A STUDY OF THE RELATIONSHIP BETWEEN GESTURE AND INTONATION IN PUBLIC SPEAKING

MOH JUN HAN

FACULTY OF LANGUAGES AND LINGUISTICS
UNIVERSITY OF MALAYA
KUALA LUMPUR

2017

A STUDY OF THE RELATIONSHIP BETWEEN GESTURE AND INTONATION IN PUBLIC SPEAKING

MOH JUN HAN

DISSERTATION SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER IN ENGLISH AS A SECOND LANGUAGE

FACULTY OF LANGUAGES AND LINGUISTICS
UNIVERSITY OF MALAYA
KUALA LUMPUR

2017

UNIVERSITY OF MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: Moh Jun Han (I.C/Passport No:

Matric No: TGB 120004

Name of Degree: Master of English as a Second Language

Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"):

A Study of the Relationship Between Gesture and Intonation in Public Speaking

Field of Study: Phonetics and Phonology

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date:

Subscribed and solemnly declared before,

Witness's Signature Date:

Name: Professor Dr Stefanie Pillai

Designation: Dekan Fakulti Bahasa dan Linguistik

ABSTRACT

Public speaking is an important skill for a person to master if he or she wants to be an effective communicator. Previous studies have investigated the relationship between speech and gesture and it was found that they enhance speakers' verbal delivery and listeners' understanding of the speech (Driskell and Radke, 2003). However, most of these studies only focused on natural speech between small groups of people or between researchers and their respondents. Moreover, these studies also mostly studied gesture and intonation separately. This study aims to investigate the relationship between gesture and intonation in public speaking. The introductory segments of four speeches which were crafted and delivered by the winners of the Toastmasters World Championship of Public Speaking were selected for this study. The videos of the selected segments of the speech were annotated using Elan and Praat. The gestures of the speakers were first annotated in Elan using Kendon's (1972) hierarchy of gestures as the analytical framework to code the gestures. The intonation of the same segments was measured and annotated in Praat using the Tone and Break Indices (ToBI) framework developed by Beckman and Elam (1997). Previous studies like Loehr's (2004) have generalised that pitch accents are always aligned with the 'apex' of a gesture in natural speech. However, the findings in the current study indicate that pitch accents in public speech are not only aligned consistently with the stroke of a gesture, but that they also occur during the post-stroke hold phrase and even when there is no gesture at all. In addition, the findings also show that a pitch accent, stroke and post-stroke hold phrase also tend to coincide more with a content word compared to a function word. Furthermore, the function words which were aligned with pitch accents also tend to be first person pronouns, confirming and expanding upon previous work. Overall, the findings seem to suggest that although gesture and intonation play various roles in public speaking, the prosodic elements of a speech seem to have a bigger influence on the quality of a speech.

ABSTRAK

Kemahiran berpidato merupakan satu daripada kemahiran yang harus dikuasai sekiranya seorang individu ingin menjadi seorang komunikator yang efektif. Sebelum ini, kajian-kajian yang mengkaji perhubungan di antara pertuturan dan gerak isyarat telah menemui bahawa perbuatan gerak isyarat meningkatkan kualiti penyampaian lisan pengucap serta membolehkan para hadirinnya untuk memahami ucapan itu dengan lebih teliti (Driskell & Radke, 2003). Namun demikian, kebanyakan daripada kajian-kajian tersebut hanya terhad kepada menganalisa perbualan secara spontan. Perbualan-perbualan tersebut hanya melibatkan beberapa orang individu yang berbual dalam kumpulan-kumpulan kecil atau di antara responden dan penyelidik. Di samping itu, kajian-kajian sebelum ini hanya memfokuskan kepada penyelidikan fungsi gerak isyarat dan intonasi dalam pertuturan secara berasingan. Kajian ini akan menyiasat hubungan antara gerak isyarat dan intonasi dalam pidato umum. Empat buah pengenalan yang dipetik daripada empat ucapan pemenang *Toastmasters World Championship of Public Speaking* telah dipilih untuk kajian ini. Keempat-empat ucapan ini telah dikarang dan disampaikan oleh pemenang-pemenang tersebut. Segmen-segmen video ucapan yang dipilih telah dianotasikan dengan Elan dan Praat. Gerak isyarat keempat-empat pemidato tersebut telah dianotasikan dahulu dengan Elan dengan menggunakan hierarki gerak isyarat yang telah dikemukakan oleh Kendon (1972) sebagai kerangka analitis. Seterusnya, intonasi dalam segmen-segmen ucapan tersebut telah dianotasikan dengan Praat. ToBI (*Tone and Break Indices*) yang dikemukakan oleh Beckman dan Elam (1997) telah digunakan sebagai kerangka analitis. Kajian sebelum ini, contohnya Loehr (2004), membuat generalisasi bahawa aksen nada akan sentiasa selari dengan apeks strok gerak isyarat dalam sesuatu perbualan spontan. Di sebaliknya, kajian ini telah mendapati bahawa aksen nada dalam pidato umum bukan sahaja selari dengan strok gerak isyarat tetapi juga selari dengan pasca strok gerak-geri. Kajian ini juga mendapati bahawa

terdapat aksen nada yang tidak selari dengan sebarang pergerakan langsung. Di samping itu, penemuan kajian ini juga menunjukkan bahawa aksen nada, strok and pasca strok gerak isyarat adalah selari dengan lebih banyak kata-kata isi berbanding dengan kata-kata fungsi. Kebanyakan kata-kata fungsi yang selari dengan aksen-aksen nada merupakan kata-kata ganti nama pertama. Oleh yang demikian, dapatan ini telah mengesahkan dan memperkembangkan penemuan kajian-kajian yang telah dijalankan sebelum ini. Pada keseluruhannya, dapatan kajian ini mencadangkan bahawa walaupun perbuatan gerak isyarat dan intonasi mempunyai pelbagai fungsi dalam pidato umum, elemen-elemen prosodi dalam sebuah ucapan nampaknya mempunyai pengaruh yang lebih besar atas kualiti ucapan tersebut.

ACKNOWLEDGEMENTS

It is said that the journey of a thousand miles begins with a single step. It is also said that the Almighty provides and brings the right people at the right time into a person's life and for that, I am forever thankful. The seed of this dissertation was planted in November 2014 after a conversation with Ms. LeAnn Tang, a renowned Toastmaster, on how to improve the quality of a person's speech. This seed, after much toil and typing, has finally flourished into the towering tree that it is today.

My thanks go to my supervisor, Prof. Stefanie Pillai for her guidance and feedback as it enabled me to refine my dissertation to the best of my ability. It was her who opened my eyes to the world of prosody and the potential things that I could learn. Ever firm, she expected nothing but the best out of me and she helped ensure that I was always spurring myself to greater heights.

My appreciation also goes to the Head of the Centre for English Language Studies, Sunway University, Prof. Stephen Hall who made time to discuss potential ideas with me for my dissertation and has always been supportive of my endeavour to develop myself professionally and as a person. Ms. Jeanette Lim, the Head of the Intensive English Programme is another person whom I would love to thank for her constant support in ensuring that I stay true to my goal of studying gestures and intonation in public speaking.

I have also been supported by numerous friends along this path. My thanks go out especially to Dr. Juergen Burkhardt, Irene Kiew and Stella Ong for they too have taken time to share their postgraduate journeys with me and encouraged me to stay the course.

I am also grateful to my family especially my parents, Michael and Sereen Moh who have continuously supported me throughout this journey as well. I appreciate that their love for me has always been constant although sometimes not always expressed explicitly. I am glad that the late nights due to my dissertation are finally coming to an end.

My greatest thanks and love is reserved for none other than my wife, Pauline See Jiasue. It was her who has always stood by my side and believed that I can run this race and finish it strong. I look forward to walking side by side with her in the decades to come. It is to her and our future family that I dedicate this work.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

ToBI	:	Tone and Break Indices System
H*	:	High pitch accent
L*	:	Low pitch accent
!H*	:	Downstepped pitch accent
H-	:	High intermediate phrase boundary
L-	:	Low intermediate phrase boundary
H%	:	High intonational phrase boundary
L%	:	Low intonational phrase boundary

CHAPTER 1: INTRODUCTION

1.1 Background of the Study

People communicate in different ways whether they realise it or not. From the politician giving a speech to a group of friends chatting in a café, they communicate to their listeners in different ways. Be it a public speech or a friendly conversation, some of the ways these speakers use to communicate with their listeners involve vocalising their words in a certain way. They use certain rhetorical devices such as alliteration and triads to make their words sound more appealing and memorable to their listeners. They also structure what they plan to say around a central message or purpose be it to inform, persuade, inspire or entertain. Moreover, they also speak at a certain pitch or volume in order to capture and keep the attention of their listeners. In addition to all of these skills, they also employ non-verbal means to communicate their message such as gesturing to reinforce their verbal content.

There are organisations which are dedicated to improving the communication and leadership skills of their members. One of these organisations is Toastmasters International. It was founded by Ralph C. Smedley on the 22nd of October 1924 where he held the first Toastmaster meeting in the YMCA building in Santa Ana, California (Toastmasters International, 2016a). Today, Toastmasters International has more than 345000 members who teach, evaluate and learn from one another to be proficient communicators, public speakers and leaders.

A typical Toastmasters meeting is divided into three sessions which will be emceed by a Toastmaster. The first session is called the Table Topic Session where both members and guests can volunteer to give a 2-minute impromptu speech. It is then followed by the Prepared Speech Session which is only open to members and they have to deliver a speech that is prepared beforehand. The speeches are crafted based on various manuals and they have to fulfil different objectives depending on their speech assignment. The speeches

are then evaluated by different speakers who will point out the strengths and weaknesses of each speech and give the speaker tips on how they can improve their speeches. In addition, there are other role players who will give their feedback on different areas of the meeting (Toastmasters International, 2016b). They are the Grammarian who gives his or her feedback on how the language and grammar was used during the meeting and the Ah-Counter who notes down the various speech crutches and filler words used by each speaker during the meeting. In addition, there is a Timer who gives his or her report on how time was managed during the meeting and the General Evaluator who will give his/her feedback on the overall conduct of the meeting.

Various researchers have tried to integrate the Toastmasters meeting format into their public speaking or EFL oral classes and they have received positive feedback from their students on how the meetings have given them a platform to improve their speaking skills and proficiency in the English language. Shahrina and Zullina (2005) incorporated the Toastmasters meeting format into their course's weekly oral presentations. Their students later reported that they felt the meetings had helped them to improve their ability to use the English language which in turn improved their confidence in using it. Another study was done by Sun (2008) where she integrated the Toastmasters approach into her EFL speech class. At the end of the course, she also received positive feedback from her students. They said that the Toastmasters approach to public speaking boosted their confidence, reduced their speech anxiety and encouraged more practice and learning of the English language. In addition, they also said that it helped to improve their public speaking skills. Therefore, in improving their public speaking skills, these students have found that their communication skills have also improved and they are able to face an audience and are able to communicate with them without fear or fervour.

Although there have been studies which have established that public speaking enhances a person's communication skills (Hairuliza & Suzana, 2001), it still warrants

further investigation. This is because public speaking involves more than just asking a person to stand in front of a crowd to deliver a speech. It comprises of many different skills from determining the purpose of the speech, organising its structure, using rhetorical devices to having purposeful gestures and effective intonation. In fact, nonverbal communication is just as important as verbal communication, as more than half of human communication takes place nonverbally (Toastmasters International, 2011b). Therefore, it is vital for speakers to be aware that their gestures and intonation need to be consistent with the message they want to convey to their audience.

1.2 Statement of the Problem

The previous section mentioned that gesture and intonation share a relationship with each other. Moreover, there are also studies that found that they reflect a unified planning process (Esposito & Marinaro, 2007). It cannot be denied that gestures and speech share a relationship with each other as gestures enhance the speaker's verbal delivery and the listener's understanding of the speech (Driskell & Radtke, 2003). In fact, Rauscher, Krauss and Chen (1996) theorised that gestures are produced as part of the speech process. In addition, McNeill (2005) found that gestures and speech are produced by the speaker simultaneously when he or she expresses an idea and they complement each other. Therefore, in order to become an effective public speaker and by extension, an effective communicator, one needs to master not only the verbal aspects of public speaking but also the nonverbal elements especially gesture and intonation.

To date, there has been a lot of research done to investigate the relationship between gestures and intonation in natural and spontaneous conversation between small groups of people. However, the dynamics of a public speech are vastly different from a natural conversation. Although a public speech and a spontaneous conversation are different ways for a speaker to communicate his or her message to their listener(s), most public

speeches are staged performances before an audience who usually consist of a large group of people. This is because most public speakers prepare their speeches extensively and their speeches are highly structured (Hairuliza & Suzana, 2001).

In contrast, natural conversations always occur spontaneously and there is usually little to no preparation before a group of friends sit down for a chat. Moreover, the language used in a public speech is usually more formal and the method of delivery in a public speech is different from a conversation within a small group of people. In a conversation within a small group, most people tend to talk quietly, adopt a casual posture and use a lot of pause fillers like 'um', 'err' and 'ah'. However, an effective public speaker adjusts the volume of his voice so the audience can hear him or her clearly, has a more erect posture and avoids distracting mannerisms and verbal habits (Lucas, 2012). Thus, as a speaker rehearses his or her speech before it is delivered before an audience, he or she also needs to ensure that every gesture made is purposeful and reflective of the message of the speech. The speaker is also encouraged to rehearse the gestures together with the speech until it looks natural to the audience (Toastmasters International, 2011b). In addition to gestures, a public speaker also needs to rehearse and vary the pitch and tone of their voice in order for them to be consistent with the message he or she wants to communicate to the audience (Toastmasters International, 2011c).

However, there is not much focus on how to use gestures and intonation to deliver a speech effectively compared to the crafting and organisation of the content of a speech (Siddens, 1998). The lack of attention paid to gestures and intonation in public speaking seems to give the impression that these two nonverbal elements are either not important to the speech as a whole or that there is a lack of understanding on how they are related to a public speech. The former can be ruled out as previous studies have established that gestures and intonation play an important role in enhancing a speech, be it prepared or spontaneous. Thus, it seems that there has been not much research done to understand the

relationship between gestures and intonation in a public speech. Most studies in the past tended to focus on natural conversation and determining the roles gestures and intonation play in enhancing a public speech. While there have been studies done to shed some light on this area, the nature of this complex relationship has not been described in detail (Loehr, 2004). In addition, many, if not most researchers and trainers tend to use a prescriptive approach to using gestures and intonation in public speaking which resulted in the vague descriptions on how to use gesture and intonation effectively. This shows that if one cannot comprehend or describe the extent of the relationship between gesture and intonation in speech, it would be difficult to suggest concrete and specific steps on how to use gestures and intonation to enhance a speech delivered in public.

1.3 Objectives and Research Questions

The objective of this study is to investigate the relationship between gestures and intonation in a public speech. In order to meet the objective of this study and to provide a clear description of the relationship between gestures and intonation in public speaking, the following questions were answered. They were:

a) What are the roles of gestures and intonation in a prepared speech?
b) To what extent is there a relationship between gestures and pitch accents?
c) Which of them appears to have a bigger influence on a public speech?

During the course of the study, the gestures and intonation of four speech introductions taken from four speakers who won the Toastmasters International World Championship of Public Speaking were analysed.

1.4 Significance of the Study

The current study aims to describe how gestures and intonation in public speaking are connected and this can be significant in some ways. First, it will provide a better understanding on how they interact with each other to enhance the quality of a speech. If the relationship between gestures and intonation can be described in detail, researchers would have a deeper understanding of the nonverbal aspects of public speaking. As a result, they will be able to prescribe more specific and effective methods and materials to improve the way a speaker communicates to his or her audience through nonverbal means i.e. gestures and intonation. Thus, when a speaker's gestures and intonation is consistent with the verbal message he or she wants to convey to the audience, the speech which they deliver in public also becomes more memorable and powerful.

This study is also significant as it can contribute to nation building. As speakers begin to improve the quality of their speeches, their ability to communicate with people will also increase as they become more confident and comfortable to express their thoughts and ideas to a group of people. These graduates will form the backbone of the Malaysian workforce and economy in the future. Hence, if they have been equipped with the ability to communicate effectively, they would be able to contribute to the nation. This is only possible if their communication skills have been honed earlier by those who have attained a deep understanding of the art of public speaking and communication.

1.5 Limitations

One of the limitations faced during the study was that the researcher had no control over how the speakers were recorded when they gave their speeches during the tournament. This is because the study focused on different speakers who had won the championship in previous years. In addition, it was impossible for any recording to be done by the researcher as Toastmasters International only allows their own videographers

to record the speeches during the tournament. These speeches will be compiled and sold in DVDs later.

Another limitation faced during the study was that despite all of the research done in the field of gesture, there is still no standardised analytical framework to describe a speaker's gestures (Loehr, 2004). Therefore, the study had to adapt McNeill's (1992) analytical framework which was originally developed to describe the gestures made by a speaker to describe a cartoon to the researcher without any preparation. Although there have been studies in the past which have used his guidelines, these studies (Loehr, 2004 & Beattie, Webster & Ross, 2014) have either used a cartoon as a stimulus for their participants to produce gestures or his guidelines had to be heavily modified to suit the context of the study.

Another concern that was raised during the study was that the amount of data might not be sufficient for a valid hypothesis or conclusive result to be formed. Nonetheless, the amount of data that was annotated is consistent with previous studies but it is always acknowledged that more data is better data. In addition, the current study only focuses on the introduction of each speech. This is because a person's first impression would most likely have the biggest influence on their opinion, i.e. the Primacy Effect (Hogg & Vaughn, 2005). However, whether annotating and analysing the body and conclusion of a prepared speech would yield similar results as the introduction is certainly worth investigating in the future. It is worth noting that every study, no matter how refined, can never be free from limitations. However, it does not discount the fact that most studies are aimed at filling a research gap in understanding the relationship between gesture and intonation and these limitations can be addressed in future studies in order to understand this complex relationship better.

CHAPTER 2: LITERATURE REVIEW

The study of gesture and intonation in speech has undergone a lot of development throughout the years. This chapter will discuss the various studies on gesture and intonation in speech and how they have built upon each other over time. The first section will review the studies on gesture in speech. The second section will describe the previous literature on intonation and the next section will highlight various studies on both gesture and intonation. Each section will also discuss how these studies on gesture and intonation were linked to the field of public speaking. The chapter will end with a discussion of the working definitions used in this study.

2.1 Studies on Gesture

The study on gestures and how they are related to speech is not a new thing. One of the earliest studies on gesture was conducted by Kendon (1972) who laid the foundation for future studies in this field. In his study, he analysed a recording of a few people having a conversation in a pub which was 90 seconds long. At the end of his study, he was able to provide a detailed description of how gestures are organised and how they are connected to the accompanying speech. He found that gestures can be organised into a hierarchical set of units. At the top of this hierarchy, the most obvious gestures made by a person is placed there followed by the less obvious ones until one reaches the different phases of a gesture or gestural phrases which are placed at the bottom. This hierarchy has enabled a gesture to be broken down into various phrases and this framework formed the basis of many gestural microanalyses which will be discussed later in the chapter.

Apart from Kendon's (1972) work, there were many other studies conducted over the years which have managed to establish that gestures enhance the quality of a speech in many ways. One of the earliest studies which found that gesture supports speech was a study done by Graham and Argyle (1975) who conducted a series of experiments on a group of English and Italian speakers. They found that both groups of speakers could communicate their message to their listeners with greater accuracy when they were allowed to use gestures together with their speech. On the other hand, they also hypothesised that if a speaker was not allowed to gesture, the speaker's performance would be affected as he or she would pay more attention to restraining their gestures and this can be a source of distraction. In addition, the need to pay extra attention to something other than communicating to the audience also takes up more memory thus increasing the cognitive load of the speaker.

This hypothesis was supported by other works such as Rauscher, Krauss and Chen's (1996) study. They theorised that gestures play a role in retrieving words from a person's lexical memory. The researchers asked 41 speakers to describe a cartoon to them under different experimental conditions and they were videotaped. In the experiment, one group was allowed to use gestures as they described the cartoon but the other group was not allowed to do so. They found that the participants used more gestures when their speech contained a lot of spatial content. They also discovered that when a speaker cannot gesture, they have more difficulty producing speech with spatial content. They realised that the act of keeping one's hands still required additional cognitive effort on the part of the speaker and it diminished their processing capacity. Apart from that, at the end of their experiment, they also theorised that the suppression of gesture affects the conceptualising stage rather than the formulating

stage of the speech process. However, they only focused on the relationship between gestures and the spatial content of a speech which could be somewhat limited in scope.

Their findings were also supported in Goldin-Meadow, Nusbaum, Kelly and Wagner's (2001) study. They theorised that gesturing may lighten the cognitive load of a person who is thinking of what to say. They hypothesised that a speaker's memory should be better when they use gestures while speaking than when they do not gesture. On the other hand, if gesturing increased the cognitive load of a person, the reverse will take place. They conducted an experiment with 40 children and 32 adults to test this hypothesis. The experiment consisted of two stages. The first stage required the participants to solve a mathematical problem individually. They are then given a list of items to remember (a few words for the children and a sequence of letters for the adults) while describing how they solved the mathematical equation at the same time. The participants were divided into two groups where one group was allowed to gesture while speaking whereas another group was not allowed to do so.

The researchers mentioned in the previous paragraph found that both children and adults remembered a significantly larger proportion of items when they used gestures when speaking compared to those who did not. The writers also found that gesturing benefitted their participants' memory regardless of their mathematical knowledge. The results were consistent with their earlier hypothesis that gesturing reduces the cognitive load of a person when they are recalling something from memory. This lends credence to Graham and Argyle (1975) and Rauscher, Krauss and Chen's findings (1996) which state that not gesturing while speaking actually increases the cognitive load of a person when they retrieve information from memory. This shows that one of the roles gesture plays in enhancing speech is it lightens the cognitive load during memory retrieval.

There were other researchers who investigated if gestures supported verbal speech in other ways apart from improving the speaker's memory. They also found that gestures help convey additional information in a verbal utterance. McNeill (1992) conducted a study where he had five subjects describe a scene from a cartoon to him. He discovered that gestures are symbols which exhibit meanings in their own right and these symbols and their meanings are created at the moment of speaking. One of the gestures used by one of McNeill's (1992) respondents was when he described how a cartoon character bent a tree to the ground. As the gesture was being produced, the respondent also clenched his fist and bent his arm backwards as though he was gripping the trunk of a tree. This gesture seemed to indicate that the cartoon character needed a lot of strength to bend that tree. This shows that gestures may contain additional information about a person or an object which is not described in the verbal utterance. In addition, McNeill's (1992) findings helped to provide an insight into the mental processes of the speakers and gestures which are used to express meaning. For example, he found that he could guess how involved a participant was in the story by observing their gestures. He could also tell whether they were relating the story from perspective of the cartoon character or as an observer looking from the outside and all this information was also not described by his respondent in their verbal utterance.

McNeill's (1992) findings were also supported by other studies conducted after his experiment. Kendon (1995) conducted an experiment to investigate how gestures complement a verbal utterance. He made several video recordings of various natural, unscripted conversations in Italian between residents of a small village in southern Italy. He found that gestures can help to provide the context to help a person interpret a verbal expression. They also help to clarify abstract concepts within the utterance itself (Kendon, 2000). Sometimes, they can be an indication that the listener is paying

more attention to certain information within the spoken utterance. Another thing he found was that certain gestures can serve as a visual representation of the intonation question marking features of an utterance.

However, one should take note that the gestures analysed in this study are culturally specific to Italian speakers and thus the results may be different if it is conducted in another setting or language. Despite that, the fact that gestures provide contextual clues for the listener was also supported even in more recent studies such as Holle and Gunter (2007). Similar to Kendon's (1995) study, they also found that gestures also contain additional information which enables the message of a speech to be delivered more effectively.

At the same time, although it needs to be noted that even though Gunter's (2007) study only focused on iconic gestures, it does not change the fact that the findings documented by different researchers have been consistent over the years. Kita (2000) theorised that iconic gestures would not be very different among people who speak different languages as they would be talking about the same spatio-motoric experience. His hypothesis was supported by McNeill's (1992) study where he found that speakers of Georgian, Swahili, Mandarin and English used similar iconic gestures when they had to describe the same scene in their own languages. Therefore, it is possible that with training, a speaker can ensure that his or her cultural background does not have a huge influence on the gestures he or she uses to communicate with the listener.

More recent studies have also found that gestures enhance the delivery of a speech in other ways. For example, in addition to a speaker being able to improve his or her memory by using gestures while speaking, gesturing can also help a speaker to organise the information in their heads before it is delivered to the audience. Kita

(2000) built upon previous studies to come up with a theory known as the Information Packaging Hypothesis. In his study, he used McNeill's (1992) concept of representational gestures as part of his hypothesis. The Information Packaging Hypothesis suggests that gesturing helps the speaker package information into manageable chunks before it is expressed linguistically to their listeners. Therefore, when a speaker is able to organise their information effectively via gesturing, they can also choose to highlight which chunk of information is more important in their speech.

This role in which gestures play in marking prominence was also supported in more recent studies such as one done by Beattie, Webster and Ross (2014). In their study, they got a group of students and staff from the University of Manchester to narrate a cartoon to them. They found that their respondents are more likely to produce gestures when a chunk of information is highly important. Furthermore, they also found that iconic gestures are more likely to be encoded with highly important information compared to the other types of gestures. However, this study did not state whether this highly important information is important to the speaker, listener or both parties. Moreover, although it was established that the speakers used gestures to accompany the important parts of their speech, the researchers got a panel of judges to determine which chunk of information contained highly important semantic information. Therefore, instead of finding out from the speakers themselves which part of the speech was deemed important to them, they used the opinions of the judges to determine the important parts of the speech before seeing if a gesture accompanied that part.

The studies mentioned in the previous paragraphs have established that gesture clearly benefits the speaker. However, the question that still needs to be answered is whether gesture provides any communicative benefit to the listener and if it does, then

the different ways in which it benefits the listener should also be identified. Driskell and Radtke (2003) did a study with 80 US Naval Reserve military personnel to investigate how much gesturing influenced listener comprehension. They also hypothesised that listener comprehension could be aided by the effect of gesture on speech production (mediation hypothesis). The participants were randomly divided into different groups under different experimental conditions. One group was allowed to use gestures while speaking but another group was not allowed to do so. The participants took part in the study two at a time where one of them had to be the speaker and the other one would be the listener. The speaker had to describe a word to the listener who would try to guess it correctly in as few attempts as possible.

At the end of the experiment, the researchers found that the listeners took fewer attempts to guess the correct word correctly when the speakers were allowed to gesture compared to when they were not allowed to. They also found that gesturing during a speech also affected the listener's comprehension on different speech content areas. This is because the researchers found that the listeners who took the highest number of attempts had to guess words that described spatial location and manipulation or movement and gestures were not allowed when those words were being described. Therefore, they concluded that gestures are most useful in conveying content that is spatial based. In addition, they only found limited evidence to support the mediation hypothesis as it was discovered that gestures seem to have a direct effect on listener comprehension regardless of the impact gesture has on speech.

In another study, Hostetter (2011) examined 63 samples from different speakers between 1951 to 2010 to investigate if gesturing benefitted the listeners' understanding of a speaker's verbal message. She also hypothesised that gestures improved listener comprehension because they are able to process the spatial

information the gestures conveyed. The author coded each video according to different criteria such as the topic of the speech and whether the gestures were spontaneous or scripted. After the samples were analysed, the writer found that the listeners generally understand a speech better when it is accompanied by visible gestures compared to when it is not.

One possibility is that the listeners pay more attention to a speaker who gestures compared to one who does not; which aided their comprehension. In fact, they also discovered that both scripted and spontaneous gestures improved listener comprehension equally. In addition, the researchers also found that gestures that accompany spatial and motor topics benefitted communication more significantly compared to gestures that accompany abstract topics. However, listeners who listen to messages that were simpler grammatically and lexically may not benefit a lot from gestures. In contrast, if a topic is more complex, gesturing is not only beneficial but also important to facilitate listener comprehension especially among those who have a lower proficiency in the language. Therefore, it can also be said that gestures benefit the listener by enabling them to understand the speaker better.

Previous studies have determined that gestures play many roles in supporting the speaker such as it helps improve the speaker's memory, makes organisation more efficient and marks important information in the speech. Apart from that, it also communicates additional information not found in the verbal utterance and helps the listener to understand the speaker better. In fact, based on Graham and Argyle (1975) and Kendon's (1995) studies, one can hypothesise that gestures seem to serve the same function among speakers regardless of their culture or language.

However, as much as one may think that a lot of work has been done in the field of gesture, there are still many issues which need to be examined. Firstly, the studies

which were mentioned analysed only natural speech which can be interpreted as somewhat limited in scope as these studies only cover one type of speech. Moreover, although Kendon (1972) developed a hierarchy for gestural analysis, these studies mainly focused on discovering the functions gestures play in speech, how often it is produced or at which part of the speech it occurs. Thus, while one may know the functions gestures play in speech, the exact moment when a gesture is produced i.e. analysing the gestural phrases to enhance the speech has not been explored in detail. Thus, the extent of how much gesture is related to speech and its many components such as its prosodic features is still not fully understood. This is because most studies on gesture did not study the verbal utterance which accompanies it in detail although researchers have claimed that gesture and speech share the same origin (McNeill, 2005). Therefore, with all of these issues and gaps in the knowledge pool, they will be discussed later in the chapter.

2.1.1 Gesture in Public Speaking

Many scholars have made various contributions to the field of public speaking. However, many of their works only focused on analysing the content of the speech which only requires a researcher to analyse the text without studying the non-verbal cues of a speaker. For example, one such study by Willyard and Ritter (2005) investigated how American vice-presidential candidates influenced the presidential victory and concessional addresses. At the end of their study, they asserted that the vice-presidential and presidential victory and concession speeches should be studied together. In addition, they also identified the similar themes which every candidate included in their speeches. In a more recent study by Slavícková (2013), she analysed four sample texts of Presidential Memorial Day speeches in order to determine their

similarities and differences in terms of word choices and themes. She also found that they mainly shared the same speech structure and style. However, she also found that Republican and Democrat Presidents emphasised different keywords in their speeches. These studies seem to indicate that many, if not most studies on public speaking tend to focus on genre analysis.

At the same time, one also cannot say that there were no studies done on gestures in the context of a public speech. However, one needs to remember that most of these studies which investigated gestures in public speaking, sparse as they are, only tried to determine whether gestures affected the speakers and the audience. For example, Whitehead III and Smith (2002) only tried to determine whether the American presidents smiled or used more hand gestures during their speeches. In their study, they analysed the inaugural speech of five American presidents (Eisenhower, Kennedy, Nixon, Bush & Clinton) which were all prepared beforehand. They found that the presidents whom they studied tended to use more hand gestures than facial movements when they addressed their audience.

On the other hand, Yeşil (2008) wanted to investigate how much students were affected by their classmates' nonverbal behaviour during a class debate. At the end of his study, he found that the students in his study were negatively affected by their classmates' facial expressions, gestures and their intonation. In a more recent study by Lempert (2011), he investigated how and when Barack Obama used the precision-grip gesture in his speeches. The gesture is made by holding his index finger and the tip of his thumb together and his other fingers are flexed to be in contact with the palm of his hand. Although Lempert (2011) found that Obama used this gesture for many functions such as stressing a point and also as a show of strength, his whole study only revolved around this gesture made by one person.

The studies mentioned in the previous two paragraphs seem to indicate that the use of gestures in public speaking has not been studied in detail. They only seem to scratch the surface on gestures in public speaking as these studies have only identified what type of gestures are used, how they affect the audience and how frequently they are used compared to other nonverbal cues. They have not addressed various questions such as when gestures are produced during a speech and the extent of their relationship with other aspects of the verbal utterance such as intonation. In addition, most of them only seem to focus on analysing the various speeches made by American Presidents and Vice Presidents. Therefore, the study of gesture in public speaking may also be somewhat limited in scope as well. Thus, it seems that the study of gesture in public speaking has progressed slowly and further investigation in this area will need to be conducted.

2.2 Studies on Intonation

The previous section showed how the study of gestures which accompany speech have evolved over the years. At the same time, one also cannot discount the role of intonation in a speech as well. Therefore, the understanding of prosody in speech and its evolution over the years also needs to be discussed. One of the earliest studies on intonation was conducted by Bolinger (1958) who came up with a theory on pitch accent. In his study, he conducted a series of tests to investigate the role of pitch and stress in speech. Before Bolinger conducted his study, ‘stress’ used to be defined as the intensity or the loudness of the speech in terms of volume (Bloomfield, 1933). However, at the end of his study, Bolinger (1958) found that ‘stress’ was actually the changes in the pitch of the voice and how prominent the changes were. Hence, he coined the term pitch accent to describe the concept of how the pitch of a speaker

changes when he or she speaks and this can be used to show that a word in a verbal utterance is prominent. However, as much as Bolinger's (1958) work had started a change in how prosody in speech is understood, a framework which could be used to describe the various prosodic features in speech had not been designed yet.

The study of intonation in speech took a huge leap forward when Pierrehumbert (1980) developed a framework which identified the different tunes which are used in the English language and how they are aligned with the semantic content of an utterance. She also identified three components which made up the phonological representation of English. Firstly, she suggested that the tunes she identified are a sequence of the high (H) and low (L) tones. These tones actually mark the changes in the pitch of the speaker i.e. the pitch accent. This is because English is a plastic language (Gut, Pillai & Mohd. Don, 2013) where pitch and intonation are used to mark certain information in a verbal utterance for different purposes. Therefore, as they are used expressively in English (Pierrehumbert, 1980), pitch accents are normally aligned with stressed syllables. In addition to the H and L tones, Pierrehumbert (1980) also identified two extra tones which form the intermediate phrase boundary and the intonational phrase boundary. In the intonational hierarchy she developed, the intonational phrase boundary is the largest unit followed by the intermediate phrase and the various pitch accents within the intonational phrase. Figure 2.1 shows the summary of Pierrehumbert's (1980:29) intonational hierarchy of the English language. The * marks the pitch accents, the – symbolises the intermediate phrase accents while the % are the intonational phrase boundaries.

The second component of her framework is a metric grid which represents the text of a speech. This grid will enable the researcher to identify the stressed and unstressed syllables in the verbal utterances. In addition, the word boundaries for each individual

word in the text can also be determined as well. The final component of her study consists of a set of rules she discovered. These rules govern how the tune should be aligned with the text based on the first two components. Her study will form the basis for many intonational frameworks which will be developed in the future and this phonemic representation was also adapted to describe the phonemic system of other languages. Finally, it also provided future researchers with a firm foundation to gain a better understanding of the prosodic elements of a speech.

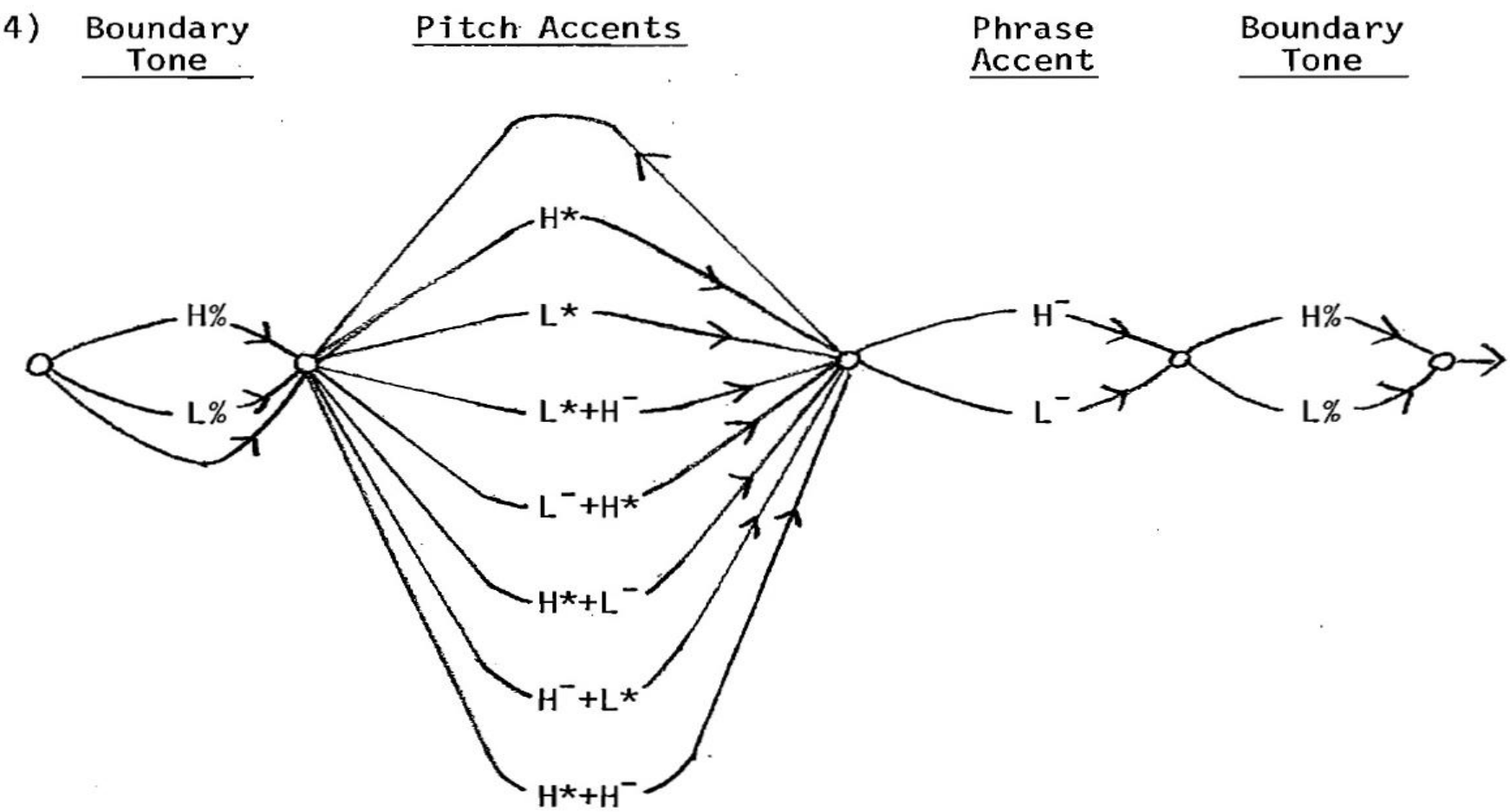


Figure 2.1: Pierrehumbert's (1980:29) diagram of the intonational hierachy in the English language.

After a formal representation of the intonation structure of the English language had been codified, researchers could accurately describe the changes in the speaker's pitch and how they are aligned with the verbal utterance. For example, Pierrehumbert and Steele (1989) were able to indicate that English speakers have two rising and falling intonation patterns and the way they are aligned with a stressed syllable is also different. In time, there were other studies that were able to identify the different roles intonation play in supporting and enhancing the quality of a speech. This is because the message the speaker intends to communicate to his or her audience can be discerned through the way the speaker varies the pitch range, accent and tune of his or her verbal utterances.

One of the most important roles intonation plays in speech is that it helps to indicate whether a piece of lexical information within an utterance is prominent. This is done by raising or lowering the pitch of the speaker. As pitch accents generally fall on the stressed syllables of a word, the speaker uses the change in the pitch to mark the word as intonationally prominent (Hirschberg & Pierrehumbert, 1986). Therefore, when a word is intonationally prominent, it can also mean the word is important semantically because it is usually a new piece of information which the speaker wants to introduce to the audience. Hence, the speaker tries very hard to ensure that the audience hears and understands that word in the speech. Moreover, if one groups a sequence of high (H) and low (L) tones together, that sequence is known as a tune (Pierrehumbert & Hirschbreg, 1990).

The tune of the speech can also convey information about the speaker's attitude and what he or she intends to communicate to their listeners without changing the meaning of the individual words in the utterance. This is because the pitch of a speaker can indicate his or her emotional state during the delivery of their speeches. For

example, Bolinger (1983) found that if a person hears a speaker talk in a high pitch, that person might infer that the speaker is tense. In addition, Hirschberg (2002) stated in her study of traditional patriarchal culture, speaking in a low pitch is normally associated with dominance and a voice with a higher pitch may indicate submission. On the other hand, a sentence can be uttered with different tunes to convey different meanings (Hirschberg, Litman, Pierrehumbert & Ward, 1987). For instance, if a person utters the sentence "This is your new cat" with a falling pitch movement, it can be interpreted as a statement. On the other hand, if the same sentence is uttered in a rising pitch movement, it can be interpreted as a question (Gut & Pillai, 2015).

Although it has been established that pitch accent plays many roles in speech, the studies mentioned earlier in this section (except for Gut & Pillai, 2015) tend to only focus their attention on the English language. Moreover, their respondents tend to be native speakers of either British or American English and these speakers usually produce 87% of all new information with a pitch accent (Brown, 1983). As English is the lingua franca of the world today, one question which can be asked is whether English speakers from other cultures use intonation in the same way as British and American English speakers. This issue was addressed by Gut, Pillai and Zuraidah (2013) when they conducted a study to investigate how Malaysian speakers of English mark new information prosodically. The researchers obtained their data from 30 university students through a game which was designed to elicit semi-spontaneous speech and they were recorded reading out a story. These students speak Malay as their first language and they only learnt English as a second language in school. Therefore, they predicted that their respondents would not prosodically mark any new information when they speak in English possibly due to interference from their first

language. This is because Malay is not a stress timed language like English and it does not use stress and intonation to mark important or new information.

At the end of Gut, Pillai and Zuraidah's (2013) study, they found that new information in Malaysian English is also marked by a pitch accent although it is not as systematic as British and American English. This is evident when new information was consistently marked by an earlier through and a larger rise. This goes to show that speakers of English as a second language tend to use different prosodic strategies to mark information structure (Gut & Pillai, 2014). Therefore, another question which can be raised is when second language learners of English speak in English, are they influenced cross-linguistically when they mark information prosodically? This is because the sentence structure, pronunciation and vocabulary of Malaysian English have been influenced by other languages spoken in Malaysia like Malay, Mandarin and Tamil. Gut and Pillai (2014) have tried to address this question in their study which involved two groups of speakers who spoke Malay as their first language. One group had to read a text in English whereas the other group had to read a text in Malay. At the end of their study, they found that prosodic patterns used by both groups are largely similar to each other. Their findings suggested that both groups of speakers appear to have applied their knowledge of their L2 (English) to their L1 (Malay) and thus, in a way have hybridised their knowledge of both languages which seem to suggest a certain extent of cross-linguistic influence on their prosodic markings of information structure.

However, further studies are still needed to shed more light on this issue. Nonetheless, these two studies show that the roles intonation play in speech are not just limited to the English language but they may also serve similar functions in different languages. Moreover, using prosody to enhance the quality of the verbal

utterance can also be taught to speakers who learn English or any other language as a second or foreign language.

The previous paragraphs have stated that intonation helps a speaker to mark important information in a verbal utterance and it also gives an indication of the speaker's attitude and the message he or she wants to convey to their listeners. This goes to show that just like gesture, intonation does not only benefit the speaker but the listener as well. Since English is a stress timed language, a listener will benefit from paying more attention to the stressed syllables where they are aligned with a pitch accent (Zheng & Pierrehumbert, 2010). This is because the words that contain stress in a speech in English normally contain the most information. Therefore, when the listener pays more attention to stressed words, it helps them to grasp the message of the speech quickly and clearly.

From creating a framework to describe the intonational structure of the English language to investigating whether intonation plays the same role across different speakers of different cultures and languages, the many studies which have been done in the field of intonation have certainly evolved over the years. However, many, if not most of these researchers only studied intonation as a separate entity from gesture as they felt there was no need to concern themselves with bodily movement (Loehr, 2004). Although this may seem like the right thing to do, it may not allow us to fully understand the mechanics behind human speech. This is because gesture and speech share the same origin (Kendon, 2004) and they are regularly coupled with intonation and they also exhibit the same ups and downs which is similar to pitch (Bolinger, 1983). Moreover, just like the various studies on gestures, all of these studies only investigated intonation in the context of natural speech which somewhat limits the

scope as well which is why the next section will discuss studies which investigated the role of intonation in the context of public speaking.

2.2.1 Intonation in Public Speaking: What Makes a 'Good' Speech?

It was only in the 20th century that researchers began to study and define what intonation is. Bolinger (1983) said that intonation is the rise and fall of the pitch as it moves along the speech chain. As researchers began to learn more about intonation and the other prosodic features that accompany speech, some of them began to investigate the role of prosody in public speaking in order to provide an objective definition of a 'good' speaker. One of these studies which investigated what makes a 'good' speaker was conducted by Strangert (2005). She compared a professional news anchor reading a piece of news on TV and a radio interview with a well-known politician. At the end of her study, she found that both the politician and the news anchor vary their speeches' tempo and they pause at certain points in order to emphasise certain portions of their speeches. In addition, their speeches are also very dynamic in terms of their tone and volume and they use this to draw the attention of their listeners to the important key words in their speech.

Strangert (2005) hypothesised that a 'good' or 'skilled' speaker is a person who is capable of attracting the interest of their listeners because he or she is able to express the message of their speech effectively in addition to having substance in the speech content. At the same time, one must also take note that both speech genres are different as the news reading is a prepared speech where the speaker can refer to the script and the interview might require the politician to speak without any preparation beforehand without any script. Hence, both speakers might use a different speaking style in these situations which might have affected the findings of the study.

On the other hand, Rosenberg and Hirschberg (2005) decided to narrow their study to investigate only one characteristic of a 'good' speaker. They investigated whether speakers who are seen as 'charismatic' share the same acoustic and prosodic elements and how these elements interact with the lexical content and syntactic form of the verbal utterance. During the study, they asked eight native speakers of American English to listen to 45 speech segments taken from various American politicians who campaigned to be the Democratic Party's nomination for President and give their opinions on each speech segment. Each segment was between 2 to 28 seconds long. At the end of the study, the researchers found that their respondents shared the same opinion on what made a speaker 'charismatic'. In their opinion, a speaker has to be enthusiastic, persuasive, charming, passionate and convincing. They also discovered that the speech segments which their respondents saw as 'charismatic' shared certain similar prosodic features. For example, they found that the speech segments which were louder in volume and had a faster speed were seen as more 'charismatic'. In addition, they also found that the speech segments which vary in pitch and intensity were also rated as 'charismatic' as the respondents might think the speakers were passionate and enthusiastic. Apart from that, the researchers also discovered that the ratio of the amount of content words to function words in each speech segment did not have any significant influence on 'charisma' but they found that speakers who used more first person pronouns in their speech segments were deemed more 'charismatic'.

However, there were also limitations in this study. One of them is that the selection of the speakers was not very clear. This is because the researchers selected them on the basis of their hypothesis that some of the politicians might demonstrate 'charismatic' qualities in their speeches. Therefore, the fact that the number of speakers used in the study who were actually 'charismatic' is very unclear and can be

called into question. Another limitation of the study is that each speech segment is so short that they can only be looked at in isolation and not in the context of their whole speech which may yield a different result. Furthermore, they are taken from various speech genres such as interviews, debates, campaign advertisements and stump speeches where the speakers may speak in a different style as they may have prepared their speeches in advance or they may to answer questions directed at them on the spot. Therefore, although both studies may have found some similarities between the prosodic features of 'good' or 'charismatic' speakers, there are still many questions which need to be answered before one can clearly define what a 'good' or 'charismatic' speaker is.

The definition of a 'good' and 'charismatic' speaker continued to be investigated in further studies after the findings mentioned in Strangert's (2005) earlier study. With this question in mind, Strangert and Gustafson (2008) expanded upon Strangert's (2005) study to further determine which prosodic features that contribute to the impression of a 'good' speaker. They found that the politician in Strangert's (2005) study displayed a greater variety of expressions and emotionally expressive acts conveyed prosodically. Therefore, Strangert and Gustafson (2008) decided to focus on studying the prosodic features of 16 speech samples (audio and video) taken from various debates between parliament members and government ministers in the Swedish parliament. These speech segments ranged between 30 to 36 seconds each and they were given to 18 Swedish students to listen to before giving their opinions on each recording. At the end of their study, they found that their respondents tended to rate a speaker as 'good' if he or she had a wide pitch range and there was a high pitch peak on the key words in their speech. In addition, their respondents also rated a speaker as 'good' if they made less mistakes such as slips of the tongue, hesitation

pauses and repetitions in their speeches. However, they also found that the speed of a speech only plays a small part in determining whether a speaker is perceived as 'good'.

As the characteristics of a 'good' speaker became more objective and quantifiable, a study was also done by Biadsy, Rosenberg, Carlson, Hirschberg and Strangert (2008) to investigate how audiences from different cultures determine what a 'charismatic' speaker is in terms of prosodic and lexical features. Their study required American, Palestinian and Swedish respondents to assess various political speech segments in Standard American English. In addition, they also got their American and Palestinian respondents to rate different speech segments in Palestinian Arabic.

