disease, 143,* 338-347.

Condon, W. C. & Ogston, R. (1967). A segmentation of behaviour. *Journal of Psychiatric Research, 5,* 221-235.

Conson, M., Cinque, F., Barbarulo, A.M. & Trojano, L. (2008). A common processing system for duration, order and spatial information: evidence from a time estimation task. *Experimental Brain Research, 2*(187), 267-274.

Cooperrider, K. & Núñez, R. (2009). Across Time, Across the Body: Transversal Temporal Gestures. *Gesture, 9*(2), 181-206.

Coulson, S. & Pagán Cánovas, C. (2014, January). Understanding Timelines: Conceptual Metaphor and Conceptual Integration. *Journal of Cognitive Semiotics, 1-2,* 198-219.

Croft, W. & Cruse, D. (2004). *Cognitive Linguistics.* New York: Cambridge University Press.

Crollen, V., Grade, S., Pesenti, M. & Dormal, V. (2013). A common metric magnitude system for the perception and production of numerosity, lenght, and duration. *Frontiers in Psychology, 4,* 1-11.

D'Andrade, R. (1989). Culturally based reasoning. In D. R. A.R.H. Gellatly, *Cognition and social worlds* (pp. 132-143). Oxford: Oxford University Press.

Davis, S. (2012). History on the Line: Time as Dimension. *Design Issues, 26*(4), 4-17.

De la Fuente, J., Santiago, J., Román, A., Dumitrache, C. & Casasanto, D. (2014). When you think about it, your past is in front of you: How culture shapes spatial conceptions of time. *Psychological Science, 29,* 1682-1690.

de L'Épée, C. (1776). *Institution des sourds et muets, par la voie des signes méthodiques.* Paris: Nyon.

De Ruiter, J. (2000). The production of gesture and speech. In D. McNeill, *Language and Gesture* (pp. 284-311). Cambridge: Cambridge University Press.

Dehaene, S., Bossini, S. & Giraux, P. (1993). The mental representation of parity and numerical magnitude. *Journal of Experimental Psychology: General, 122*(3), 371-396.

Di Bono, M., Casarotti, M., Prifitis, K., Gava, L., Umiltà, C., & Zorzi, M. (2012). Priming the mental time line. *Jornal of Experimental Psychology: Human Perception and Performance,* 838-842.

Dolscheid, S., Shayan, S., Ozturk, O., Majid, A., & Casasanto, D. (2012). Language shapes mental representations of musical pitch: Implications for metaphorical language processing. *Cognition and Language Workshop 2012.* University of California, Santa Barbara.

Dormal, V. & Pesenti, M. (2013). Preocessing numerosity, lenght and duration in a three-dimensional Stroop-like task: towards a gradient of processing automaticity? *Psychological Research, 77*, 116-127.

Dormal, V., Seron, X. & Pesenti, M. (2006). Numerosity-duration interference: A Stroop experiment. *Acta Psychologica, 121*, 109-124.

Droit-Volet, S., Clément, A. & Fayol, J. (2003). Timing and number discriminations in a bisection task with sequence of stimuli: A developmental approach. *Journal of Experimental Child Psychology, 84*(1), 63-76.

Duffy, S. & Feist, M. (2014). Individual differences in the interpretation of ambiguous statements about time. *Cognitive Linguistics*, 29-54.

Duffy, S. & Evans, V. (2016). The top trumps of time: Factors motivating the resolution of temporal ambiguity. *Language and Cognition*, 1-23.

Duffy, S. & Evans, V. (2017). The top trumps of time: factors motivating the resolution of temporal ambiguity. *Language and Cognition*, 293-315.

Duffy, S. & Feist, M. (2016). Power in time: The influence of power posing on metaphoric perspectives on time. *Language and Cognition*, 1-11.

Duffy, S. (2014). The Role of Cultural Artifacts in the Interpretation of Metaphorical Expressions About time. *Metaphor and Symbol*, 94-112.