The researchers found various similarities and differences between these three groups of speakers at the end of their study. They found that listeners from all three groups perceive a speaker as 'charismatic' if he or she has a high pitch range, a greater and more varied intensity within the speech segment and if there are many words which are accented with a downstepped pitch accent (!H*). They also found various differences in perceiving 'charisma' between all three groups. For instance, the Swedish students tend to perceive a speech (in Standard American English) which has a higher pitch in a compressed range to be more 'charismatic' but Americans are more likely to think that a speaker (English and Arabic) is 'charismatic' if he or she speaks at a higher rate and the speaker's pitch range is lower.

On the other hand, Palestinian listeners tend see a speaker (Arabic) as 'charismatic' if he or she speaks with a varied pitch range and lower pitch peaks compared to the American students who heard the Arabic speakers. Their opinion on what makes a speaker who speaks in English 'charismatic' are quite similar to their American counterparts except for the fact they are not as forgiving of speakers who have pause fillers and self-repairs in their speeches compared to the Americans. These two studies

have shown that prosody plays a big part in determining whether a speaker is ‘good’ or ‘charismatic’. The main prosodic features that researchers should focus on in future studies is the pitch range and intensity of a speech. At the same time, one should not discount the relationship between the lexical items of a speech and its prosodic features as both studies found that speeches (in English) which contained more first person pronouns were seen as ‘good’ and ‘charismatic’.

2.2.2 Intonation in Public Speaking: A Change in Direction

The studies in Section 2.2.1 managed to establish that prosody, especially the pitch range and intensity, plays a huge role in enhancing the quality of a speech. However, these studies mainly focused on one genre of public speech and the method of selecting the ‘good’ and ‘charismatic’ speakers for analysis may not be very objective. Recently, some researchers decided to shift the direction of their research of prosody in public speaking where they focused more of their attention on analysing the prosodic features of a speech and they also investigated other genres of public speeches as well.

One of these studies was conducted by Tsai (2015) where he compared a group of TED Talk speakers with a group of lecturers from the University of California Berkeley. His aim was to identify the differences in prosodic features between the speakers of both groups. Compared to previous studies, his method of selecting the TED Talk speech segments for analysis was more objective (albeit not perfect) as he selected the TED Talk speaker based on the number of views their speeches had and each of these speakers had more than one million views at the time of the study. This is because it could be interpreted that the more popular the speaker, the more likely he or she would be seen as ‘good’ or ‘charismatic’. In contrast, the method of selecting the academic lectures for analysis could be seen as not very objective. This is because

the speech segments were extracted from the first available lecture from various courses throughout the university as the researcher hypothesised that they are a representative of how the speakers speak in public.

At the end of the study, it was found that TED Talk speakers tend to have a lower pitch range compared to the academic lecturers. The difference in this finding compared to the previous studies could be affected by the fact that all of the TED Talk speakers (and the academic lecturers) selected in this study were all male so the average pitch range may be lower than the other studies mentioned in the previous section as they had a selection of speakers from both genders. In addition, the study also found that TED Talks contained less silence and more high energy speech. However, the writer did not really elaborate on what high energy speech was; be it the intensity of the speech or some other prosodic feature. Another thing which needed to be taken into account was that the lectures contained a lot of technical subjects and they were between 50 to 80 minutes long whereas the subject of each TED Talk is different and they typically last for about 10 to 20 minutes. Furthermore, the objective of both speech genres were different so the data selection could be seen as imbalanced as both groups of speakers had to fulfil different roles which could have affected the way they delivered their speeches. Moreover, the study mainly focused on reporting the findings of the TED Talk Speakers whereas the findings of the lectures were hardly mentioned so Tsai's (2015) study does not really shed much light on the prosodic differences between various genres of public speeches.

In contrast, another study was conducted around the same time to investigate the prosodic characteristics of public speaking. This time, the researcher only aimed to describe the prosodic features of academic lectures. Freydina (2015) selected 25 British lecturers (15 men and 10 women) and she recorded them (audio only) as they

delivered their lecture to a group of students. During the acoustic analysis, she measured the pitch, duration and intensity of their speeches. At the same time, she also took into account the social and cultural context and the rhetorical ethos of her speakers as these factors could affect her findings. Therefore, at the end of her study, she found that the lecturers in her study tend to alternate between a formal prepared speech (academic style) and an informal, spontaneous dialogue (conversational style) with their students. This is because although an academic lecture is supposed to be a formal monologue, dialogue and interaction with the audience is sometimes integrated into it to make the class more interactive in addition to being a common strategy in rhetorical discourse.

Freydina's (2015) findings also uncovered various prosodic differences between the academic style and the conversational style of the speakers' presentations. These differences are shown in Table 2.1 which is a chart that was taken from Freydina's (2015) study. It described the differences between the prosodic features of a prepared speech and a spontaneous conversation. Moreover, she also discovered that when the speakers are speaking in the academic style, they tend to have a significant increase in their pitch level when they want to emphasise the semantic value of a lexical unit. This finding is similar to the previous studies on intonation which were mentioned in Section 2.2 (pg. 34) where pitch accents can be used to mark prominence. In addition, she found that when an utterance contains new information (key utterance), the key words in the utterance would also have a high pitch level. Apart from that, the pitch range of the utterance would be broader, the tempo would be slower and the volume would be increased.

On the other hand, the other utterances which preceded or followed the key utterance has a narrow pitch range, faster tempo and the volume is reduced. Any pitch

accents within those utterances were also lower compared to the pitch accents in the key utterance. In fact, the speakers also employed pauses as a rhetorical device as they were also used to mark new and important information. However, the study also found that some of the prosodic characteristics of the academic and conversational style may have overlapped with one another from time to time. Therefore, it goes to show that prepared and spontaneous speech may interact with one another when they are used simultaneously and they may also share certain prosodic characteristics.

Table 2.1: Freydina's (2015:18) summary of the prosodic differences between a prepared speech and a spontaneous conversation

Table 2
Prosodic markers of style

Prosodic features	Academic style	Conversational style
Key	Greatly varied, mostly high and medium	Low, medium
Pitch ranges	Greatly varied, mostly broad and medium	Medium, narrow
Terminal tones	High Falls, Mid Falls, Fall-Rises, compound tones	Mid Level Tones, Low Falls, Fall-Rises
Pre-nuclear patterns (Heads)	High Level Head, Falling Head, Stepping Head, Sliding Head	Low Level Head, Mid Level Head, Falling Head
Segmentation	Medium and long intonation groups	Varied length of intonation groups with a high proportion of short intonation groups
Rate of speech	Average, decreasing on important parts of utterance	Fast
Pauses	Syntactic pauses, medium and long: emphatic and rhetorical pauses	Considerable number of hesitation pauses (flled and unflled)

The studies discussed in this section traces the development of the study of intonation in public speaking. Overall, it seems that the studies which investigated the role of intonation in public speaking have been conducted in greater quantity and depth compared to the previous studies which tried to investigate the role of gestures in public speaking. Orators from as early as ancient Rome and Greece realised the importance of a good voice and tone in the making of a good speaker. It was only in modern times where researchers have attempted to measure the prosodic elements of the human voice in order to accurately determine what makes a 'good' or 'charismatic' speaker. Studies have also been done across various speech genres and it was also found that there are some similarities and differences between speeches from different genres or cultures. In addition, the prosodic differences between a formal, prepared speech and an informal conversation has also been documented.

Nonetheless, there are issues which still need to be addressed. For instance, one has to remember that some of these studies analysed speakers who spoke from a prepared script (nomination and stump speeches) and also speakers who spoke without any prior preparation together (interviews and debates) in the same category. Thus, the method the speeches were categorised may affect the accuracy of the findings to a certain extent. Moreover, Freydina (2015) mentioned that when a presenter speaks from a prepared script and interacts with the audience spontaneously, the prosodic characteristics of both styles of speaking tend to overlap each other which also affects the accuracy of the findings. This shows that there is a need to ensure that the analysis of prepared speeches and spontaneous conversation needs to be done separately before comparing them side by side in order to accurately study and understand their prosodic characteristics. Furthermore, one also has to acknowledge that using an appropriate variation of pitch does not in itself make a memorable speech as a speaker also needs

to be aware of his or her body language (Hincks, 2004). The various studies on gestures and intonation have so far been discussed separately, which in no doubt has uncovered many new findings. Thus, the next section will discuss studies which attempt to analyse these two elements side by side in order to gain a better understanding of how these nonverbal communicative cues go hand in hand with speech.

2.3 Studies on Gesture and Intonation

Prior to the studies which attempted to study gesture and intonation together, earlier research had established that gesture and intonation come from the same origin (Bolinger, 1983). In fact, he believed that it is possible to conduct a microanalysis of the intonational patterns of speech based on its pitch movements. Although not much has been said about the possibility of conducting a microanalysis of gesture, McClave (1991) was one of the earliest researchers to investigate the extent of the relationship between gesture and intonation via microanalysis. She recorded a spontaneous conversation between two dyadic pairs of students (male-male and female-female) before analysing the gestures and intonation used by the students during their conversation. She drew on Kendon's (1972) gestural framework and McNeill's (1992) guidelines for transcribing gestures to annotate the gestures in her study. She also used McNeill's (1992) system of categorising gestures into iconic, metaphoric, deictic and beat gestures which will be discussed in the Methodology. On the other hand, she used some of Cruttenden's (1997) criteria to categorise the various intonational groups from pauses to pitch in her study.

At the end of her study, she found that the strokes of iconic, metaphoric and deictic gestures tend to coincide with the stressed syllables of a word and they do not usually cross the intonational phrase boundary. In addition, her data showed that the words which accompany iconic, metaphoric and deictic gestures are stressed unless they are pronouns. She also went on to say that her study did not find any significant correlation between intonation patterns and gestural movement. Her data showed that there was no evidence that a falling final pitch is matched by a gesture moving downward as there were also downward gestures which coincided with a rising pitch pattern.

McClave's (1991) findings dismissed earlier claims made by previous researchers which moved the research in this area in a new direction. However, one must also remember that this study was limited by the technology of that time as the researcher did not have access to software which would allow her to get a very precise measurement of the intonational patterns in her respondents' conversation. Moreover, although there have been studies which have identified the various tones in the English language (Pierrehumbert, 1980), there was also the lack of a standardised framework like the Tone and Break Indices (ToBI) framework (Beckman and Elam, 1993). Thus, the researcher did not also have a framework which could be used to organise and annotate the pitch movements of the speakers systematically which could have yielded a different result. Furthermore, her study was also confined to analysing natural speech which only gives a partial understanding of the relationship between gesture and intonation. However, despite all of these limitations, McClave (1991) still managed to set the foundation in the study of gesture and intonation and it enabled other researchers to build upon the findings of her study.

Building upon McClave's (1991) work, there was one such study which attempted to conduct a more detailed study of the relationship between gesture and intonation in speech. This was made possible with the development of various video and phonetic annotation software like Anvil (Kipp, 2016) and Praat (Boersma & Weenik, 2016) which allows a researcher to obtain a precise measurement of the gestural and intonational elements of a speech. In his study, Loehr (2004) hypothesized that an analysis of the unit boundaries of both gesture and intonation will reveal various parallels with each other. He studied four people conversing with their friends in a natural environment on a variety of topics. The speakers spoke American English as their first language and all of them were good friends. This was to ensure the conversation flowed as naturally as possible. He then annotated the gestures and intonation using the gestural and phonetic annotation software mentioned earlier.

Although his study was among the first to study this relationship in such detail, he still drew upon previous work on gesture and intonation as a guide to measure the gestural and intonational elements his subjects used during their conversation. For his intonational annotation, he used the Tone and Break Indices framework which was developed by Beckman and Elam in 1993. On the other hand, he had to adapt a set of guidelines published by McNeill (1992) as his framework to annotate his gestures. This is because unlike the area of intonation, there is still a lack of a standardised analytical framework which can be used to annotate the gestures used by a speaker.

At the end of his study, one of the significant things he found was that gesture and intonation are related in terms of timing, structure and meaning. This is because the 'apex' of a gesture is generally aligned with a pitch accent. This finding is a more precise expansion of McClave's (1991) finding which found that gestural strokes are often aligned with stressed syllables. In addition, Loehr's (2004) study also debunked

Bolinger's (1983) Parallel Hypothesis which predicted that pitch and body movements move in the same direction which reflects increased or decreased tension. These findings also lent credence to McClave's (1991) study where she came to the same finding despite the technological limitation at that time. He also found that most of the gestural phrases in his data were also aligned with the intermediate phrases of the speech segments in his data. Furthermore, his research also found that gesture and intonation work together to serve various pragmatic functions in speech such as marking prominence which was also consistent with previous studies that investigated gesture and intonation separately. These findings also strengthened the hypothesis that gesture and intonation are of the same origin even though they express the same idea in different ways.

However, as much as Loehr's (2004) work supported earlier studies, some of his findings could also be questioned as well. One of his main findings was that he found that pitch accents are usually aligned with the 'apex' of a gesture. Although he chose to use a seemingly more accurate measure of 'stress' to identify a significant change in the pitch i.e. the pitch accent, his attempt to provide a more accurate measure of a gestural stroke did not seem to be as clear as the concept of a pitch accent. He defined the 'apex' as part of the stroke phrase where the 'apex' of the stroke is the exact moment when the 'kinetic goal' of the gesture is expressed i.e. the 'peak of the peak'. He did not elaborate or provide any clear criteria to identify this 'peak of the peak' of the stroke. Thus, compared to Kendon's (1972) concept of the gestural stroke, the means of identifying this 'apex' seem to be based on very subjective guesswork. Moreover, this concept does not take other factors into account such as the fact that when an idea is expressed by a gesture, it may not always come down to one moment but it may take a certain amount of time. Moreover, there are also other factors which

the gestural ‘apex’ cannot account for and they will be discussed later in the study. Thus, all of the other findings were also affected so it seems that what was found in this study may need further investigation.

There were other researchers who continued to build upon Loehr’s (2004) study. Mendoza-Denton and Jannedy (2011) analysed a speaker’s speech and gestures at a public congressional town hall meeting in Tucson, Arizona. Similar to Loehr’s (2004) experiment, this was also natural speech where the speaker spoke spontaneously. The writers wanted to investigate the gesture-intonation timeline of a speech and they used one speech at the meeting as a case study. The speaker was engaged in a dialogue with a Congressman during the town hall meeting. They also analysed the speech using microanalysis of the speaker’s gestures and intonation. They hypothesised that speech and gesture are parallel to each other and they both carry meaning in structure, content and social meaning.

At the end of their study, the writers found that the speaker they analysed tends to align pitch accents with syllables of words which carry important information of her speech. In addition, they also found the gestural ‘apices’ in the speech always occur together with the pitch accent which is similar to Loehr’s (2004) findings. Thus, they concluded that gestures in speech are not spontaneous movements but are finely coordinated structures of movement which are aligned with semantic content. This is because they discovered that accented words are usually accompanied by a gesture. Therefore, they theorised that gestures together with intonation help reinforce the important parts of a speech and it also highlights new content that is brought into a speech.

However, as much as this study supported previous research in this area, they also used Loehr's (2004) concept of the gestural 'apex' to investigate the relationship between speech and gesture. In addition, they did not provide any concrete guidelines to accurately identify the gestural 'apex' of a stroke. Apart from that, one also has to remember that the gestural 'apex' also does not take into account whether one gesture can be used to express more than one idea in a verbal utterance and it is also silent on whether a gesture may have more than one 'kinetic goal'. This shows that although there seems to be some light shed on the relationship between gesture and intonation and the roles they play in supporting a speech, there are still some issues which need to be addressed especially in obtaining a clear definition of the 'apex' of the stroke of a gesture and providing clear guidelines to identify it.

Researchers have also attempted to investigate the relationship between gestures and intonation with speakers who come from another culture via microanalysis. Brentari, Marotta, Margherita and Ott (2013) compared how gestures and intonation worked together in American and Italian speakers. They made a few recordings of their participants telling them about an event that happened in their lives and their description of a cartoon which was also used in McNeill's (1992) study. They analysed the intonational elements in the recordings with Praat (Boersma & Weenik, 2016) and the gestures were analysed with Elan (Wittenburg, Brugman, Russel, Klassman & Sloetjes, 2006).

At the end of the study, it was found that the majority of the pitch accents occurred within the stroke phrase of the gesture, regardless of the cultural background of the speakers which confirms previous work in gesture and gesture and intonation. Apart from that, this study also confirms Loehr's (2004) findings which showed that a gesture tends to coincide with its co-occurring word or slightly precedes it. This is

because the study found that if the pitch accent does not occur within the stroke, it usually follows the stroke rather that precede it, also regardless of the speaker's cultural background. On the other hand, the researchers also found some differences between both groups of speakers. For example, they found that the Italian speakers gestured more with their hands compared to their American counterparts. The study also showed that gesturing had a greater effect on the vowels of the American speakers compared to the Italian ones.

One thing that stood out in Brentari, Marotta, Margherita and Ott's study (2013) was that they found that most gestures which contribute meaning to a verbal utterance are largely produced by the speaker's hands, regardless of culture. This is similar to Whitehead III and Smith's (2002) findings (refer to section 2.1.1, pg. 29) where they found that their subjects gestured mostly with their hands. Therefore, it is also possible that a lot, if not most of the nonverbal communication in either a natural conversation or a prepared speech will also be communicated through the speakers' hands.

Another prominent feature of their study was they did not use Loehr's (2004) concept of the gestural 'apex'. Instead, they used Kendon's (1972) concept of the gestural stroke and this would likely make their findings more accurate. In addition, this study used the term representational gestures to define the meaningful gestures which accompanied the speeches in their study. However, the study did not provide a clear definition of some of the gestures which were annotated. For instance, prosodic gestures, prosodic grooming gestures and emblems were not clearly defined in the text but instead the study only included one visual example of these gestures. Although the study stated that these gestures are often produced at the end of a prosodic unit, it did not clarify whether these gestures are actually beat gestures or they only form one type of beat gestures. Moreover, the sample size of the study was small and the amount of

data that was annotated (out of 36 minutes and 5 seconds) was not mentioned either. Thus, it was also unclear whether there was sufficient data to form a credible hypothesis from the findings. Nonetheless, this study and the ones which preceded it has revealed that research in the field of gesture and intonation is moving in a certain direction and there are certain similarities which run through all of them.

The research in the field of gesture and intonation have taken many steps forward over the years. One of the main things these studies have established is that gesture and intonation share a complex relationship. They also found that the main elements to watch out for when analysing gesture and intonation are the strokes of a gesture and any significant changes in the pitch movement of a speaker. This is because all of these studies found that these pitch accents and the stroke phrases tend to cluster together when one analyses them in detail. Apart from that, it has also been established that one of the main functions gesture and intonation serve in speech is that they are used to mark prominent lexical items in a verbal utterance. In addition to serving other pragmatic roles in speech, the studies mentioned in this section have also debunked earlier claims that gesture and pitch move in the same direction. In fact, these findings also seem to be similar in speakers from other cultural and linguistic backgrounds other than American English speakers albeit with a few differences.

One thing which has remained constant is the method used to analyse the data gathered during a study. Many, if not most of these studies used the method of microanalysis in their research to study gestures and the relationship between gestures and intonation. However, the earlier studies in these fields were limited by the technology of those times which made annotating the gestures and the individual syllables in minute detail a very challenging task. Thus, earlier researchers were only able to analyse a small amount of data in their studies. Conden and Ogston (1967)

microanalysed 5½ seconds of data which comprised of 24 words spoken by one person. Kendon (1972) came up with his theories on gesture based on 90 seconds of data which consisted of an excerpt from a single individual's conversation turn in a pub in London. It was McClave (1991) who built upon these initial findings and studied gesture and intonation in greater detail despite the various limitations she faced. With the advent of video and phonetic analysis software, microanalysis of gestures and intonation was able to be conducted with more precision so the findings became more detailed and specific as researchers were able to isolate the gestural phrases and individual pitch accents down to the microsecond.

However, there are also still areas which remain unexplored despite all of these advancements. The scope of the research still seems limited as all of these studies mentioned have mainly concentrated on analysing one type of speech which are natural, spontaneous conversations. These conversations are usually between pairs or small groups of people or these speakers were asked to describe a cartoon to the researchers which is similar to McNeill's (1992) experiment. In fact, there seems to be little to no studies done on gesture and intonation in other types of speeches, namely public speeches, which are usually prepared and rehearsed beforehand. There is a need for research into this type of speech as Freydina's (2015) findings has indicated that prepared and spontaneous speeches are different prosodically. Therefore, gestural and intonational research into public speaking could yield different findings or even reinforce the findings from the previous studies.

Another thing which needs to be looked at is although previous studies have discovered that there is a relationship between gesture and intonation in speech, the extent of this relationship is still not fully understood. For instance, Mendoza-Denton and Jannedy (2011) mentioned that people can speak without gesturing but they rarely

gesture without speaking. Thus, this brings up the question of whether gesture or intonation has the bigger influence on the quality of a speech. As much as the current study acknowledges the advancements made by all of these studies mentioned in this section, it also aims to address these questions that were raised or left unanswered in order to help contribute to a better understanding of how gesture and intonation are related.

2.4 Working Definitions

The current study will use certain terms to describe the relationship between gesture and intonation in public speaking. However, these terms will need to be clearly defined in the context of the study in order to prevent any misunderstanding.

Public speaking requires the use of nonverbal cues such as gesturing. A gesture usually refers to the spontaneous bodily movements particularly of the hands, fingers and arms that accompany speech (McNeill, 2005). Gestures are visible actions together with an utterance or part of an utterance and it does not refer to movements that people make when they are nervous (Kendon, 2004).

Gestures can be classified into several types such as iconic gestures, metaphoric gestures, deictic gestures and beats (McNeill, 1992). An iconic gesture is a gesture which is used to represent a concrete object, place or person during a speech. For example, if a speaker wants to describe the action of bending a tree, he stretches out his hands and pulls them towards his chest as though bending the tree backwards. On the other hand, a metaphoric gesture is similar to an iconic gesture except that it is used to describe abstract concepts to the audience. For instance, a speaker can describe the concept of love by putting his or her hands over their heart. Deictic gestures are used when the speaker points at a place on the stage, the audience or him or herself.

These gestures can be used to point at a concrete place, an object or an abstract idea. In contrast, beats are different from the other types of gestures as they are repetitive gestures that move in rhythm with the speech. Beats are normally performed by the hand and they are usually of the same form regardless of the content of the speech. They also move in the same direction in time with the speech be it up and down, in and out or left and right. However, unlike the other type of gestures which are produced with the clear purpose of communicating an idea to the audience, a beat gesture is empty of meaningful content (Loehr, 2007).

At the same time, one also needs to be aware of Hostetter and Alibali's (2010) definition of representational gestures. These gestures refer to the movements a speaker makes to depict the objects, events and experiences the speaker is describing. In other words, Hostetter and Alibali's (2010) concept of representational gestures also synthesises McNeill's (1992) classification of iconic, metaphoric and deictic gestures under one comprehensive definition. In addition, it was also found that during the pilot study, which will be discussed in the next chapter, some gestures can be classified in more than one gestural category. For example, when one of the speakers stretched his hands to the side and raised them to indicate that the bright lights of the petrol station were shining on him, it could be interpreted as a metaphoric or deictic gesture. This is shown in Figure 2.2 where one of the speakers (Vasilev) performed this gesture. This could be interpreted to represent the lights themselves (metaphoric) or him pointing at the lights (deictic). This is possible because some speakers may use one gesture to represent two or more different ideas in their speeches.

As McNeill's (1992) classification of gestures did not account for this possibility, the current study will use Hostetter and Alibali's (2010) definition of representational gestures where any gesture (regardless of its classification) that contributes meaning

to the message of the speech is classified as a gesture. Hence, as beat gestures do not contain any semantic content, the study will not annotate them as they are not considered representational gestures. Furthermore, it was found that there were no beat gestures in any of the speech segments analysed in this study as every Toastmaster is trained to only use purposeful gestures to communicate their ideas to their audience.

As it is with gesture, one also needs to understand what intonation is. Intonation can be defined as the rise and fall of the pitch as it occurs along the speech chain (Bolinger, 1983). These changes in pitch may be prominent or otherwise. Sometimes, these tonal patterns may occur at prominent or stressed syllables (Pierrehumbert, 1980). If there is a significant height or pitch movement that is connected to a stressed syllable, it is known as a pitch accent (Gut, Pillai & Zuraidah, 2013). Therefore, if a pitch accent occurs on a stressed or prominent syllable, they also can be used to mark lexical items which are prominent in an utterance (Pierrehumbert & Hirschberg, 1990). This shows that the way intonation has been defined across many studies over the years seems to focus on the changes in the pitch of the speaker when he or she is delivering a speech. In fact, most studies have shown that it is the changes in the pitch which carries the most prosodic information in a verbal utterance. This is because words are made intonationally prominent to convey information such as contrast, focus, topic and information status (Rosenberg & Hirschberg, 2009). Moreover, pitch movements also serve to distinguish the various meanings a phrase or a sentence may have without changing the meaning of the individual words in the utterance (Gut & Pillai, 2015).

Therefore, the current study has chosen to define intonation as any significant movement, regardless of direction, in the pitch pattern of the speaker's verbal utterances, i.e. pitch accent. Moreover, the intermediate phrases and intonational

phrases also need to be clearly defined in this study. An intermediate phrase consists of one or more pitch accents and a simple high or low tone which marks the end of the phrase. On the other hand, intonational phrases consist of one or more intermediate phrases (Pierrehumbert & Hirschberg, 1990). The current study will use Pierrehumbert & Hirschberg's (1990) definition of the intermediate and intonational phrases when the speeches are annotated and analysed. This is because these definitions have been used in many studies after them and this definition has been able to account for almost every possible scenario a researcher might face when he or she annotates the intonational elements of a speaker's verbal utterance.

When the hand gestures and intonation in the study are annotated, one can see that they always accompany certain lexical items in the speech segments. These lexical items are divided into content words and function words. Content words are lexical items that carry a high information load and they usually consist of nouns, verbs, adjectives and adverbs. Function words (or grammatical words), on the other hand, contribute to the grammatical structure of the sentence. They comprise of prepositions, determiners, conjunctions and pronouns (Thornbury, 2002). These definitions of content and function words are also used in other studies on public speaking (Rosenberg & Hirschberg, 2005) so the current study will use these definitions as well.

The key terms in this study whose definitions were discussed earlier are *gestures, representational gestures, intonation, pitch accent, intermediate phrase, intonational phrase, content words* and *function words*. They will be used in this context throughout the rest of this study.

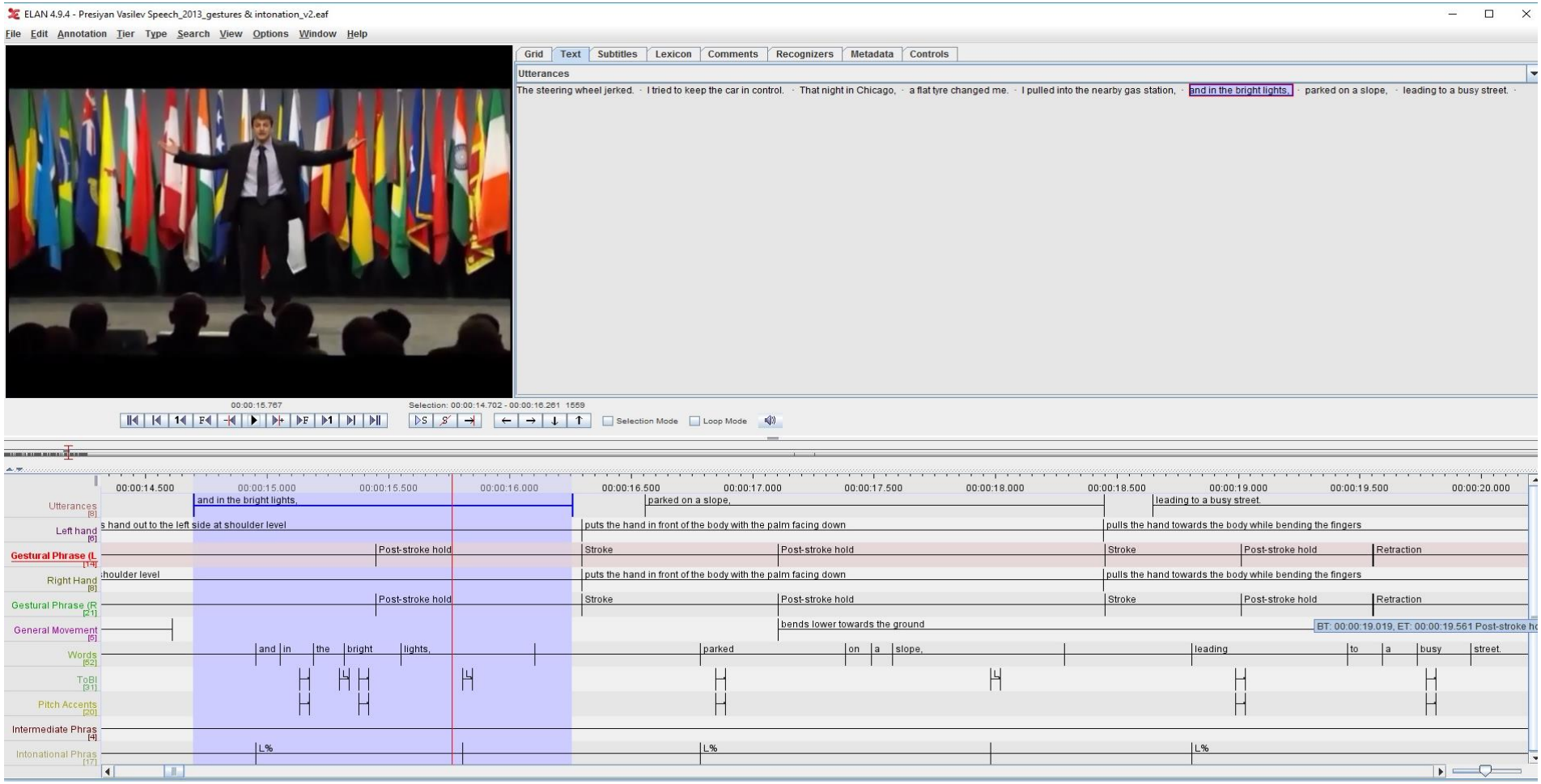


Figure 2.2: Vasilev raising his hands in the air during his speech.

CHAPTER 3: METHODOLOGY

This chapter is divided into three sections. The first section will discuss the analytical framework used in this study. The next section will then describe how the four speech segments were selected for this study. Then, the following part will describe how the hand gestures and intonation of each speech segment were annotated and analysed. This chapter will end with a discussion of a pilot study which was conducted before the four speeches were selected.

3.1 Analytical Framework

The study will use two analytical frameworks as a guide when the data from the four speech segments were analysed. The framework used during the gestural analysis is a hierarchy developed by Kendon (1972). On the other hand, the Tone and Break Indices (ToBI) framework developed by Beckman and Elam in 1993 was used during the intonational analysis.

3.1.1 Analytical Framework (Gestures)

Kendon (1972) developed a hierarchy to classify gestures and its various stages and this hierarchy is shown in Figure 3.1. This framework was selected for this study as it clearly captures the gestures made by the hands of the speaker. In addition, it breaks down the stages of a hand gesture from the macro to the micro level. Although this framework does not contain guidelines on how to head or body movements should be defined or annotated, this does not affect the study as the focus is only on the hand gestures of the speakers.

The most obvious unit of Kendon's (1972) hierarchy is the movement of the arms and body posture. The second most obvious unit was the movement of the head. It was

then followed by the gestural unit and finally the gestural phrase. A gestural unit is the stage when the arms of the speaker move from a rest position to the time it returns to the rest position. This rest position may not be the same position as the original rest position used by the speaker.

At the bottom of the hierarchy is the gestural phrase which contains three main phrases. They are the preparation, stroke and the retraction phrase. However, there are also two optional gestural phrases that the speaker may use which are the pre-stroke hold and the post-stroke hold. The preparation stage is when the hands begin to move from the rest position towards the place where the stroke will take place. Next, the stroke is the stage where the gesture is performed and the final stage is the retraction. It is the time when the speaker's hands move back to a rest or a neutral position. The pre-stroke hold is the moment where a speaker 'holds' a gesture just before the moment he or she executes it i.e. the stroke phrase. The post-stroke phrase, on the other hand, is the time when a speaker maintains the same gesture after the stroke has been performed.

As mentioned in Section 2.3 (pg. 51), Loehr (2004) defined the 'apex' of a gesture as the exact moment when the stroke of a gesture is performed. However, the current study will use McNeill's (1992) definition of the stroke. He defined the stroke as the moment when the gesture is executed and its message is expressed to the audience. This is because in Loehr's (2004) study, the gestures his subjects make usually consist of the preparation, stroke and retraction phrase. In contrast, the speakers in this study occasionally maintain a gesture after they have performed it, i.e. the post-stroke hold. If Loehr's (2004) definition of a gestural 'apex' refers to a single moment when the gesture is performed during the stroke phrase, it does not account for the time when the speakers maintain the 'peak' of their gesture for a period of time during the post-

stroke hold phrase. Moreover, sometimes the same gesture may be used to express more than one idea in a speech. Thus, it can be quite difficult to accurately identify the gestural ‘apex’ which expresses both ideas.

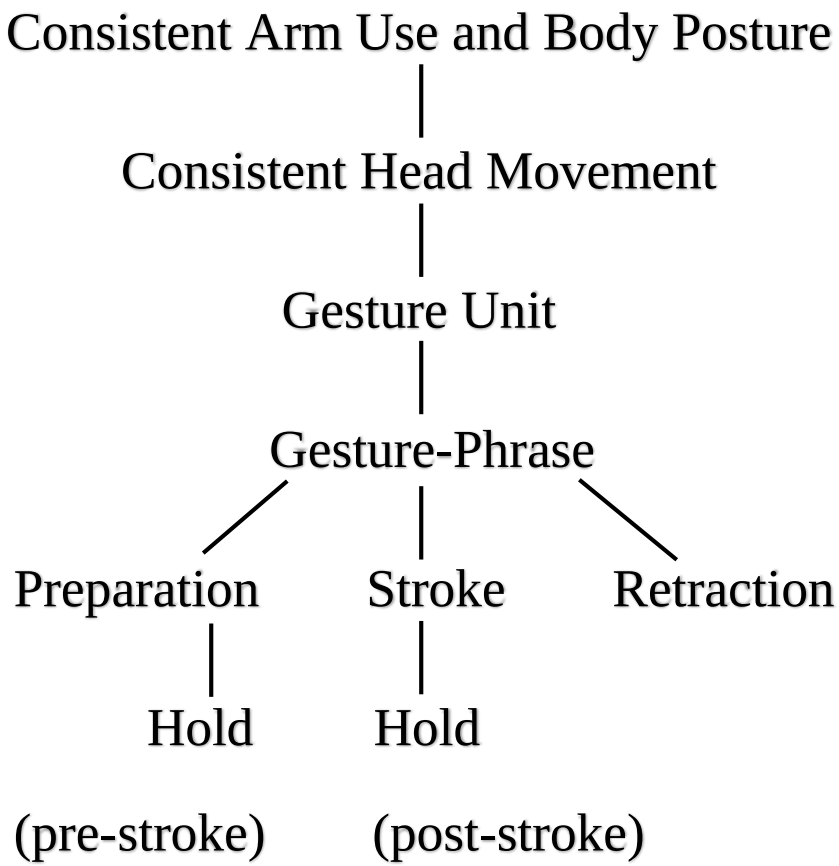


Figure 3.1: Kendon's (1972) gestural hierarchy [taken from McNeill (1992:82)]

3.1.2 Analytical Framework (Intonation)

On the other hand, The Tone Break and Indices (ToBI) system (Beckman and Elam, 1993) was used as a guideline to annotate the intonational features of the speakers. This framework was based on Pierrehumbert's (1980) study which identified the different tones in the English language and how they are aligned with different texts. The annotation based on the ToBI framework usually consists of four tiers which are the tone tier, the orthographic tier, the break index tier and the miscellaneous tier. The tone and the break indices tiers form the core of the entire prosodic annotation and analysis in the ToBI framework.

The tone tier is used to keep track of the different prosodic units and the intonation patterns in the verbal utterance. Therefore, this tier can be used to annotate any significant changes in the pitch of the four speech segments used in this study. Beckman and Elam (1993) devised a way to mark these changes in a speaker's pitch by using certain symbols. For example, if there is a rise in the pitch of a speaker, that rise is marked with a H*. On the other hand, a significant drop in a speaker's pitch is marked with a L*. In addition, a downstepped tone is marked as !H*. In addition, the tone tier can also be used to identify if a verbal utterance is rising or falling. For instance, if the tone of an intermediate phrase is falling, it is marked with the symbol (L-) and a falling intonational phrase is marked with L%. If these phrases are rising, they are marked with the H- or H% symbol instead. One example of their annotation is shown in Figure 3.2 below where the symbols mentioned earlier are used to mark the syllables where the change in the pitch occurred.

Mariana made the marmalade.
H* H* L-L%

Figure 3.2: Sample of Beckman and Elam's (1993:11) annotation.

The break index tier is used to indicate the beginning and ending i.e. of the boundaries of a verbal utterance. In this tier, the boundaries between individual words can be annotated. In addition, the intermediate and intonational phrase boundaries can also be identified. There are 5 levels of break indices between words and phrases which are:

Level 0: Minimal juncture in clear phonetic marks of clitic groups such as flapping

Level 1: Typical inter-word boundaries

Level 2: a juncture marked by a pause which indicates a phrase boundary but with no tonal marks

Level 3: an intermediate phrase boundary

Level 4: an intonational phrase boundary

One way these levels can be annotated in an utterance is that they are written as numbers below the individual words or at the end of a verbal utterance in accordance with Beckman and Elam's (1993) guidelines. Figure 3.3 below shows an example of how the break index tier can be combined with the tone tier mentioned in the previous paragraph. The 1s are written in between each word in Figure 3.3 indicate the boundaries between each word and the 4 at the end of the sentence shows the intonational phrase boundary.

Mariana	made	the	marmalade.
H*			H* L-L%
1	1	1	4

Figure 3.3: Another sample of Beckman and Elam's (1993:11) annotation.

The orthographic tier in ToBI is used to capture the transcription of the entire verbal utterance in the conventional English orthography (Beckman and Elam, 1993). This is the stage when the verbal utterances in a study is transcribed into conventional English orthography. However, one may argue that Harvey and Henderson's speech segments contained sounds like the barking of a dog and the crackling of a radio which are not meaningful words in the English language. These sounds are part of a group of words called onomatopoeia which are words that are used to describe how a sound is made such as bang, whizz and sizzle. Thus, these sounds made by both speakers were also annotated in this study.

The final tier used in the ToBI framework is the miscellaneous tier. This tier is usually used to capture elements which are not part of the prosodic analysis such as coughs or pause fillers which do not serve any purposeful functions such as marking prominence or reinforcing the meaning of the verbal utterance. Therefore, in the context of this study, this tier will not be included as all four speech segments did not contain any pause fillers, coughs or other unnecessary utterances. This is because every member of Toastmasters International is taught that every gesture and sound they make in their speeches must always be done with a purpose in mind and they must contribute to reinforcing the message of their speech.

Although there have been previous studies (Breen, Dilley, Kraemer & Gibson, 2012) which suggested that the ToBI framework contained certain limitations, it was still chosen as a guideline for the intonational annotation. This is because ToBI was developed to annotate the prosodic features of mainstream American English (Breen, Dilley, Kraemer and Gibson, 2012). Three of the speakers in this study are Americans and they are Harvey, Henderson and Avery. Although the fourth speaker (Jhingran) is a native of India, he is a graduate of the Massachusetts Institute of Technology and he

works in America. Moreover, all of them are members of Toastmasters clubs based in America so it is likely all four speakers can speak American English fluently.

Another reason why this framework is used is because the researchers who developed ToBI have provided guidelines to enable a user to annotate any ambiguous audio signal (Brugos, Veilleux, Breen & Shattuck-Hufnagel, 2008). Furthermore, it can also be used to clearly mark the word and intonational boundaries during the annotation process. In addition, the invention of annotation software such as Praat version 6.0.21 (Boersma & Weenik, 2016) has made the task of using ToBI to annotate the prosodic elements such as the changes in a speaker's pitch a lot more precise and objective. This is because the software enables the researcher to see the pitch track of a speaker so the changes can be clearly identified compared to only listening for the changes in a speaker's pitch. Furthermore, the software even allows a word in a speech segment to be isolated to its individual phoneme which allows for the various boundaries in the break index tier to be marked clearly and accurately.

3.2 Gathering of Data

Every year, Toastmasters International organises the World Championship of Public Speaking. 10 speakers from around the world are selected from more than 345000 members to compete in this tournament and they are required to deliver a 5 to 7-minute speech before a large audience. These speeches are prepared a day in advance and the speakers will have rehearsed the speech multiple times before they delivered it to the audience. Therefore, what the audience will see is a scripted performance and the likelihood of any spontaneous interaction between the speaker and the audience is very low. In fact, some of these speakers could also have gone as far as rehearsing their gestures before delivering their speech to look as natural as

possible to the audience (Toastmasters International, 2011b). This ensures that the study will only annotate and analyse the gestural and intonational features of a prepared speech.

Four video recordings of speakers who gave their winning speeches at the Toastmasters World Championship of Public Speaking in 2004, 2007, 2010 and 2012 were selected. The 2004 (Randy Harvey) and 2012 (Ryan Avery) champions are Caucasian males from North America and the 2010 (David Henderson) champion is an African American male. On the other hand, the 2007 (Vikas Jhingran) champion is a native of India but he is working in North America. After obtaining their speeches, they were converted into different formats before they were exported to Elan (Wittenburg, Brugman, Russell, Klassman & Sloetjes, 2006) and Praat for gestural and intonational annotation. Both software will be further discussed in this chapter.

Each of the four speakers spoke for about seven minutes and the length of each of their entire speeches is shown in Table 3.1. Although the length of each speech is different, it did not affect the study in any way as the study only annotated and analysed the hook or introduction of each speech. This study only focused on the hook of the speech because of the Primacy Effect (refer to Section 1.5, pg. 19). Therefore, it is possible that the introduction of the speech will have the greatest influence on how the audience and judges evaluate the speaker as their initial impression of the speaker may significantly influence how they view the rest of the speech.

During the selection of each of the speech segments, there were a few issues which needed to be addressed. The speakers' different cultural backgrounds may raise the question of whether it would affect the accuracy of the annotation and analysis of the study. However, the training which speakers receive in Toastmasters International

may minimise any influence their cultural backgrounds may have on the crafting and delivery of their speeches (refer to Section 2.1, pg. 24).

One evidence of this training is reflected in the similarities of the structure and hook of the speech of each of the four speakers. Although they come from different cultural backgrounds, all of them are members of Toastmasters International and they use the same manuals produced by the organisation to hone their public speaking skills. Therefore, each speaker is trained to craft a speech which consists of a clear hook or introduction, body and conclusion along with a clear message for the audience to take home. Therefore, each speech segment i.e. the hook, is mostly similar in terms of structure. Although the length of their hooks may differ, all of them used a personal anecdote to grab the attention of the audience as this is considered to be one example of a good introduction (Toastmasters International, 2011a). Then, they moved on to the body of the speech by saying "Mr. Contest Chair…" which is a clear transition marker that the introduction had ended. The duration of each speaker's introduction is also shown in Table 3.1.

There was also the concern in previous studies which found that the behaviour of the subjects would change when they know they are being recorded (McClave, 1991). She encountered this issue during her study when one of her subjects froze up and did not gesture at all during the entire session when she realised her conversation was being filmed. In contrast, the speakers in the current study have been trained to manage their nervousness and emotions when they are speaking in front of an audience. In fact, it is also likely that their earlier speeches in the earlier rounds have also been recorded so they would also be more comfortable with speaking in front of a camera during the competition. Furthermore, they have crafted their speeches and rehearsed them before they stepped up to the stage to deliver them to the audience. Thus, it is

likely that speaking in front of the cameras would not have a huge influence on their behaviour.

Table 3.1: The total length of each speech and its introduction.

Speech	Duration	
	Hook	Whole Speech
Harvey (2004)	36.82 seconds	7 minutes 27 seconds
Jhingran (2007)	51.06 seconds	6 minutes 58 seconds
Henderson (2010)	27.5 seconds	7 minutes 11 seconds
Avery (2012)	50.61 seconds	6 minutes 45 seconds
Total	165.99 seconds	28 minutes 21 seconds

The total amount of data annotated amounted to 166 seconds out of slightly more than 28 minutes of speech (refer to Table 3.1). This is consistent with previous microanalytic studies on gesture and intonation. Loehr (2004) only looked at 164 seconds out of an hour of recording and Mendoza-Denton and Jannedy (2011) analysed 130 seconds out of 2 hours of recording. Furthermore, 138 gestures (when the speeches from the pilot study were included) were annotated during the study which is almost the same amount in previous studies as well. McClave (1991) annotated only 125 gestures and Loehr (2004) annotated 147 gestures in their studies. The more recent studies (Brentari, Marotta, Margherita & Ott, 2013) did not mention the duration of their annotation or the amount of gestures annotated.

3.3 Gestural Annotation

The focus of the study was on the hand gestures made by each speaker and how they are related to the changes of the pitch in each speech segment. The reason why the study chose to pay attention to the hand gestures made by these speakers was because most nonverbal communication is channelled through the hands (refer to Section 2.3, pg. 54).

The gestures used in each speech were annotated with Elan (Wittenburg, Brugman, Russell, Klassman & Sloetjes, 2006) which is a linguistic software which is used to analyse video clips. A total of 11 tiers used in the pilot study was needed to annotate the words, gestures and intonation in each speech. Six tiers were used when the gestures were annotated in Elan. The first tier was used to annotate the total number of utterances in the hook of each speech.

The second tier is used to describe the gestures performed by the right hand during the speech. However, up to date, there has been also no specific framework which could be used as a guideline to code gestures produced during public speeches. Therefore, the method used to describe the gestures of the left and right hand had to be adapted from McNeill's (1992) guidelines. A sample of his guidelines are shown in Figure 3.4. However, only certain parts of McNeill's (1992) framework were suitable for the context of this study. This is because the framework only described the hand gestures of the respondent in one frame but this study requires the hand gestures of the speakers to be described over a duration of time to accurately capture the location where the pitch accent occurs within the gesture. In addition, the annotation for the left and right hands of the speaker were not separated in McNeill's (1992) framework. However, in the current study, the movements of the speaker's left and right hands were separated into different tiers.

The second tier was used to annotate the movements of the speakers' right hands. After that, the different gestural phrases for each gesture performed by the right hand were annotated in the third tier using Kendon's (1972) hierarchy as a guideline. Next, the fourth tier will be used to capture the gestures the speakers made with their left hands. The fifth tier was used to annotate the gestural phrases for each gesture performed by the left hand. During the annotation of the gestural phrases, each gesture is viewed frame by frame in Elan (Wittenburg, Brugman, Russell, Klassman & Sloetjes, 2006) to determine the correct gestural phrase within each gesture.

The sixth tier was labelled General Movement. It was used to describe any other obvious gestures such as moving around the stage or looking in a certain direction. This tier was created to assist in describing the gestures of both hands accurately. For example, when Vasilev tilted his body to the left and right when his hands were in front of his body as though grasping a steering wheel, it was initially unclear whether his hands were moving or his body movements were giving the illusion that his hands were moving. By adding this tier, the general body movements and hand gestures could be isolated thus confirming that his hands were moving during this section of his speech. Figure 3.5 shows a screenshot of the six tiers in Vasilev's speech segment. The sixth tier i.e. General Movement was used to isolate Vasilev's body and hand gestures so they can be annotated accurately and it is indicated by the arrow in Figure 3.5.

4. If there are exactly two movement phases,[6] is the initial phase in the same place as the second phase?

A score of zero means a beat on formal grounds; higher scores reflect increasing iconicity, on formal grounds.

VI. Format

Use the following format for transcribing the gestures accompanying speech:

Frame number
Speech (bracketed, in boldface type, and underscored with dots to show movement phases and post-stroke holds)
Gesture type
Which hand, shape, palm/finger orientation, hand place, motion shape, motion place
H (hand meaning); M (motion meaning); B (body meaning); S (space meaning, if any)
Gloss of entire gesture
Viewpoint (and which character in C-VPT)
Beat filter score
Confidence score.

VII. Examples

A. *An iconic gesture in which the two hands make rings in front of the eyes* (see fig. A.2):

Figure A.2. An iconic gesture on the narrative level with "so he's [looking at] Tweety Bird."