Duffy, S., Feist, M. & McCarthy, S. (2014). Moving Through Time: The Role of Personality in Three Real-Life Contexts. *Cognitive Science*, 1662-1674.

Durante, A. (1997). *Linguistic Anthropology.* Cambridge: Cambridge University Press.

Efron, D. (1972). *Gesture, Race and Culture.* The Hague: Mouton and Co.

Ekman, P. &. (1969). The repertoire or nonverbal behavior: categories, origins, usage, and coding. *Semiotica*, 49-98.

ELAN (Version 4.8) [Computer software]. (2015). Nijmegen: Max Planck Institute for Psycholinguistics. Retrieved from https://tla.mpi.nl/tools/tla-tools/elan/

Evans, V. & Green, M. (2006). *Cognitive linguistics: an introduction.* Edinburgh: Edinburgh University Press.

Evans, V. (2004). *The Structure of Time.* Amsterdam; Philadelphia: J. Benjamins.

Evans, V. (2007). *A Glossary of Cognitive Linguistics.* Edinburgh: Edinburgh University Press.

Evans, V. (2013). *Language and time: a cognitive linguistics approach.* Cambridge: Cambridge University Press.

Faucconier, G. & Mark, T. (1998a). Principles of conceptual integration. In J.-P. K. (ed.), *Discourse and Cognition* (pp. 269-283). Stanford: Center for the study of language and information.

Faucconier, G. & Turner, M. (2002). *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities.* New York: Basic Books.

Fauconnier, G. & Turner, M. (1998). Conceptual Integration Networks. *Cognitive Science, 22*(2), 133-187.

Fauconnier, G. (1994). *Mental spaces: aspects of meaning construction in natural language.* Cambridge; New York, NY, USA: Combridge University Press.

Fauconnier, G. (1997). *Mappings in Thought and Language.* New York: Cambridge University Press.

Fedden, S. & Boroditsky, L. (2012). Spatialization of time in Mian. *Frontiers in Psychology, 3*, 1-9.

Feist, M. I& Duffy, S. (2015). Moving beyond Next Wednesday: The interplay of lexical semantics and constructional meaning in an ambiguous metaphoric statement. *Cognitive Linguistics*, 633-656.

Fillmore, C. (1982). Frame Semantics. In L. S. (ed.), *Linguistics in the morning calm* (pp. 111-137). Seoul: Hanshin Publishing.

Fillmore, C. (1982). Towards a descriptive framework for spatial deixis. In R. J. Klein, *Speech, Place, and Action* (pp. 31-59). London: John Wiley.

Floyd, S. (2016). Modally hybrid grammar? Celestial pointing for time-of-day reference in Nheengatú. *Language*, 31-64.

Fricke, E. (2007). *Origo, Geste und Raum. Lokaldeixis im Deutcher.* Berlin/New York: Mouton de Gruyter.

Fuhrman, O., McCormick, K., Chen, E., Jiang, H., Shu, D., Mao, S. & Boroditsky, L. (2011). Cross-Cultural Differences in Mental Representations of Time: Evidence From an Implicit Non-linguistic Task. *Cognitive Science, 35*, 1305-1328.

Galton, A. (2011). Time flies but space does not: Limits to the spatialization of time. *Journal of Pragmatics*, 695-703.

Gibson, J. (1966). The problem of temporal order in stimulation and perception. *The Journal of Psychology*, 141-149.

Glenberg, A. & Robertson, D. (2000). Symbol grounding and meaning: A comparison of high-dimensional and embodied theories of meaning. *Journal of Memory and Language*, 379-401.

Goffman, E. (1963). *Stigma: notes on the management of spoiled identity.* N.J.: Prentice-Hall.

Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us think.* Cambridge, MA: Belknap.

Grady, J. (1997). *Foundations of meaning: primary metaphors and primary scenes.* Berkeley: PhD Dissertation, University of California.

Guillaume, G. (1929). *Temps et verbe: theorie des aspects, les modes et des temps.* Paris.