Frame number:	0199
Speech:	[so he's **looking** at Tweety Bird]
	....
Gesture type:	iconic
Which hand:	2SH
Shape:	O-hand
Palm/finger orientation:	PTC and FAB
Hand place:	at eyes

6. A movement phase is one of the following: preparation, stroke, or retraction.

Motion shape and place:	no motion
Hand meaning	H = S's hands holding binoculars
Motion meaning	M = no motion
Body meaning	B = Sylvester
Space meaning	S = space between Sylvester and Tweety Bird
Gloss:	S is looking through binoculars at TB
Viewpoint:	C-VPT (Sylvester)
Beat filter:	4 (Q1 = yes, Q2 = 2, Q3 = yes, Q4 = nonapplicable)

Confidence: 5 (metaphysical certitude)

B. *A metaphoric gesture in which one hand appears to support an object and hand it to the listener* (see fig. A.3):

Figure A.3. A metaphoric gesture on the metanarrative level with "[some of the ways that he goes about trying to catch Tweety Bird.]"

Frame number:	2741
Speech:	[some of the ways that he goes about trying to]
	.....
Gesture type:	metaphoric
Which hand:	RH
Shape:	B-spread
Palm/finger orientation:	PTU
Hand place:	lower center periphery
Motion shape and place:	no motion
Hand meaning	H = container/"some of the ways"
Motion meaning	M = no motion
Body meaning	B= none
Space meaning	S = none
Gloss:	presenting "some of the ways" to the listener
Viewpoint:	O-VPT

Figure 3.4: A sample of McNeill's (1992: 382-383) guidelines.

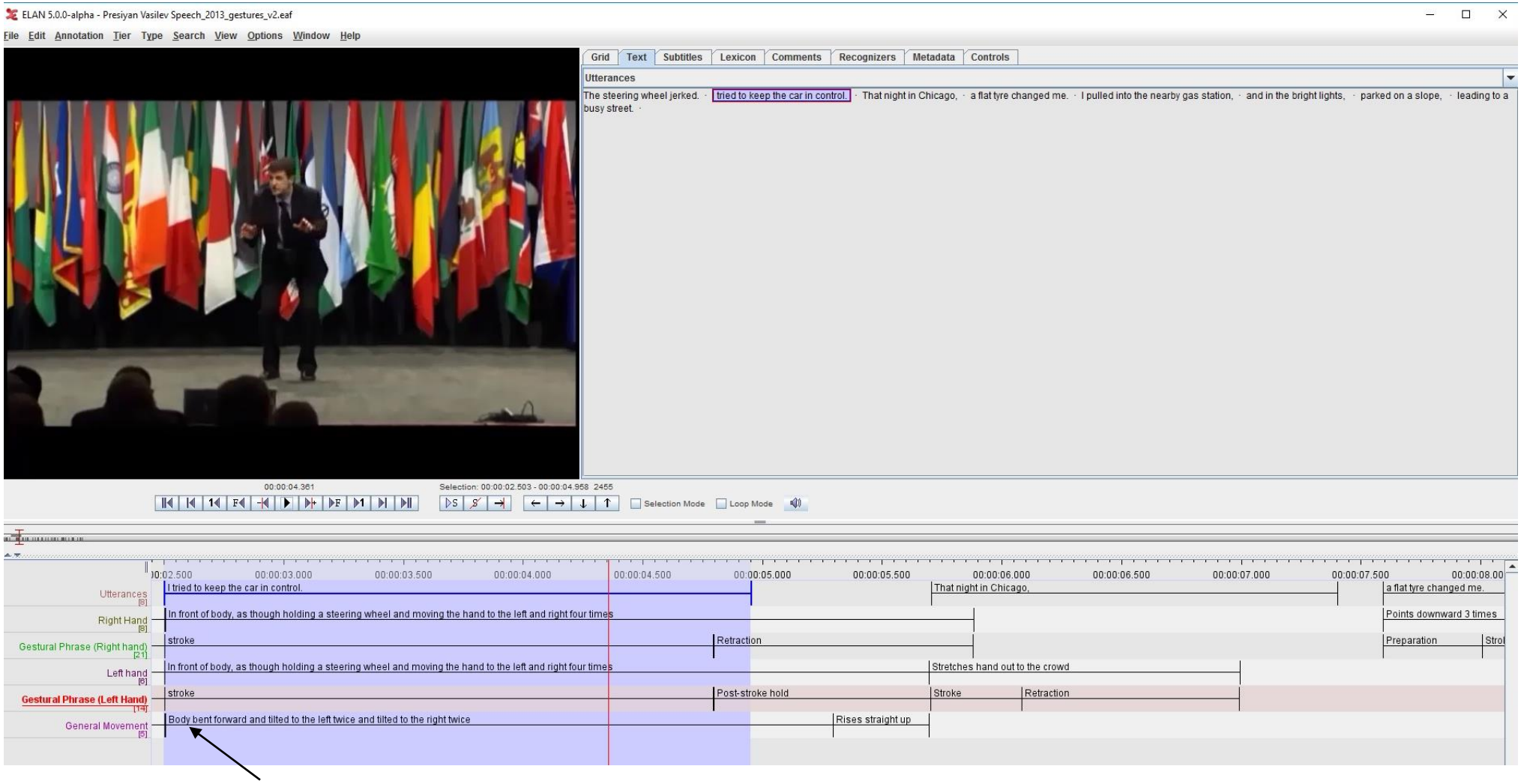


Figure 3.5: The General Movement Tier.

3.4 Intonational Annotation

After the gestures had been annotated, the Utterance tier was exported to Praat (Boersma & Weenik, 2016) to annotate the intonational elements of the speech segments. At this stage, the audio file was extracted from the video and imported into Praat as well. It was decided that the intonational annotation of the speech segments should be done separately without looking at the gestures as to prevent the gestural annotation from influencing the intonational annotation of the speech segments.

Altogether there were six tiers used during the intonational annotation and they were adapted from the tone, break index and orthographic tiers from the ToBI system (refer to Section 3.1.2, pg. 66-68). Before the different types of tones could be identified, the word boundaries needed to be established first. This was done in Praat where the Utterance tier was exported from Elan (Wittenburg, Brugman, Russell, Klassman & Sloetjes, 2006). The Utterance tier formed the first tier in the intonational annotation which was adapted from the orthographic tier from the ToBI system as the words from each speech segment were transcribed. Then, the individual words in each utterance were identified, and the time which each word took to be verbalised was also annotated. This duration formed the word boundaries which made up the second tier in the intonational annotation and it was labelled as Words. These inter-word boundaries formed the second level (Level 1) of the break index tier of the ToBI framework.

After the Utterance and Words tiers had been created, the location and the different types of tones were then identified and annotated at the tone level which was based on the tone index from ToBI. This was done by using the spectrogram and the pitch track in Praat to identify the changes in the tone and pitch of the speakers. This formed the third tier in the intonational annotation which was labelled ToBI. At this stage, one

needed to determine whether the prominent tones were high (H) or low (L) in addition to whether they were pitch accents (*), phrase boundaries (-) and boundary tones (%). If there was a downstepped tone, a (!) was written before the (H) or (L).

As the focus of the study was on the prominent changes in the pitch of each speech segment, a fourth tier called Pitch Accents was created to separate these pitch accents be it high (H*), low (L*) or a downstepped (!H*) pitch accent. The study chose to focus on the pitch movements of the speakers' verbal utterances was because pitch accents indicate a lot, if not most of the nonverbal communicative elements of a verbal utterance (refer to Section 2.4, pg. 59).

The last two tiers of the intonational analysis were based on Levels 3 and 4 of the break index tier of the ToBI framework. The fifth tier was labelled Intermediate Phrases. This was the stage where the intermediate phrases were separated from the ToBI tier. These intermediate phrases are part of an intonational phrase. These phrases were annotated in a separate tier to help identify any pauses within any verbal utterance of each speech segment. The sixth and final tier was labelled as the Intonational Phrase. This tier allowed the intonational phrase boundaries in each speech segment to be identified. An intonational phrase is usually about the length of an utterance in a speech segment.

After the intonational annotation had been completed, all the tiers except the Utterance tier were exported back to Elan so the relationship between gestures and intonation could be analysed (refer to Figure 3.11). Figure 3.6 shows a screenshot from Avery's speech segment where of all of the tiers mentioned earlier which were used when the intonational elements in each speech segment was annotated. The Pitch Accent tier and the Intonational Phrase tier are marked by arrows in the screenshot.

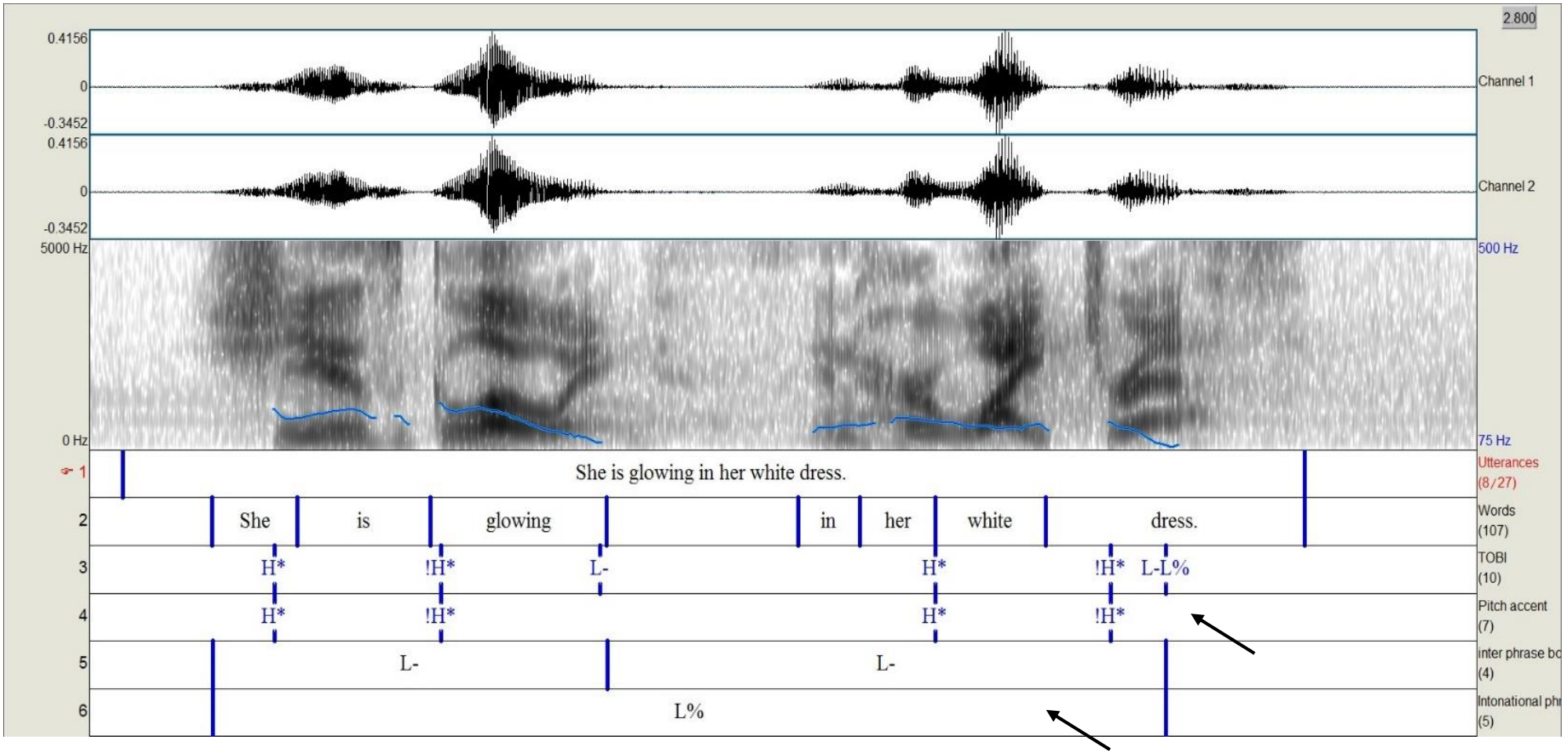


Figure 3.6: The intonational tiers of Avery's speech segment.

3.5 Pilot Study

Before the four speech segments were annotated, a pilot study was conducted. It was hypothesised that the cultural background or native language of the speakers may not have a big influence on the speakers as they have undergone the same training in Toastmasters.

Two speech segments from two different speeches from the Toastmasters World Championship of Public Speaking were annotated and analysed during the pilot study. The speeches were crafted and delivered by the winners of this tournament. They are Jim Key who won the tournament in 2003 and Presiyan Vasilev who was the 2013 champion. Key is a Caucasian male from North America whereas Vasilev is a native of Bulgaria. The pilot study also focused on the hook of their speeches. The total length of both speech segments was 69 seconds long out of 13 minutes and 24 seconds. The length of each speech segment and the total length of both speeches in the pilot study is shown in Table 3.2.

Table 3.2: The length of each speech and its hook analysed in the pilot study.

Speech	Duration	
	Hook	Whole Speech
Key (2003)	49 seconds	7 minutes 12 seconds
Vasilev (2013)	20 seconds	6 minutes 12 seconds
Total	69 seconds	13 minutes 24 seconds

The hook of each speech was annotated to study the gestures made by each speaker. At this stage, each speech segment was exported to Elan (Wittenburg, Brugman, Russell, Klassman & Sloetjes, 2006) and six tiers were used to annotate and analyse the hand gestures made by both speakers. A screenshot of the six tiers used during the gestural annotation (Vasilev's speech) is shown in Figure 3.7. Figure 3.8 is a close-up of the earlier screenshot and the gestural tiers are circled in the screenshot.

The first tier was used to divide the hook into separate utterances or phrases and the software also enabled it to be transcribed. The next tier described the gestures each speaker made with his right hand. The third tier described the different gestural phrases of each gesture (right hand) according to Kendon's (1972) hierarchy of gestural phrases. The fourth tier contained the annotation of the gestures each speaker made with his left hand. Next, the fifth tier described the gestural phrases of each gesture made by the speaker's left hand. The sixth tier was used to capture the general movement each speaker made during his speech such as walking from one end of the stage to another or sitting down on a chair.

Once the gestural annotation was completed, both speech segments were imported into Praat so the intonational characteristics of each speech segment could be annotated. There were six tiers used during the annotation of the audio tracks of both speech segments. These tiers, which are circled, are shown in Figure 3.9.

The first tier contained the annotation of the verbal phrases which were transcribed during the gestural annotation before it was exported to Praat and it was labelled Utterances. The second tier was labelled Words where the word boundaries of each word was determined and annotated. The annotation of the intonation of the speech segments was done according to the Tone and Break Indices (ToBI) framework which was developed by Beckman and Elam in 1993. The third tier was used to identify and

annotate the different tones in the hook of both speeches. The next tier was used to annotate the pitch accents in each speech segment. The fifth and sixth tiers were used to annotate the break indices in the speech segment. The fifth tier was used to identify the intermediate phrases and the final tier contained the annotation of the intonational phrases in the speech segment. This intonational phrase is marked by an arrow in Figure 3.9.

When the intonational annotation was completed, the TextGrid file containing the six tiers was exported back to Elan so that the gestural and intonational annotations could be analysed as a whole. A screenshot of all of the tiers used to annotate the gestural and intonational tiers are shown in Figure 3.10. Figure 3.11 is a close-up of all of these tiers and they are circled in the screenshot. In total, there were 11 tiers used during the gestural and intonational annotation and analysis of each speech segment in the pilot study. The Utterance tier was used during the gestural and intonational annotation so it was the only tier to be exported back and forth from Praat (Boersma & Weenik, 2016) and Elan (Wittenburg, Brugman, Russell, Klassman & Sloetjes, 2006) (refer to Figure 3.8, 3.9 and 3.11).

The pilot study also found that both hooks share the same structure and style despite the differences in the speakers' backgrounds and the speech content. This is because Key and Vasilev began their speeches with a personal anecdote to capture the attention of the audience. Then, as was mentioned in the earlier section (refer to Section 3.2, pg. 71), both speakers used the same transition phrase to end the introduction of their speeches which is "Mr. Contest Chair….".

Another indication that the speakers' culture and native language may not have a large influence on their speeches is that the findings from both speech segments in the pilot study were consistent with findings from previous studies (Loehr, 2004 &

Mendoza-Denton & Jannedy, 2011). They found that pitch accents usually tend to be aligned with the stroke phrase of a gesture in natural speech. The pilot study also found that the majority of the pitch accents in Key and Vasilev's speech introductions occurred during the stroke phrase of both their left and right hands. Table 3.3 shows the number and percentage of pitch accents which occurred during each gestural phrase (or the lack of it) in Jim Key's speech segment. On the other hand, Table 3.4 shows the total number of pitch accents which were aligned in each of the gestural phrases (or the lack of it) in Vasilev's speech segment. Thus, the findings from the pilot study also seem to support the strong that the speakers' training in Toastmasters has overridden any influence the speakers' culture and native language might have on their speeches. Therefore, this variable will be discounted in this study.

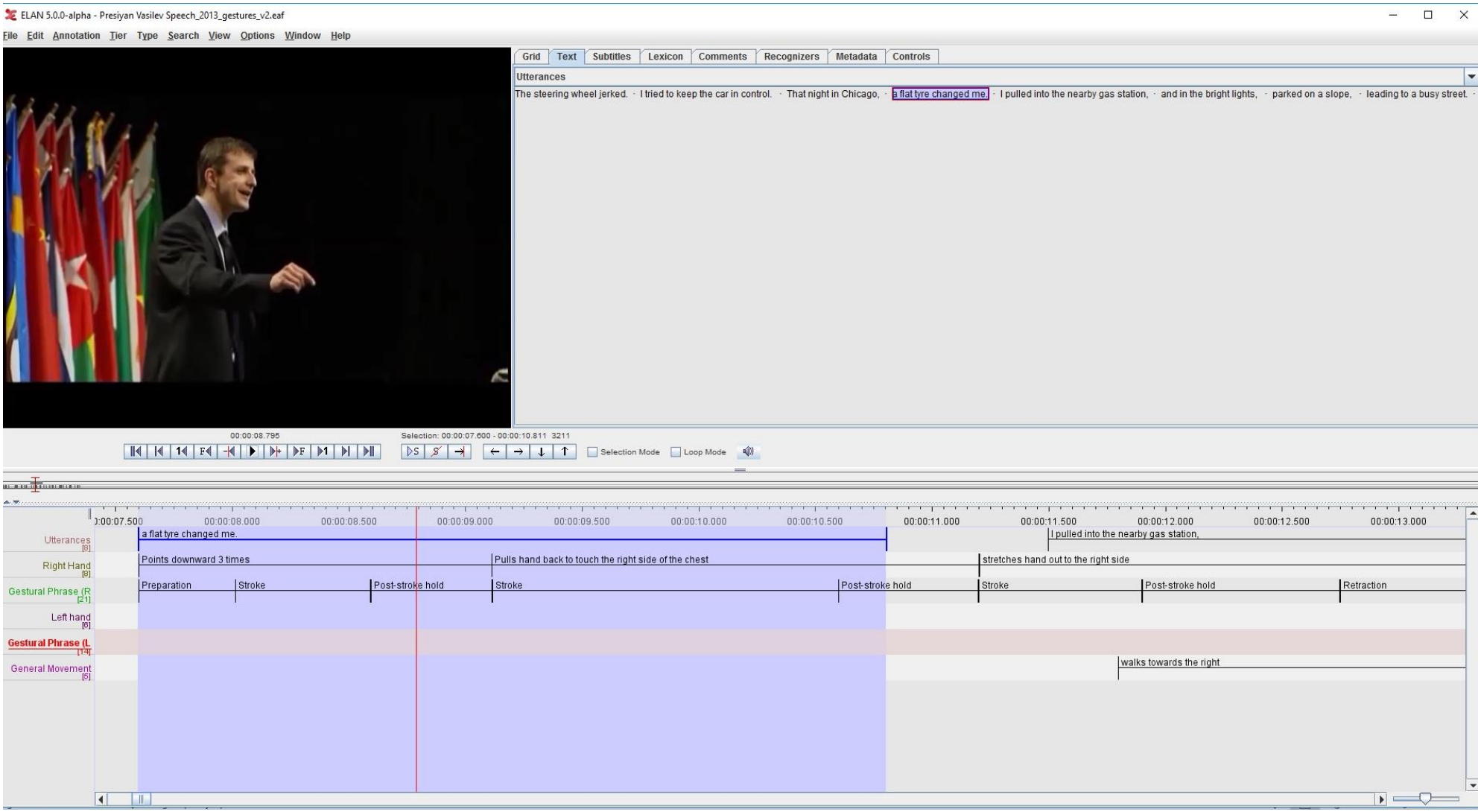


Figure 3.7: The tiers used during the gestural annotation of Vasilev's speech.

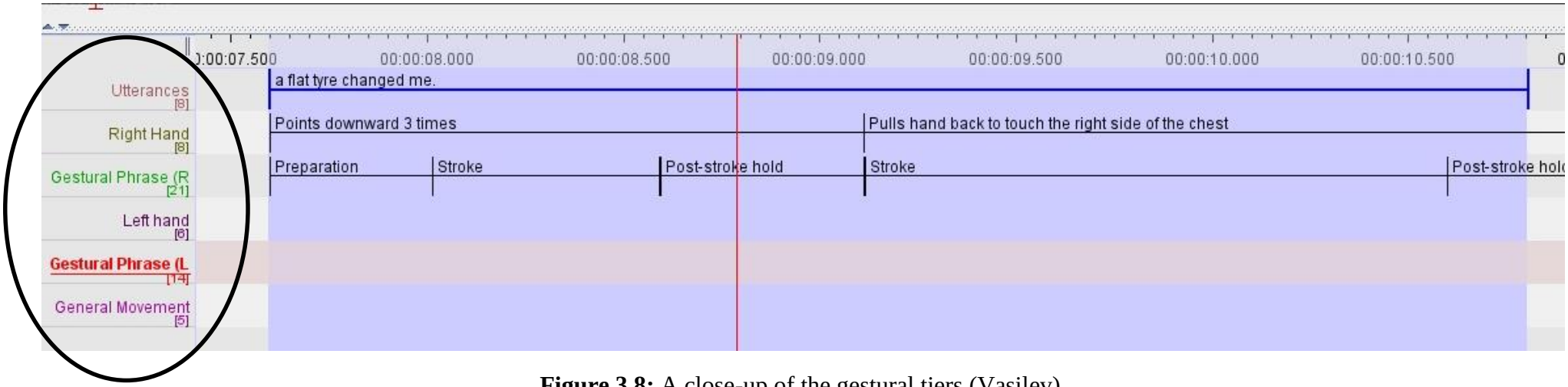


Figure 3.8: A close-up of the gestural tiers (Vasilev).

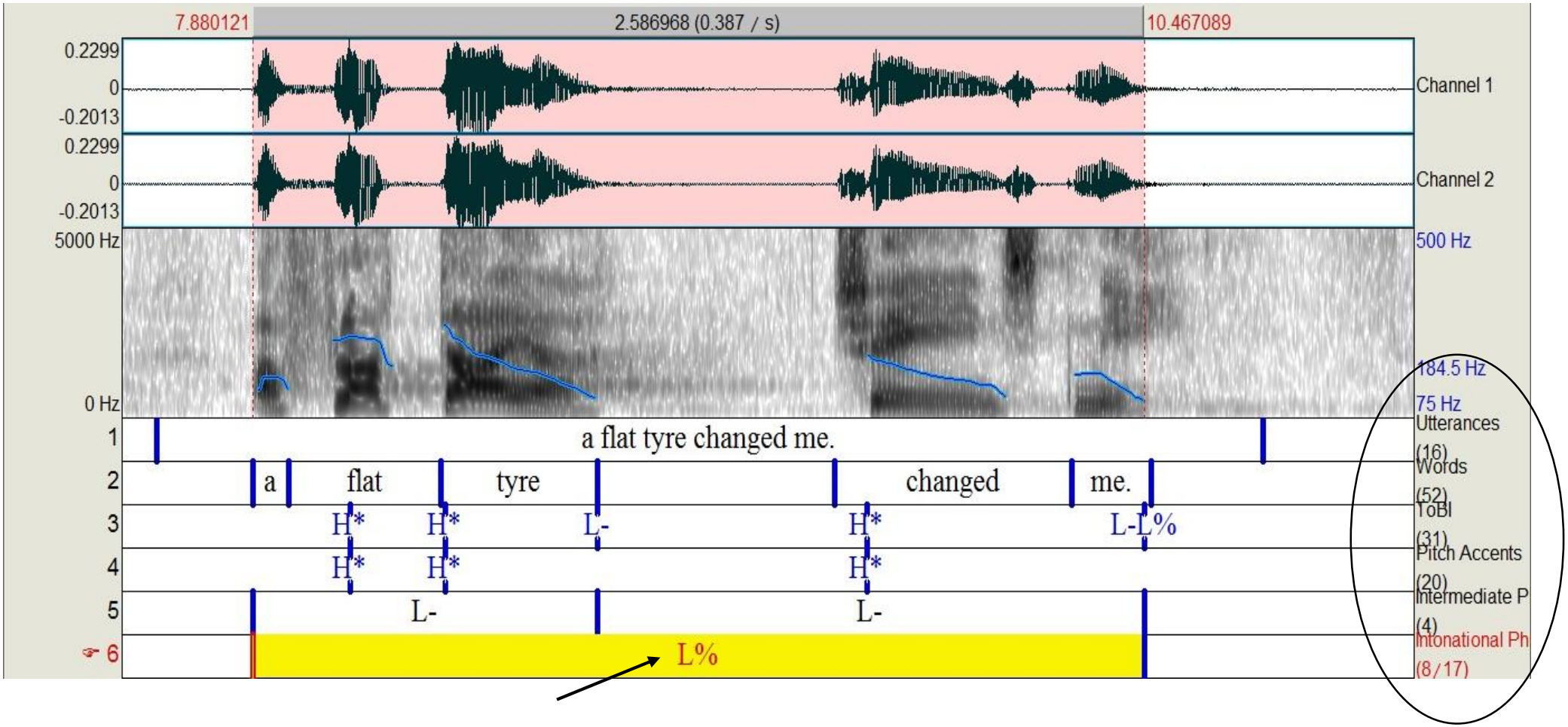


Figure 3.9: The tiers used to annotate the intonational elements (Vasilev).

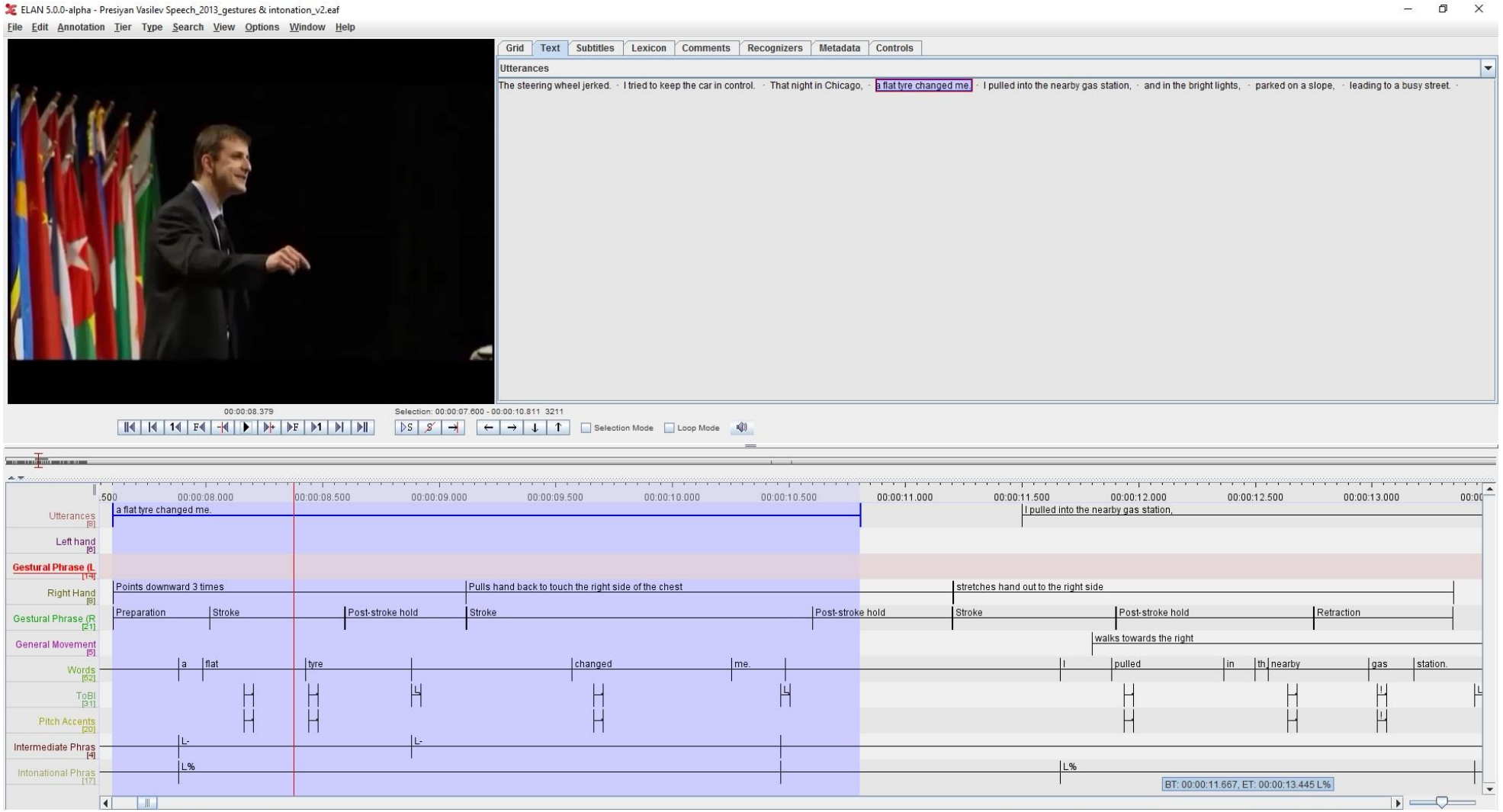


Figure 3.10: The tiers used in the gestural and intonational annotation (Vasilev).

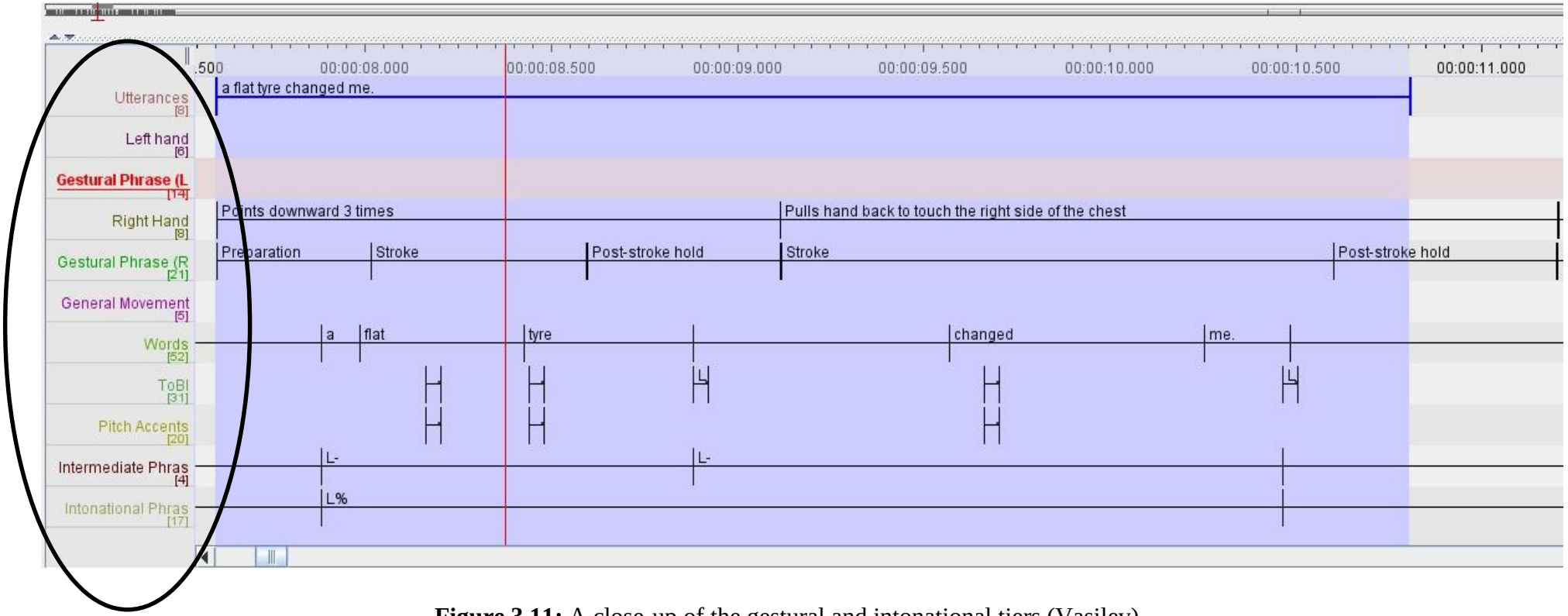


Figure 3.11: A close-up of the gestural and intonational tiers (Vasilev).

Table 3.3: The number of pitch accents and their gestural phrases (Key).

Hand / Gestural Phrase	Left Hand		Right Hand	
	Pitch Accent			
	Number	Percentage (%)	Number	Percentage (%)
Preparation	2	6.25	3	9.4
Pre-Stroke Hold	2	6.25	2	6.25
Stroke	12	37.5	11	34.37
Post-Stroke Hold	2	6.25	7	21.88
Retraction	6	18.75	5	15.6
No Gesture	8	25	4	12.5
Total	32	100	32	100

Table 3.4: The number of pitch accents and their gestural phrases (Vasilev).

Hand / Gestural Phrase	Left Hand		Right Hand	
	Pitch Accent			
	Number	Percentage (%)	Number	Percentage (%)
Preparation	2	10	2	10
Stroke	8	40	11	55
Post-Stroke Hold	1	5	3	15
Retraction	3	15	3	15
No Gesture	6	30	1	5
Total	20	100	20	100

CHAPTER 4: RESULTS

The findings from the annotations of all the four speech segments will be presented in the first part of this chapter. This chapter will describe the number of pitch accents which coincide with the different gestural phrases of each individual speech segment. Then, the different lexical items and how they are related to the gestures and intonation of each speech segment will be described. Finally, an overview of the annotation of all the four speeches will be presented in this chapter.

4.1 Harvey's Speech

There were 10 utterances in Harvey's speech. However, one of the utterances was not taken into account although the gestures in that utterance were annotated because Harvey imitated a dog. This is because the study only analysed the gestures and pitch accents that were used with words that carried meaning. This segment of Harvey's speech contained four falling (L%) intonational phrases and three rising (H%) intonational phrase boundaries. All of the rising intonational phrases in his introduction did not end in a question. There were also two intonational phrases which contained intermediate phrase boundaries. One of these intonational phrases was rising (H%) and it contained two intermediate phrases and another falling (L%) intonational phrase which contained 6 intermediate phrases. Most of these intermediate phrases were falling (L-) with only three of them in Harvey's hook were rising (H-), which also did not end in a question. Figure 4.1 is a screenshot from Harvey's speech segment which shows the location of the utterances and the intermediate phrases mentioned earlier. One of the rising (H-) intermediate phrases (and its co-occurring words) which did not end in a question is circled in an enlarged

version of the screenshot which only shows the various tiers (Figure 4.2). Some of these intermediate phrases also occurred when Harvey paused within an utterance and sometimes in between utterances as indicated by the arrows in Figure 4.2. However, this is the only time something like this happened in any of the speech segments analysed in this study.

The hook of Harvey's speech contained 32 pitch accents in total which occurred together with the gestures of his left and right hands. Almost all of them (31 out of 32) were high (H*) ones except for a downstepped (!H*) high pitch accent. However, two of them had to be taken out of the equation as they were located on the border of two gestural phrases. Thus, it is difficult to clearly tell which gestural phrase the pitch accent is clearly aligned to so only 30 pitch accents (for both hands) in Harvey's speech segment will be analysed to see which gestural phrase they coincide with. Figure 4.3 shows one of the pitch accents which is located in between two gestural phrases in Harvey's speech segment. Figure 4.4 is a close-up of the two gestural phrases and the pitch accent which is located in between them. The pitch accent mentioned earlier and the two gestural phrases are circled in the screenshot. Most of the pitch accents in this segment of Harvey's speech occurred when he did not gesture with either of his hands and this amounted to more than a third (40%, left hand) to half (50%, right hand) of the total number of pitch accents that were annotated as indicated in Table 4.1. The table also shows the number of pitch accents (regardless of whether they are H*, !H* or L*) that occur with each gestural phrase and when there is no gesture together with their matching percentage.

Table 4.1: The percentage of the pitch accents and gestural phrases (Harvey).

Hand / Gestural Phrase	Left Hand		Right Hand	
	Pitch Accent			
	Number	Percentage (%)	Number	Percentage (%)
Preparation	0	0	1	3.3
Stroke	10	33.3	8	26.7
Post-Stroke Hold	6	20	5	16.7
Retraction	2	6.7	1	3.3
No Gesture	12	40	15	50
Total	30	100	30	100

However, if one does not take into account the number of pitch accents that occurred when Harvey did not gesture with his hands, it is found that most of the pitch accents occur during the stroke phrase of Harvey's gestures be it his left or right hand. In fact, the proportion of the pitch accents which occur during the stroke phrase accounts for slightly over half of the total number of pitch accents that occur in all of the four gestural phrases that were annotated. 55.6% (10 out of 18) of the total number of pitch accents occurred during the stroke phrase when Harvey gestured with his left hand and 53.3% (8 out of 15) of all of the pitch accents occurred when he gestured with his right hand. Table 4.2 shows the total number of pitch accents which were aligned with all of the gestural phrases after the *No Gesture* row has been removed.

Table 4.2: Breakdown after the No Gesture row has been removed (Harvey).

Hand / Gestural Phrase	Left Hand		Right Hand	
	Pitch Accent			
	Number	Percentage (%)	Number	Percentage (%)
Preparation	0	0	1	6.7
Stroke	10	55.6	8	53.3
Post-Stroke Hold	6	33.3	5	33.3
Retraction	2	11.1	1	6.7
Total	18	100	15	100

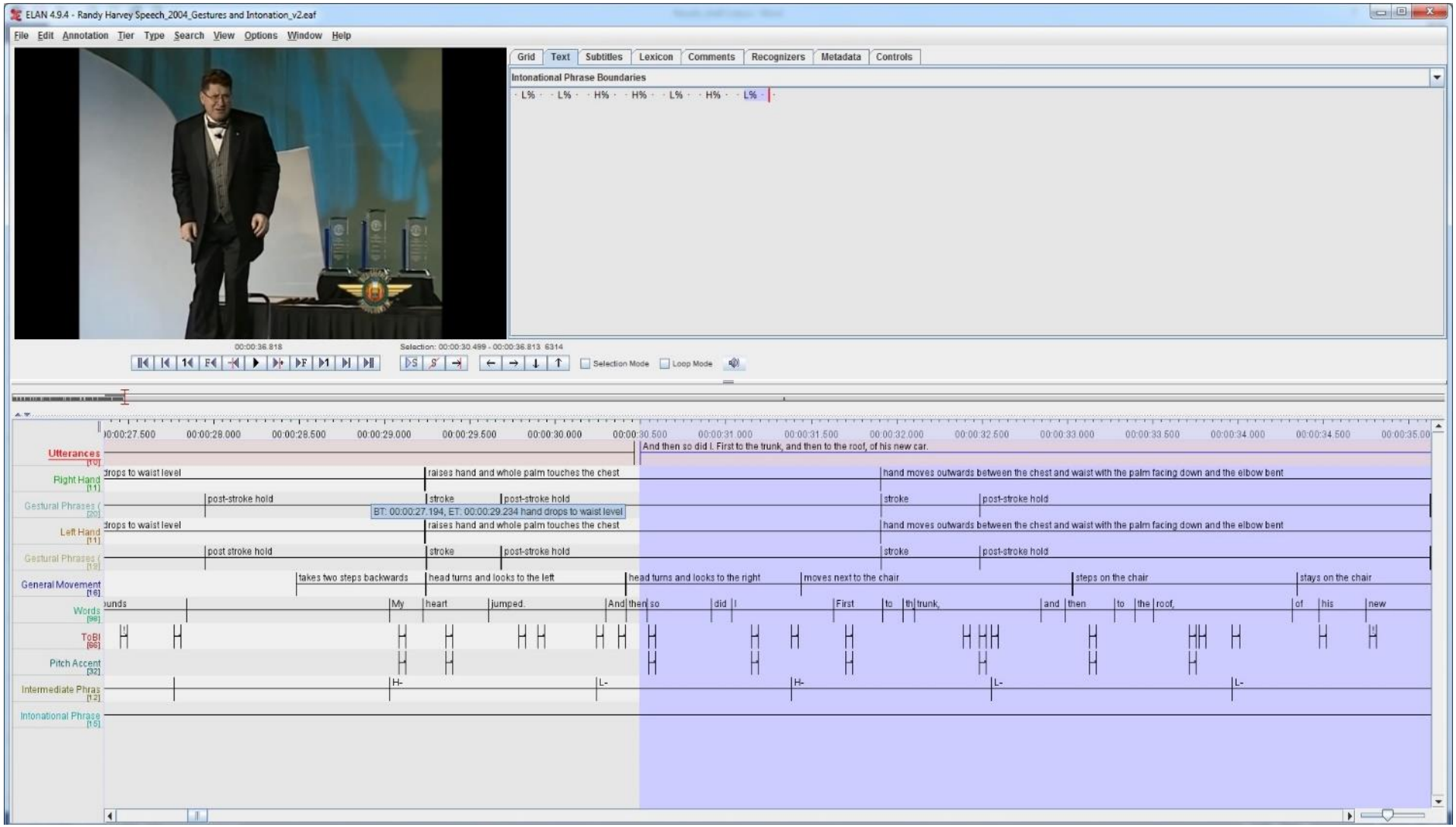


Figure 4.1: A segment of Harvey's speech.

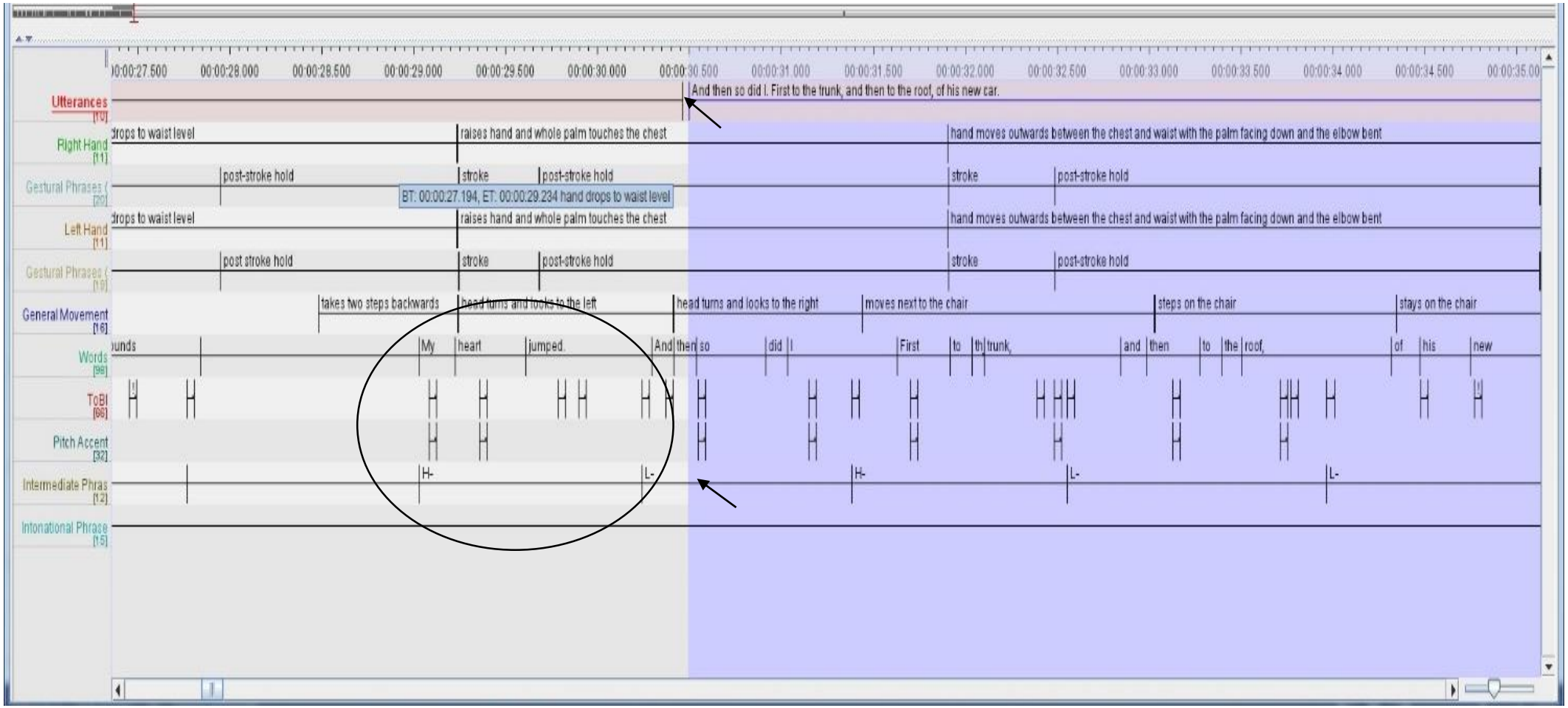


Figure 4.2: The tiers in Harvey's speech segment.

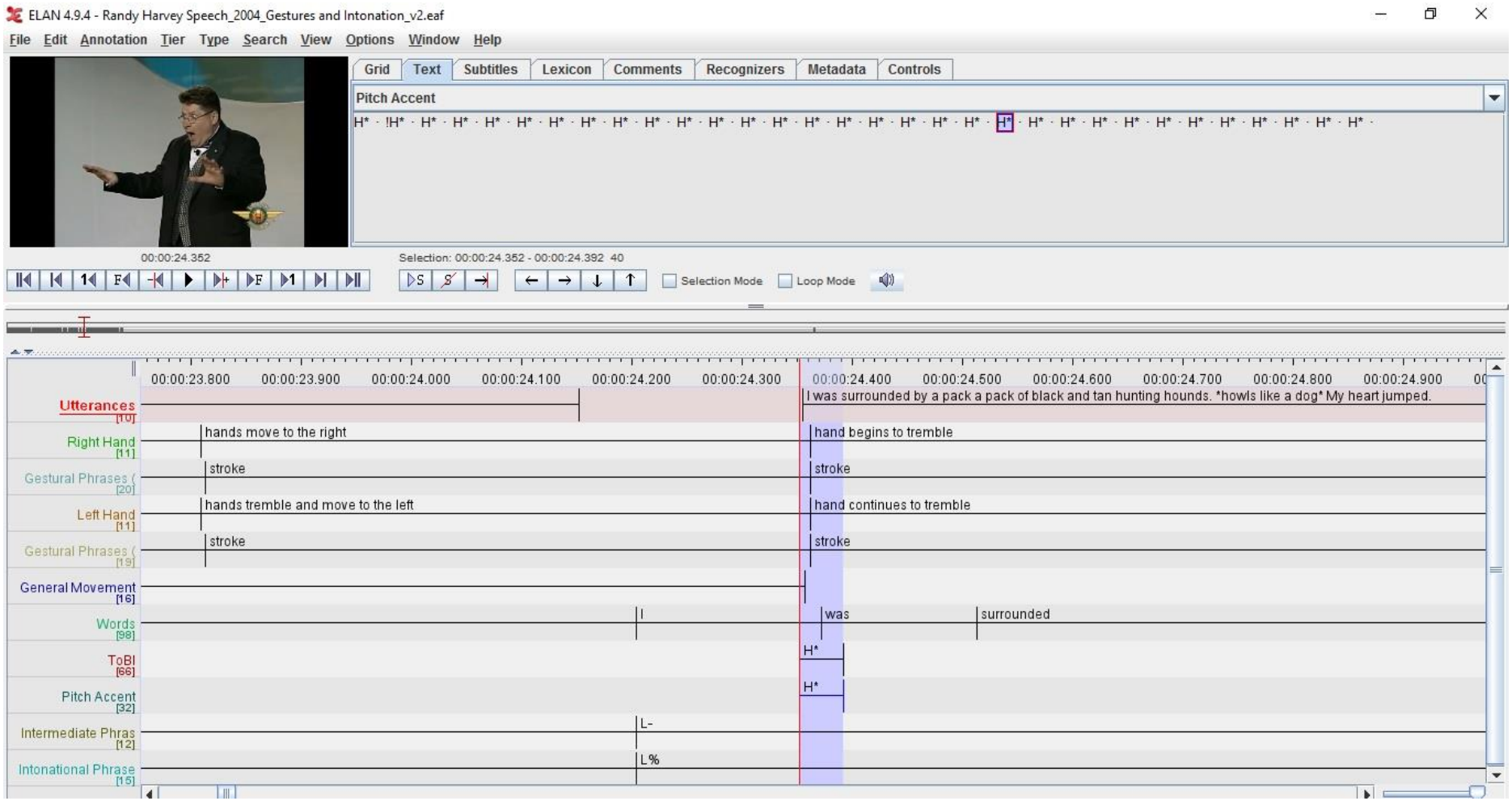


Figure 4.3: A pitch accent located in between two gestural phrases (Harvey).

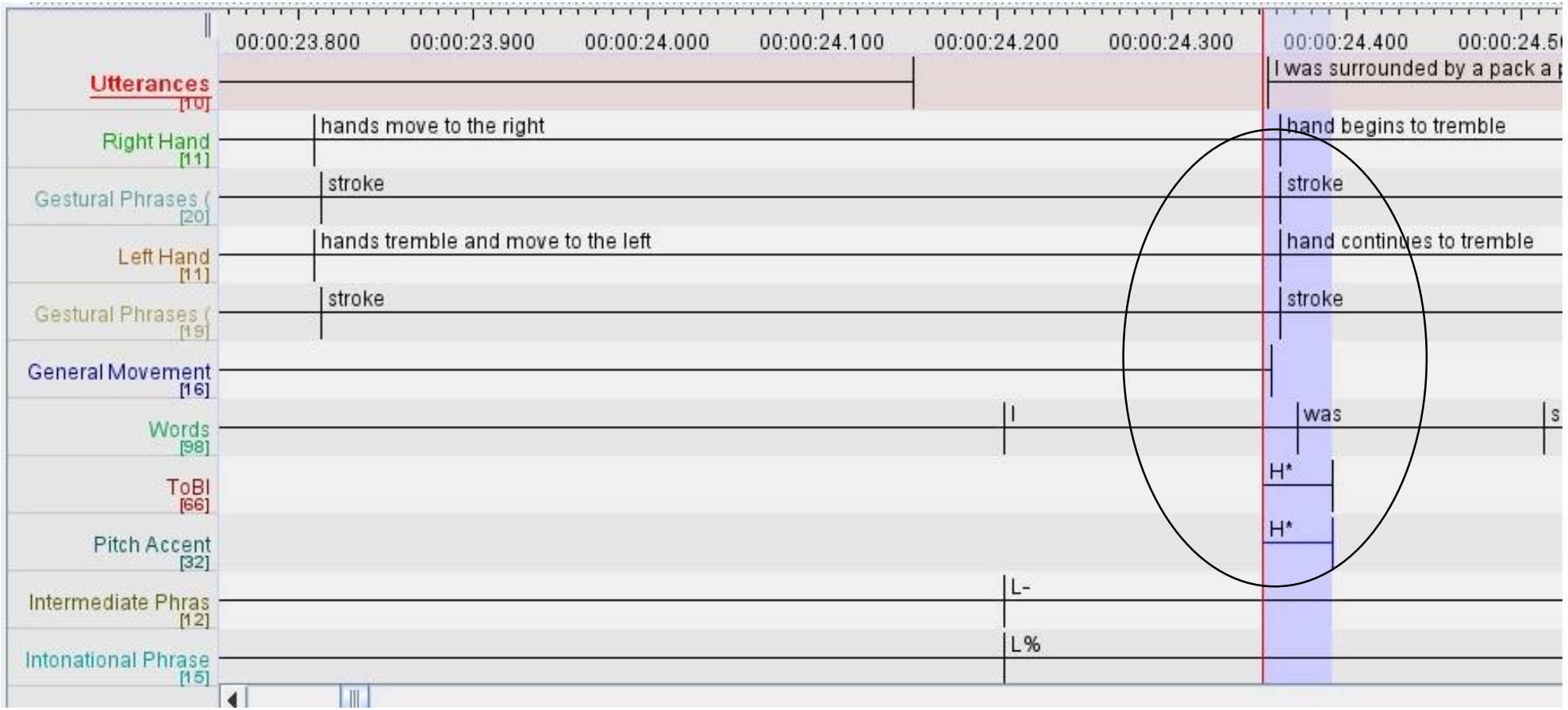


Figure 4.4: A close-up of the pitch accent in Harvey's speech segment.

After one has looked at which gestural phrases that are aligned with the pitch accents, one also has to add the lexical items from Harvey's introduction into the equation as well. His introduction consisted of 89 words and there were 32 of them which coincided with a pitch accent. However, as mentioned earlier, two of these pitch accents had to be taken out of the equation as both were located on the border of two words (refer to Figure 4.4). Therefore, it is unclear which word both of these pitch accents were aligned to so only 30 of these words were analysed. 70% (21 out of 30) the pitch accents were aligned with the content words in this speech segment and only slightly less than a third (30%) of the pitch accents occurred with a function word. At the same time, the data shows that 3 out of 9 of the function words which contained a pitch accent were first-person pronouns.

In addition, the study also took note of the lexical items which occurred with the stroke and post-stroke hold phrase as well. The study only will focus on these two gestural phrases as the analysis of the four speech segments has shown that most of the pitch accents tended to occur with the stroke and post-stroke hold phrase. In addition, the pitch accents which did not occur with any gesture were also calculated as well because many pitch accents were also aligned with many words which do not contain any gestural phrase. Table 4.3 below shows the number of content and function words which coincided with the two gestural phrases (both hands) mentioned in this paragraph and the pitch accents which were not attached to any gestural phrase. It also shows the ratio of the pitch accents that occurred with the content and function words in Harvey's speech segment. One needs to remember that not all of the words which occurred with these two gestural phrases (or the lack of them) contained a pitch accent. This was the same for all the speech segments in this study. Table 4.4 below shows the percentage of the different function words which coincided with a pitch accent.

Table 4.3: The ratio of the words to their gestural and intonational elements (Harvey).

Gestures / Speaker	Stroke		Post-Stroke Hold		No Gesture		Pitch Accents	
	Left	Right	Left	Right	Left	Right	Content	Function
Harvey	C:19 (51.4%) F:18 (48.6%)	C:16 (47.1%) F:18 (52.9%)	C:4 (23.5%) F:13 (76.5%)	C:4 (23.5%) F:13 (76.5%)	C:15 (48.4%) F:16 (51.6%)	C:15 (44.1%) F:19 (55.9%)	21 (70%)	9 (30%)
Total	37 words	34 words	17 words	17 words	31 words	34 words	30 pitch accents	

C represents the *content words* and F represents the *function words*

Table 4.4: Ratio of the types of function words with a pitch accent (Harvey).

Function Words / Speaker	Type							
	Pronoun (1st Person)	Pronoun (2nd Person)	Pronoun (3rd Person)	Possessive Pronoun	Articles	Preposition	Conjunction	Others
Harvey	3 (33.4%)	–	–	1 (11.1%)	–	1 (11.1%)	2 (22.2%)	2 (22.2%)
Total	9 words							

Out of the 89 words, there were 10 words which overlapped more than one gestural phrase. One example is circled in Figure 4.5 and a close-up of this can be seen in Figure 4.6. Almost all of them (9 out of 10) were content words. Five of these words coincided with a pitch accent and all of them are content words. The majority (3 out of 5) of these words overlapped the stroke and retraction phrases. Table 4.5 below shows the different gestural phrases which each word coincides with, regardless of the order of the gestural phrases. It can be seen that some of these words (4 out of 10) overlapped the stroke and retraction phrases. This shows that Harvey was in the midst of completing a gesture as he was speaking. In addition, there were also two words which overlapped two stroke

phrases and they are marked with a double tick in the box (✓✓). In fact, almost all of these words (9 out of 10) overlapped the stroke of a gesture.

Table 4.5: Words which overlapped gestural phrases and their pitch accents (Harvey).

Speaker: Harvey / Words	Word Type		Gestural Phrases				Pitch Accent
	Content	Function	Preparation	Stroke	Post-Stroke Hold	Retraction	
seven	✓			✓		✓	✓
1960	✓		✓	✓			✓
backseat	✓			✓		✓	✓
sleeping	✓			✓		✓	✓
to		✓		✓		✓	
car	✓			✓✓			
hunting	✓			✓✓			
jumped	✓			✓	✓		
trunk	✓			✓	✓		✓
new	✓				✓	✓	
Total	9	1	1	11	3	5	5

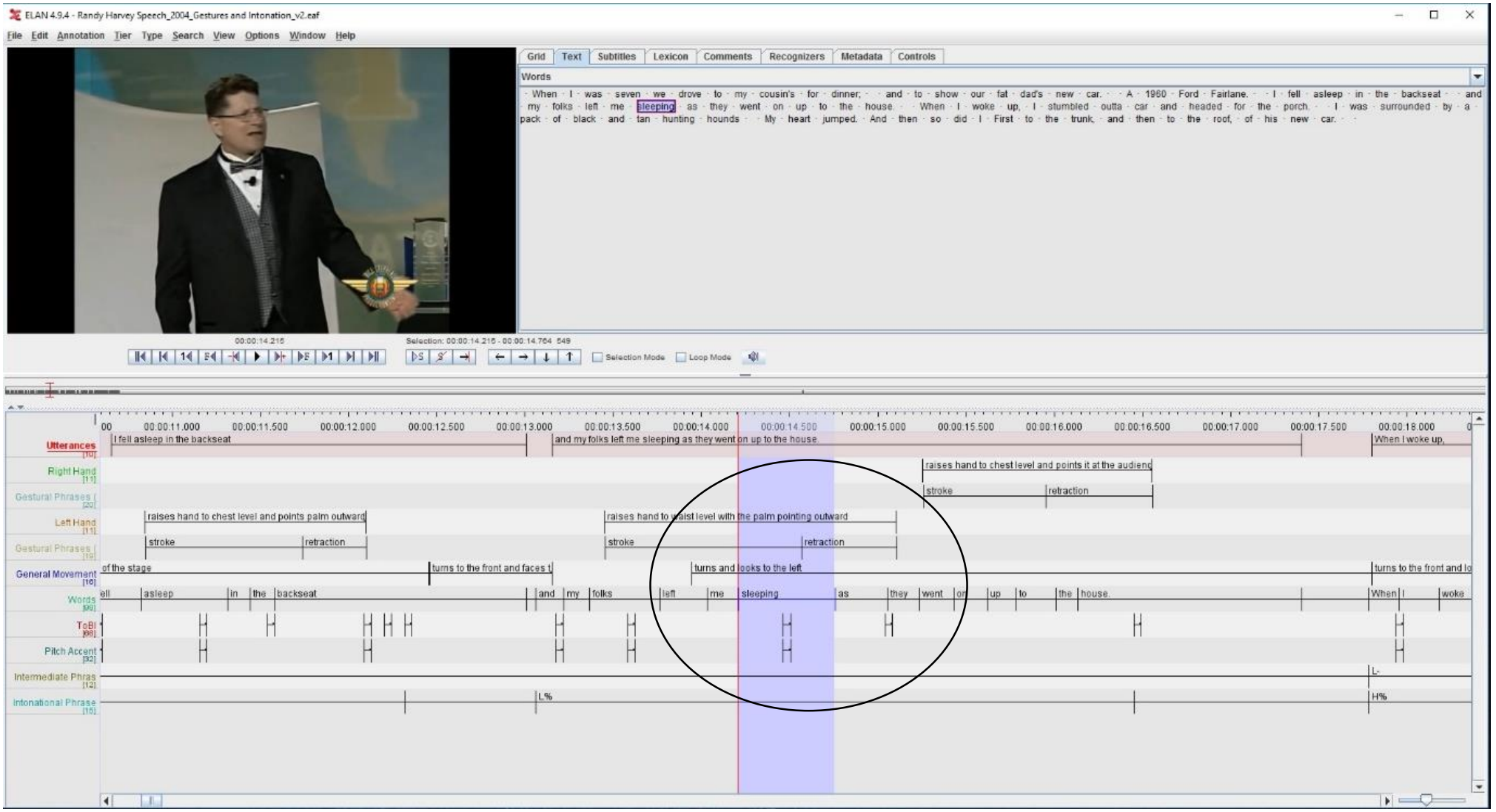


Figure 4.5: Example of a word overlapping two gestural phrases (Harvey)

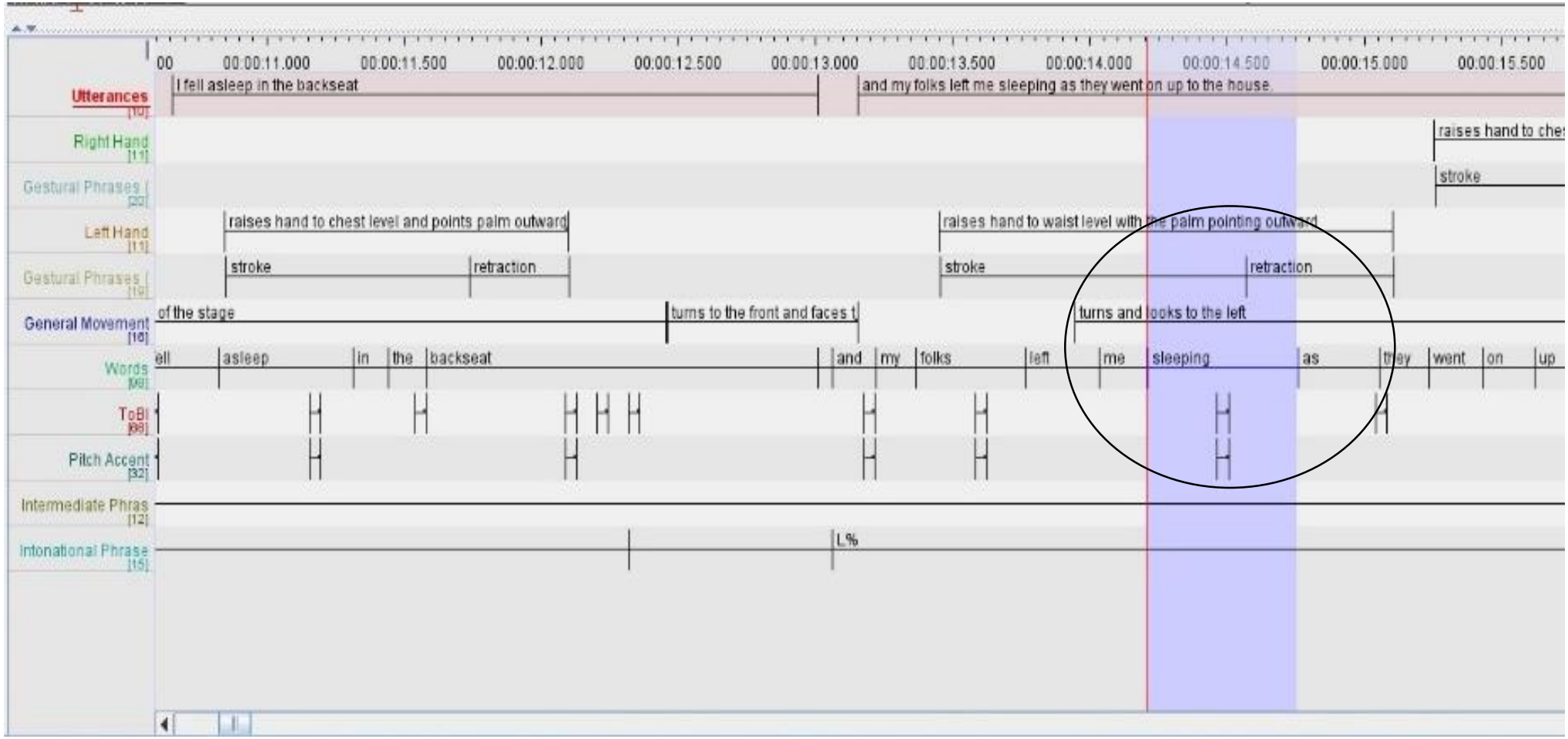


Figure 4.6: A close-up of the earlier screenshot (Harvey).

4.2 Jhingran's Speech

The introduction of Jhingran's speech contained 15 intonational phrases. 13 of the intonational phrases were falling (L%) but two of them were rising (H%). These two intonational phrases were also not questions but they were statements such as when Jhindgran stated the answer was inside his letter for the second time and when he told his audience when his mind drifted back to the past. The screenshot (Figure 4.7) shows the two rising (H%) intonational phrases in the speech segment together with their accompanying verbal utterances. The part where Jhingran delivered the two statements mentioned earlier are circled in the enlarged tiers of the earlier screenshot (see Figure 4.8). Only one of his utterances contained intermediate phrases and this was the longest utterance in his hook. There were four intermediate phrases in that utterance and all of them were falling (L-). These intermediate phrases are indicated by the arrows in Figure 4.9 when Jhingran paused three times in the middle of this utterance.

This segment of Jhingran's speech contained a total of 48 pitch accents that coincided with the gestures of his left and right hands. Almost all of them (44 out of 48) were high (H*) except for three downstepped high pitch accents (!H*) and one low pitch accent (L*). However, one of the pitch accents was located in between two gestural phrases (right hand) so it was not included in the analysis. Thus, 48 pitch accents were analysed when Jhingran gestured with his left hand but only 47 pitch accents were analysed when he gestured with his right hand. Figure 4.10 shows a screenshot of the pitch accent which was not included in the analysis and Figure 4.11 is a close-up of the earlier screenshot. The pitch accent together with its corresponding gestural phrases were circled in the screenshot. In addition, one can see that the pitch accent only coincides with only one gestural phrase of the left hand which is the post-stroke hold phrase and this is indicated by an arrow. Most of the pitch accents in

Jhingran's speech occurred during the post-stroke hold phrase when he gestured with his left hand and during the stroke phrase when he gestured with his right hand. The pitch accents which occurred during these two stages consisted of more than half of the total number of pitch accents that were annotated. 58.3% of the total number of pitch accents (28 out of 48) occurred during the post-stroke hold phrase when Jhingran gestured with his left hand and 63.8% of them (30 out of 47) occurred during the stroke phrase when Jhingran gestured with his right hand. The number of pitch accents (regardless of whether they are H*, !H* or L*) that occur with each gestural phrase and their matching percentages are shown in Table 4.6. The same table also shows the number of pitch accents which are not attached to any hand gestures together with their matching percentages.

Table 4.6: Breakdown of Jhingran's pitch accents and gestural phrases.

Hand / Gestural Phrase	Left Hand		Right Hand	
	Pitch Accent			
	Number	Percentage (%)	Number	Percentage (%)
Preparation	1	2.1	3	6.4
Stroke	9	18.7	30	63.8
Post-Stroke Hold	28	58.3	10	21.3
Retraction	3	6.3	3	6.4
No Gesture	7	14.6	1	2.1
Total	48	100	47	100

Even though the number of pitch accents that occur when Jhingran did not gesture with his hands are small, it can still be removed from the equation. When these pitch accents are not taken into account, the number of pitch accents that occur with the post-stroke hold (left hand) and the stroke (right hand) account for about two-thirds of the

total number of pitch accents that were annotated for both of his hands. The pitch accents which occurred during the post-stroke hold phrase of Jhingran's left hand amounted to 68.3% (28 out of 41) while 65.2% (30 out of 46) of the total number of pitch accents coincided with the stroke phrase of his right hand. The total breakdown of the number of pitch accents (regardless of whether they are H*, !H* or L*) that occurred with each gestural phrase are shown in Table 4.7 with their matching percentages. The pitch accents that occur when Jhingran did not gesture with his hands have been removed in this table.

Table 4.7: Breakdown without the No Gesture row (Jhingran).

Hand / Gestural Phrase	Left Hand		Right Hand	
	Pitch Accent			
	Number	Percentage (%)	Number	Percentage (%)
Preparation	1	2.4	3	6.5
Stroke	9	22	30	65.2
Post-Stroke Hold	28	68.3	10	21.8
Retraction	3	7.3	3	6.5
Total	41	100	46	100

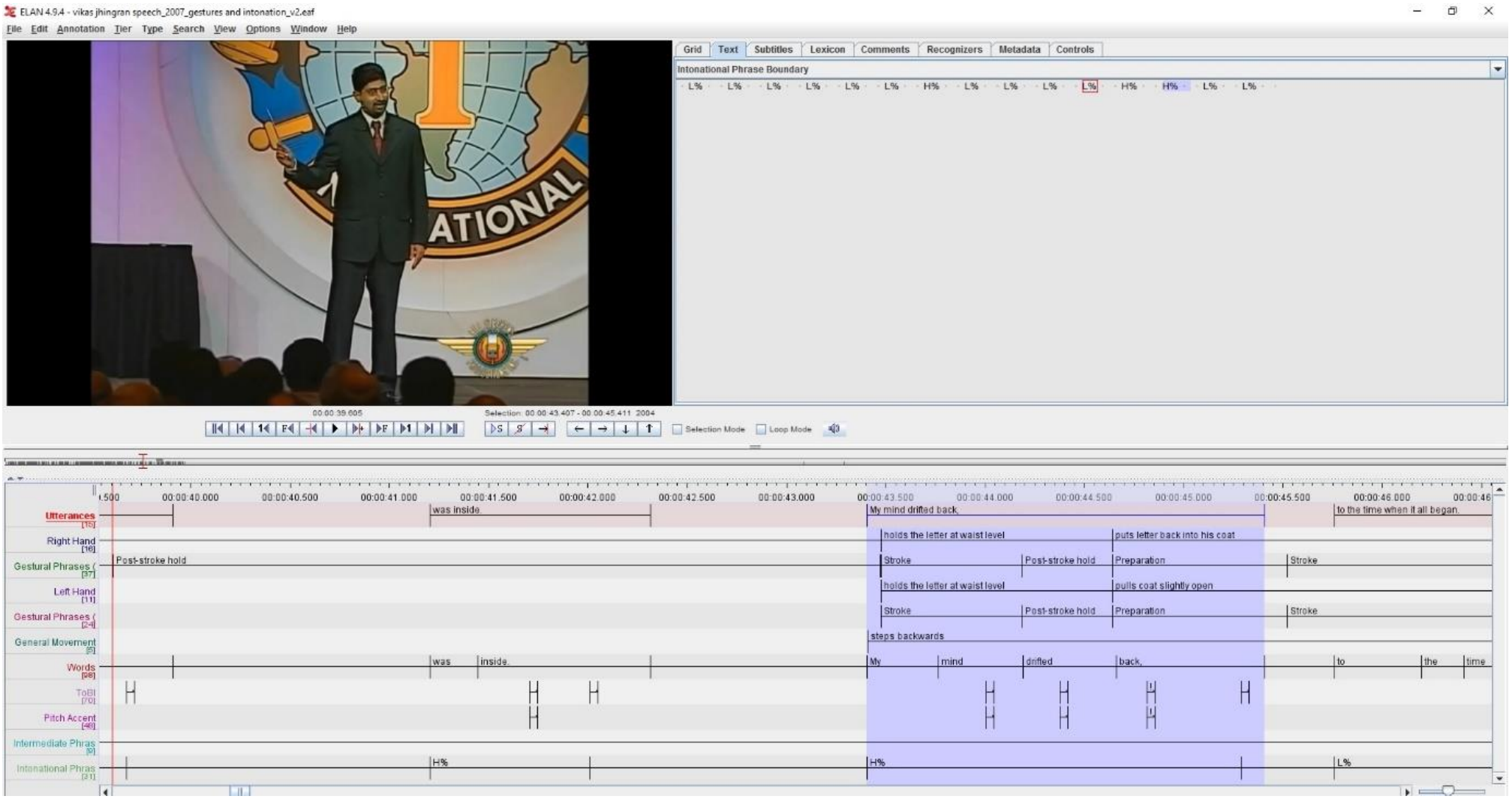


Figure 4.7: Two rising (H%) intonational phrases in Jhingran's speech segment.

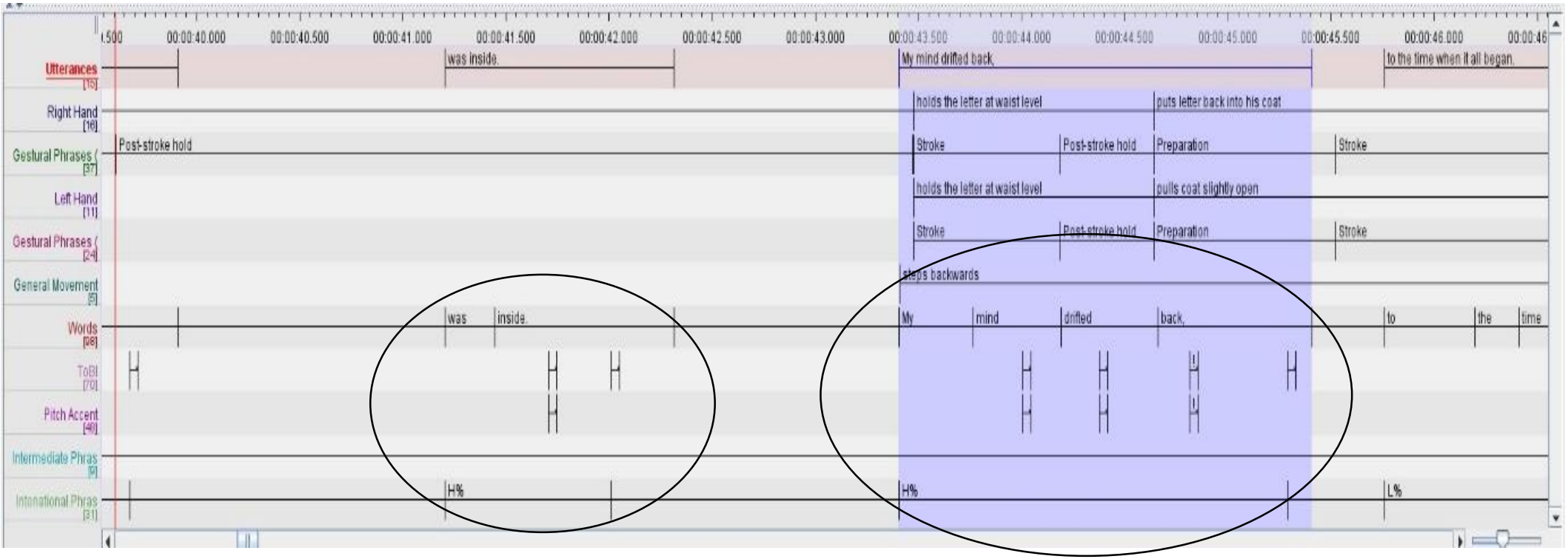


Figure 4.8: The rising intonational tiers which are statements (Jhingran).

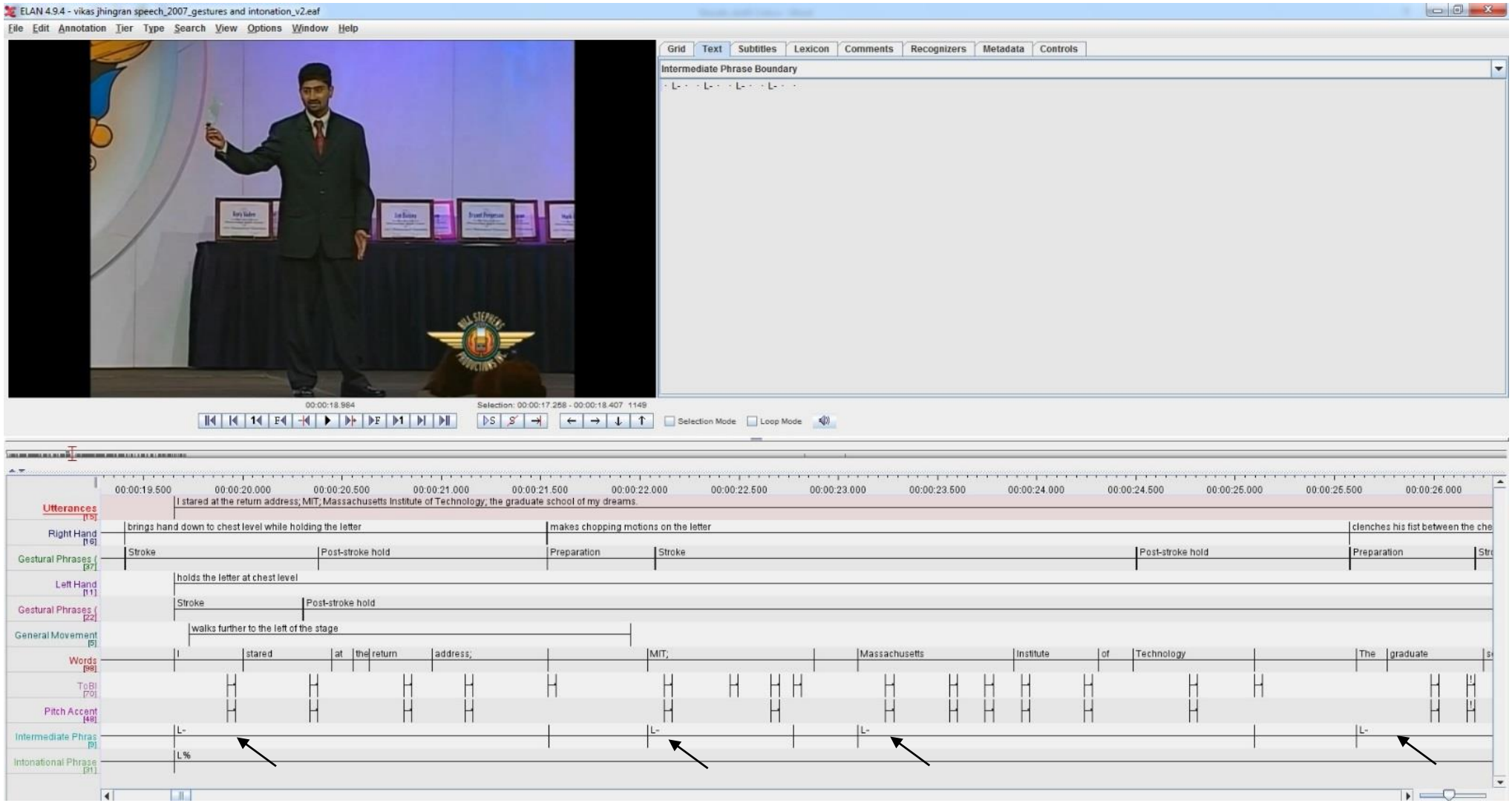


Figure 4.9: The four intermediate phrases in Jhingran's utterance.

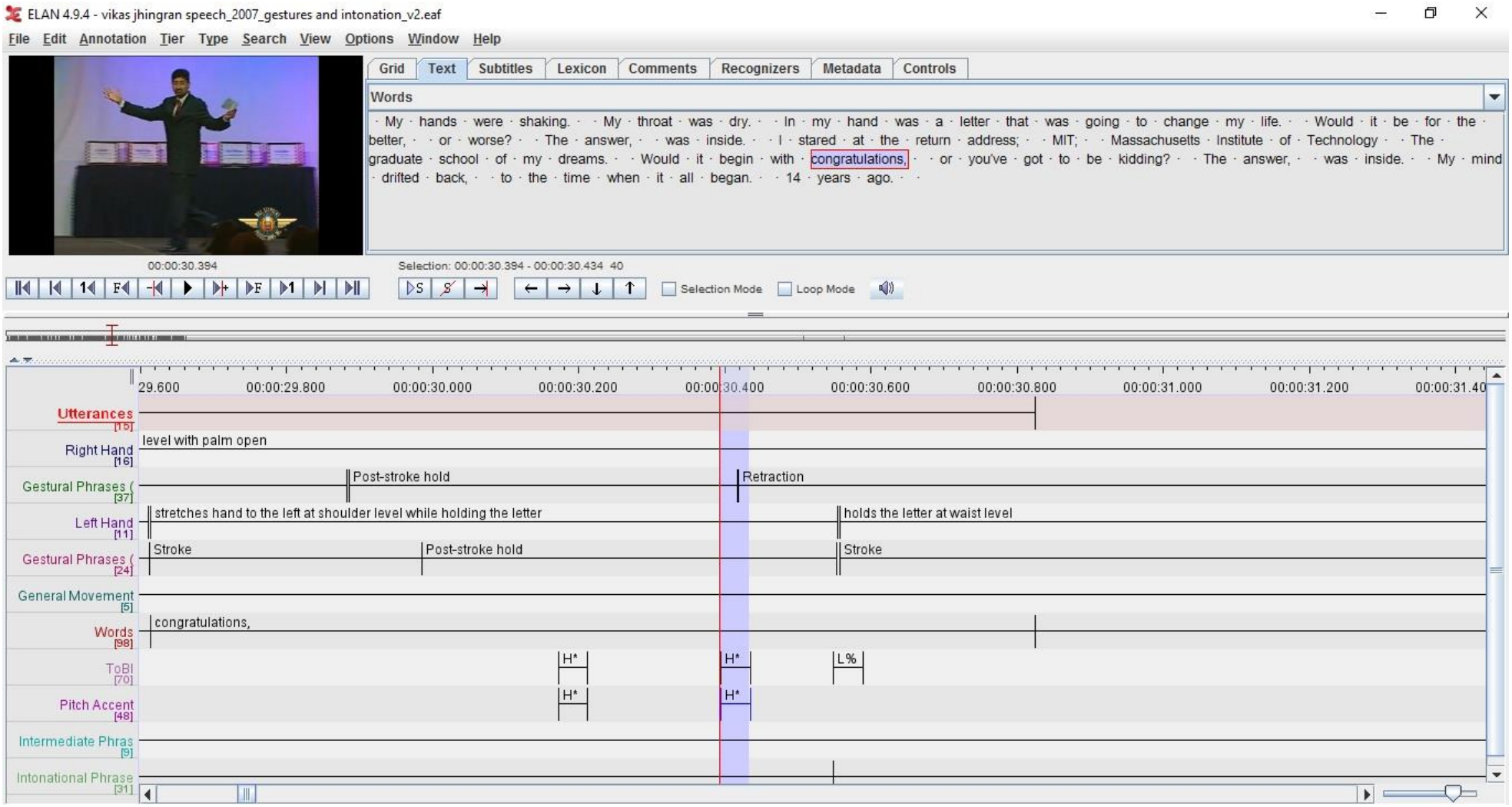


Figure 4.10: A pitch accent located in between two gestural phrases (Jhingran).

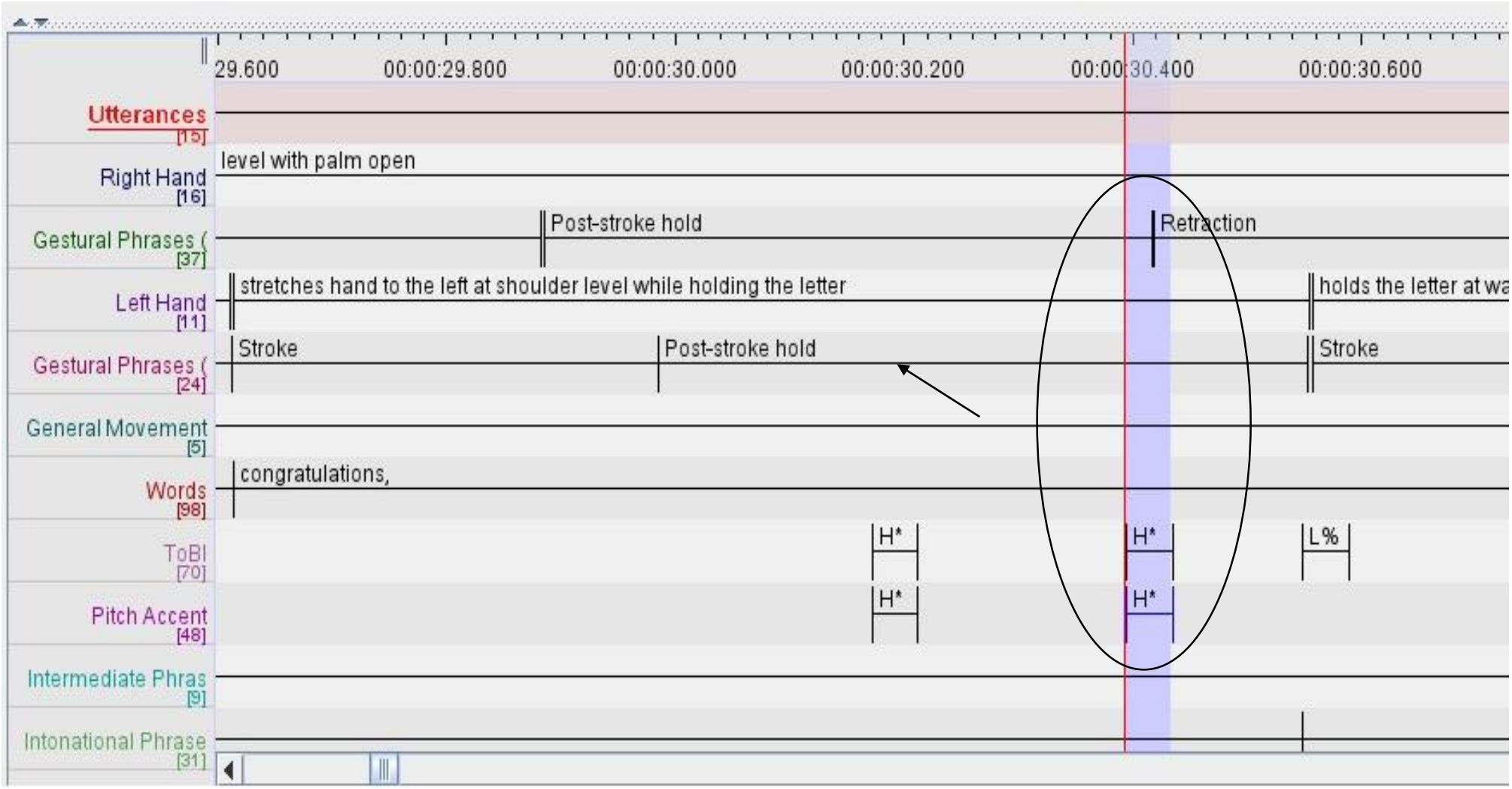


Figure 4.11: A close-up of the earlier screenshot in Jhingran's speech segment.

Jhingran's introduction consisted of 79 words and there were 41 words which coincided with the pitch accents in his speech segment. Some the words in Jhingran's speech segment contained more than one pitch accent. For example, when he told the audience that he received a letter from the Massachusetts Institute of Technology, there were three pitch accents in the word *Massachusetts* as seen in Figure 4.12, which is indicated by an arrow. The three pitch accents aligned to the word are circled in Figure 4.13. Therefore, as the study focused on comparing the ratio of content to function words that are aligned with pitch accents, this word only counts as one even though there are three pitch accents aligned to it. This occurrence only happened in Jhingran's and Avery's speech segment which will be discussed later in the chapter. All of the words in the other speech segments only had one pitch accent attached to them.

Out of the 42 words which are aligned with pitch accents, 29 of them or 69% are content words whereas the balance is made up of function words. However, in contrast to Harvey's speech, the majority of function words which contain a pitch accent were not first person pronouns. Instead, Jhingran tends to attach many of his pitch accents to his prepositions (3 out of 13 or 23.05%). Table 4.8 below shows the pitch accents which are aligned with the different types of function words.

Table 4.8: The ratio of the types of function words with a pitch accent (Jhingran)

Function Words / Speaker	Type							
	Pronoun (1st Person)	Pronoun (2nd Person)	Pronoun (3rd Person)	Possessive Pronoun	Articles	Preposition	Conjunction	Others
Jhingran	1 (7.7%)	–	1 (7.7%)	2 (15.4%)	2 (15.4%)	3 (23.05%)	1 (7.7%)	3 (23.05%)
Total	13 words							

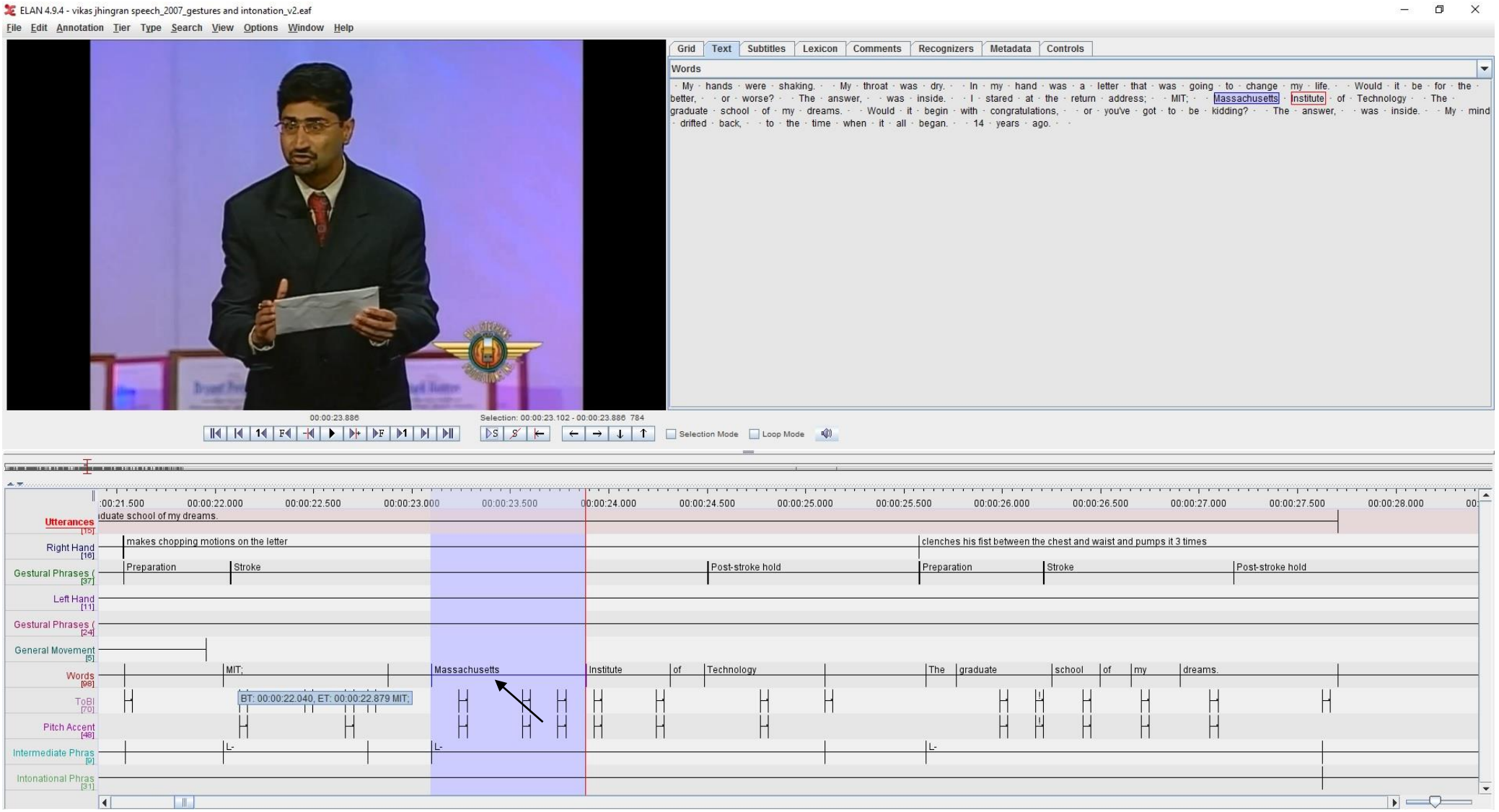


Figure 4.12: A word which is aligned with more than one pitch accent.

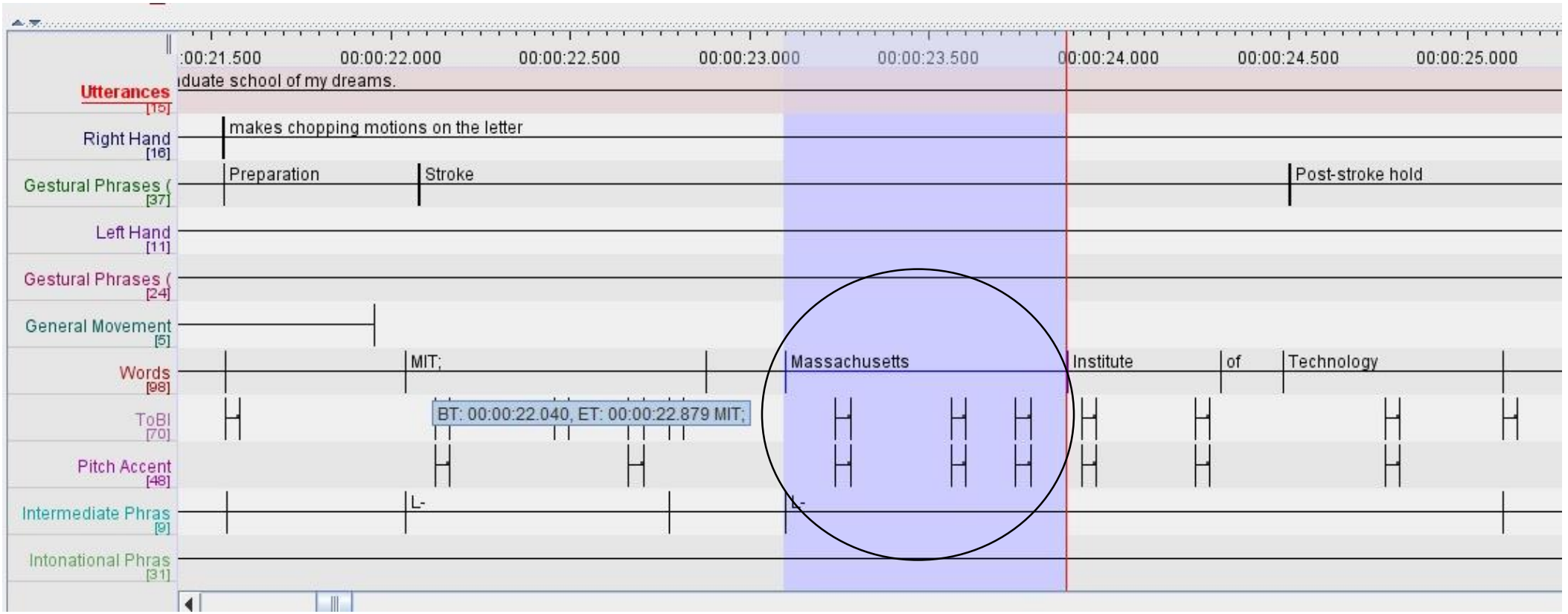


Figure 4.13: The word *Massachusetts* which is aligned with three pitch accents.

In addition, the words which were attached to Jhingran's strokes and post-stroke holds were almost similar to Harvey's data (refer to Table 4.3). There were more function words that were attached to both of the gestural phrases (and no gestural phrases) for both hands in Jhingran and Harvey's speech segments compared to the number of content words in their introductions. There were more content words which coincided with Jhingran's post-stroke hold (right hand) and an equal number of content and function words which were aligned to the stroke phrase of Jhingran's left hand. However, Harvey's data only showed that the content words only exceeded the number of function words during the stroke of his left hand. Furthermore, the number of content words which did not contain any gestural phrase also exceeded the number of function words as well in both speech segments. Table 4.9 shows the complete breakdown of the number of words aligned to the strokes and post-stroke holds (or the lack of a gestural phrase) and their percentages as well for Jhingran's speech segment. Moreover, the table also contains information on the proportion of content and function words which contain a pitch accent as well.

Table 4.9: Ratio of the words to their gestural and intonational elements (Jhingran).