Hartmann, M. & Mast, F. (2012). Moving along the mental time line influences the processing of future related words. *Consciousness and Cognition, 21,* 1558-1562.

Hauser, D., Carter, M., & Meier, B. (2009). Mellow Monday and furious Friday: The approach-related link between anger and time representation. *Cognition and Emotion,* 1166-1180.

Hewson, J. (2012). Tense. In R. I. Binnick, *The Oxford Handbook of Tense and Aspect.* Oxford: Oxford University Press.

Hutchins, E. (2005). Material anchors for conceptual blends. *Journal of Pragmatics*(37), 1555-1577.

Huumo, T. (2015). Temporal frames of reference and the locative case marking of the Finnish adposition ete- 'in fron of/ ahead'. *Lingua*, 45-67.

Ishihara, M., Keller, P., Rossetti, Y. & Prinz, W. (2008). Horizontal spatial representations of time: evidence for the STEARC effect. *Cortex, 44*(4), 454-461.

Iverson, J. & Goldin-Meadow, S. (1998). Why people gesture when they speak. *Nature, 396,* 228.

Iverson, J. & Thelen, E. (1999). Hand, mouth, and brain: The dynamic emergence of speech and gesture. *Journal of Consciousness Studies, 6,* 19-40.

January, D. & Kako, E. (2007). Re-evaluating evidence for the linguistic relativity hypothesis: response to Boroditsky (2001). *Cognition, 104*(2), 417-426.

Johnson, M. (1987). *The body in the mind: the bodily basis of meaning, imagination, and reason.* Chicago: University of Chicago Press.

Katz, J. (1972). *Semantic Theory.* New York: Harper & Row.

Kendon, A. (1972). Some relationships between body motion and speech. An analysis of an example. In A. S. Pope, *Studies in Dyadic Communication* (pp. 177-210). New York: Pergamon Press.

Kendon, A. (1980). Gesture and speech: two aspects of the process of utterance. In M. R. Key, *Nonverbal Communication and Language* (pp. 207-227). The Hague: Mouton.

Kendon, A. (2000). Language and Gesture: Unity or Duality. In D. McNeill, *Language and Gesture: Window into Thought and Action* (pp. 47-63). Cambridge: Cambridge University Press.

Kendon, A. (2004). *Gesture: Visble Action as Utterance.* Cambridge: Cambridge University Press.

Kita, S. & Essegbey, J. (2001). Pointing left in Ghana: How a taboo on the use of the left hand influences gestural practice. *Gesture, 1*, 73-94.

Kita, S. (2000). How representational gestures help speaking. In D. Mcneill, *Language and gesture* (pp. 162-185). Cambridge: Cambridge University Press.

Kövecses, Z. (2003). *Metaphor and Emotion: Language, culture, and body in human feeling.* Cambridge: Cambridge University Press.

Kövecses, Z. (2010). *Metaphor: A practical introduction.* Oxford: Oxford University Press.

Kranjec, A. & Chatterjee, A. (2010). Are temporal concepts embodied? A challenge for cognitive neuroscience. *Frontiers in Psychology*, 1-9.

Kranjec, A. & McDonough, L. (2011). The implicit and explicit embodiment of time. *Journal of Pragmatics, 43*, 735-748.

Ladewig, S. (2010). Beschreiben, auffordern und suchen-Varianten einer rekurrenten Geste. *Sprache und Literatur*, 89-111.

Ladewig, S. (2011). Putting the cyclic gesture on a cognitive basis. *CogniTextes*.

Ladewig, S. (2014). Recurrent Gestures. In A. C. C. Müller, *Body-Language-Communication: An International Hand-book on Multimodality in Human Interaction* (pp. 1558-1575). Berlin, Boston: De Gruyter: Mouton.

Lakoff, G. & Johnson, M. (1999). *Philosophy in the flesh: the embodied mind and its challenge to Western thought.* New York: Basic Books.

Lakoff, G. & Johson, M. (1980). *Metaphors we live by.* Chicago: Chicago University Press.