Gestures and Intonation / Speaker	Stroke		Post-Stroke Hold		No Gesture		Pitch Accents	
	Left	Right	Left	Right	Left	Right	Content	Function
Jhingran	C:14 (50%) F:14 (50%)	C:25 (48.1%) F:27 (51.9%)	C:19 (48.7%) F:20 (51.3%)	C:13 (56.5%) F:10 (43.5%)	C:4 (26.7%) F:11 (73.3%)	C:1 (33.3%) F:2 (66.7%)	29 (69%)	13 (31%)
Total	28 words	52 words	39 words	23 words	15 words	3 words	42 pitch accents	

C represents the *content words* and F represents the *function words*

Out of the 79 words in Jhingran's speech, 16 of them overlapped more than one gestural phrase. All of them except two are content words which were almost similar to the findings in Harvey's speech segment (refer to Table 4.5) where the majority of words which overlapped more than one gestural phrase were also content words (15 out of 16 words). 13 of these words coincided with a pitch accent and all of them except for one were content words. This was also almost similar to Harvey's speech segment as well where all of the words like these were also content words. Another thing which is consistent with the findings in Harvey's speech segment was that almost all of the words (15 out of 16) in this speech segment overlapped at least one stroke phrase. In contrast with Harvey's speech segment, the majority of the words (9 out of 16) overlapped the stroke and post-stroke hold phrase instead of the stroke and retraction phrase, regardless of order. Moreover, 6 out the 13 words which overlapped the stroke and post-stroke hold phrase coincided with a pitch accent. In fact, there was even one word which overlapped three gestural phrases (congratulations) and the gestural phrases were the stroke, post-stroke hold and the retraction. In addition, there was one word (better) that overlapped two stroke phrases and this is marked with a (✓✓). This is shown in Table 4.10 which also contains the various gestural phrases which overlaps the words in Jhingran's speech segment together with the pitch accents that coincide with them.

Table 4.10: Words that overlapped gestural phrases and their pitch accents (Jhingran).

Speaker: Jhingran / Words	Word Type		Gestural Phrases				Pitch Accent
	Content	Function	Preparation	Stroke	Post-Stroke Hold	Retraction	
dry	✓				✓	✓	✓
change	✓			✓	✓		✓
life	✓			✓	✓		
better	✓			✓✓			✓
stared	✓			✓	✓		✓
MIT	✓		✓	✓			✓
Technology	✓			✓	✓		✓
graduate	✓		✓	✓			✓
dreams	✓			✓	✓		✓
begin	✓			✓		✓	✓
congratulations	✓			✓	✓	✓	✓
got	✓			✓	✓		
The		✓	✓	✓			✓
answer	✓			✓	✓		✓
14	✓			✓	✓		✓
ago	✓			✓	✓		
Total	15	1	3	16	11	3	13

4.3 Henderson's Speech

The hook in Henderson's speech contained five utterances. The last utterance was not taken into account as Henderson did not speak during that utterance. This is because Henderson made various sounds such as machine gun fire, the crackling of a radio and he whooped in joy which did not contain any meaningful words. This is similar to one of the utterances in Harvey's speech which was also not taken into account as he barked like a dog. Therefore, only four intonational phrases were annotated in this segment of his speech and all of them were falling (L%) intonational phrase boundaries. There was only one utterance which contained intermediate phrases. This particular utterance contained three intermediate phrases and all of them were falling (L-) ones. However, Henderson also mixed the words in one particular utterance of his speech with certain sounds such as an aeroplane flying and the crackling of a radio. Harvey's speech also shares this similarity where he barks like a dog in in the middle of one of his utterances. The utterances where Henderson and Harvey only produced sounds in between the words of their speeches are shown in Figure 4.14 and 4.15 below. The sounds are circled in the Words tier in both screenshots. In addition, the words which were located next to the sounds are indicated by arrows in both screenshots.

The hook in Henderson's speech contained 15 pitch accents. Almost all of the pitch accents were high (H*) except for one downstepped high pitch accent (!H*). Most of the pitch accents occurred during the post-stroke hold phrase of his right hand which is 11 out of 15 pitch accents or 73.3%. However, he did not make many gestures with his left hand as most of the pitch accents (7 out of 15 or 46.7%) did not occur within any gestural phrase of that hand. Table 4.11 shows the number of pitch accents, regardless of whether they are high (H*), low (L*) or downstepped (!H*), which occurred with the various gestural phrases (or the lack of it) in Henderson's speech segment.

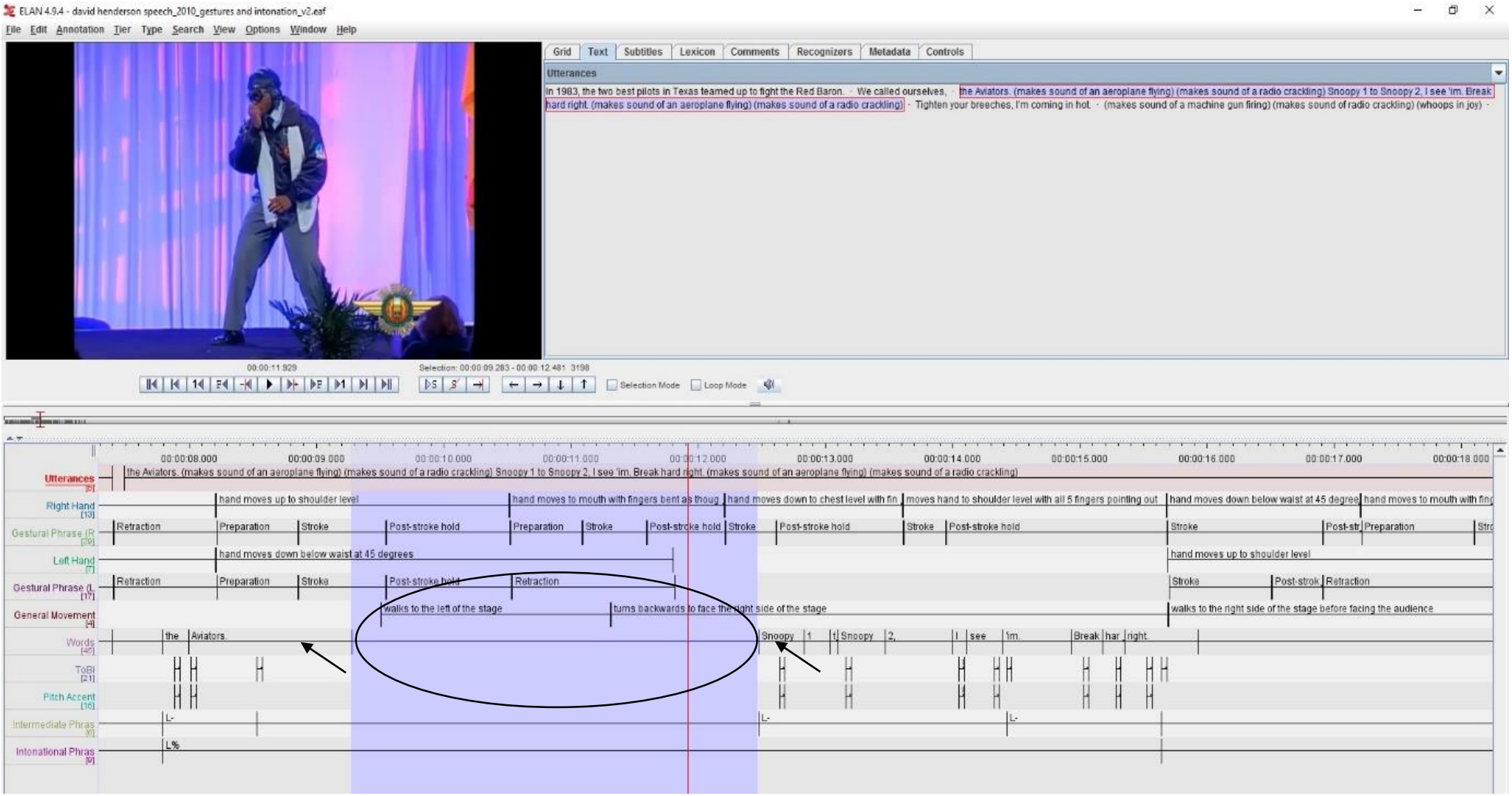


Figure 4.14: Henderson making the sound of a crackling radio.

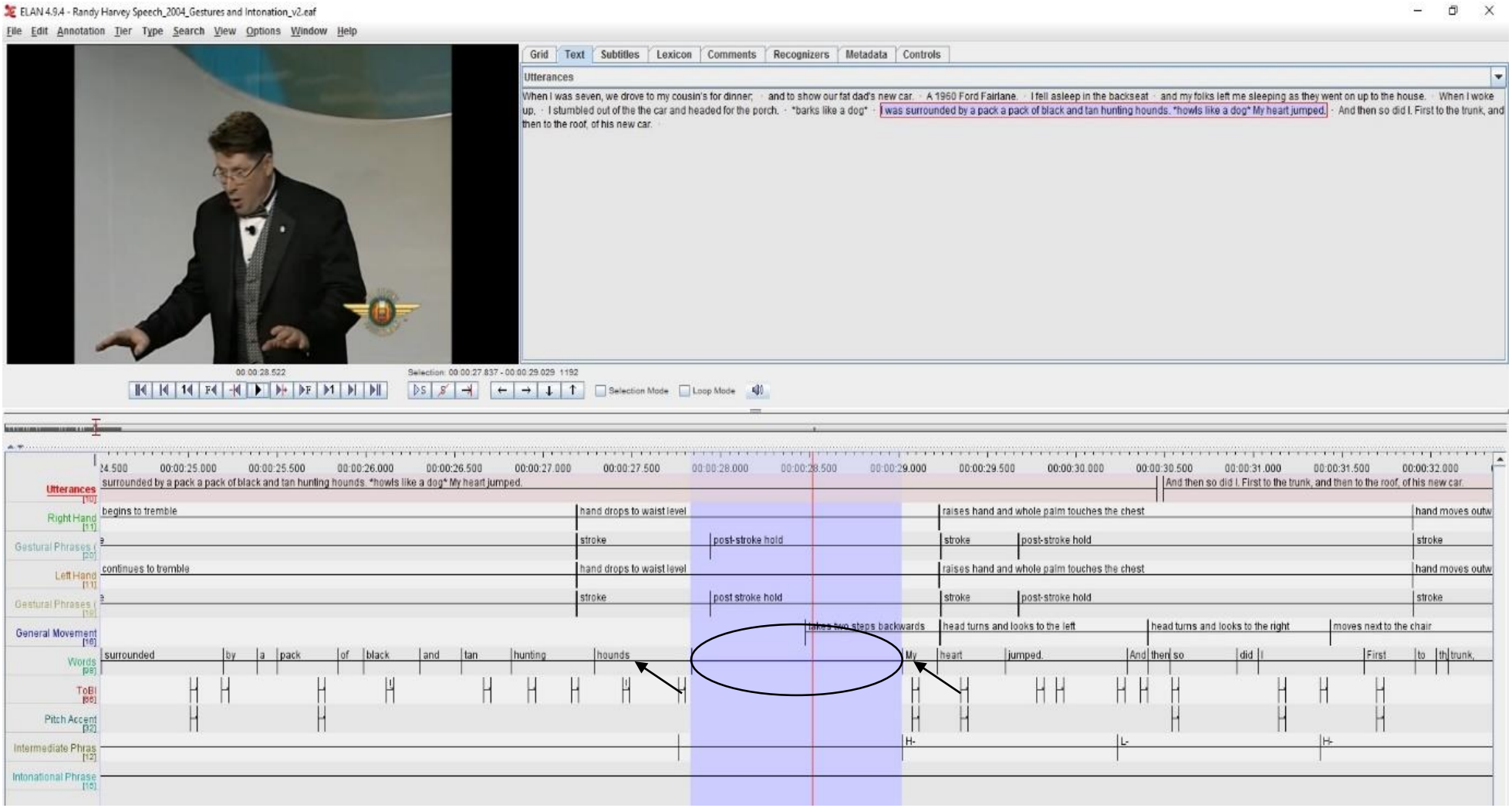


Figure 4.15: Harvey barking like a dog in between the words *hounds* and *My*.

Table 4.11: Proportion of Henderson's pitch accents and their gestural phrases

Hand / Gestural Phrase	Left Hand		Right Hand	
	Pitch Accent			
	Number	Percentage (%)	Number	Percentage (%)
Preparation	1	6.7	1	6.7
Stroke	1	6.7	1	6.7
Post-Stroke Hold	4	26.6	11	73.3
Retraction	2	13.3	2	13.3
No Gesture	7	46.7	0	0
Total	15	100	15	100

However, when the pitch accents that occurred without a hand gesture were not taken into account, most of the pitch accents were found to occur during the post-stroke hold phrase of Henderson's hand gestures be it his left or right hand. 50% of the total number pitch accents (4 out 8) occur during the post-stroke hold phrase when Henderson gestures with his left hand and 73.3% of the total number of pitch accents (11 out of 15) occur during the post-stroke hold phrase when he gestures with his right hand. Table 4.12 below contains the total breakdown of the ratio of pitch accents to gestural phrases in Henderson's speech segment after the *No Gesture* row has been removed.

Table 4.12: Proportion without the No Gesture row (Henderson).

Hand / Gestural Phrase	Left Hand		Right Hand	
	Pitch Accent			
	Number	Percentage (%)	Number	Percentage (%)
Preparation	1	12.5	1	6.7
Stroke	1	12.5	1	6.7
Post-Stroke Hold	4	50	11	73.3
Retraction	2	25	2	13.3
Total	8	100	15	100

Henderson's hook contained 38 words and 15 of them coincided with a pitch accent. Most of the pitch accents in Henderson's introduction coincided with the content words of his speech segment as well. Out of the 15 words, 11 of them (73.3%) were content words whereas the balance were function words. The percentage of this finding was also almost similar to the findings of the two earlier speech segments (Harvey and Jhingran) where 70% (21 out of 30) of the words in Harvey's hook and 69% (29 out of 42) in Jhingran's introduction were also content words. In addition, many of the function words in this speech segment which contained a pitch accent were articles. They made up 50% (2 out of 4) of the total and the other two were the first person pronouns and possessive pronouns. Table 4.13 shows the breakdown of the type of function words in Henderson's speech segment together with their respective percentages.

Table 4.13: The types of function words with a pitch accent (Henderson)

Function Words / Speaker	Type							
	Pronoun (1st Person)	Pronoun (2nd Person)	Pronoun (3rd Person)	Possessive Pronoun	Articles	Preposition	Conjunction	Others
Henderson	1 (25%)	–	–	1 (25%)	2 (50%)	–	–	–
Total	4 words							

The stroke and post-stroke hold phrases of Henderson's speech segment also mainly occurred with his content words. For example, 70% (7 out of 10) of the words in his speech segment which coincided with the stroke of his right hand were content words and 66.7% (20 out of 30) of the words that occurred with the post-stroke hold (right hand) were also content words. In addition, the majority of the words which were not

attached to any gestural phrases (left hand) were also content words (70% or 7 out of 10). On the other hand, the data showed that Henderson gestured continuously with his right hand and this is reflected in the fact that all of his words and pitch accents occurred with at least one gestural phrase made by his right hand. Table 4.14 below shows the total amount of content and function words which occurred with the strokes and post-stroke holds (or the lack of gestural phrase) in Henderson's speech segment. Their respective percentages and the ratio of content words aligned with pitch accents to function words which are attached to pitch accents are also included as well.

Table 4.14: Ratio of words to their gestural and intonational elements (Henderson).

Gestures and Intonation / Speaker	Stroke		Post-Stroke Hold		No Gesture		Pitch Accents	
	Left	Right	Left	Right	Left	Right	Content	Function
Henderson	C:5 (62.5%) F:3 (37.5%)	C:7 (70%) F:3 (30%)	C:12 (63.2%) F:7 (36.8%)	C:20 (66.7%) F:10 (33.3%)	C:7 (70%) F:3 (30%)	C:0 F:0	11 (73.3%)	4 (26.7%)
Total	8 words	10 words	19 words	30 words	10 words	No words	15 pitch accents	

C represents the *content words* and F represents the *function words*

As it was with the other two previous speech segments, there were also words in Henderson's speech segment which overlapped more than one gestural phrase. 15.8% (6 out of 38) of the total words in his hook coincided with two gestural phrases except for one word (aviators) which overlapped three gestural phrases. This was similar to the findings in Jhingran's introduction where one of the words in his speech segment overlapped three gestural phrases as well. Moreover, almost all of the words (6 out of

7) in Henderson's speech segment that overlapped more than one gestural phrase are content words. In addition, there were also five words (out of the seven) which contained a pitch accent and all of them were also content words. The majority of those words (4 out of 5) also overlapped the stroke and post-stroke hold phrases, regardless of order. In contrast to Harvey and Jhingran's hooks, Henderson's speech segment did not contain any words which overlapped two stroke phrases. One should take note that unlike the other three speech segments, Henderson's introduction did not contain any pitch accents which were located between any words or gestural phrases. Table 4.15 contains all of the information mentioned in the paragraph and it also shows which words in Henderson's hook overlapped with more than one gestural phrase.

Table 4.15: Words that overlapped gestural phrases and pitch accents (Henderson)

Speaker: Henderson / Words	Word Type		Gestural Phrases				Pitch Accent
	Content	Function	Preparation	Stroke	Post-Stroke Hold	Retraction	
1983	✓			✓	✓		✓
ourselves		✓	✓	✓			
aviators	✓		✓	✓		✓	✓
Snoopy	✓			✓	✓		✓
2	✓			✓	✓		
right	✓			✓	✓		✓
breeches	✓			✓	✓		✓
Total	6	1	2	7	5	1	5

4.4 Avery's Speech

The hook in Avery's speech consisted of 13 utterances. Most of the intonational phrase boundaries were falling (L%) and except for three of them which were rising (H%). These rising intonational phrases were also not questions which were similar to the rising (H%) intonational phrases in Harvey and Jhingran's speech segments. In addition, there were only two utterances in the hook of his speech that contained intermediate phrases and all of these intermediate phrases were falling (L-). These phrases occurred when Avery paused for a short while in the middle of an utterance. In total, there were four falling intermediate phrases in Avery's speech segment. For example, when he said, "She is glowing in her white dress," he paused after saying *glowing* before he completed the utterance. This utterance with two intermediate phrases is shown in the screenshot (Figure 4.16) and Figure 4.17 is a close-up of the earlier screenshot where both intermediate phrases in that particular utterance is indicated by an arrow and the pause is circled in the *Words* tier.

It was found that the hook of Avery's speech contained 45 pitch accents which were aligned with the gestures of his left and right hands. Most of the 45 pitch accents in this segment of the speech were high (H*). There were also 5 downstepped high pitch accents (!H*) that were annotated and only one low pitch accent (L*). However, one pitch accent was located between two gestural phrases of his right hand and two pitch accents were also located between two gestural phrases of his left hand so these pitch accents were not taken into account (refer to Table 4.16). Figure 4.18 shows one example of a pitch accent located between two gestural phrases and Figure 4.19 is a close-up of the earlier screenshot in Figure 4.18. The pitch accent and the gestural phrases which were aligned to it are circled in Figure 4.19.

Despite that, the majority of the pitch accents in Avery's hook occurred during the stroke phrase for both of his hands. Out of the 45 pitch accents, 17 of them occurred with the stroke phrase when Avery gestured with his left and right hands. The pitch accents that occurred with the stroke when he gestured with his hands accounted for 39.5% (17 out of 43, left hand) and 38.6% (17 out of 44, right hand) of the total amount. At the same time, the times when Avery did not gesture should also be taken note of. This is because there were 16 pitch accents out of a total of 43 (37.2%) which did not coincide with any gestural phrase in his left hand and 15 pitch accents (34.1%) out of a total of 44 which also did not occur with any gestural phrase in his right hand. In addition, there were very few pitch accents which occurred with the post-stroke holds of both of his hands. Table 4.16 shows the breakdown of the pitch accents which coincided (or not) with all of the gestural phrases in Avery's speech segment regardless of their type. Their matching percentages are also shown in the table as well.

Table 4.16: The pitch accents in Avery's hook and their gestural phrases.

Hand / Gestural Phrase	Left Hand		Right Hand	
	Pitch Accent			
	Number	Percentage (%)	Number	Percentage (%)
Preparation	5	11.6	7	15.9
Stroke	17	39.5	17	38.6
Post-Stroke Hold	3	7	3	6.8
Retraction	2	4.7	2	4.6
No Gesture	16	37.2	15	34.1
Total	43	100	44	100

However, when the pitch accents that occur without a hand gesture were not taken into account, the number of pitch accents that occurred with the stroke phrase of a gesture made up for more than half of the total amount for both of his hands. 63% (17 out of 27) of the total number of pitch accents that occurred during the stroke phrase when he gestured with his left hand and 58.6% (17 out of 29) of the gestures of the pitch accents occurred during the stroke phrase of his right hand. In contrast to the previous three speech segments, only a small number of pitch accents occurred within the post-stroke hold phrase. There were only four pitch accents which coincided with the post-stroke hold phrase for both his left (11.1%) and right (10.3%) hands. Instead, it was the preparation phrases which had the second-highest amount of pitch accents attached to them. This was the same for both of his hands where there were five pitch accents that were attached to the preparation phrases performed by his left hand (18.5%) and seven pitch accents occurred with the preparation phrases performed his right hand (24.1%). Table 4.17 shows the various gestural phrases which contain a pitch accent and their corresponding percentages without the *No Gesture row.*

Table 4.17: Proportion without the No Gesture row (Avery).

Hand / Gestural Phrase	Left Hand		Right Hand	
	Pitch Accent			
	Number	Percentage (%)	Number	Percentage (%)
Preparation	5	18.5	7	24.1
Stroke	17	63	17	58.6
Post-Stroke Hold	3	11.1	3	10.3
Retraction	2	7.4	2	7
Total	27	100	29	100

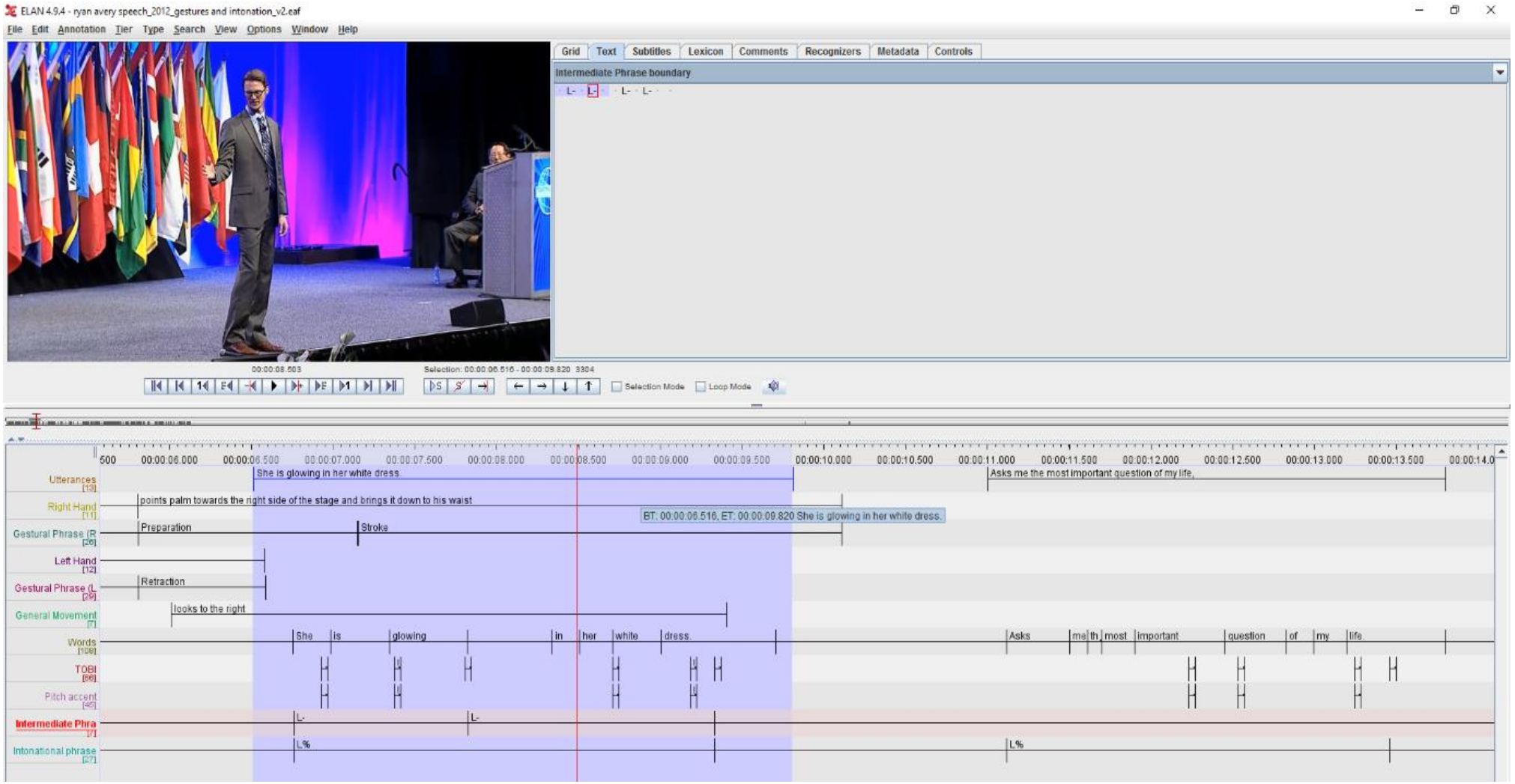


Figure 4.16: The utterance that contained two intermediate phrases (Avery).

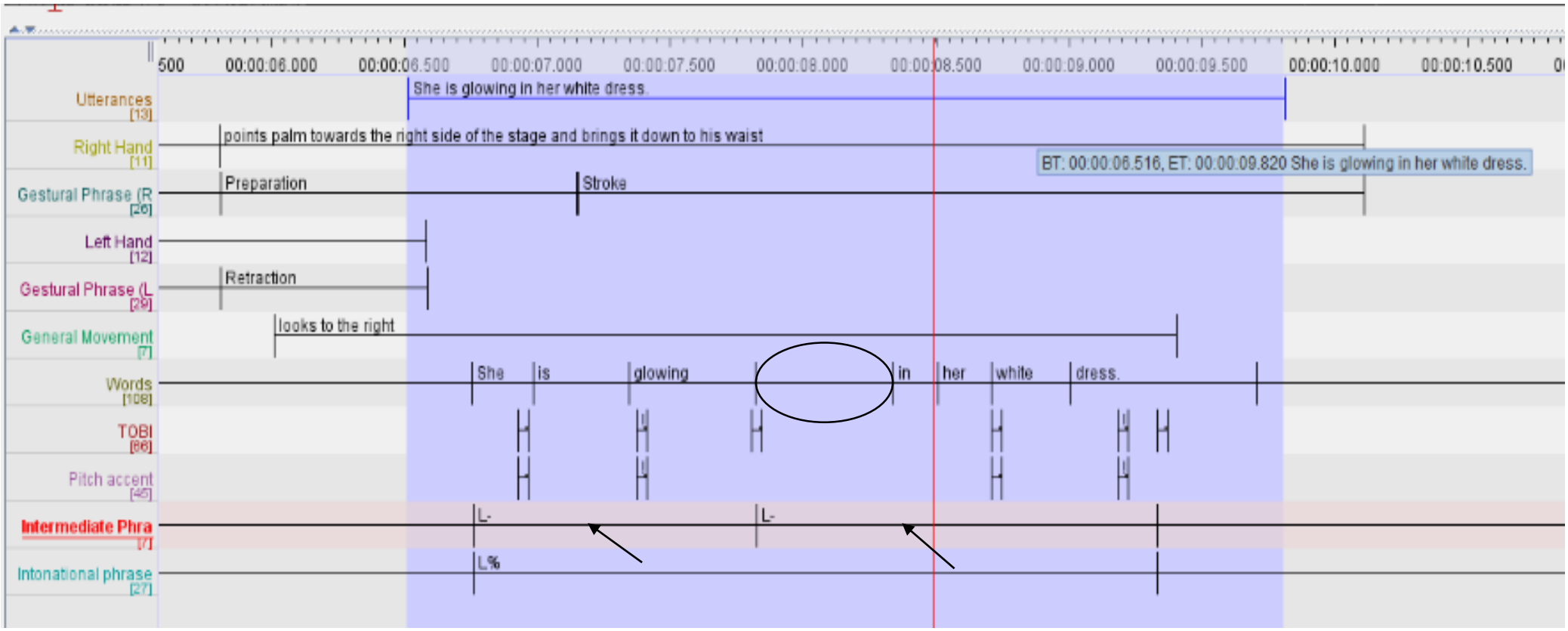


Figure 4.17: A close-up of the intermediate phrases in the utterance and the pause between them.

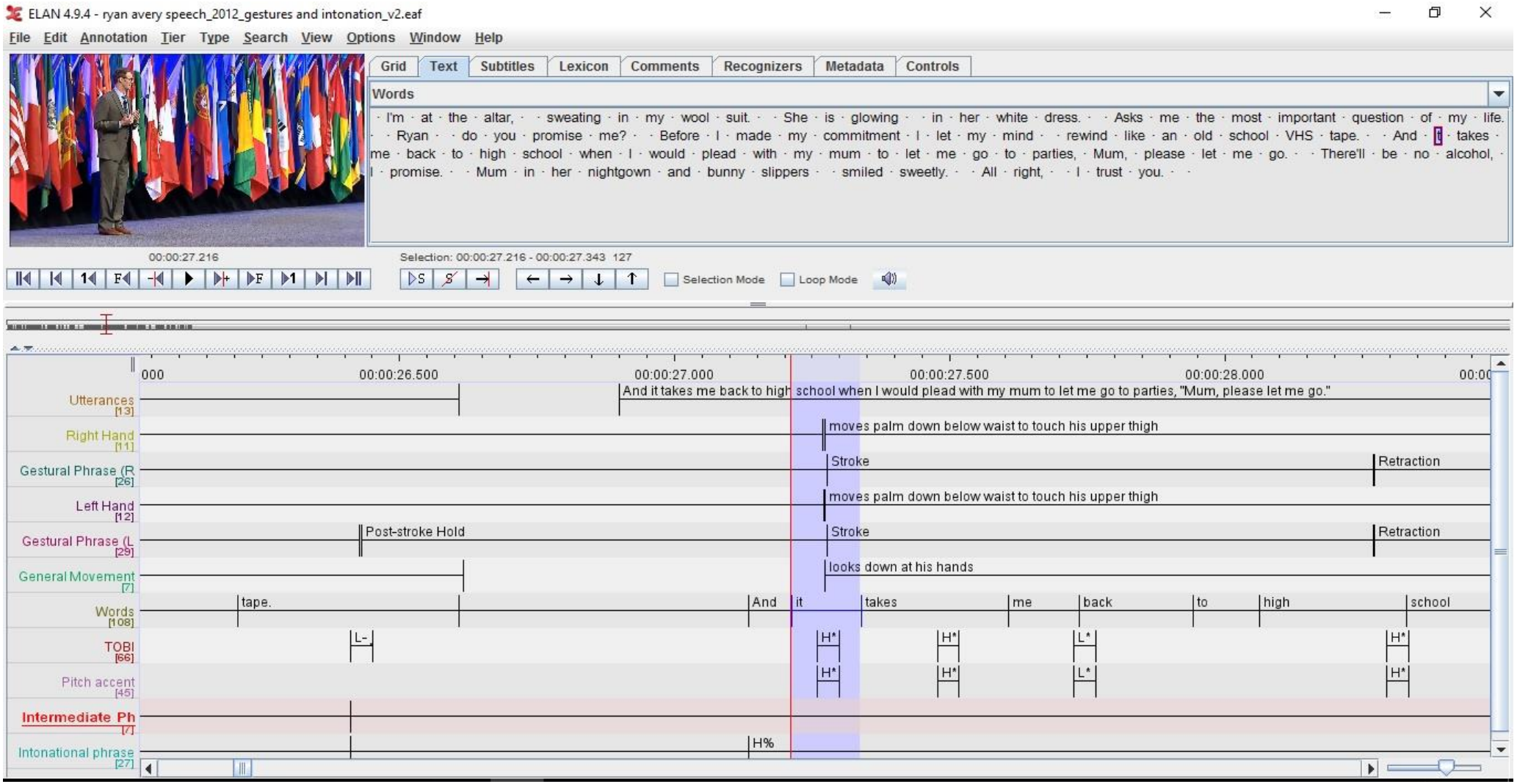


Figure 4.18: A pitch accent located between two gestural phrases (both hands)

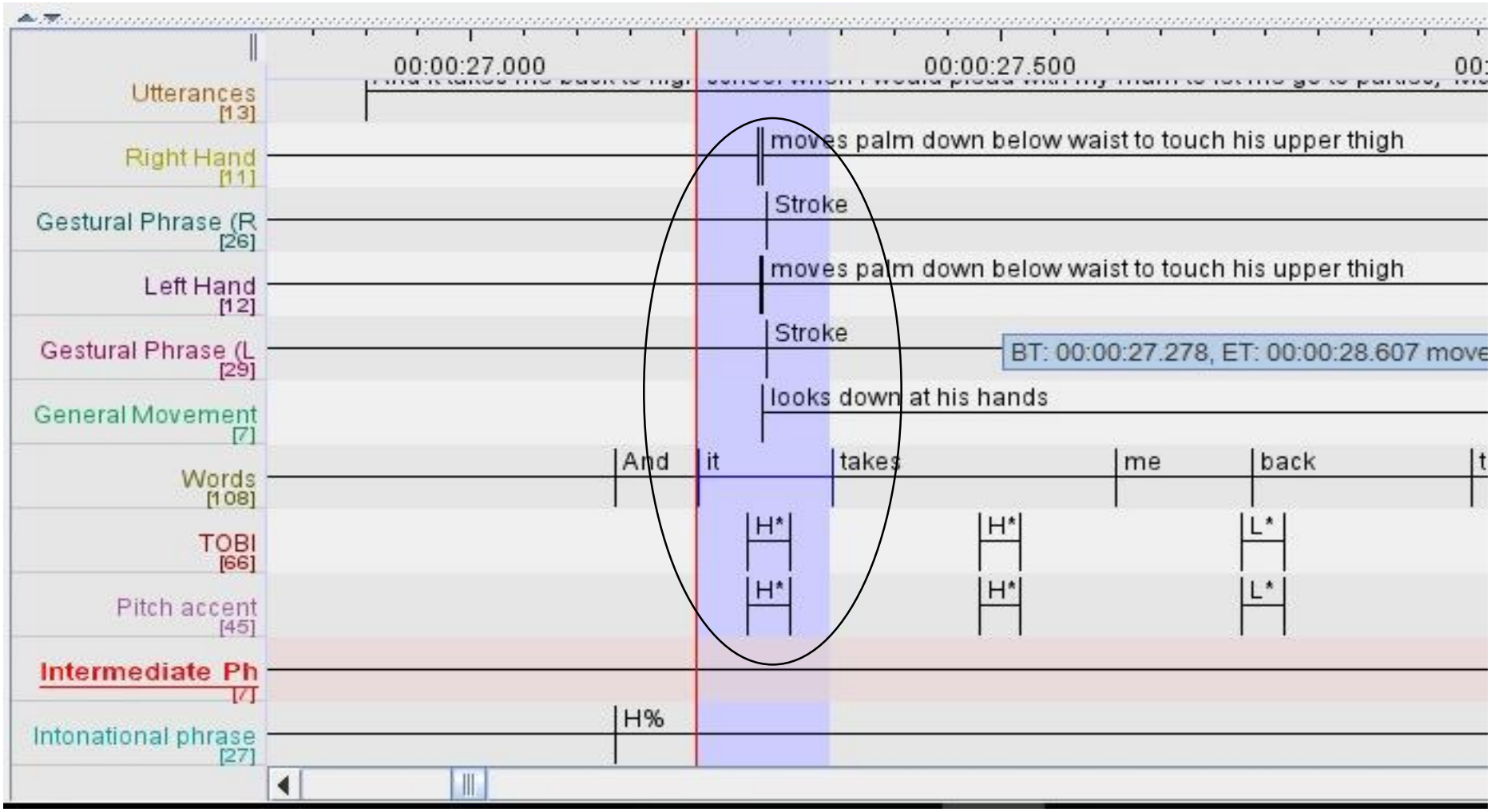


Figure 4.19: A close-up of the pitch accent between two gestural phrases (Avery).

As mentioned earlier in this section, there were 45 pitch accents in Avery's speech segment. However, five of the pitch accents occurred at a word boundary. Thus, these pitch accents were located in the middle of two words making it impossible to clearly tell which word the pitch accent is clearly aligned to. As a result, these pitch accents and their co-accompanying lexical items will not be taken into account during the analysis. In addition, the second word (at) in his speech segment contained two pitch accents. Thus, it was only counted as one resulting in only 39 pitch accents that were taken into account. Figure 4.20 shows one example of one of the pitch accents that was aligned to two lexical items. Figure 4.21 is a close-up of the earlier screenshot (Figure 4.20) where the location of the pitch accent is circled. The lexical items which were aligned with the pitch accent are also shown in Figure 4.21 and they are indicated with arrows. Out of the 39 words, 32 of them (82.1%) were content words and the rest were function words. Out of the 7 function words which occurred with a pitch accent, three of them (42.9%) were pronouns (first, second and third person) while the rest consisted of other types of function words. Table 4.18 shows the breakdown of the function words containing a pitch accent together with their matching percentages.

Table 4.18: The types of function words with a pitch accent (Avery).

Function Words / Speaker	Type							
	Pronoun (1st Person)	Pronoun (2nd Person)	Pronoun (3rd Person)	Possessive Pronoun	Articles	Preposition	Conjunction	Others
Avery	1 (14.3%)	1 (14.3%)	1 (14.3%)	2 (28.5%)	–	1 (14.3%)	–	1 (14.3%)
Total	7 words							

The amount of content words which occurred within the stroke and post-stroke hold phrases in Avery's introduction also exceeded the number of function words that coincided with those two gestural phrases. For example, there were 23 content words (53.5%) which occurred with the stroke of Avery's left hand but the number of function words attached the same gestural phrase numbered at 20 (46.5%). In addition, there were 25 (55.6%) content words which coincided with the stroke of Avery's right hand and only 20 (44.4%) function words that were attached to the stroke of his right hand. However, the only exception to this finding was the number of function words that were not attached to any gestures performed by Avery's left hand exceeded the number of content words of that same hand. There were 17 function words (53.1%) which did not occur with any gesture (left hand) and 15 content words (46.9%) which were not attached to any gesture made by his left hand. Table 4.19 shows the number of content and function words which coincided with the stroke and post-stroke hold phrases (and no gestures) in Avery's speech segment. The table also included the ratio of content to function words that were attached to a pitch accent as well.

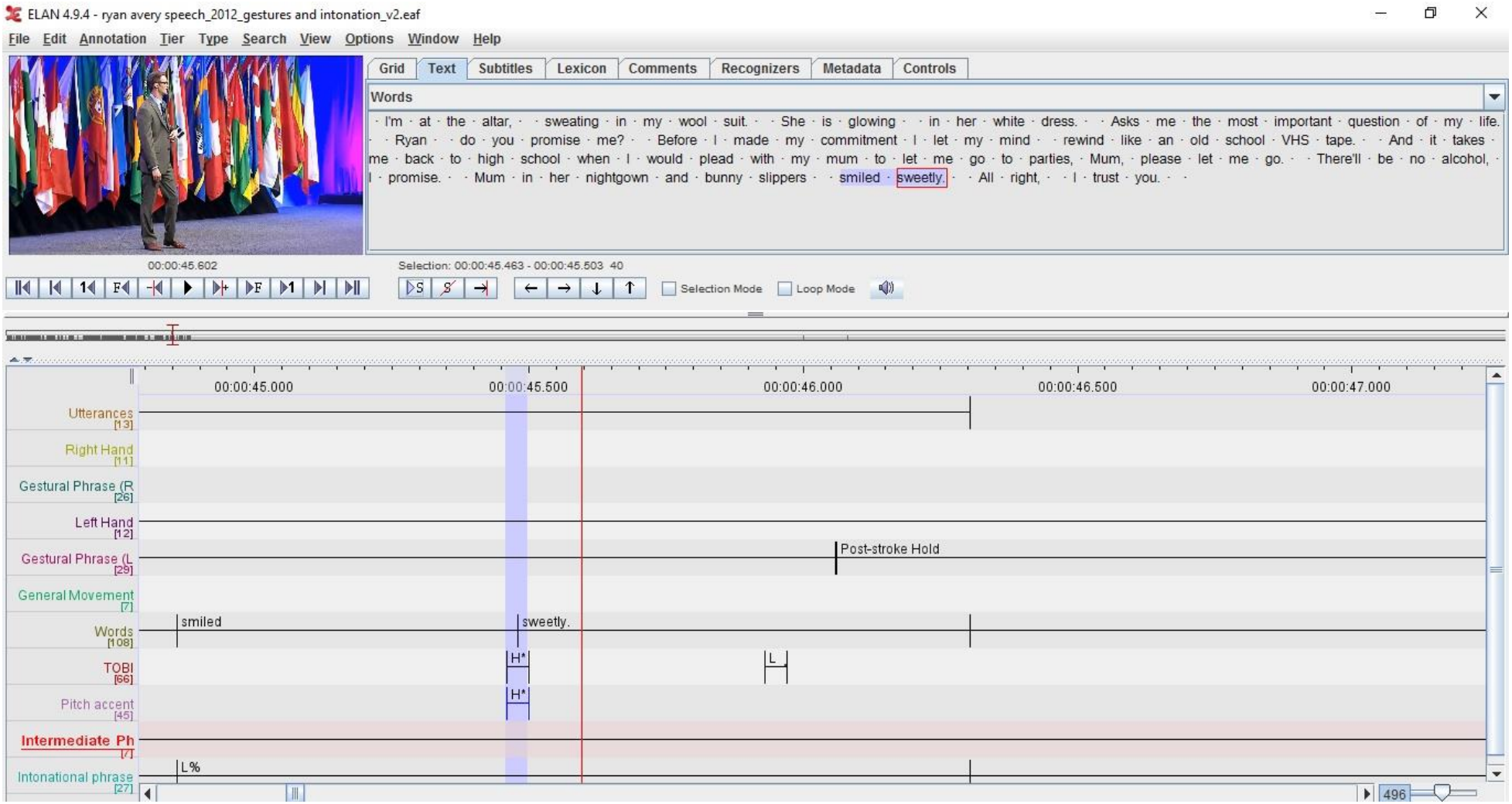


Figure 4.20: A pitch accent located between two words (Avery).

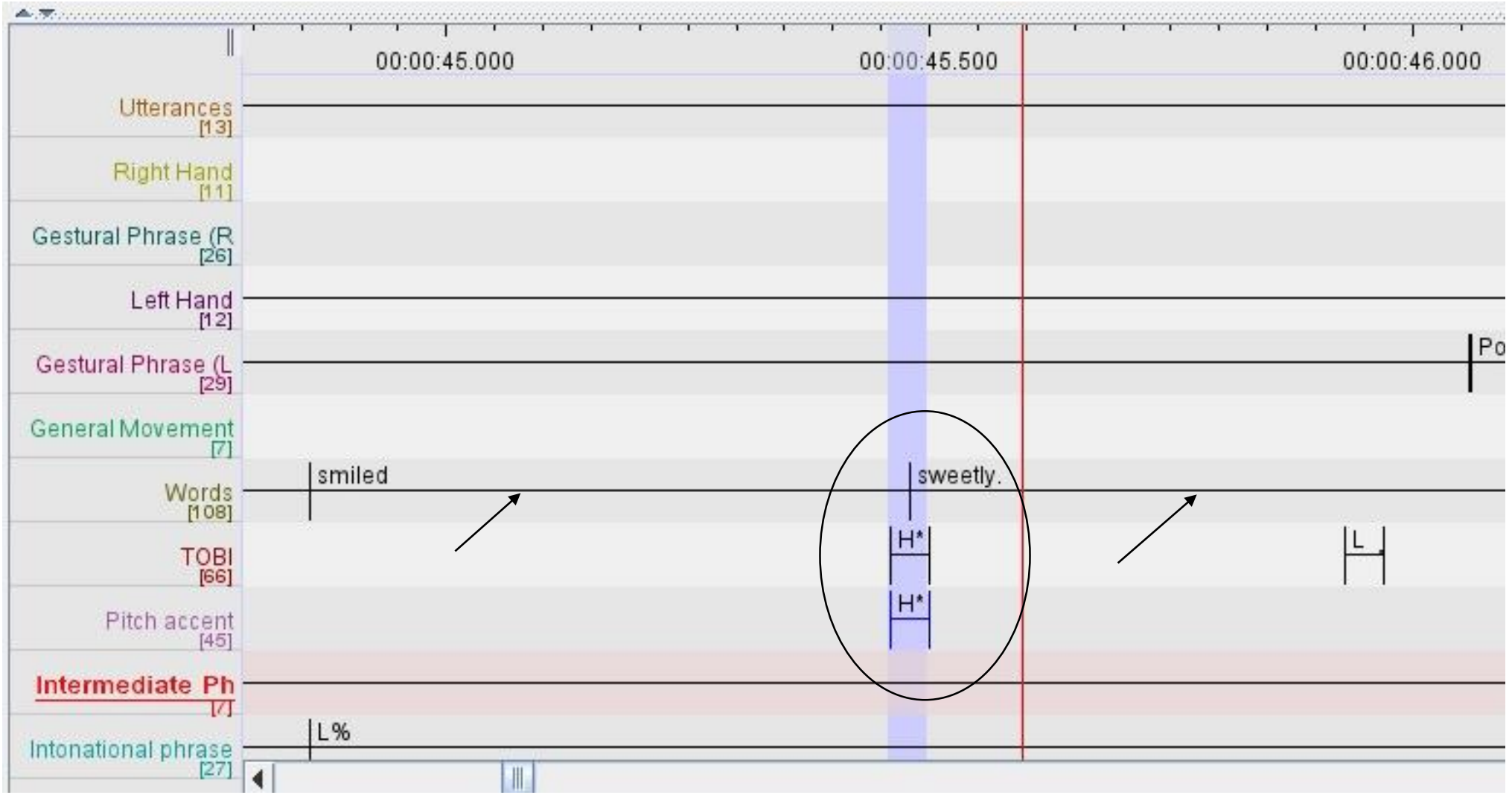


Figure 4.21: A close-up of the pitch accent and its lexical items (Avery).

Table 4.19: Ratio of the words to their gestural and intonational elements (Avery).

Gestures and Intonation / Speaker	Stroke		Post-Stroke Hold		No Gesture		Pitch Accents	
	Left	Right	Left	Right	Left	Right	Content	Function
Avery	C: 23 (53.5%) F:20 (46.5%)	C:25 (55.6%) F:20 (44.4%)	C:7 (63.7%) F:4 (36.3%)	C:5 (55.6%) F:4 (44.4%)	C:15 (46.9%) F:17 (53.1%)	C:18 (54.5%) F:15 (45.5%)	32 (82.1%)	7 (17.9%)
Total	43 words	45 words	11 words	9 words	32 words	33 words	39 pitch accents	

C represents the *content words* and F represents the *function words*

Avery's hook contained 21 words (out of 92) which overlapped more than one gestural phrase. As it was with the other three speech segments, there were more content words which overlapped more than one gestural phrase compared to the function words in Avery's speech segment. 16 (76.2%) out of the 21 words which coincided with more than one gestural phrase were content words whereas the rest were function words. 13 out of these 21 words contained a pitch accent and 11 (84.6%) of these words were also content words. In addition, all of these words except two coincided with at least one stroke phrase. There was one word (alcohol) which was aligned with two stroke phrases but there were no words which coincided with three gestural phrases. The majority of the words coincided with the stroke and post-stroke hold phrase (8 words) and the preparation and stroke phrase (7 words) regardless of order. Table 4.20 shows the gestural phrases that were aligned with each word and whether they have a pitch accent attached to them. The word which contained two stroke phrases was marked with a (✓✓).

Table 4.20: Words that overlapped gestural phrases and their pitch accents (Avery).

Speaker: Avery / Words	Word Type		Gestural Phrases				Pitch Accent
	Content	Function	Preparation	Stroke	Post-Stroke Hold	Retraction	
altar	✓			✓		✓	✓
sweating	✓		✓	✓			✓
in		✓	✓	✓			
suit	✓		✓	✓			
is		✓	✓	✓			
promise	✓			✓	✓		✓
VHS	✓			✓	✓		✓
tape	✓			✓	✓		
it		✓		✓	✓		✓
high	✓			✓		✓	✓
school	✓			✓		✓	
mum	✓		✓			✓	✓
to		✓		✓	✓		
parties	✓		✓		✓		✓
please	✓		✓	✓			✓
be		✓	✓	✓			
alcohol	✓			✓✓			✓
I		✓		✓	✓		✓
promise	✓			✓	✓		✓
nightgown	✓		✓	✓			✓
sweetly	✓			✓	✓		
Total	15	6	9	20	9	4	13

4.5 Overview of the Speech Segments

If one looks at all of the four segments of the speeches as a whole, there is a total number of 140 pitch accents which occurred with a gestural phrase or otherwise for the speakers' left and right hands. However, four pitch accents which occurred with the left and right hands of each speaker had to be taken out of the equation. This is because these pitch accents were located in between two gestural phrases so was not possible to tell which gestural phrase they were clearly aligned to. Many of the pitch accents occurred during the stroke phrase and they also sometimes occur when the speakers do not gesture with their hands. However, one also needs to take note of the pitch accents that occur during the post-stroke hold phrase as they form quite a significant number out of the total number of pitch accents that were annotated in all of the four speech segments. The number of pitch accents that occured during the post-stroke hold phrase consist of 21.3% (29 out of 136, right hands) and 30.1% (41 out of 136, left hands) of the total number of annotated pitch accents. Table 4.21 shows the number of pitch accents, regardless of whether they are H*, !H* or L*, that occur with each gestural phrase and when there are no hand gestures together with their matching percentage for all of the four speech segments.