Lakoff, G. (1987). *Women, fire, and dangerous things: what categories reveal about the mind.* Chicago: University of Chicago Press.

Lakoff, G., Espenson, J. & Schawartz, A. (1991). *Master Metaphor List, Second Edition.* California: Berkeley: Cognitive Linguistics Group-University of California at Berkeley.

Langacker, R. (1991). *Foundations of Cognitive Grammar.* Stanford: Stanford University Press.

Langacker, R. (2008). *Cognitive Grammar: An introduction.* Oxford: Oxford Scholarship Online.

Le Guen, O. (2011). Speech and Gesture in Spatial Language and Cognition Among the Yucatec Mayas. *Cognitive Science, 35*(5), 905-938.

Lee, D. (2001). *Cognitive linguistics: an introduction.* Melbourne; Oxford; New York: Oxford University Press.

Lehrer, A. (1974). *Semantic fields and lexical structure.* Amsterdam; New York: North-Holland; American Elsevier.

Lausberg, H., & Sloetjes, H. (2009). Coding gestural behavior with the NEUROGES-ELAN system. Behavior Research Methods, Instruments, & Computers, 41(3), 841-849. doi:10.3758/BRM.41.3.591.

Levelt, W. (1989). *Speaking.* Cambridge, MA: MIT Press.

Levinson, S. (1996). Frames of reference and Molyneux's question: Crosslinguistic evidence. In P. P. Bloom, *Language and Space* (pp. 109-169). Cambridge: MIT Press.

Levinson, S. (2003). *Space in Language and Cognition.* Cambridge: Cambridge University Press.

Magnani, B., Oliveri, M., Mancuso, G., Galante, E. & Frassinetti, F. (2011). Time and spatial attention: effects of prism adaptation on temporal deficits in brain damaged patients. *Neuropsychologia,* 1016-1023.

Mayberry, L. & Jaques, J. (2000). Gesture production during stuttered speech: Insights into the nature of gesture-speech integration. In D. M. (Ed.), *Language and gesture* (pp. 199-213). Cambridge: Cambridge University Press.

McClave, E. (2000). Linguistic functions of head movements in the context of speech. *Journal of Pragmatics*, 855-878.

McNeill, D. & Duncan, S. (2000). Growth points in thinking-for-speaking. In D. McNeill, *Language and Gesture* (pp. 141-161). Cambridge: Cambridge University Press.

Mcneill, D. & Levy, E. (1982). Conceptual Representations in Language Activity and Gesture. In R. J. Klein, *Speech, Place, and Action: Studies in Deixis and Related Topics* (pp. 271-295). Chinchester, England: Wiley.

McNeill, D. (1979). *The Conceptual Basis of Language.* Hillsdale, N. J.: Erlbaum.

McNeill, D. (1985). So you think gestures are nonverbal? *Psychological Review, 92,* 350-371.

McNeill, D. (1992). *Hand and Mind.* Chicago: University of Chicago Press.

McNeill, D. (2005). *Gesture and Thought.* Chicago: University of Chicago Press.

Meck, W. & Church, R. (1983). A mode of control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behaviour Processes, 9*, 320-334.

Miles, L., Karpinska, K., Lumsden, J. & Macrae, C. (2010b). The meandering mind: Vection and mental time travel. *PLOS One*.

Miles, L., Nind, L. & Macrae, C. (2010a). Moving Through Time. *Psychological Science, 21*(2), 222-223.

Miles, L., Tan, L., Noble, G. D., Lumsden, J. & Macrae, C. (2011). Can a mind have two timelines? Exploring space-time mapping in Mandarin and English speakers. *Psychon Bull Rev, 18*(3), 598-604.

Mittelberg, I. (2006). Methodology for multimodality: One way of working with speech and gesture data. In M. Gonzalez-Marquez, *Methods in Cognitive Linguistics* (pp. 225-249). Amsterdam/Philadelphia: John BenjaminsPublishing Company.