Table 4.21: The number of pitch accents that occur with every gestural phrase.

Hand / Gestural Phrase	Left Hand		Right Hand	
	Pitch Accent			
	Number	Percentage (%)	Number	Percentage (%)
Preparation	7	5.1	12	8.8
Stroke	37	27.2	56	41.2
Post-Stroke Hold	41	30.1	29	21.3
Retraction	9	6.6	8	5.9
No Gesture	42	31	31	22.8
Total	136	100	136	100

When the number of pitch accents that occur without a hand gesture is removed from the equation, one can see that the majority of the gestural phrases which contains a pitch accent are still the stroke and the post-stroke hold phrases for both of the speakers' left and right hands. The number of pitch accents that occur within these two phrases number from slightly over a quarter to slightly more than half of the total number of pitch accents that were taken into account (see Table 4.22). For example, 53.3% (56 out of 105) of the total pitch accents which occurred when the speakers gestured with their right hands were aligned with the stroke phrase. On the other hand, 43.6% (41 out of 94) of the pitch accents which occurred when the speaker gestured with their left hands were aligned with the post-stroke hold phrase.

This trend where most of the pitch accents occur within the stroke and post-stroke hold phrase is mostly similar throughout all the four speech segments whether they are analysed individually or as a whole. The only difference that could be seen within the various speech segments was the pitch accents in some of the speech segments may occur more frequently within the post-stroke hold phrase compared to the stroke phrase (see Tables 4.7 & 4.12). In addition, the number of pitch accents that occurred during the preparation and retraction phrases in all of the four speech segments were very small. Table 4.22 shows the total number of pitch accents (regardless of type) that were aligned with their respective gestural phrases and matching percentages after the pitch accents that did not occur within a gestural phrase have been removed from the equation.

Table 4.22: The pitch accents of the four hooks and their gestural phrases.

Hand / Gestural Phrase	Left Hand		Right Hand	
	Pitch Accent			
	Number	Percentage (%)	Number	Percentage (%)
Preparation	7	7.4	12	11.5
Stroke	37	39.4	56	53.3
Post-Stroke Hold	41	43.6	29	27.6
Retraction	9	9.6	8	7.6
Total	94	100	105	100

The study annotated 298 words in the four speech segments. There were also 140 pitch accents which were also identified during this study. However, only 126 words were analysed during this stage of the study. This is because there were some words which contained more than one pitch accent (see Figure 4.12 and Figure 4.13) so they were counted as one word as long all of the pitch accents were clearly aligned within the word. On the other hand, there were also seven pitch accents which were located at the border of two words making it impossible to tell which word it was clearly aligned with (see Figure 4.3 and Figure 4.4). The data showed 93 out of the 126 (73.8%) words which were aligned with a pitch accent were content words. In contrast, the other 33 words (26.2%) which contained a pitch accent were function words. The data also showed that the four speakers attached pitch accents to at least one first person pronoun and a possessive pronoun. In total, there were six first person pronouns (out of 33) and six possessive pronouns (out of 33) which had a pitch accent attached to them. In addition, out of these 33 words, 9 of these words (27.2%) were personal pronouns (first, second and third person) and this group of words make up the majority of the function words in all of the speech segments which have a pitch accent aligned with them. Table

4.23 shows the pitch accents which were aligned with the different types of function words in all of the four speech segments.

Table 4.23: The pitch accents with various types of function words.

Function Words / Speaker	Type							
	Pronoun (1st Person)	Pronoun (2nd Person)	Pronoun (3rd Person)	Possessive Pronoun	Articles	Preposition	Conjunction	Others
Harvey	3	–	–	1	–	1	2	2
Jhingran	1	–	1	2	2	3	1	3
Henderson	1	–	–	1	2	–	–	–
Avery	1	1	1	2	–	1	–	1
Total	6 (18.2%)	1 (3%)	2 (6%)	6 (18.2%)	4 (12.1%)	5 (15.2%)	3 (9.1%)	6 (18.2%)

The number of content words which were aligned with the stroke phrases for both hands slightly exceeded the number of function words when all of the words from the four speech segments were added up. 61 content words (out of 116) or 52.6% which coincided with the stroke phrase of the speakers' left hands and 73 out of the 141 (51.8%) words which were aligned with the stroke of the speakers' right hands were also content words. In addition, there were also more content words which occurred with the post-stroke hold phrase of the speakers' right hands. In contrast, there were more function words which coincided with the post-stroke hold phrases of the speakers' left hands and when they did not gesture with either of their hands. Table 4.24 shows the total number of the content and function words which occurred with both of the gestural phrases and when the speakers did not gesture together with their respective percentages. It also shows the number of content and function words which were aligned to a pitch accent in all of the four speech segments.

Table 4.24: Ratio of the words to their gestural and intonational elements (All).

Gestures and Intonation / Speaker	Stroke		Post-Stroke Hold		No Gesture		Pitch Accents	
	Left	Right	Left	Right	Left	Right	Content	Function
Harvey	C:19 F:18	C:16 F:18	C:4 F:13	C:4 F:13	C:15 F:16	C:15 F:19	21	9
Jhingran	C:14 F:14	C:25 F:27	C:19 F:20	C:13 F:10	C:4 F:11	C:1 F:2	29	13
Henderson	C:5 F:3	C:7 F:3	C:12 F:7	C:20 F:10	C:7 F:3	C:0 F:0	11	4
Avery	C: 23 F:20	C:25 F:20	C:7 F:4	C:5 F:4	C:15 F:17	C:18 F:15	32	7
Total	C: 61 (52.6%) F: 55 (47.4%)	C: 73 (51.8%) F: 68 (48.2%)	C: 42 (48.8%) F: 44 (51.2%)	C: 42 (53.2%) F: 37 (46.8%)	C: 41 (46.6%) F: 47 (53.4%)	C: 34 (48.6%) F: 36 (51.4%)	93 (73.8%)	33 (26.2%)
	116 words	141 words	86 words	79 words	88 words	70 words	126 pitch accents	

C represents the *content words* and F represents the *function words*

In total, there were 54 words (out of 298) in all of the four speech segments which overlapped more than one gestural phrase. 45 of these words (83.3%) were content words while the rest were function words. 36 of these words (66.7%) contained a pitch accent and 32 of them were also content words. In fact, the majority of these words (24 out of 54 words or 44.4%) overlapped the stroke and post-stroke hold phrase, regardless of order. In addition, 12 words out of the total (22.2%) coincided with the preparation and stroke phrase which is the second highest number out of the total amount of words that overlapped more than one gestural phrase. Almost all of the words in all of the speech segments which overlapped more than one gestural phrase contained at least one stroke except for four words. Table 4.25 shows the frequency of the number of words in each speech segment which overlap the different gestural phrases.

Table 4.25: Words which overlapped different gestural phrases.

Number of Words / Speaker	Gestural Phrases							
	Stroke & Post-Stroke Hold	Preparation & Stroke	Stroke & Retraction	Two Strokes	Three Gestural Phrases	Preparation & Post-Stroke Hold	Post-Stroke Hold & Retraction	Preparation & Retraction
Harvey	2	1	4	2	–	–	1	–
Jhingran	9	3	1	1	1	–	1	–
Henderson	5	1	–	–	1	–	–	–
Avery	8	7	3	1	–	1	–	1
Total	24	12	8	4	2	1	2	1
	54 words							

CHAPTER 5: DISCUSSION

The previous chapter described the frequency of the pitch accents that occurred in each gestural phrase and how they interact with the lexical items in each speech segment. It was found that many pitch accents were attached to the stroke and post-stroke hold phrase in all of the four speech segments. In addition, there were also many pitch accents which did not occur with any gestural phrases too. The study also found that many pitch accents also coincided with many content words in all of the four speech segments. Based on these findings, this chapter will discuss the research questions asked in the Introduction. They were: (1) What are the roles of gestures and intonation in a prepared speech? (2) What is the extent of the relationship between gestures and pitch accents in a prepared speech? and (3) Which of them appears to have a bigger influence on a public speech? The chapter will end with a discussion of any other findings which were discovered alongside the answers to the study's research questions.

5.1 What are the roles of gestures and intonation in a prepared speech?

One of the aims this study seeks to achieve is to identify the roles gestures and intonation play in a prepared speech. The previous chapters mentioned that the study will only pay attention to the hand gestures made by the speakers. The study found that in all of the four speech segments, all of the speakers used representational gestures as one of the main means of nonverbal communication with their audience. The gestures annotated in all of the speech segments were consistent with Hostetter and Alibali's (2010) definition of representational gestures which depicts the semantic content of a speaker's description.

The study found that in a prepared speech, gestures can be used to describe objects, events, experiences and people to the audience. They could also be used by a speaker to describe his or her speech in greater detail and it even contained information which was not mentioned in their verbal utterances. For example, in the hook of Avery's speech, he uses various gestures to describe the objects, events, experiences and the people in his wedding. For example, when he told the audience that he was sweating in his wool suit, he wiped his whole left forearm across his brow. Figure 5.1 shows a screenshot of the moment Avery wiped his arm across his brow and a close-up of the accompanying verbal utterance, which was circled, is shown in Figure 5.2. The utterance showed Avery telling his audience that he was sweating in his wool suit but he did not mention the extent to which he was sweating. Therefore, the gesture he made while verbalising that utterance seemed to indicate that he wanted to show his audience that he was sweating profusely.

The fact that gestures were used to convey information that was not available in the verbal utterance was also consistent in the other speech segments other than Avery's introduction. In Henderson's speech segment, he told his wingman to break hard right because he saw the Red Baron. However, his verbal utterance did not mention the location of the Red Baron but he stretched his arm out in front of him as if to indicate to the audience that the Red Baron was right in front of them. Figure 5.3 shows the time when Henderson stretched out his arm while telling his wingman that he saw the Red Baron and Figure 5.4 is a close-up of the verbal utterance and the tier (Words) containing the utterance has been circled. These findings were consistent with previous studies like McNeill (1992) and Kendon (1995) as they also found that speakers used gestures to convey additional information apart from their verbal utterances. This study also expanded upon Holle and Gunther's (2007) work which found that iconic gestures

provide additional information such as contextual clues apart from the speaker's verbal utterance in natural speech. The current study found that representational gestures, which comprises of iconic, metaphoric and deictic gestures, are also used by a speaker to communicate additional information apart from their verbal speeches.

The findings of the current study also indicated that gestures are also used to reinforce an idea besides being used to describe objects, experiences, events and people in a public speech. For example, in Harvey's introduction, he told the audience how scared he felt when he was surrounded by a pack of hunting hounds. When he told the audience that his heart jumped, he raised his hands and put them on his chest (stroke phrase) to indicate his heart jumping. After that, he told the audience that he jumped after his heart jumped. He maintained the same gesture of keeping both of his hands on his chest after he described his heart (post-stroke hold) to describe himself jumping.

When one looks at this whole exchange, Harvey first described his fear by gesturing that his heart jumped in fear when he saw the hounds. Then, he further reinforces the fact he was so afraid that he jumped by maintaining the same gesture. This gesture of putting both of his hands to his chest represented two things: his heart and himself. On the other hand, this gesture only represents one emotion which is his fear but when the gesture is maintained, it reinforced the idea of how afraid he was of the hounds that his whole body (besides his heart) jumped. Figure 5.5 shows the moment when Harvey put his hands on his chest to indicate to the audience that his heart jumped. The utterances which accompanied the gesture is circled in the screenshot. Figure 5.6 shows multiple screenshots of Harvey maintaining the same gesture throughout the verbal utterance even though he moves his head and body in different directions. The screenshots are also numbered and the time stamps in each screenshot are indicated by arrows.

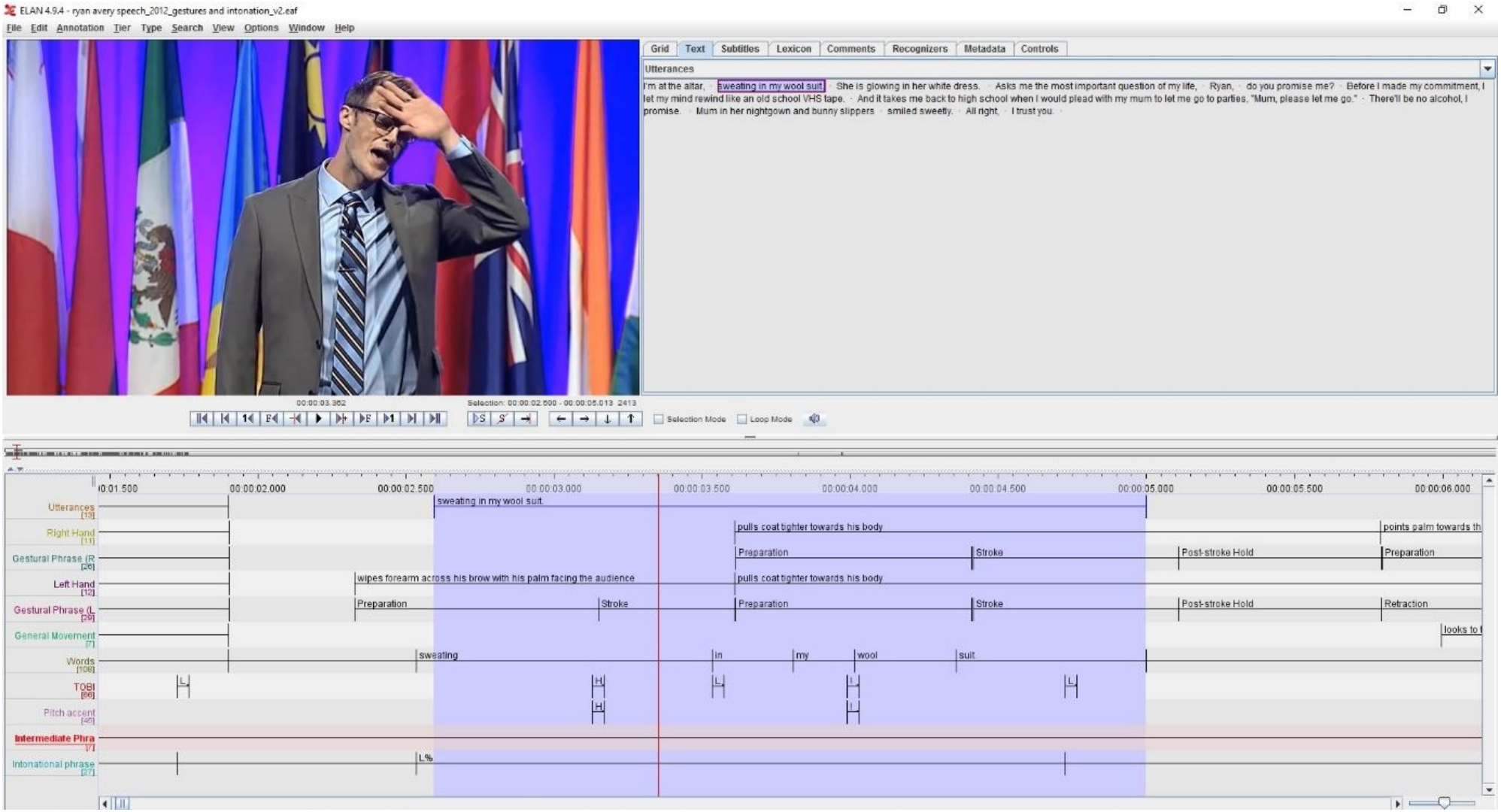


Figure 5.1: Avery wiping his whole arm across his brow.

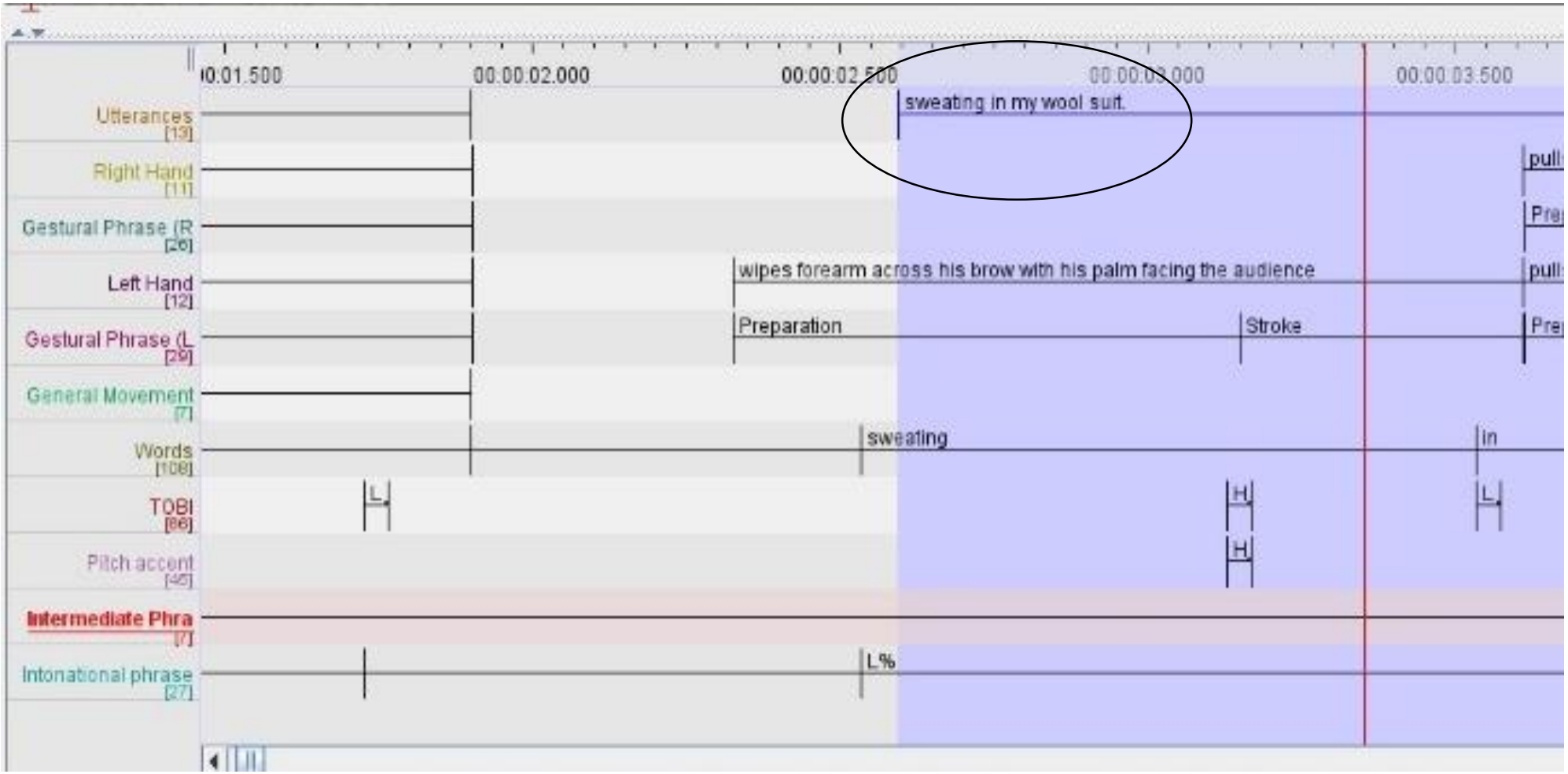


Figure 5.2: A close-up of the verbal utterance and its gesture (Avery).

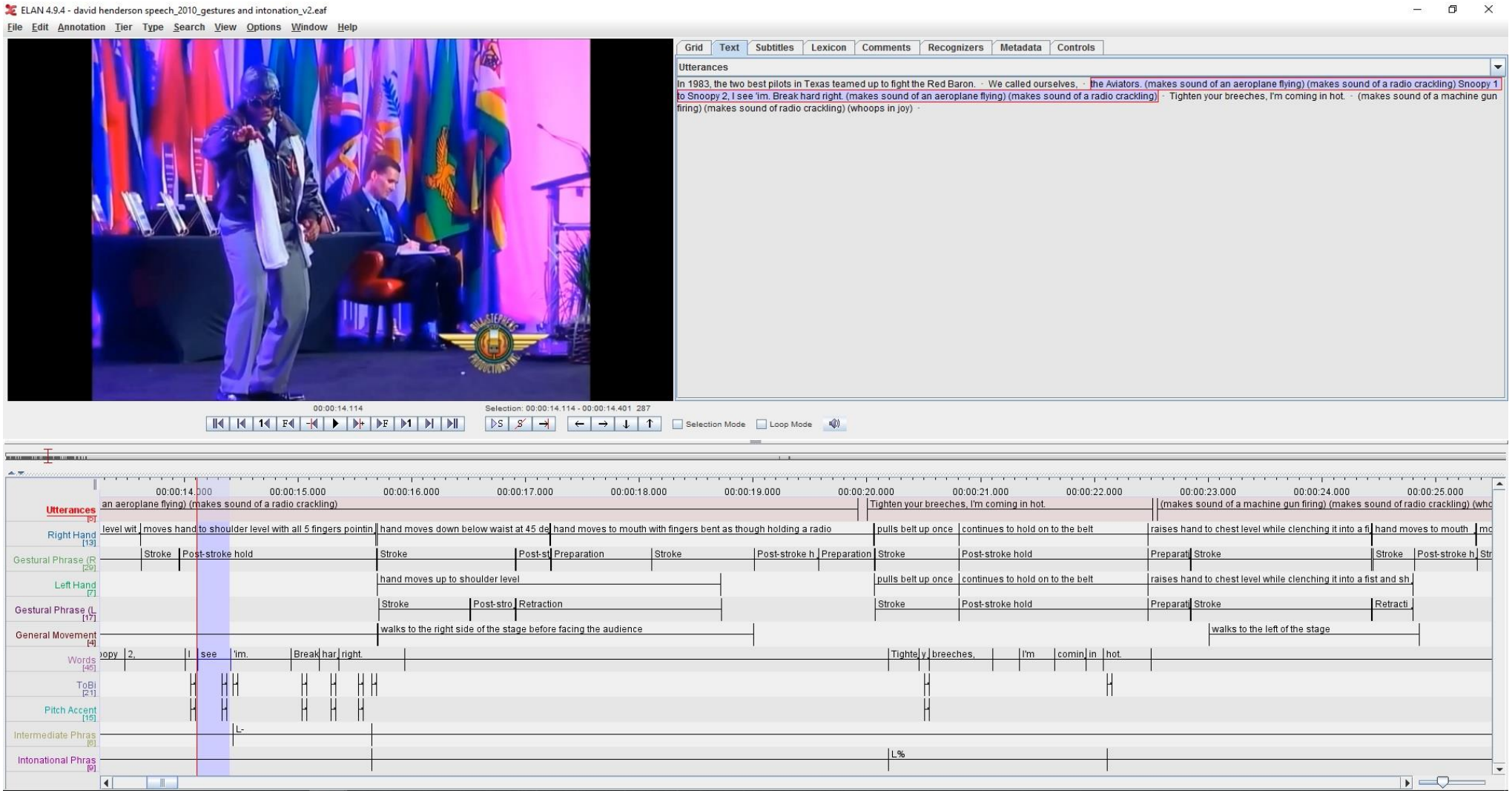


Figure 5.3: Henderson telling his wingman he saw the Red Baron.

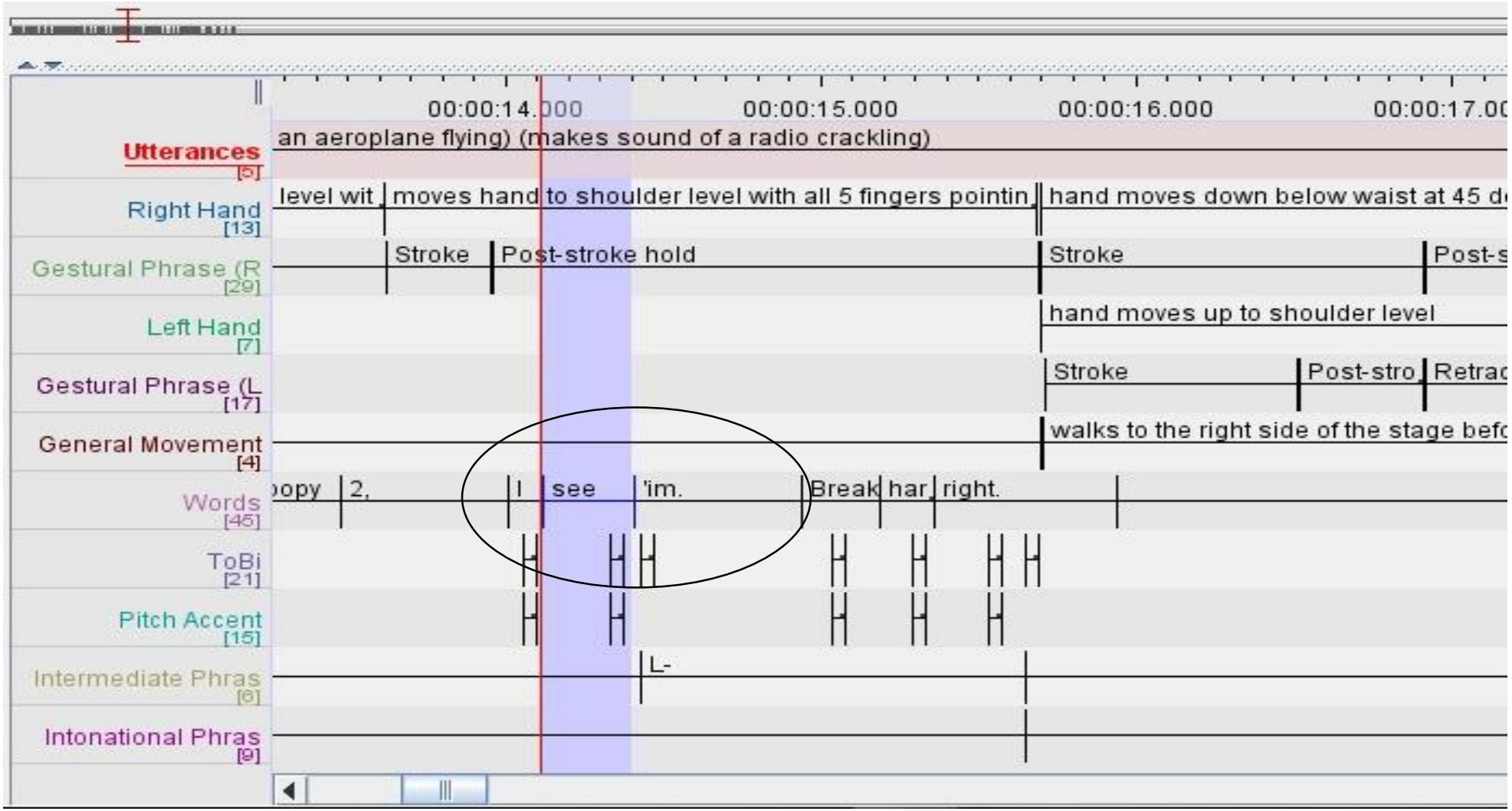


Figure 5.4: A close-up of the verbal utterance and its gesture (Henderson).

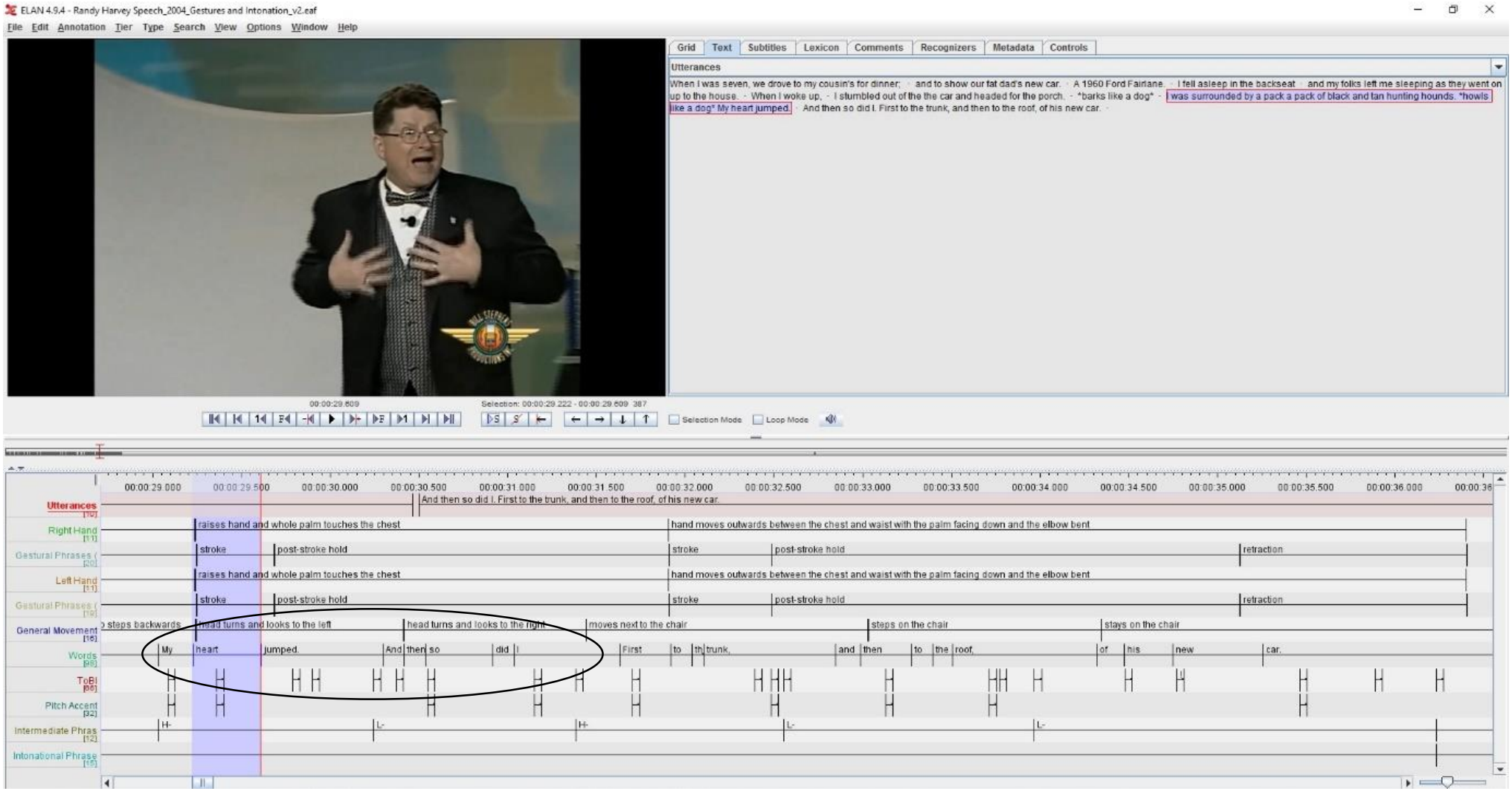
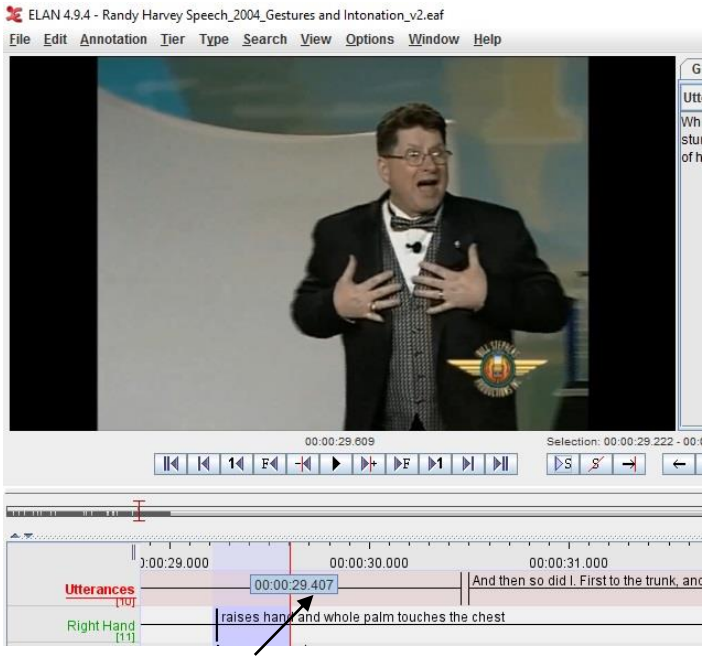
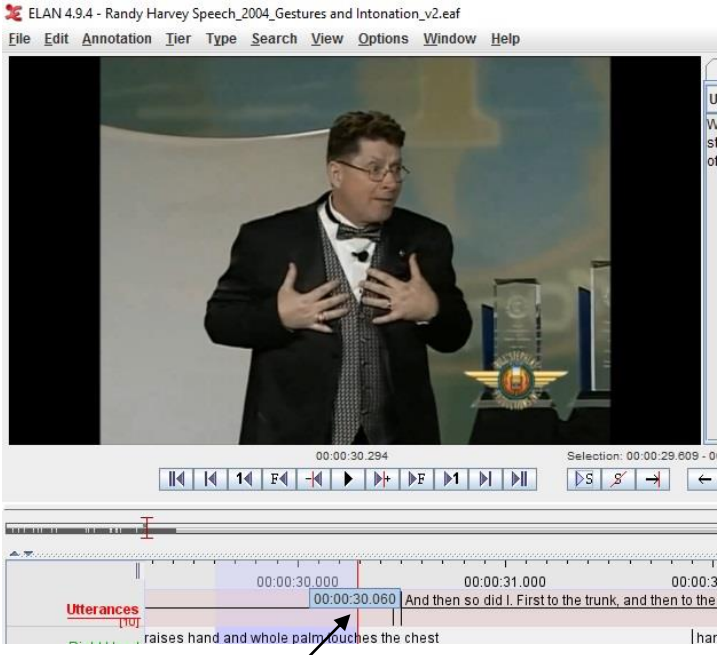


Figure 5.5: Harvey's heart jumped when he was surrounded by the hounds.

1



2



3

4

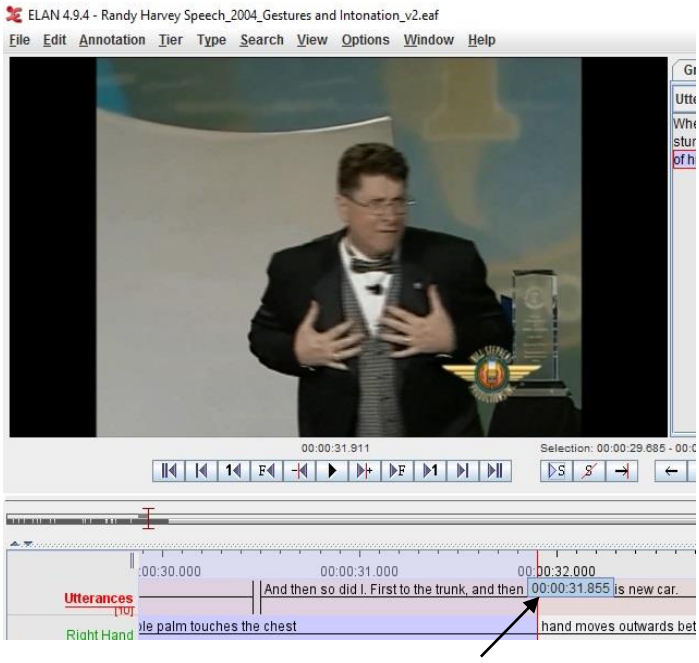


Figure 5.6: Harvey using the same gesture to reinforce a point.

In the current study, the intonation in all of the four speech segments also played various roles as well. One of the roles intonation plays in enhancing a prepared speech is it helps the speaker to convey his or her emotions to the audience. This finding is consistent with Bolinger (1983) who found that the pitch of a speaker can help the audience infer his or her emotion at the point of speaking (refer to Section 2.2, pg. 35). However, his study only focused on natural speeches where most speakers may not be aware of the relationship between their pitch and their emotional state. In contrast, a public speaker will likely be aware of the fact that pitch can be used to communicate their emotions to their audience. For example, one of the Toastmasters training manuals states that a speaker's voice reflects their psychological and emotional state of mind (Toastmasters International, 2011a) and it also tells the reader about how to improve the way they speak to the audience.

This training was reflected in Harvey's speech segment when he told the audience that he and his heart jumped when he was surrounded by a pack of hounds (refer to Figure 3). This is because in addition to his gestures, his voice reached a high pitch as he described his heart jumping and how he then jumped on his father's car. The rise in the pitch in his voice was similar to Bolinger's (1983) findings (refer to Section 2.2, pg. 35) where Harvey's tension and fear was reflected in his pitch. These pitch accents told the audience that he was so afraid of the hounds that he jumped onto his father's car. Besides giving an indication of his emotional state at that time, it is also possible that his intonation also helped to reinforce the message of how fearful he was. This shows that gestures and intonation likely share many similar functions in enhancing a prepared speech. Moreover, as Harvey had undergone the training in Toastmasters, it is also highly possible that he was aware that raising his pitch could be used to hammer his

message home. Thus, he may have consciously used a high pitch to communicate his fear to the audience during the competition.

This shows that a person can be trained to use the right intonation during a speech to convey his or her message to the audience. Furthermore, this also indicates that gestures and intonation work in sync with the words of the verbal utterance to convey the same message but with increasing intensity and prominence. For instance, Harvey said that he jumped when he was surrounded by the hounds. This could indicate fear but his pitch and gestures confirmed it, thus reinforcing the message and giving the audience a clearer picture of the extent of his fear. Figure 5.7 shows the pitch track of the utterance in Harvey's speech segment and the rises in his pitch were marked by arrows. These peaks were also annotated in the Pitch Accent tier as they were also the pitch accents in that utterance. The utterance in the screenshot was the same utterance in Figures 5.5 and 5.6 which showed how the gestures and intonation in the speech segment were synchronised with the words in a verbal utterance to convey the same message.

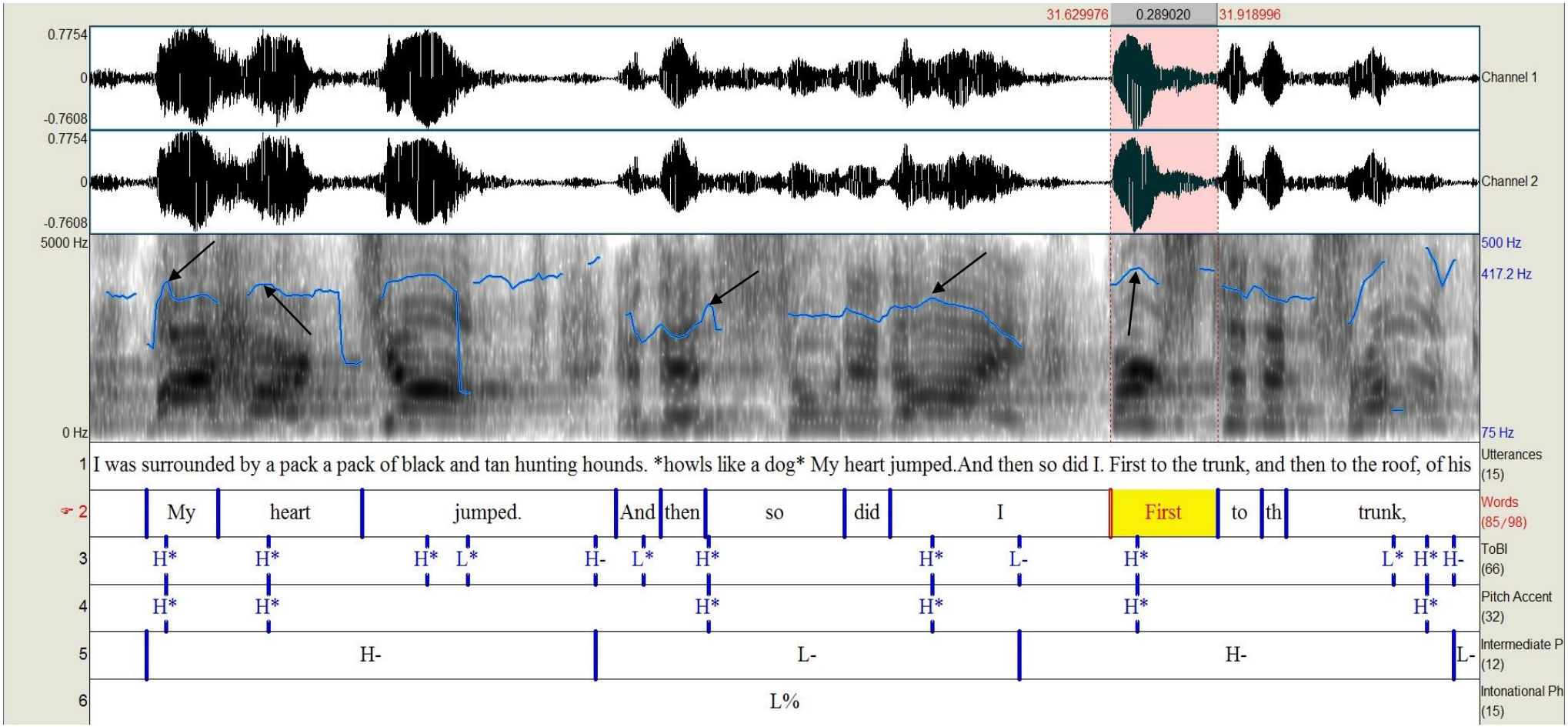


Figure 5.7: The rises in Harvey's pitch track.

Another feature which the study found was that the speakers in all of the speech segments used intonation to mark the prominent parts of their speeches. This was reflected in the number of pitch accents which were aligned with the different lexical items in each speech segment. Each speech segment consisted of content and function words. Content words are words like nouns, verbs, adjectives and adverbs whereas function words consist of words like pronouns, articles and conjunctions. The data in the current study showed that the majority of the pitch accents (93 out of 126) were found to be aligned with the content words of each speech segment. Thus, it showed that each speaker mostly stressed at least one syllable of most of the content words in their speech segments. Based on the data, one can also say that the content words in a prepared speech are generally more prominent than its function words because pitch accents are used to mark words which are important (Hirschberg & Pierrehumbert, 1986). This is because these words introduced new information to the audience and they were also used to capture their attention as well. These findings were consistent with previous studies that showed that the content words in a speech contained the most information (Zheng and Pierrehumbert, 2010).

One example of how pitch accents were used to mark prominence in an utterance could be seen in Jhingran's speech segment when he told the audience that he received a letter from the Massachusetts Institute of Technology. This university was the graduate school of his dreams and as he introduced this piece of information to his audience, the pitch of his voice changed significantly i.e. rose at the words like *graduate*, *school*, *my* and *dreams*. This is because Jhingran felt that getting into this school was very important to him and he wanted his audience to know that too. Thus, these words had pitch accents attached to them to signal to the audience that they were important. In fact, three of these words were content words but there was only one

function word (*my*). Figure 5.8 shows a screenshot of the pitch track in Jhingran's speech segment as he shared that piece of information mentioned earlier in this paragraph. The pitch accents which were attached to the prominent words are indicated by arrows.

Gestures and intonation in prepared speeches serve to convey additional information which is not stated in the speaker's verbal utterance. This is because gestures help to show the intensity of an action as shown in Figure 5.1 and both gestures and intonation can also be used simultaneously to communicate the speaker's emotions to the audience (refer to Figure 5.5). In addition, gestures were also used to reinforce the speaker's message to the audience (refer to Figure 5.6). On the other hand, intonation in a prepared speech is used to mark prominent lexical items so the audience would know that they should pay more attention to them as they carry the most information which would help them to grasp the message of the speech.

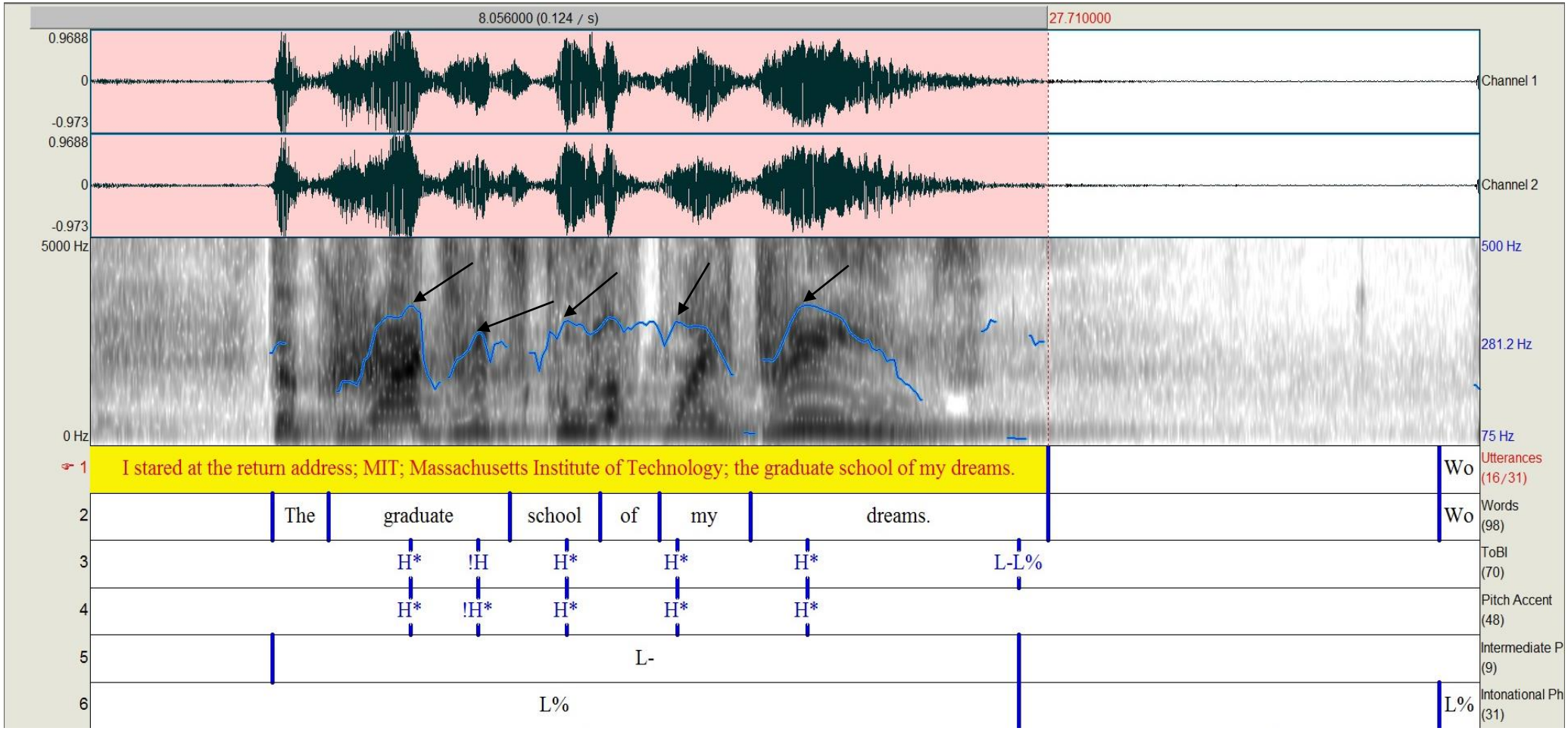


Figure 5.8: The pitch accents which marked the prominent words (Jhingran).

5.2 What is the extent of the relationship between gestures and pitch accents in a prepared speech?

Previous studies have investigated the relationship between intonation and gesture in natural speech. They have found that gestures and speech share the same system (Hostetter & Alibali, 2008) and they are co-expressive (McNeill, 2005) as they express the same idea but in different ways. However, the previous chapters mentioned that the extent of this relationship has not been fully explored. This study found that gestures and intonation in a prepared speech are also co-expressive as they both express an idea in their own way. This is because a speaker gestures with his or her arms in many ways to communicate a piece of information to their listeners. On the other hand, pitch accents are the changes in the pitch of a person's speech. These changes in the pitch usually happen when a syllable in a word is made prominent by the speaker because he or she wants to draw the audience's attention to that word or to convey a certain message to them.