Moore, K. (2000). *Spatial Experience and Temporal Metaphors in Wolof: Point of View, Conceptual Mappings, and Linguistic Practice.* Doctoral dissertation, University of California, Berkeley.

Moore, K. (2006). Space-to-time mappings and temporal concepts. *Cognitive Linguistics, 17*, 199-244.

Moore, K. (2011). Ego-Perspective and field-based frames of reference: Temporal meanings of FRONT in Japanese, Wolof and Aymara. *Journal of Pragmatics*, 759-776.

Moore, K. (2014). *The spatial language of time metaphor, Metonymy and Frames of Reference.* Amsterdam; Philadelphia: John Benjamins.

Müller, C. (1998). *Redebegleitende Gesten: Kulturgeschichte, Theorie, Sprachvergleich.* Berlin: Arno Spitz.

Müller, C. (2010). Wie Gesten bedeuten. Ein kognitiv-linguistische und sequenzanalytische Perspektive. *Sprache und Literatur*, 37-68.

Müller, D. & Schwarz, W. (2008). "1-2-3": is there a temporal number line? Evidence from a serial comparison task. *Experimental Psychology, 55*, 143-150.

Nilsson, A. (2016). Expressing time through space: Embodying conceptual metaphors in an L2. *ISGS7.* Paris.

Núñez, R. & Cooperrider, K. (2013, May). The Tangle of Space and Time in Human Cognition. *Trends in Cognitive Sciences, 17*(5), 220-229.

Núñez, R. & Sweetser, E. (2001). Spatial embodiment of temporal metaphors in Aymara: Blending source domain gesture with speech. In *Abstracts from the 7th International Cognitive Linguistics Conference* (pp. 249-250). Santa Barbara: University of California, Santa Barbara.

Núñez, R. & Sweetser, E. (2006). With the Future Behind Them: Converge Evidence from Aymara Language and Gesture in the Crosslinguistic Comparison of Spatial Construals of Time. *Cognitive Science*, 1-49.

Núñez, R. (2008). Proto-numerosities and Concepts of Number: Biologically Plausible and Culturally Mediated Top-down Mathematical Schemas. *Behavioral and Brain Sciences, 3*, 665-666.

Núñez, R., Cooperrider, K., Doan, D. & Wassmann, J. (2012). Contours of Time: Topographic Construal of Past, Present and Future in the Yupno Valley of Papua New Guinea. *Cognition, 124*(1), 25-35.

Oliveri, M., Vicario, C., Salerno, S., Koch, G., Turriziani, P., Mangano, R., Chillemi G, & Caltagirone C. (2008). Perceiving numbers alters temporal perception. *Neuroscience Letters, 432,* 308-311.

Ouellet, M., Román, A. & Santiago, J. (2012). A multisensory interaction effect in the conceptual realm of time. *Experimental Psychology, 59,* 236-242.

Ouellet, M., Santiago, J., Funes, M. J. & Lupiáñez, J. (2010a). Thinking about the future moves attention to the right. *Journal of Experimental Psychology: Human, Perception and Performance, 36,* 17-24.

Ouellet, M., Santiago, J., Israeli, Z. & Gabay, S. (2010b). Is future the right time? *Experimental Psychology, 57,* 308-314.

Pagán Cánovas, C. & Jensen, M. (2013). Anchoring time-space mappings and their emotions: The timeline blend in poetic metaphors. *Language and Literature, 22*(1), 45-59.

Pagán Cánovas, C., Valenzuela, J. & Santiago, J. (2015). Like the Machete the Snake: Integration of Topic and Vehicle in Poetry Comprehension Reveals Meaning Construction Processes. *Psychology of Aesthetics, Creativity, and the Arts, 9*(4), 385-393.

Pagán-Cánovas, C. & Piata, A. (Forthcoming). The Way Time Goes By: Generic Integration Templates and the Poetics of Time. In S. Csábi, *Expressive Minds and Artistic Creations: Studies in Cognitive Poetics.* Oxford: Oxford University Press.

Pagán-Cánovas, C. & Teuscher, U. (2012). Much more than money: Conceptual integration and the materialisation of time in Michael Ende's Momo and the social sciences. *Pragmatics & Cognition,* 546-569.

Piata, A. & Pagán Cánovas, C. (2017). Powerful Rhyme and Sluttish Time: A Cross-linguistic Study of Time Personification in Poetic Discourse. *Language and Literature, 26*(1), 18-33.

Piata, A. (2012). Conventionality and Creativity in the Conceptualisation of Time: Evidence from Modern Greek. In G. Fragkaki, *New Perspectives in Greek Linguistics.* Newcastle: Cambridge Scholars Publishing.

Powell, H. (2012). *Stop the Clocks!: Time and Narrative in Cinema.* London: I.B. Tauris.

Raden, G. & Dirven, R. (2007). *Cognitive English Grammar.* Amsterdam: John Benjamins.

Román, A., Santiago, J., Jasmin, K. & Casasanto. (2014). The hands of time in Spaniards and Moroccans. *Manuscript in preparation*.

Rosch, E. & Mervis, C. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7,* 573-605.

Rosh, E. & Lloyd, B. (1978). *Cognition and categorization.* New York: Halsted Press.

Rosh, E. (1973). On the Internal Structure of Perceptual and Semantic Categories. In T. Moore, *Cognitive Development and the Acquisition of Language.* New York: Academic Press.

Saj, A., Fuhrman, O., Vuilleumier, P. & Boroditsky, L. (2014). Patients With Left Spatial Neglect Also Neglect the "Left Side" of Time. *Psychological Science*, 207-214.

Santiago, J, Román, A., Ouellet, M., Rodríguez, N. & Pérez-Azor, P. (2008). In hindsight, life flows from left to right. *Psychological Research, 74*(1), 59-70.

Santiago, J., Lupiáñez, J., Pérez, E. & Funes, M. (2007). Time also flies from left to right. *Psychonomic Bulletin & Review, 14*(3), 512-516.

Schank, R.C& Abelson, R. (1977). *Scripts, plans, goals, and understanding: an inquiry into human knowledge structures.* New York: Halsted Press Division of John Wiley and Sons.

Scott, A. (1989). The vertical dimension and time in Mandarin. *Australian Journal of Linguistics, 9*, 295-314.

Selvik, K. (2006). *Spatial Paths Representing Time. A cognitive analysis of temporal expressions in Norwegian Sign Language.* Oslo: University of Oslo: Faculty of Humanities.

Semino, E. (2008). *Metaphor in Discourse.* Cambridge: Cambridge University Press.

Sinha, C., Da Silva Sinha, V., Zinken, J. & Sampaio, W. (2014). When time is not space: The social and linguistic construction of time intervals and temporal event relations in an Amazonian culture. *Language and Cognition*, 137-169.

Smith, B. (ed.). (1988). *Foundations of Gestalt Theory.* Munich and Vienna: Philosophia Verlag.

Srinivasan, M. & Carey, S. (2010). The long and the short of it: On the nature and origin of fuctional overlap between representations of space and time. *Cognition*, 217-241.

Stamenkovic, D. & Jacevic, M. (2015). Time, Space, and Motion in Braid: A Cognitive Semantic Approach to a Video Game. *Games and Culture, 10*(2), 178-203.

Stoianov, I., Kramer, P., Umiltà, C. & Zorzi, M. (2008). Visuospatial priming of the mental number line. *Cognition*, 770-779.

Stokoe, W. (1960). *Sign Language Structure: An Outline of the Visual Communication Systems of the American Deaf.* Buffalo: Department of Anthropology and Linguistics, University of Buffalo.

Sweetser, E. (2006). Looking at space to study mental spaces: Co-speech gesture as a crucial data source in cognitive linguistics. In M. Gonzalez-Marquez, *Methods in Cognitive Linguistics* (pp. 201-224). Amsterdam/Phildadelphia: John Benjamics Publishing Company.