In all of the speech segments that were annotated, the pitch accents, regardless of type, frequently occur during the stroke phrase of the gestures each speaker makes with both of their hands once the pitch accents that were not aligned to any gestural phrase were taken out of the equation (refer to Table 4.22, pg. 138). Although every speech was prepared and they were staged performances, the findings in this study were consistent with previous studies by Loehr (2004), Mendoza-Denton & Jannedy (2011) and Brentari, Marotta, Margherita & Ott (2013) who also found that many, if not most pitch accents are aligned with the stroke phrase in a speech. This is because the stroke is the most important phrase of a gesture and a speaker will only gesture if the information or an idea of a speech is important enough (Hostetter, 2008). One example which reflects this was when Harvey wanted to tell the audience that his heart jumped

(refer to Figure 5.5, pg. 149), he put both of his palms on his chest to signify his heart. The data showed that the stroke phrase of his gesture occurred at the word *heart* and there was also a pitch accent attached to that lexical item. In addition to gestures, the words in a speech can also be made intonationally prominent by changing the pitch (be it high or low) in order to show crucial information such as the main topic of a speech and its focus (Rosenberg & Hirschberg, 2009). This was shown in the earlier section (refer to pg. 154) and it goes to show that gestures and intonation can work together to tell the audience which piece of information is important in a speech and they are also used to draw the attention of the audience to certain parts of the speech.

At the same time, one also needs to take note that there was a significant number of pitch accents that occurred during the post-stroke hold phrase in each speech segment. The number of pitch accents that occurred during this phrase were the highest after the pitch accents which occurred within the stroke phrase of the right hand and they even exceeded the number of pitch accents within the stroke phrase of left hand (refer to Table 4.22, pg. 138). This did not happen in any of the previous studies on gestures and intonation mentioned earlier (refer to pg. 157) as all of them have always found that pitch accents were mostly aligned with the stroke phrase and only Loehr (2004) briefly mentioned that pitch accents sometimes occurred with other phrases other than the stroke. However, he did not specify which gestural phrases the pitch accents were aligned with. As all of the previous studies on gestures and intonation (refer to Section 2.3, pg. 48-57) only investigated natural speech within a small group of people, it seemed that in the context of a public speech in front of a large audience, one also needs to pay attention to the pitch accents that are aligned with the post-stroke hold phrase.

As there were a significant number of pitch accents which were aligned to the post-stroke hold, this gestural phrase may play an important role in public speaking. Thus,

one also needs to consider the role the post-stroke hold phrase plays in a public speech. When the speaker held the same gesture at the end of the stroke, it could mean that the speaker wanted to reinforce a point that has been made earlier in the speech. For instance, when Jhingran told the audience that the answer to his application to MIT was inside the envelope he was holding, he lifted the letter to shoulder level with his right hand to show it to the audience (stroke phrase). However, he maintained the gesture he made earlier for a few moments before he put his right hand down which is the post-stroke hold phrase. When he used the same utterance, "*The answer was inside*," he made the same gesture of lifting his right hand to the level of his shoulder again. He also maintained the same gesture for a few moments again before he put his right hand down (post-stroke hold stage).

There were also pitch accents which occurred within the stroke and post-stroke hold phrase of those particular gestures mentioned in the previous paragraph. All of these pitch accents were attached to the same lexical items in both utterances and they were *The*, *answer* and *inside*. In fact, most of the pitch accents consisted of the same type and they occurred at almost similar locations within those words mentioned in the last paragraph. Figure 5.9 shows the moment when Jhingran told the audience that the answer was inside his letter and Figure 5.10 is a screenshot where Jhingran used the same utterance again together with the same gestures. The utterances are circled and the pitch accents, which were attached to the same words, are indicated by arrows in both screenshots.

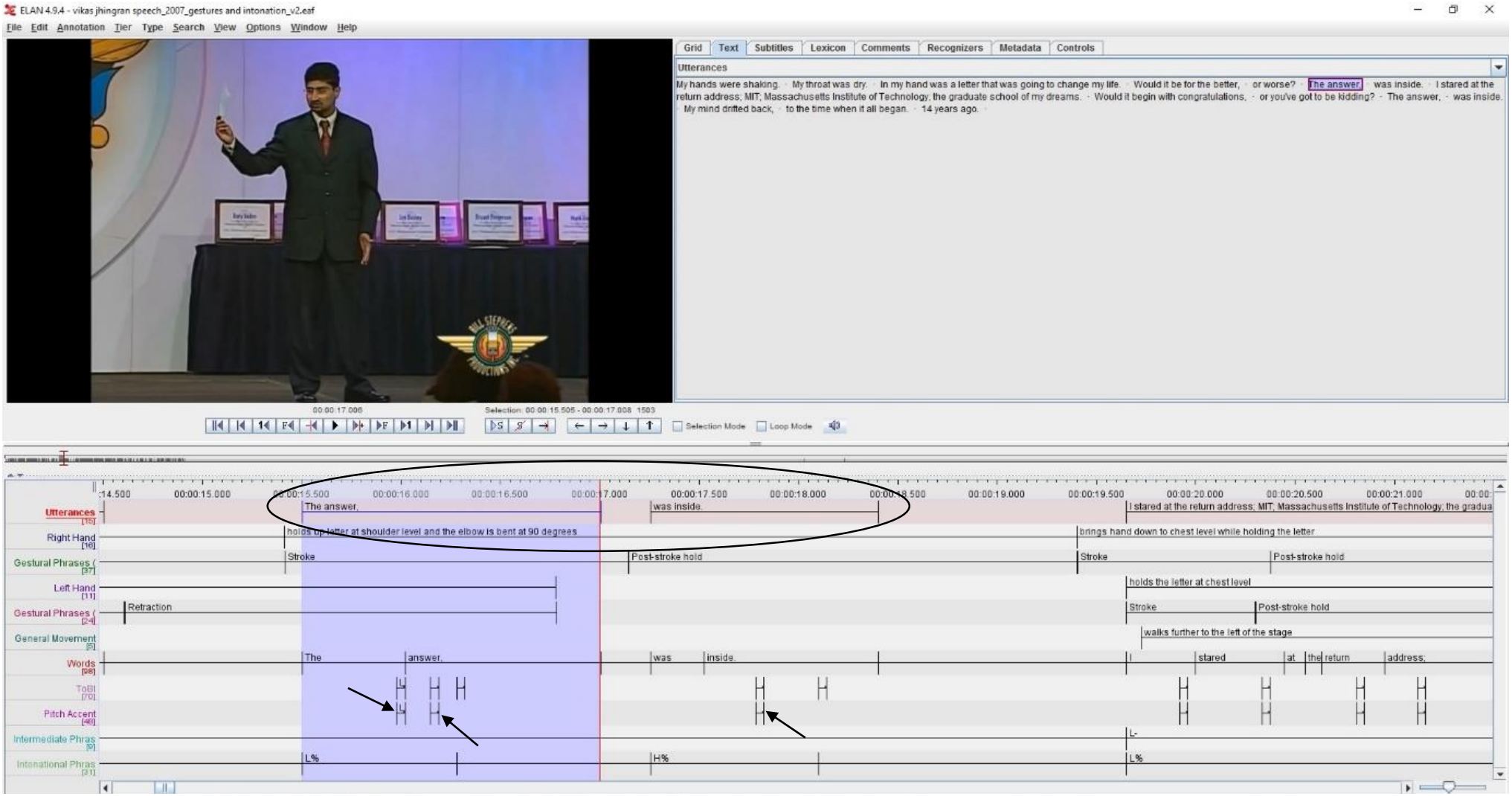


Figure 5.9: Jhingran holding up the letter for the first time.

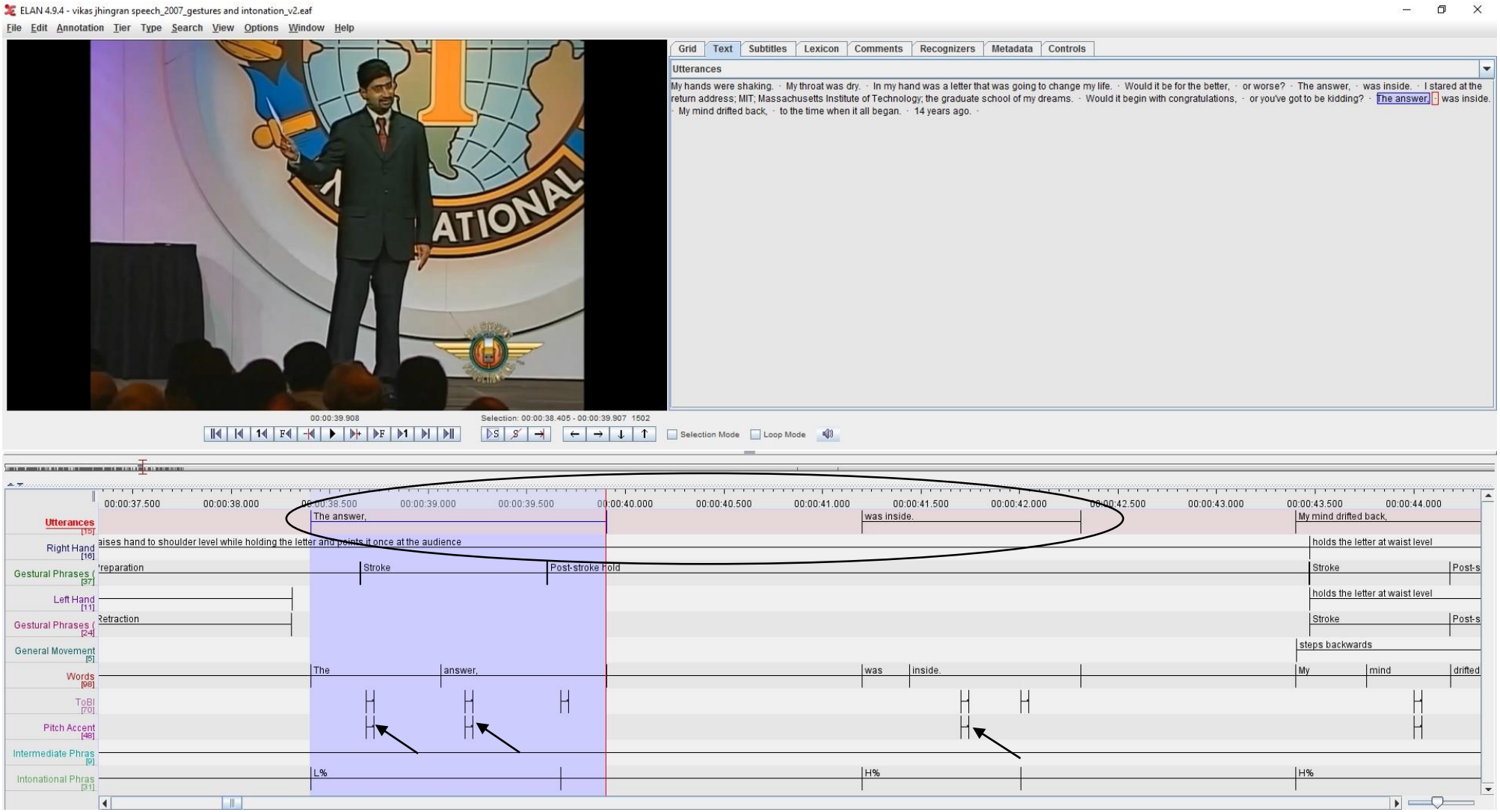


Figure 5.10: Jhingran holding up the letter for the second time.

This act of maintaining a gesture after it has been made has hardly been mentioned in any of the previous studies on gesture and intonation in natural speech. However, the current study found that this happened quite frequently in all of the prepared speech segments. This could be because a speaker needs to attract and maintain the attention of a large crowd of people. Hence, the gestures that are made must be clear enough for the audience to see them (Toastmasters International, 2011b). Apart from that, gestures and intonation can also be used together for dramatic effect to hold the attention of the audience. When a gesture is maintained for a few moments, it serves as a tool to reinforce a point that was introduced to the audience earlier and to stress the importance of that particular idea. When a few pitch accents occur when the gesture is maintained i.e. during the post-stroke hold, it increases the effectiveness of that gesture and it sends a stronger signal to the audience of the significance of its co-accompanying verbal utterance.

In contrast, the act of maintaining a gesture is not important in natural speech so the pitch accents rarely occur during the post-stroke hold phrase, that is if the post-stroke hold phrase is even used during a spontaneous conversation. This is because the speaker does not need to hold a gesture as it is visible to a small group of people the moment it is executed. Moreover, the act of holding a gesture is usually inappropriate within a small group of people as the speaker is not doing a staged performance for their listeners unlike a public speech which is not only a way of communicating to a large group of people but it is also a rehearsed performance by the speaker.

It was shown earlier that gesture and intonation share the same roles in enhancing a public speech and they can also work together to mark a prominent part of the speech. In addition, they were also used to reinforce the message of a speech. However, one thing which needs to be noted is that although gesture is always accompanied by speech,

a person can also speak without gesturing (Mendoza-Denton & Jannedy, 2011). The study found that this also happened with gestures and pitch accents which confirmed Mendoza-Denton and Jannedy's (2011) findings. Even though the stroke and post-stroke hold phrases were mostly aligned with pitch accents, the same could not be said about pitch accents. The study found that close to a quarter (22.8%, right hand) to almost a third (31%, left hand) of the total number of pitch accents that were annotated occurred when the speakers did not gesture with their hands (refer to Table 4.21, pg. 136). This seemed to indicate that pitch accents do not need to occur with a gesture in a speech.

On example of this could be seen when Avery told the audience that his wife asked him the most important question of his life. During that part of his speech, he did not make any gestures with his hands and they were in a neutral position at his sides. However, he made three words prominent within that utterance by attaching a pitch accent to them. These words were *important*, *question* and *life*. Therefore, it seems that a speaker only needs to use pitch accents to draw the attention of the audience to important information in his or her speech. Moreover, if a speaker gestures excessively in a speech, be it prepared or spontaneous, it would be very difficult for the listener to pay attention let alone understand the message the speaker is trying to convey to him or her. Moreover, gesturing excessively in public speaking is also considered inappropriate as speaker speaking in public should always ensure that every gesture is made with a purpose (Toastmasters International, 2011a). If not, it may result in the audience losing interest in the speaker and what he or she wants to say. Figure 5.11 shows a screenshot of the utterance when Avery did not use any gestures to communicate his message to the audience. The pitch accents that were attached to the words mentioned earlier are indicated by arrows.

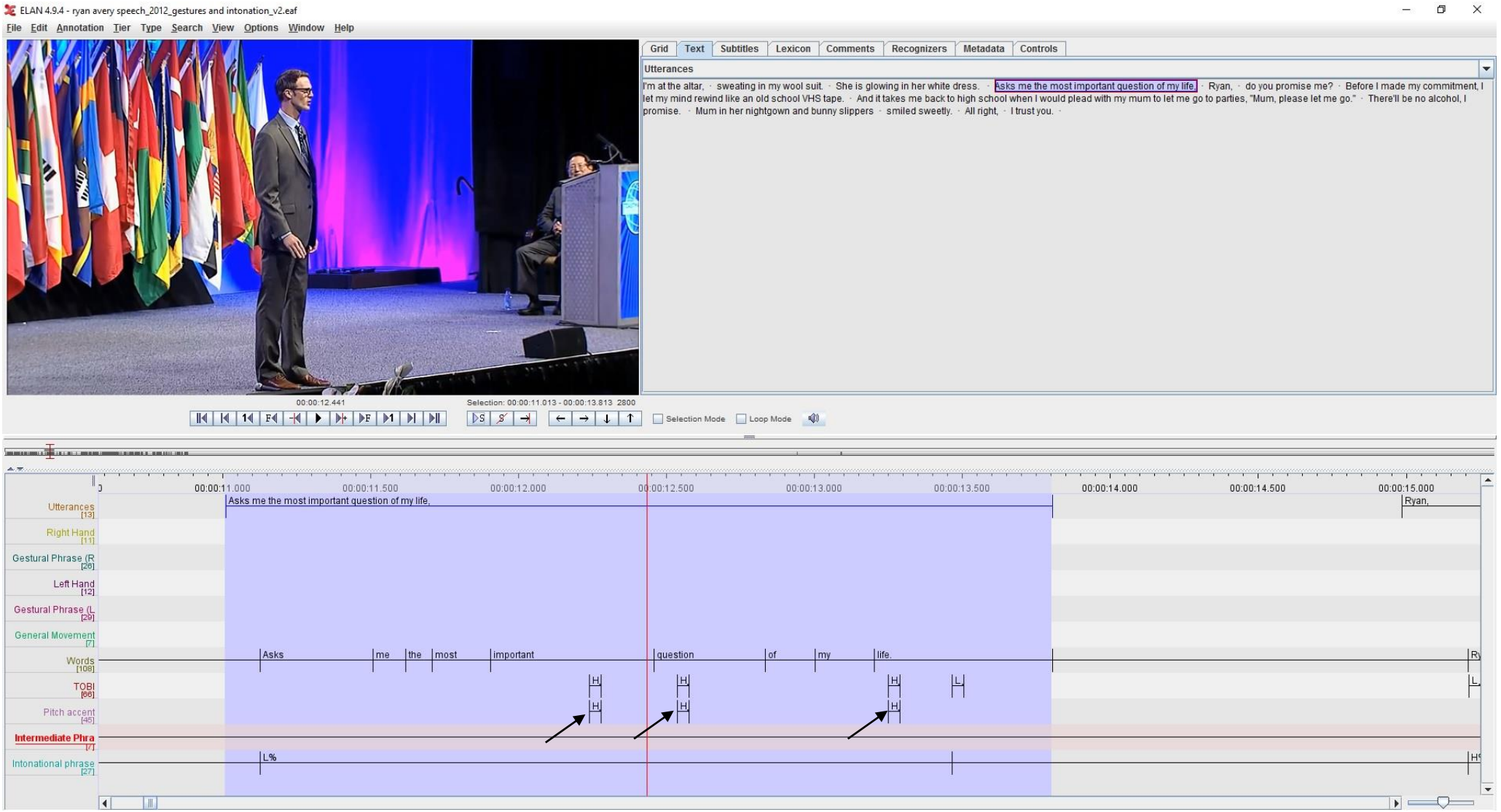


Figure 5.11: Avery using only intonation to mark prominence.

The data of the study has shown that gestures and intonation share a relationship which is consistent with the findings of many previous studies (McClave, 1991, Loehr, 2004, Mendoza-Denton & Jannedy, 2011). This can be seen especially when they are used simultaneously to mark the prominent parts of a speech and to reinforce a message which the speaker has already mentioned to the audience at an earlier. At the same time, the study also expanded on these findings as well. It was mentioned earlier that they found that the strokes of a gesture tend to coincide with a pitch accent or in the context of McClave's (1991) study, a stressed syllable. This study found that pitch accents did not only occur during the stroke of gesture but also during the post-stroke hold phrase. This showed that the act of maintaining the same gesture during a speech could be just as important as the execution of the gesture itself. Apart from that, the data in the study also showed that pitch accent can occur without a gesture but the reverse rarely happens. This raises the next question which needs to be answered in the next section.

5.3 Which of them appears to have a bigger influence on a public speech?

The earlier section explored the extent of the relationship between gesture and intonation in a prepared speech. At a first glance, the data analysed in the study seems to show that intonation may have a bigger influence on a prepared speech compared to gestures. This is due to the number of pitch accents which occurred without any hand gestures in all of the four speech segments. As mentioned in the previous section (refer to pg. 163), there were 42 pitch accents (out of 136) or 31% which were not attached to any gestures made by the speakers' left hands. On the other hand, 31 pitch accents out of 136 or 22.8% also were not attached to any gestural phrases of the speakers' right hands.

It was also mentioned earlier that a pitch accent does not need to be aligned with a gestural phrase (refer to Section 5.2, pg. 163). This is because gestures normally cannot exist in isolation without being accompanied by speech as it would very difficult for the audience to understand what the speaker wants to say. For example, if one were to look at the gestures at the beginning of Henderson's speech segment in isolation i.e. without any sound, they would not be able to guess that he used both of his hands to represent two pilots when he raised them to his waist. They had to watch his gestures and listen to the words in his utterance to know that his hands represented the two pilots. In contrast, the pitch accents in that utterance occurred at the words *1983*, *two*, *teamed* and *the* which marked them as important. In fact, if this were to be done without any gestures, the audience would still be able to know that the two best pilots in Texas teamed up to fight the Red Baron in 1983 based on his verbal utterance. This was the same in Avery's speech segment (refer to Figure 5.11, pg. 164) where the audience still understood him even though he did not use any gestures to tell them that his wife asked him the most important question of his life. Figure 5.12 shows a screenshot of the moment when Henderson raised his hands to represent the two pilots in his speech. The pitch accents in that utterance are marked with arrows and their co-accompanying words are circled in Figure 5.13 which is a close-up of the screenshot in Figure 5.12.

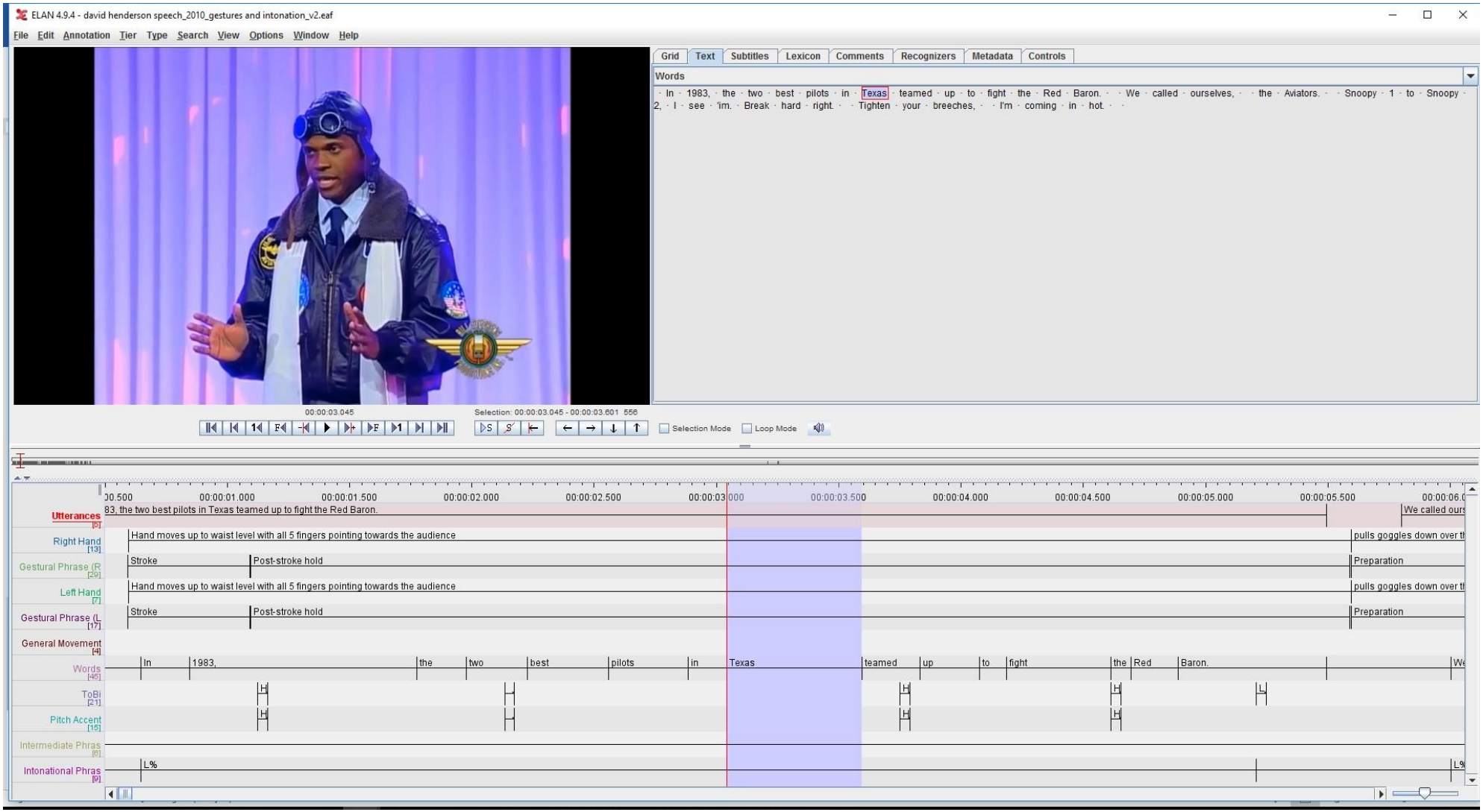


Figure 5.12: Henderson using his hands to represent the two best pilots in Texas

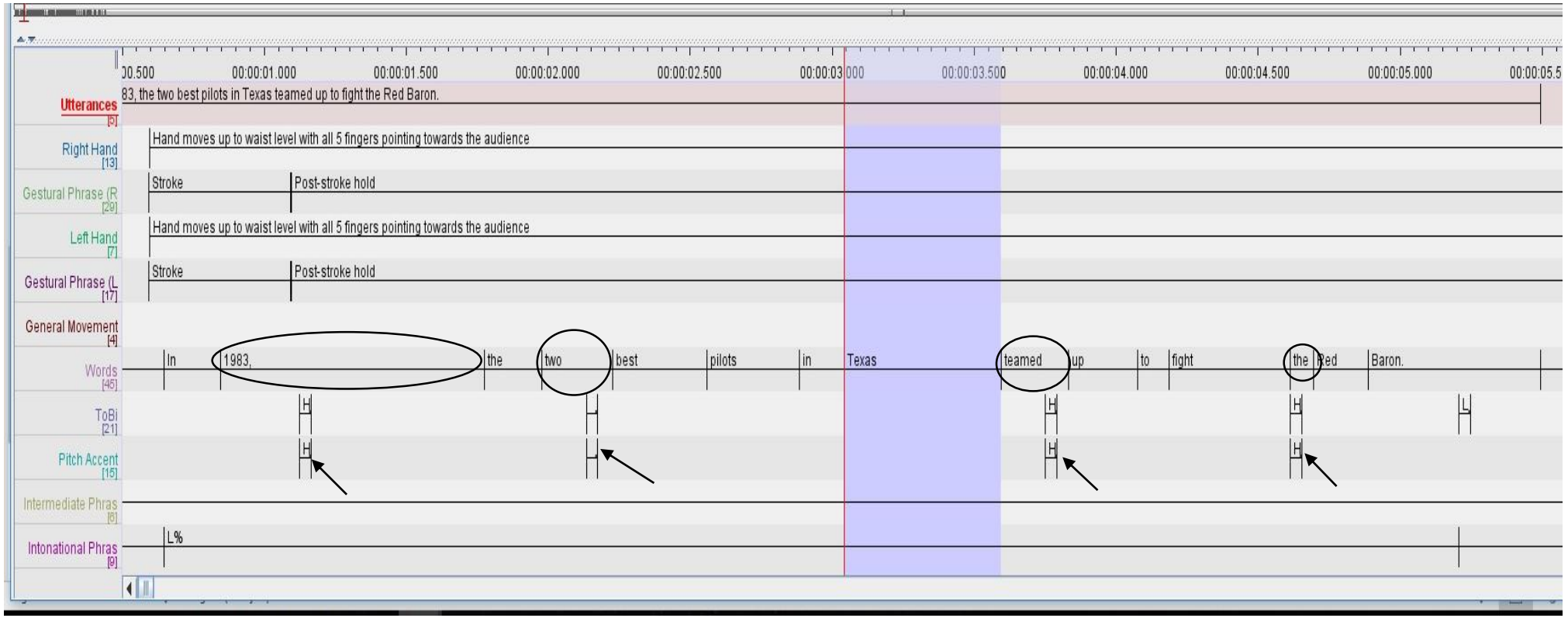


Figure 5.13: A close-up of the screenshot and its gestures and pitch accents (Henderson).

Another finding which seem to indicate that intonation has a bigger influence than gesture in a prepared speech is because of the type of lexical items they were attached to in the speech segments. Each speech segment consisted of numerous content and function words. The earlier section (refer to Section 5.1, pg. 154) mentioned that the pitch accents were mainly aligned with the content words in each segment. As previous studies have shown that content words carry the most information in a speech (Zheng & Pierrehumbert, 2010) so it would only make sense that more content words should also be aligned with the important gestural phrases in all of the speech segments but this was not the case.

It was found during the course of the study that most of the words, whether they were content or function words, tended to occur with the stroke of a gesture (refer to Table 4.24, pg. 140). In addition, it was only the stroke that was generally aligned to more content words compared to function words. Even then, they numbered only slightly higher (52.6%, left hand and 51.8%, right hand) than the words which were aligned with the function words in all of the speech segments. Furthermore, the post-stroke hold phrase and even the words which were not aligned to any gestural phrases tended to have more function words attached to them compared to the content words.

Although this may seem like a contradiction at first, it was actually not the case when the gestural phrases were studied closely. This was because the study annotated each gestural phrase as an interval of time instead of identifying the 'apex' of a stroke like in Loehr's (2004) study where he tried to annotate the 'peak of the peak' of a gesture i.e. its 'exact' moment. Therefore, when the gestural phrases in each speech segment were analysed, there were many cases where a gestural phrase was aligned with more than one word. This is because the speaker was speaking as well as gesturing at the same time. One example can be seen in Avery's speech segment when he told the

audience about one of his memories in high school. As he was talking, he moved both of his hands below his waist which was the stroke phrase. This gestural phrase alone was aligned with four function words and two content words. This finding showed that intonation seemed to be a clearer indication of word prominence compared to gestures in a prepared speech as a pitch accent is able to pinpoint the exact moment when a syllable is stressed compared to a gestural phrase which may have more than one word attached to it. Figure 5.14 shows the utterance in Avery's speech segment which contained the stroke and the six lexical items mentioned earlier. Figure 5.15 is a close-up of the screenshot in Figure 5.14 and the strokes of both hands are circled while the six words are indicated by arrows.

However, there were also other findings and examples in the study which may seem to suggest that gestures and pitch accents are interdependent on each other and even play an equal role in enhancing a prepared speech. When new or important information was presented to the audience, the speaker would change the pitch of his or her voice to get their attention and to show that the piece of information is a prominent part of the speech. In addition to the speaker's pitch accent, he or she would gesture to further show the audience that they should pay attention to this idea and remember it. At the same time, it is also acknowledged the speaker does not need gestures to highlight important and new information in a speech but when an idea needs to be reinforced in a speech, using pitch accents on their own may not be sufficient to achieve the desired result. This is because pitch accents are only able to show the audience that a piece of information or idea is important and they should pay attention to it but it does not clearly show whether an idea is being reinforced or introduced. Moreover, it is not able to provide an extensive range of extra information compared to a gesture as it is limited by the words and grammar of a speech (Mendoza-Denton & Jannedy, 2011).

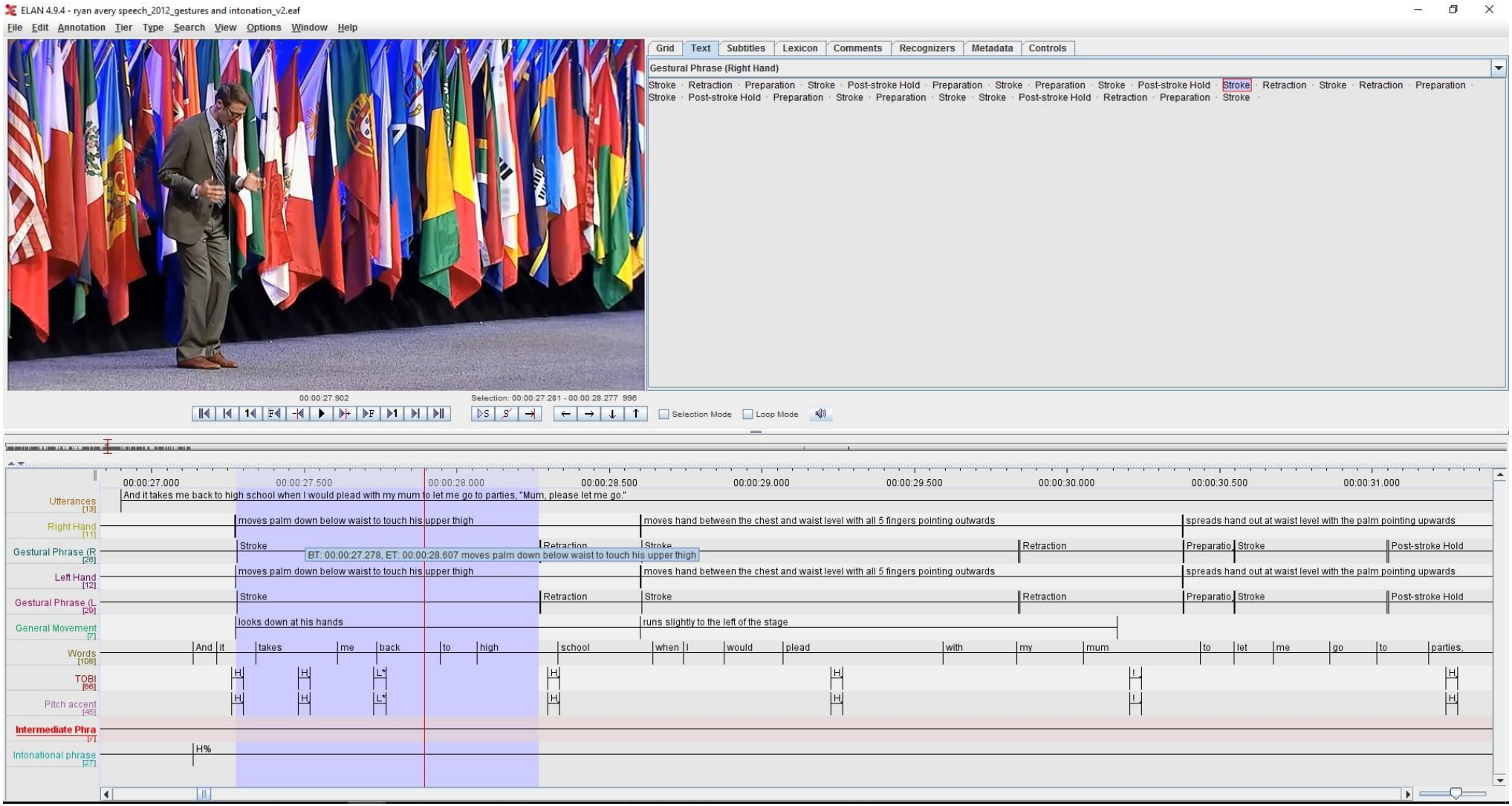


Figure 5.14: Avery's gesture which contained six lexical items.

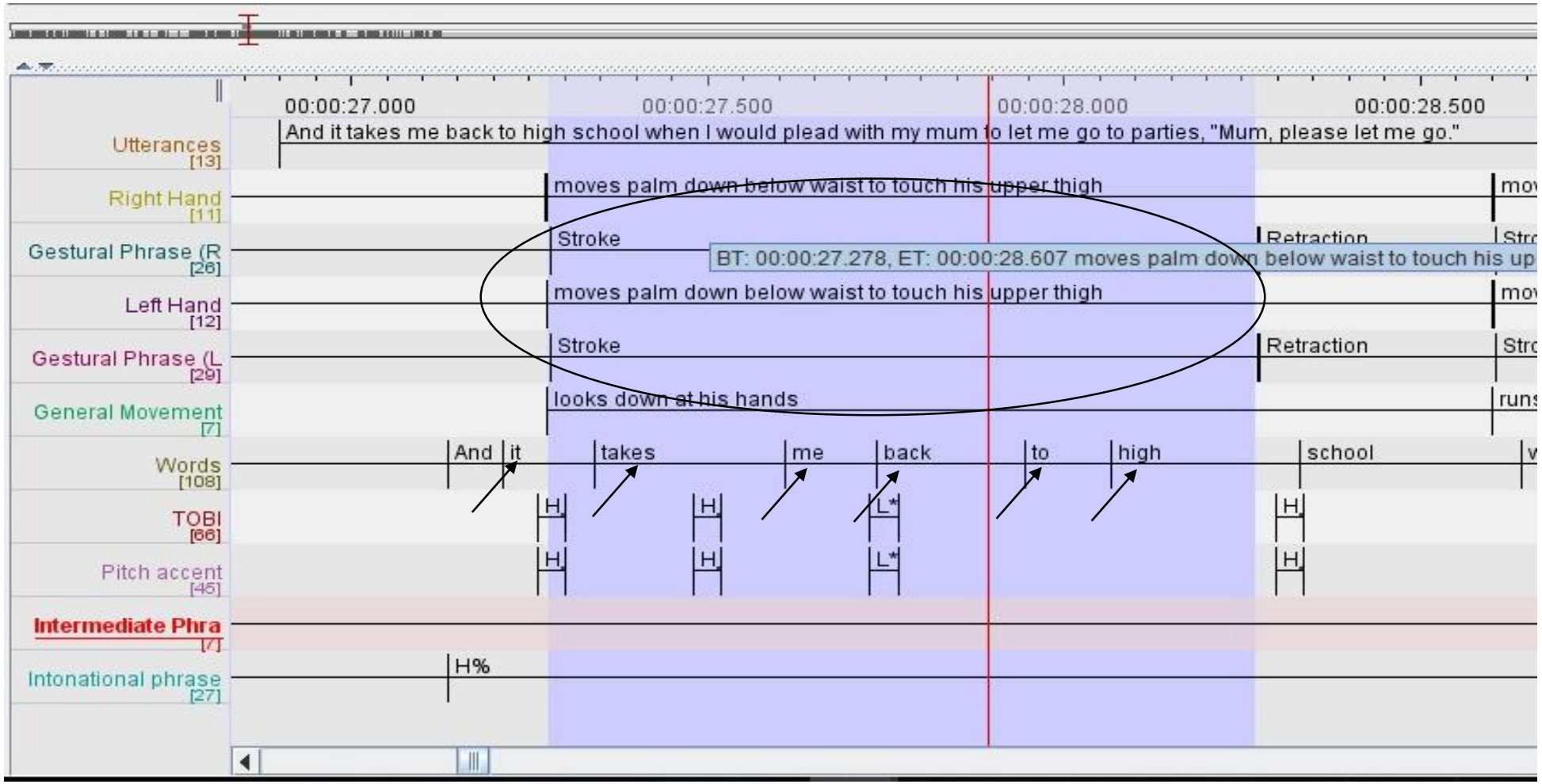


Figure 5.15: A close-up of the strokes and lexical items.

In Figures 5.5 and 5.6, Harvey changed the pitch of his voice to show that he was afraid of the hounds and he also indicated his fear by marking certain lexical items in his verbal utterance with pitch accents. On the other hand, he also used the same gesture to show that his heart and later his whole body jumped in fear, reinforcing the idea of how fearful he was of the hounds. In fact, by utilising both gestures and intonation, he also showed the audience that he grew increasingly fearful without explicitly mentioning it in his verbal utterance. This shows that gestures and intonation in a public speech needs to be investigated further but they should be studied together as they work hand in hand to communicate an idea to an audience.

5.4 Other Theoretical Implications

The question of what makes a 'good' and 'charismatic' speaker has been explored by researchers in the past decade (refer to Section 2.2.1, pg. 38-42). From researchers like Rosenberg and Hirschberg (2005) to Biadsy, Rosenberg, Carlson, Hirschberg and Strangert (2008), their studies found that listeners tend to see a speaker as 'good' or 'charismatic' if he or she shares these characteristics. Firstly, a speaker is seen as 'good' or 'charismatic if he or she has a varied pitch and intensity when they speak. Strangert and Gustafson (2008) also found that a speaker was considered 'good' if they had a high pitch accent on the important words in their speech. In addition, the data from these studies also showed that speakers who used more first-person pronouns in their speech segments tend to be considered a 'good' or 'charismatic' speaker. Furthermore, their respondents also felt that a 'good' or 'charismatic' speaker should have little to no mistakes in their speeches such as hesitations, pause fillers, self-repairs and repetitions which are not purposeful. However, the previous studies could not agree on whether

the speed at which a speaker spoke had any significant influence on whether a speaker is seen as ‘good’ or ‘charismatic’.

As mentioned in the Literature Review (refer to Section 2.2.1, pg. 39-40), many previous studies had several limitations because they did not have an objective method to select speakers that could be considered genuinely ‘good’ or ‘charismatic’. The current study addressed this issue by only analysing speech segments taken from speakers who had won the World Championship of Public Speaking which was organised by Toastmasters International. These speakers could be considered ‘good’ and ‘charismatic’ as they had competed and won many rounds of competition from the club level to the world stage. Moreover, it is highly likely that they were judged by different teams of judges as they progressed through each round. In addition, the study only annotated and analysed the introductions, regardless of length, given by each speaker during the tournament. This is to ensure that the format and type of the speech segments which were analysed were as similar to each other as possible so that the accuracy and validity of the findings would not be compromised as Freydina (2015) found that spontaneous and prepared speeches had many different prosodic characteristics between them.

The study found that based on the speech segments of each speaker, they shared many characteristics which the earlier researchers found in speakers who were seen as ‘good’ or ‘charismatic’. Firstly, all of the speakers in the current study varied their pitch as they delivered their speeches. Moreover, they also placed mostly high (H*) pitch accents on the key words in their introduction. This was reflected in the data which showed that they placed pitch accents on the content words (93 out of 126) in their speech segments. In fact, these findings also supported Freydina’s (2015) study which also found that her respondents also marked the lexical items which have a high

semantic value in their speeches with a significant rise in the pitch of their voice i.e. high pitch accent. In addition, all of the speakers did not have any pause fillers, hesitations, self-repairs or repetitions which served no purpose in their speeches. Thus, it can be said that these speeches were likely close to being perfect.

It was mentioned in the earlier paragraph (refer to pg. 173) that speakers who used a lot of first-person pronouns in their speeches could be seen as 'good' or 'charismatic'. However, one issue that could be raised was that none of those studies mentioned the number of first-person pronouns each speaker in their study used. Thus, they did not answer the question of how many first-person pronouns must a speaker use before it could be considered as 'a lot'. This does not provide a neither clear nor objective measurement of how frequent first-person pronouns should appear in a speech for a speaker to be seen as 'good' or 'charismatic'. Furthermore, it seemed that those studies did not take into account that the subject and the genre of the speech may affect the number of the first-person pronouns used by the speakers in their studies.

The current study took this variable into account as it only analysed speakers who only used personal anecdotes as part of their speeches. Although the subject of each speech was different, each speaker shared a personal experience which they lived through as their introduction to grab the attention of the audience. Therefore, they had a reason to use as many first-person pronouns as possible in their speech segments. Despite the fact that these speakers became champions of a very competitive tournament and thus they could be perceived as 'good' and 'charismatic', all of them only used a few first-person pronouns in their speech segments if they are compared alongside the number of words in their speech segments. For example, Avery's speech segment contained 92 words but it only contained 11 first-person pronouns, regardless of whether they were subject or object pronouns. Moreover, only one of these pronouns

had a pitch accent attached to it which seems to show that there were other types of words or information in his speech segment which were considered more important.

The study also looked at the number of first-person possessive pronouns in each speech segment. For instance, Jhingran only used one first-person pronoun, which was aligned to a pitch accent, but his speech segment contained six first-person possessive pronouns. However, only two of these possessive pronouns had a pitch accent attached to them. Therefore, this seems to indicate that the number of first-person pronouns a speaker uses while delivering his or her speech has little to no influence on determining whether a speaker who delivers a prepared speech is considered 'good' or 'charismatic'. Table 5.1 shows the number of first-person pronouns (regardless of subject or object pronouns i.e. *I*, *we*, *me* and *us*) and first-person possessive pronouns (regardless of type i.e. *my* and *our*) used by each speaker in their speech segments. The number of these function words which coincided with a pitch accent is also shown in the table.

Table 5.1: Total of first-person pronouns and possessive pronouns (All).

Function Words / Speaker	First-Person Pronouns		First-Person Possessive Pronouns	
	Normal	Pitch Accent	Normal	Pitch Accent
Harvey (89 words)	5	3	3	1
Jhingran (79 words)	−	1	4	2
Henderson (38 words)	2	1	−	−
Avery (92 words)	10	1	3	2
Total (298 words)	17	6	10	4

Loehr (2004) mentioned in his study that a pitch accent in a spontaneous conversation is generally aligned with the 'apex' of a gesture. As mentioned in the Literature Review (refer to Section 2.3, pg. 51), the 'apex' is located within the stroke phrase and it is the exact moment when the 'kinetic goal' of the gesture is expressed. However, it was also mentioned that using the 'apex' was not clearly defined in Loehr's (2004) research. The current study also seemed to also lend credence to the fact that locating and analysing the 'apex' from Loehr's (2004) study may not be the best way to investigate the relationship between gesture and intonation.

This is reflected in the findings which showed that apart from the stroke phrase, the post-stroke hold phrase also plays an important role in public speaking as many pitch accents were also aligned to this phrase (refer to Table 4.22, pg. 138). Moreover, the fact that many pitch accents were also not aligned with any gestural phrase should also be taken into account (refer to Table 4.21, pg. 136). Therefore, this shows that the stroke may not be the only important gestural phrase in a prepared speech. This causes the 'apex' to be redundant as it will not enable the 'kinetic goal' of the post-stroke hold phrase to be identified. Moreover, focusing on mainly locating the 'apex' and the pitch accents aligned to them will not answer the question of why there also were many pitch accents which were not attached to any gestural phrases in the speech segments.

The 'apex' according to Loehr (2004) only captured one moment which is the 'peak of the peak' of the stroke. However, the current study contained gestures whose 'peak of the peak' did not lie in one moment but over a duration of time. For example, when Jhingran told the audience that his hands were shaking, the stroke phrase of his right hand covered two verbal utterances. In fact, identifying the 'kinetic goal' or the 'peak of the peak' of this gestural phrase will be impossible because his hands were shivering at a constant and in an identical manner. Figure 5.16 shows a screenshot where

Jhingran's hands were shivering as he held a letter he had received from the Massachusetts Institute of Technology. Figure 5.17 is a close-up of the earlier screenshot where the stroke phrase of his right hand is circled and the two utterances i.e. intonational phrases are indicated by arrows.

Thus, the only way to identify the 'apex' of this gesture was to analyse it over a duration of time and how it was related to the pitch accents and lexical items attached to it. Moreover, when a speaker uses one gesture to convey two ideas (refer to Figure 5.6), the 'peak of the peak' cannot be identified as the gesture will be maintained which brings it into the post-stroke hold phrase. This is reflected in the fact that there were many lexical items which overlapped more than one gestural phrase and the majority of them (24 out of 54, refer to Table 4.25, pg. 141) overlapped the stroke and post-stroke hold phrase. This again shows that the 'kinetic goal' of a gesture cannot be identified by narrowing it down into one exact moment but it must be viewed as a phrase which moves over a period of time.

Previous studies like McClave (1991) found that the gestural phrases do not only share a relationship with the pitch accents but also the intonational phrases as well. McClave (1991) found that the stroke phrases of iconic, metaphoric and deictic gestures generally do not exceed the intonational phrase boundary of its verbal utterance in the context of a natural speech. In fact, she even hypothesised that the end of a verbal utterance might also act as a boundary for a gesture. This boundary may also serve to ensure that a gesture which follows a stroke or the stroke itself will also not extend over its intonational boundary.