Talmy, L. (2000a). *Toward a cognitive semantics. Volume I, Concept structuring systems.* Cambridge: MIT Press.

Talmy, L. (2000b). *Towards a Cognitive Semantics: Vol.II: Typology and Process in Concept.* Cambridge: MIT Press.

Taylor, J. (2002). *Cognitive Grammar.* Oxford: Oxford University Press.

Tenbrink, T. (2007). *Space, time, and the use of language: an investigation of relationships.* Berlin; New York: Mouton de Gruyter.

Tenbrink, T. (2011). Reference frames of space and time in language. *Journal of Pragmatics*, 704-722.

Torralbo, A., Santiago, J. & Lupiañez, J. (2006). Flexible conceptual projection of time onto spatial frames of reference. *Cognitive Science 30*, 745-757.

Turchyn, S., Olza-Moreno, I., Pagán-Cánovas, C., Steen, F., Turner, M., Valenzuela, J. & Soumya, R. (2018). Gesture Annotation with a Visual Search Engine for Multimodal Communication Research. *13th Conference on Innovative Applications of Artificial Inteligence*.

Turim, M. (1989). *Flashbacks in film: memory and history.* New York: Routledge.

Tversky, B., Kugelmass, S. & Winter, A. (1991). Cross-cultural and developmental trends in graphic productions. *Cognitive Psychology, 23*, 515-557.

Vallesi, A., Binns, M. & Shallice, T. (2008). An effect of spatial-temporal association of response codes: understanding the cognitive representations of time. *Cognition 107, 107*(2), 501-527.

Vallesi, A., Mclntosh, A. & Stuss, D. (2011). How time modulates spatial responses. *Cortex*, 148-156.

Vicario, C., Caltagirone, C. & Oliveri, M. (2007). Optokinetic stimulation affects temporal estimation in healthy humans. *Brain and Cognition, 64*(1), 68-73.

Vicario, C., Pecoraro, P., Turriziani, P., Koch, G., Caltagirone, G. & Oliveri, M. (2008). Relativistic compression and expansion of experiential time in the left and right space. *Plos One 3*.

Walker, E. & Cooperrider, K. (2016). The continuity of metaphor: Evidence from temporal gestures. *Cognitive Science, 40*(2), 481-495.

Walker, E., Bergen, B. & Núñez, R. (2013). Investigating spatial axis recruitment in temporal reckoning through acoustic stimuli and non-spatial responses. *Center for Research in Language Technical Report, University of California, San Diego, 25*, 1-10.

Walker, E., Bergen, B. & Núñez, R. (2014). Disentangling Spatial Metaphors for Time Using Non-spatial Responses and Auditory Stimuli. *Metaphor and Symbol*, 316-327.

Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences, 7*, 483-488.

Weger, U. & Pratt, J. (2008). Time flies like an arrow: space-time compatibility effects suggest the use of a mental timeline. *Psychonomic Bulletin & Review, 15*(2), 426-430.

Wilcox, P. (2002). The iconic mapping of space and time in signed languages. In L. Albertazzi, *Unfolding Perceptual Continua* (pp. 255-281). Amsterdam: Benjamins.

Williams, R. (2008). Gesture as a conceptual mapping tool. In C. Müller, *Metaphor and Gesture* (pp. 55-92). Amsterdam/Philadelphia: John Benjamins.

Winter, B., Marghetis, T. & Matlock, T. (2015). Of magnitudes and metaphors: Explaining cognitive interactions between space, time, and number. *Cortex, 64*, 209-224.

Zorzi, M., Bonato, M., Treccani, B., Scalambrin, G., Marenzi, R. & Priftis, K. (2012). Neglect impairs explicit access to the mental number line. *Frontiers in Human Neuro-science, 6*(125), 1-12.

Zorzi, M., Priftis, K. & Umiltà, C. (2002). Neglect disrupts the mental number line. *Nature, 417*, 138-139.