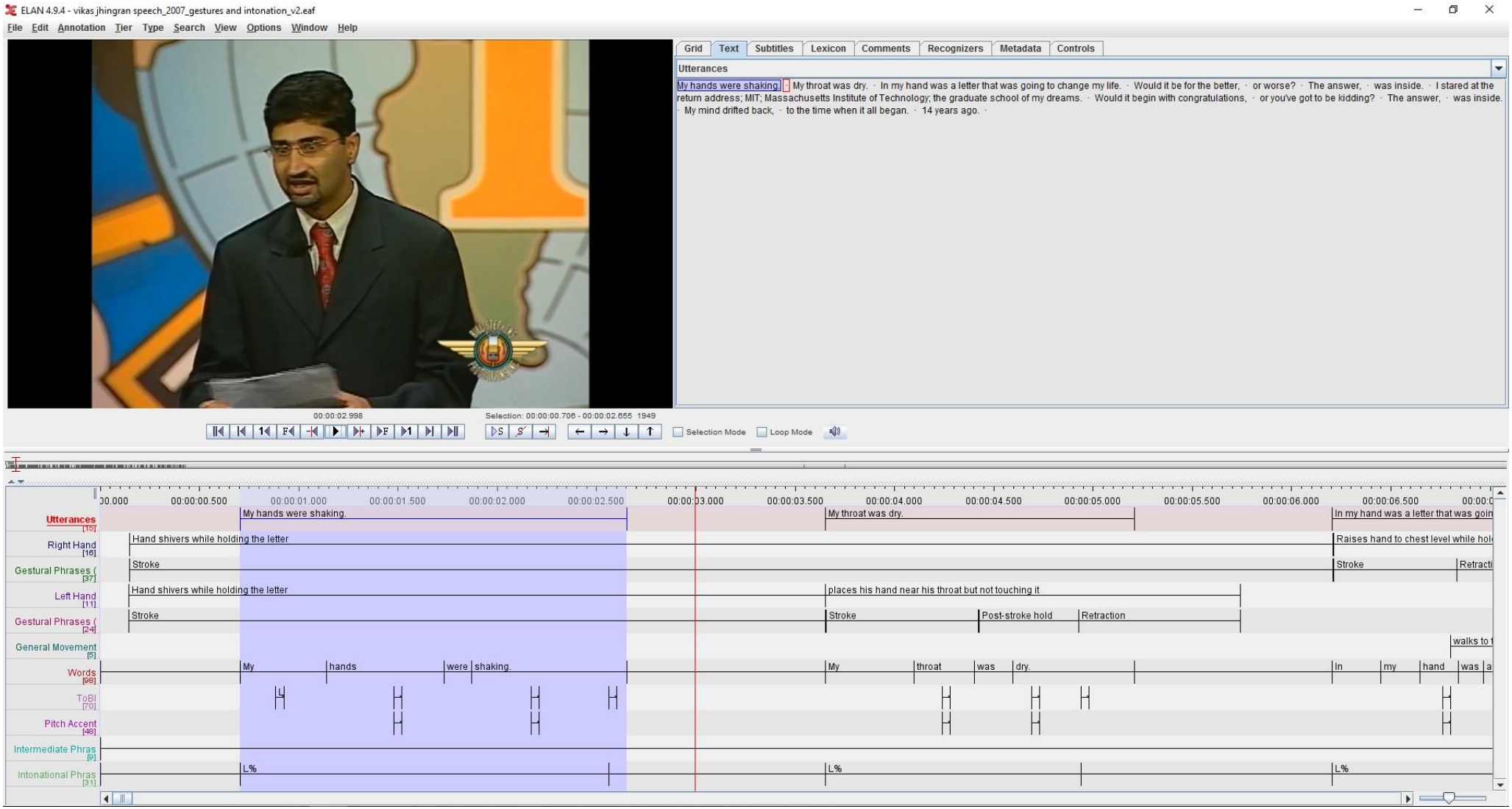


Figure 5.16: Jhingran's hands shivering continuously.

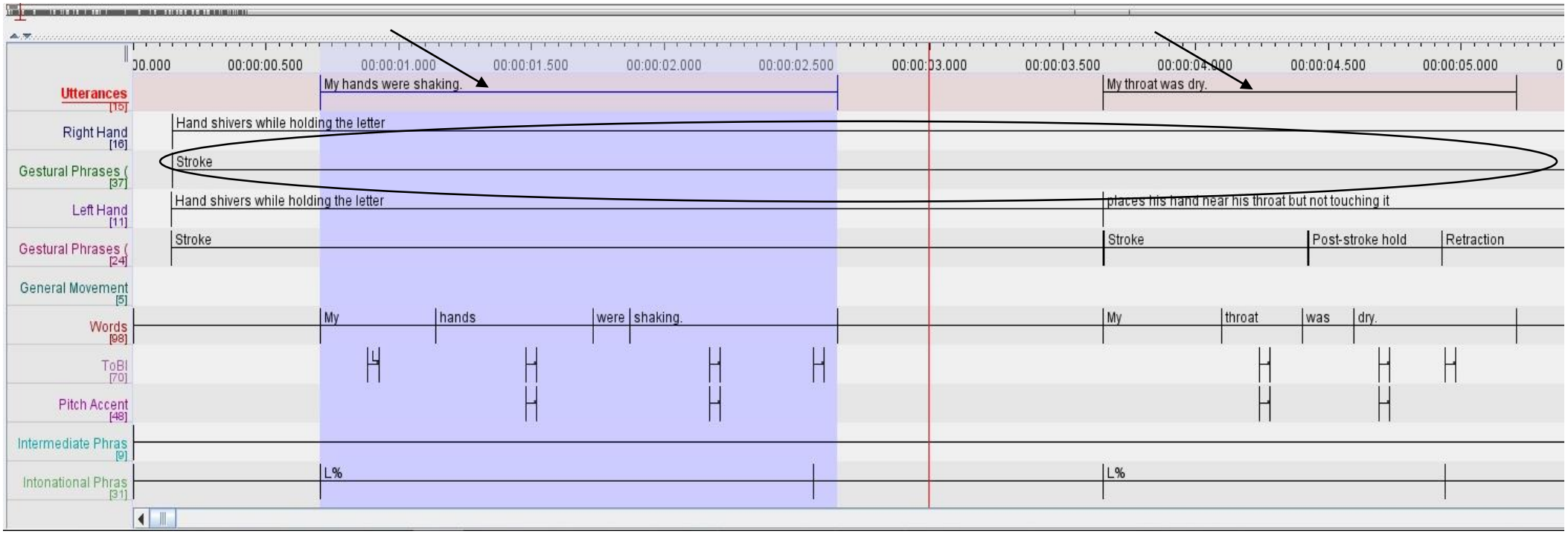


Figure 5.17: The stroke phrase (right hand) spanning two utterances.

However, the findings in the study seem to indicate that McClave's (1991) hypothesis (refer to Section 5.4, pg. 178) might not be applicable in the case of a prepared speech. The current study analysed the representative gestures in all of the speech segments which consists of iconic, metaphoric and deictic gestures. It was found that McClave's (1991) findings were only reflected in Harvey and Avery's speech segments where the number of stroke phrases which did not exceed the intonational phrases in their speech segments outnumbered the ones that did. For example, there was a total of 11 strokes performed by Avery's right hand and 8 of them did not exceed the intonational phrase they were found in, regardless of direction. It was also the same case with his left hand where 8 out of 12 strokes did not exceed the intonational phrase they coincided with. Figure 5.18 is a close-up of the annotation tiers in Avery's speech segment which shows a stroke phrase located within its intonational phrase. On the other hand, Figures 5.19 and 5.20 show two examples of stroke phrases in Avery's speech segment which exceeded their intonational phrase boundaries. Both screenshots are close-ups of Avery's annotation tiers as well. Figure 5.19 shows a stroke which continued to be performed after its intonational phrase had been completed while Figure 5.20 shows a stroke which preceded its intonational phrase as the gesture began to be performed before its verbal utterance. The intonational phrases are circled whereas the strokes are marked with an arrow.

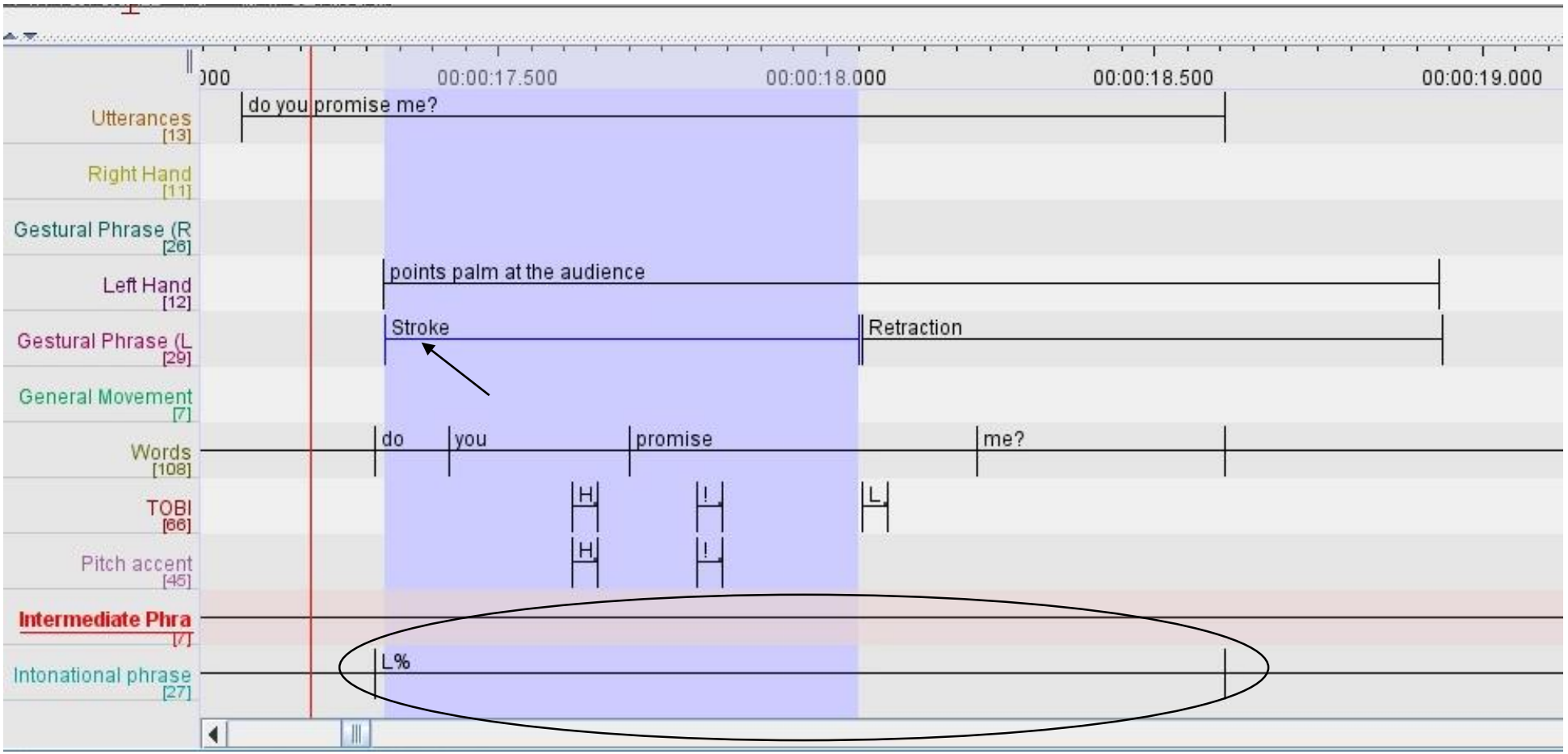


Figure 5.18: A stroke within its intonational phrase (Avery).

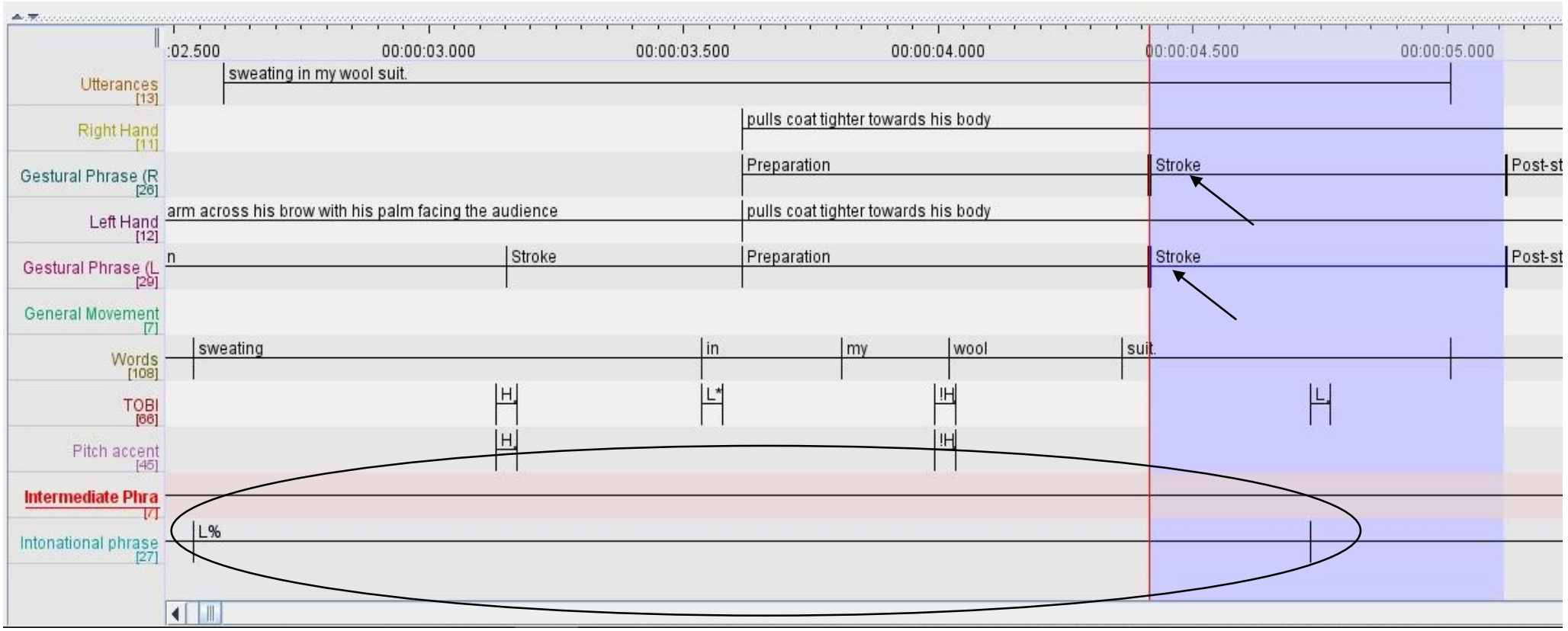


Figure 5.19: A stroke which exceeded its intonational phrase (Avery).

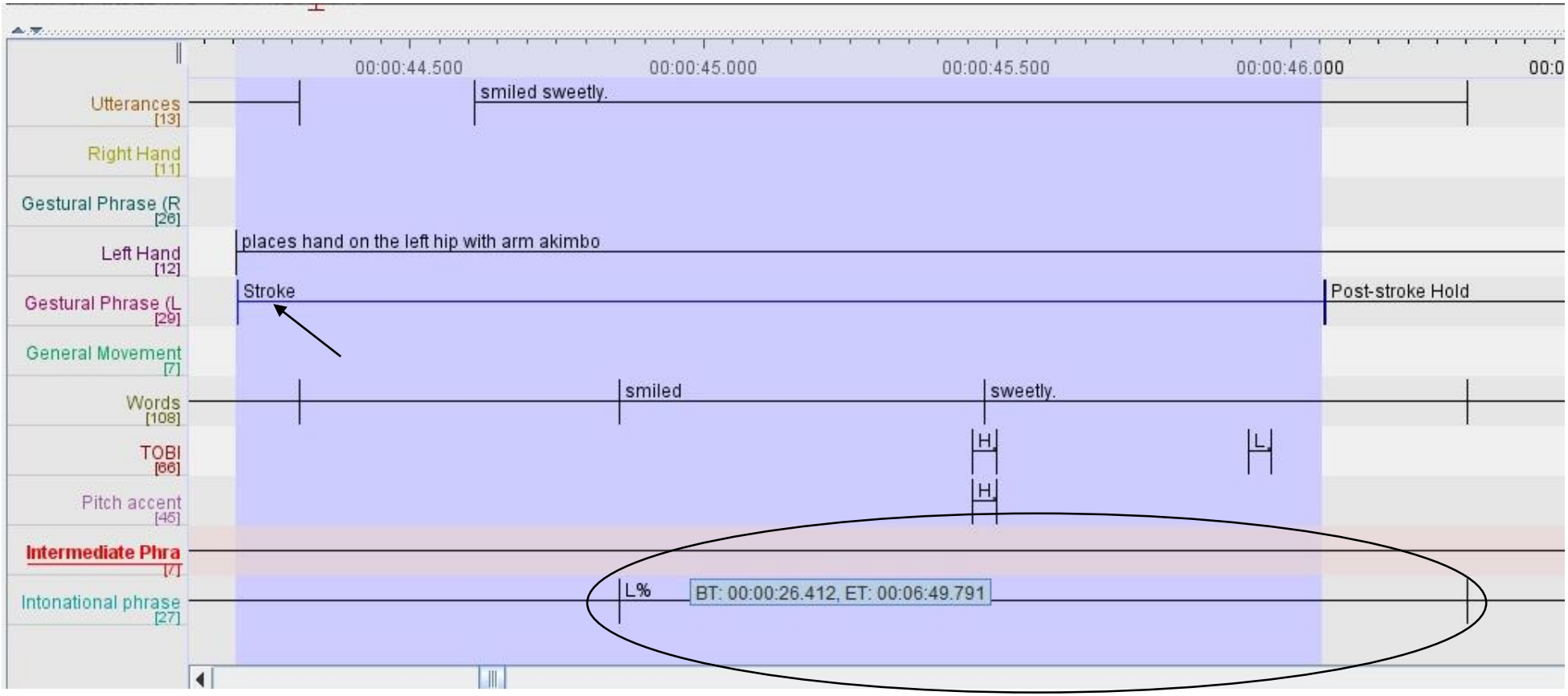


Figure 5.20: A stroke which preceded its intonational phrase (Avery).

In contrast, McClave's (1991) hypothesis (refer to Section 5.4, pg. 178) did not apply to Jhingran and Henderson's speech segments. For instance, 4 out of the 5 strokes performed by Henderson's left hand exceeded the intonational phrase it was located in whereas 5 out of 9 of the strokes performed by his right hand also exceeded its intonational phrase, regardless of their direction. However, in Jhingran's case, it was only the number of strokes performed by his left hand (6 out of 11) which exceeded their intonational phrase.

This finding stood out as it indicated that it was only the strokes of the Caucasian speakers which shared the same characteristics as McClave's (1991) findings despite the fact that Henderson also spoke English as his first language and all of the speakers went through the same training in Toastmasters International. This shows that a speaker's cultural background may still have a certain degree of influence on his or her speech despite he or she going through a standardised training system and thus may warrant further investigation. Table 5.2 shows the number of stroke phrases which crossed over their intonational phrase boundaries in all of the four speech segments, regardless of their direction. There were also a few strokes which were not taken into account in Harvey and Henderson's speech segments as they did not occur with any meaningful words as Harvey barked and Henderson made mechanical sounds during certain portions of their introductions.

Table 5.2: The strokes which exceeded their intonational phrases.

Speakers / Strokes	Harvey		Jhingran		Henderson		Avery	
	Left	Right	Left	Right	Left	Right	Left	Right
Exceeded its intonational phrase	1	2	6	7	4	5	4	3
Located within its intonational phrase	8	7	5	9	1	4	8	8
Total	9	9	11	16	5	9	12	11

In the earlier paragraphs, McClave (1991) hypothesised that a stroke or the gesture that follows it will not extend over its intonational phrase boundary in a natural speech. However, the study has shown that the stroke of a gesture may exceed its intonational phrase boundary from time to time. Moreover, the study also found that a speaker who delivers a prepared speech may sometimes maintain the same gesture after it has been performed so the stroke will transition into the post-stroke hold phrase (refer to Figure 5.20, pg. 184). If this happens while a speaker is speaking, the gesture may likely exceed its intonational phrase boundary.

The study has also found that a stroke which follows an earlier one in a prepared speech may sometimes also exceed its intonational phrase boundary. One example of this can be seen in Jhingran's speech segment where he performs two different gestures with his right hand i.e. two strokes within the same verbal utterance. Both strokes were almost the same in length and the second one exceeded the intonational phrase where both strokes were located. The first stroke was 1.103 seconds long while the second stroke in the utterance was 1.099 seconds long. This shows that this part of McClave's (1991) theory may not be applicable to a prepared speech as a speaker may not shorten

his or her stroke so that their gestures would fit into an intonational phrase. In fact, one should expect that the gestures within a prepared speech to cross over its intonational phrase boundary or a gesture would be performed just before the speaker begins to speak. Figure 5.21 shows a screenshot where both gestures i.e. strokes were annotated within their intonational phrase. Figure 5.22 shows a close-up of the earlier screenshot. The first stroke and its co-accompanying gesture are indicated by arrows while the following stroke and its gesture are circled. The intonation phrase where both gestures are located is underlined. Figure 5.23 shows multiple screenshots of Jhingran performing two different gestures i.e. strokes with his right hand within the same intonational phrase even though his head may move in a different direction. The screenshots are numbered and their time stamps are indicated by arrows.

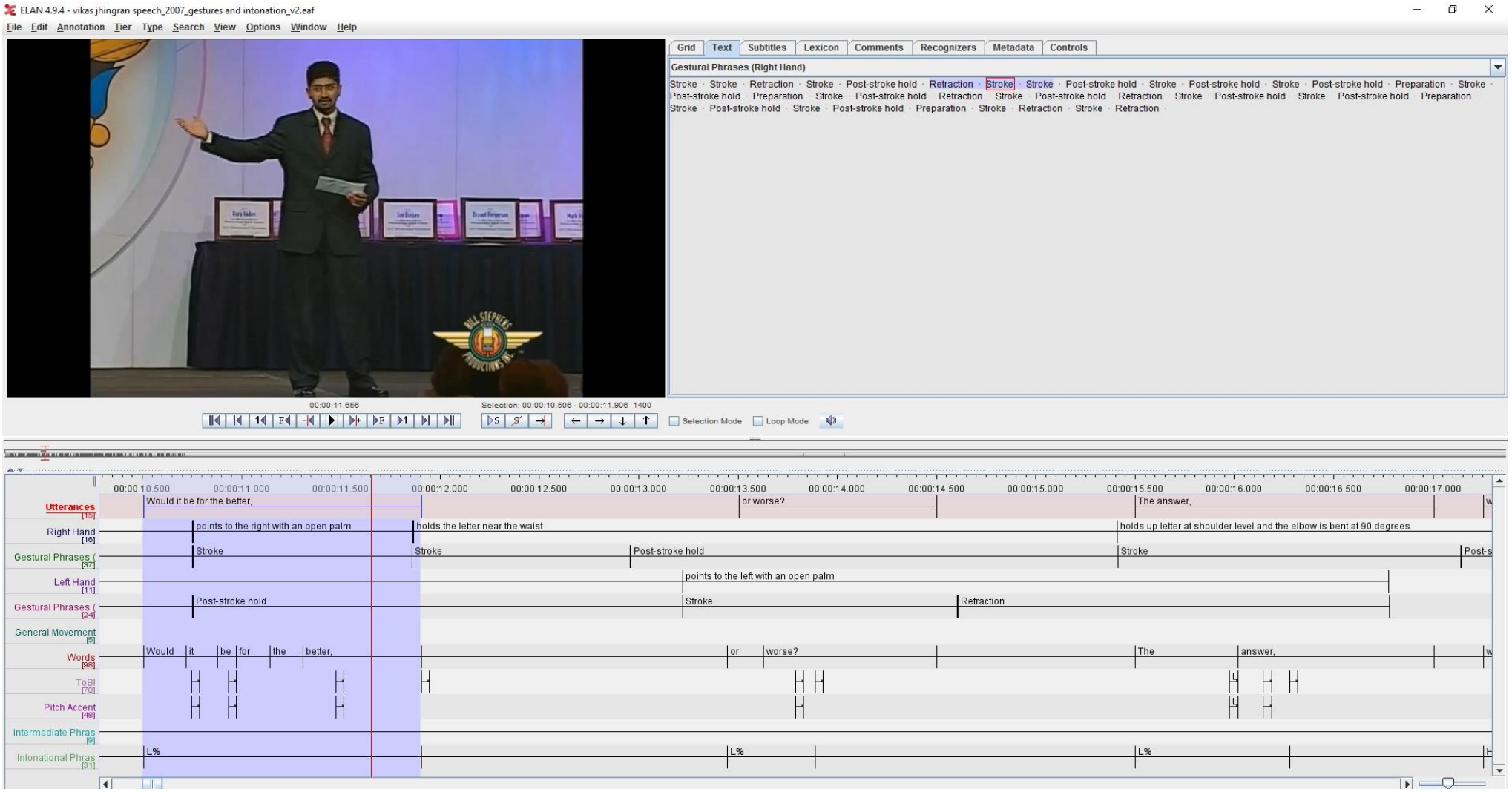


Figure 5.21: Jhingran performing two different gestures in the same utterance.

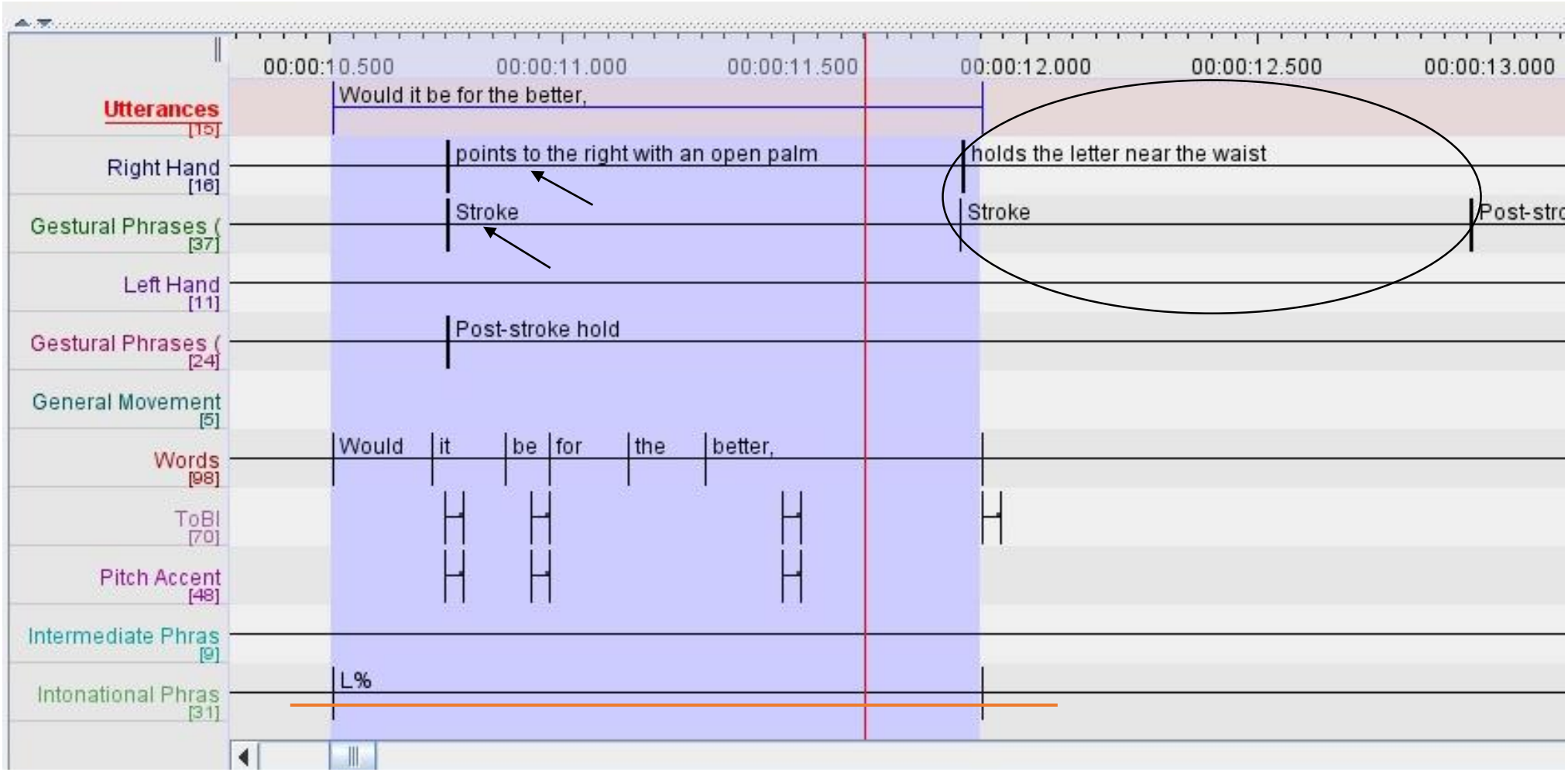
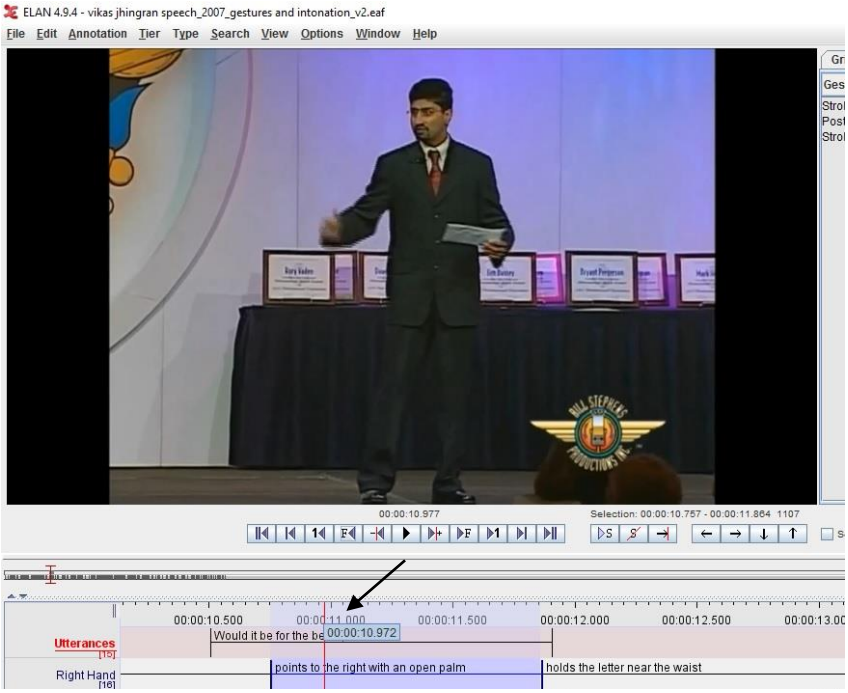


Figure 5.22: The close-up of the annotation tiers with the two gestures (Jhingran).

1



2

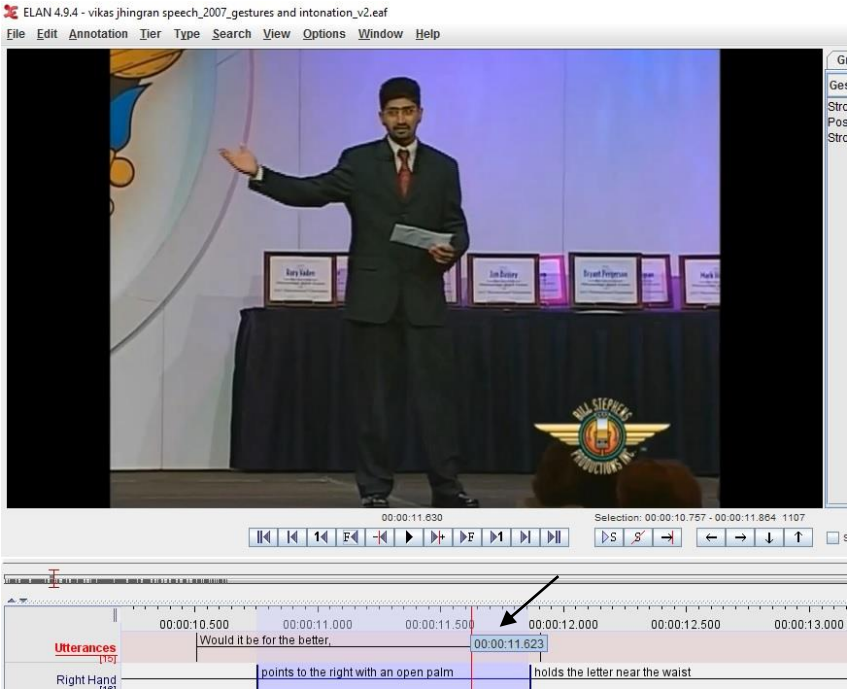
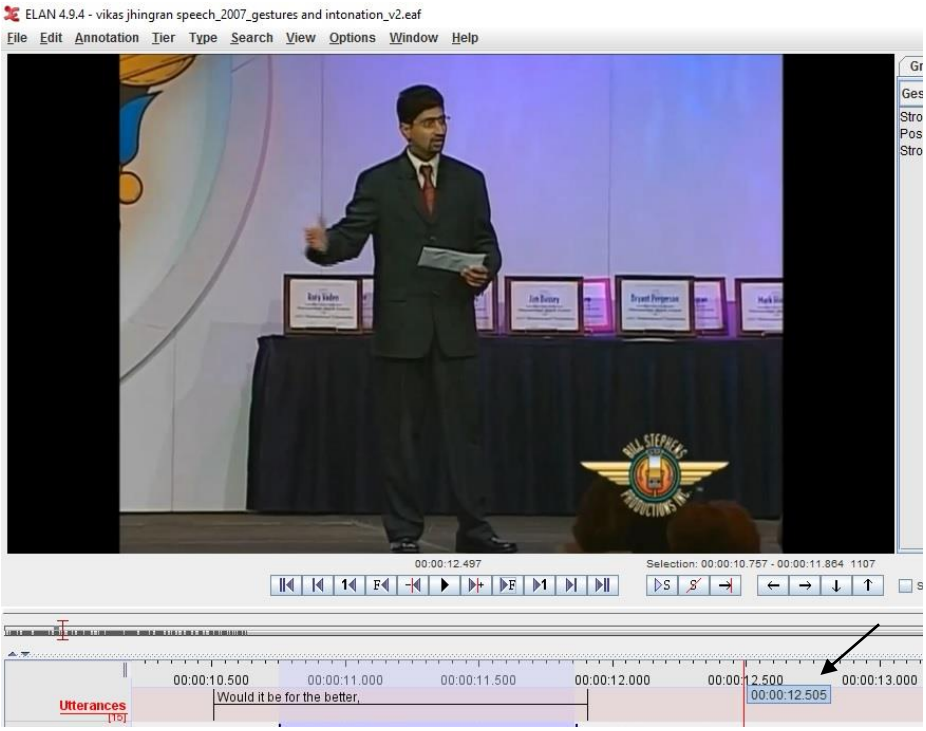


Figure 5.23: Jhingran's two strokes in the same intonational phrase.

3



4

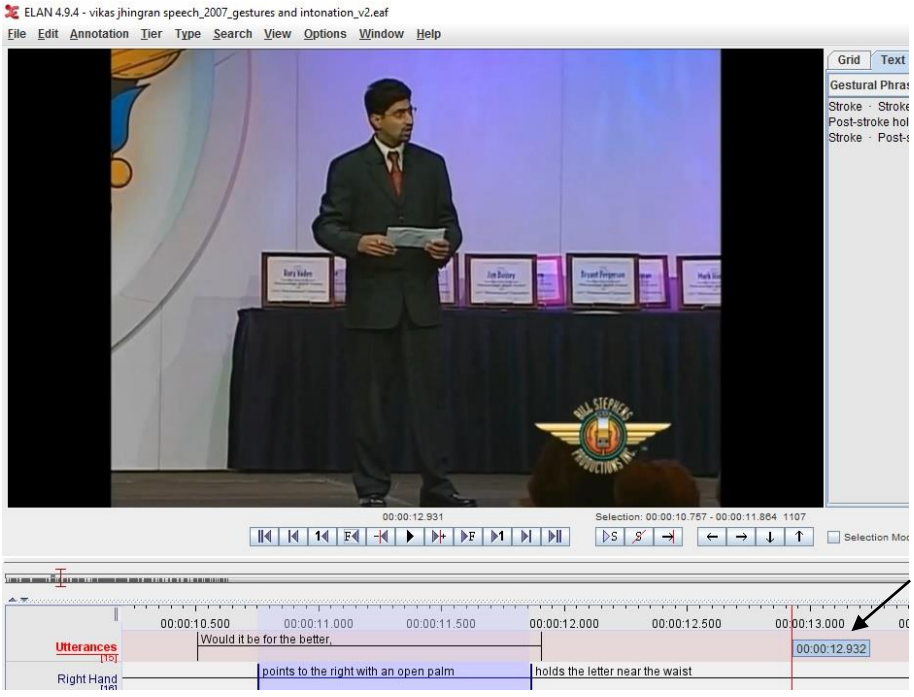


Figure 5.23: Jhingran's two strokes in the same intonational phrase (continued).

CHAPTER 6: CONCLUSION

This chapter provides a summary on the study on the relationship between gestures and intonation in public speaking. The summary will also review the questions asked at the beginning of this study and the findings it arrived at. This chapter will conclude with a discussion on future work which can be conducted following the completion of this study.

6.1 Summary

This dissertation aims to investigate the relationship between gestures and intonation in the context of public speaking. This is because the nature of this relationship has not been fully understood even though numerous studies have done in these fields in the past. One of the main reasons is that many of these studies seemed to focus on studying gestures and intonation separately. In addition, they mostly concentrated on natural, spontaneous conversations between small groups of people or only between the participant and the researcher.

On the other hand, in their quest to define what made a 'good' or 'charismatic' speaker, researchers in the past also conducted studies in the field of public speaking. Once again, they mainly studied gestures and intonation separately and many of these studies which tried to define a 'good' and 'charismatic' speaker seemed to mainly focus on the speaker's intonation. Furthermore, the majority of studies on public speaking in the past seemed to focus more on analysing the content and the words used in speeches.

To date, there has been very few studies which investigated the role of gesture and intonation in public speaking. In fact, the studies on gesture and intonation in speech only seemed to increase in the 21st century due to the advancements in technology which enabled researchers to conduct video and speech annotations i.e. microanalysis of

conversations and speeches. However, although these studies have managed to establish the roles gesture and intonation play in enhancing the quality of a speech and how they benefit the listeners, the extent of this relationship is still not fully comprehended and the question on who has the bigger influence on speech has yet to be answered. In addition, the scope of these studies may be limited as they mainly studied natural and spontaneous conversations between groups of friends of the same gender.

With all these issues in mind, three questions were chosen to be answered during the course of this study. This study sought to find out what the roles of gestures and intonation in a prepared speech are. It also wanted to know the extent of the relationship between gestures and pitch accents in a prepared speech and whether gesture or intonation has the bigger influence in enhancing a prepared speech.

It was found that gestures play many roles in speech such as they are used to describe people, places, events and experiences. In the context of a public speech, they are also used to reinforce an idea the speaker had mentioned earlier. In addition, they also contain additional information which is not mentioned in the speaker's verbal utterances. On the other hand, intonation in a prepared speech is used to mark prominent lexical items in the speech so the audience will be able to take note of the key information in the speech. However, it should also be noted that gestures can also be used to mark prominence to a certain extent. Furthermore, intonation is also used by the speaker to convey his or her emotions to the audience.

Previous studies have also established that there is a relationship between gestures and intonation but the extent of this relationship has yet to be fully determined. They have found that pitch accents are consistently aligned with the stroke phrases of a gesture in a spontaneous conversation. However, this study has found that in a prepared speech, pitch accents do not only consistently occur within the stroke of a gesture but they also frequently appear within the post-stroke hold phrase of a gesture. This is because post-

stroke hold phrases are used to reinforce a point which a speaker has made earlier in his or her speech. This study also confirms previous work that gestures and intonation are synchronised and they work together to communicate meaning to the listener, be it a natural or public speech. On the other hand, many pitch accents in all of the speech segments also did not coincide with any gestural phrases which brings up the matter of who has the bigger influence in enhancing a prepared speech.

The study seems to indicate that intonation may have a greater influence on enhancing the quality of a prepared speech. This could be due to the frequency of the pitch accents occurring without a gestural phrase. Moreover, it has been established that pitch accents can occur with their co-accompanying verbal utterances without any gestures attached to them but the reverse rarely, if never, happens. Furthermore, the findings also showed that the pitch accents in all of the speech segments were mostly attached to the content words of each segment. On the other hand, the number of strokes and post-stroke holds which coincided with the content and function words were almost equal in number, which seemed to suggest that intonation plays a bigger role in marking the important words of a speech. However, some of the findings within the study seem to suggest that gestures and intonation are equally important and they are interdependent. This is because although intonation can be used as a clear way to mark prominent words in a prepared speech, gestures may be able to give the listener more information on why a particular utterance or lexical item is important to the speech. Thus, more study will be needed to shed more light on the relationship between gestures and intonation in public speaking.

The findings in this study also lent credence and at the same time questioned the findings of other studies on gesture and intonation in speech. It showed that the speech segments in the study shared many characteristics with other ‘good’ and ‘charismatic’ speeches from previous studies such as Rosenberg and Hirschberg (2005) and Biadsy, Rosenberg, Carlson, Hirschberg and Strangert (2008). In addition, the study also

addressed several of the limitations which those studies faced by selecting only to analyse the introductions presented by speakers who had won the Toastmasters World Championship of Public Speaking. Furthermore, unlike the previous studies mentioned in this paragraph, the findings indicated that first-person pronouns do not have a large influence on whether a speech is deemed 'good' or 'charismatic'.

The study also suggested that annotating the 'apex' of the stroke (Loehr, 2004) may not be the best way to investigate the relationship between gesture and intonation in a prepared speech. This is because Loehr's (2004) 'apex' of a stroke cannot account for many of the findings in the current study. For example, the study that a message conveyed by a gesture does not necessarily boil down to one moment within the stroke but the speaker may use an entire gestural phrase to communicate his or her message to the audience. In addition, the findings have also shown that the stroke is not the only gestural phrase that is deemed important in the context of a public speech.

Finally, it also showed that McClave's (1991) hypothesis (refer to Section 5.4, pg. 178) on the location of the strokes and their respective intonational phrases may still warrant further investigation. This is because her hypothesis was only reflected in Harvey and Avery's speech segments despite all four speakers going through the same training. Therefore, the question of how much a speaker's cultural background may affect his or her prepared speech will still need to be addressed in the future. Furthermore, the findings of the study have also indicated that a speaker delivering a prepared speech may not necessarily shorten his or her strokes so that they will fit into their intonational phrase boundary.

6.2 Future Studies

Although the current study has managed to answer several questions on the extent of the relationship between gesture and intonation in a prepared speech, it also found there are still many issues which need to be addressed. It was found that one of the main challenges faced by researchers in this field is the amount of data which was analysed at any one time. This was reflected in previous studies (refer to Section 2.2.1, pg. 39-40 & Section 3.2, pg. 72) where the researchers were only able to annotate and analyse extracts from various parts of a speech or conversation. Although the current study only analysed the introduction of a public speech due to the likelihood of it having the greatest influence on the audience, it is proposed that further studies should be conducted on the relationship between gestures and intonation in the body and conclusion of a prepared speech given in public to see if it would yield findings which are similar or otherwise to the ones in this study. This is because a speaker may use more gestures to make the introduction of their speech more dramatic to capture the attention of their audience. Thus, further study on the different parts of a speech is needed in order to get a more complete picture on how gestures and intonation work together to enhance a public speech.

Another thing which can be further investigated is how different types of gestures are related to intonation in public speaking. As representational gestures refer to the gestures which a speaker makes to describe the objects, events and experiences in his or her speech, it is still a combination of iconic, metaphoric and deictic gestures. Although the study has pointed out that certain gestures may fall into more than one of these categories, it may be still possible to isolate other gestures which only fall into one category i.e. iconic, metaphoric and deictic gestures, and investigate their relationship with intonation. However, it must be ensured that there is a clear and objective guideline to help determine the criteria of an iconic, metaphoric or deictic gesture to ensure that there will be little to no dispute on the category a gesture falls into.

The previous studies which investigated the prosodic features in a public speech also stated that the intensity i.e. volume of a speech may have a certain degree of influence on whether a speech is seen as 'good'. Thus, future studies can also annotate this feature alongside the gestures and pitch accents of a speech in order to determine the extent of the influence intensity has on the quality of a public speech. This is because the previous studies (refer to Section 2.2.1, pg. 41) only mentioned in passing the role of intensity in speech which showed that this prosodic feature may not have been investigated fully.

The study also found that although all of the four speakers went through the same training in Toastmasters International, only the Caucasian speakers (Harvey and Avery) had strokes which generally did not exceed their intonational phrases, as per McClave's (1991) hypothesis, compared to Jhingran, who is a native of India, and Henderson, who is African-American. This shows that it is possible that their cultural backgrounds still may have influenced the findings to a certain degree even though the study found that the speakers' training in Toastmasters had likely overwritten most of the influence from their different cultures. Thus, future studies on the relationship between gesture and intonation in public speaking may also need to consider the cultural backgrounds of their speakers too.

In fact, as most studies in these two fields such as Loehr (2004) mainly analysed speakers who spoke American English as their first language, future studies could also focus on speakers who do not speak English as their first language and also those who come from a different culture or country. This is because speakers from different cultures and language backgrounds may use gesture and intonation differently when delivering a speech. Hence, when these speakers deliver a prepared speech in front of an audience, the dynamics between gestures and intonation may change and thus affect the relationship between them. If this happens, it may yield a different finding which in turn will provide

a better understanding on how gesture and intonation across different cultures work together to enhance a prepared speech.

When public speaking is mentioned, it seems one tends to mostly associate this art with political speeches. Hence, it could explain why most of the studies done on gestures and intonation in public speaking mainly focused on speeches delivered by politicians from various countries. It was only recently that researchers began to study the various characteristics of different types of public speeches from TED Talks to academic lectures. Thus, in order to attain a better understanding of this art, the current study focused on analysing the gestures and intonation of public speeches delivered in a tournament organised by Toastmasters International. However, there are still many genres of public speeches such as storytelling to delivering an after-dinner speech which have not been studied in detail. Thus, one may find that the gestural and prosodic features of those speeches may be similar or even different from each other due to their different methods of delivery. In fact, even the message the speaker wants to convey to his or her audience might affect the gestures and intonation he or she uses during the delivery of their speech.

In addition to influencing the type of gestures used by a speaker, the message of the speech may also influence the hand which a speaker chooses to gesture with. This hypothesis was formed during Casasanto and Jasmin's (2010) study where they analysed the gestures made by four US presidential candidates. They found that all of them, regardless of political affiliation, tended to gesture more with their dominant hands, be it left or right, when they were speaking about a positive topic and vice versa. Thus, the researchers hypothesised that a speaker may unconsciously associate positive things with his or her dominant hand and negative ones with their non-dominant hand.

However, there are still some questions this hypothesis has left unanswered which could be investigated in the future. Firstly, it is likely that the speakers analysed may have been coached on how to speak and gesture before the debate. Moreover, as a debate

normally consists of presenting a prepared speech and answering impromptu questions from the opponent, there is a possibility that the speakers may have switched between both speech types and they may have been analysed together as the study did not mention the part of the debate which had been analysed. Hence, the hypothesis mentioned in the earlier paragraph may not be very accurate nor valid. Therefore, future studies can be done to validate this hypothesis by analysing spontaneous conversations or impromptu speeches delivered in public. One example of an impromptu speech would be the Table Topics speeches delivered during a Toastmasters meeting or competition. During the Table Topics session, a speaker would be given a topic and he or she would have to deliver a two-minute speech on the spot without any time given for preparation. Thus, compared to the study mentioned in the earlier paragraph, it would be more likely that the gestures made by the speakers would be as natural as possible.

Due to the fact that gestures and intonation varies from culture to culture, further study would be needed to gain a deeper understanding on how positive and negative speech topics affect the gestures and intonation of a speaker. Hence, the hypothesis mentioned on pg. 198 can be further examined by analysing speeches delivered by orators who come from a different linguistic and cultural background. This is because certain cultures such as Malaysian culture are not as expressive as other cultures in Europe and the Americas. In addition, one also has to take the genre of the speech into account as well and also ensure that spontaneous and prepared speeches do not overlap each other during the analysis.

One can also extend the current study to analyse how the message in a prepared speech would influence the prosodic features of a speaker. Although it has been established that prosody is used to mark prominence in a speech, there is still very little work done to investigate how a positive or negative topic in a speech would affect the pitch and

intensity of the speaker. In fact, another question which can be studied further is whether prominence would be marked differently based on the topic of a speech.

Although the technology of the day has enabled studies on gesture and intonation in public speaking to be conducted with precision, it would still require more than a few studies before one can say an adequate understanding of these two elements has been attained. This is because the areas of human language and behaviour are so vast and varied that despite the numerous studies which have been done to investigate and analyse them, there are still so many areas left to explore. In addition, gestures and intonation are one of the main nonverbal means of communication and they share a very unique and complex relationship. Thus, a further study of these two elements will provide us with a glimpse or maybe more into the heart and mind of a speaker, be it an orator or just your average joe and jane; having a conversation with their friends.

REFERENCES

Beattie, G., Webster, K. A., & Ross, J. A. D. (2014). Do Speakers Really Unconsciously and Imagistically Gesture About What Is Important When They Are Telling a Story? *Semiotica, 202,* 41-79. doi: 10.1515/sem-2014-0033

Beckman, M. E., & Elam, G. A. (1993). *Guidelines for ToBI Labelling*. Retrieved from http://www.ling.ohio-state.edu/research/phonetics/E_ToBI/

Biadsy, F., Rosenberg, A., Carlson. R., Hirschberg, J., & Strangert, E. (2008). A Cross-Cultural Comparison of American, Palestinian and Swedish Perception of Charismatic Speech. In Barbosa, P. A., Madureira, S., & Reis, C. (Eds.), *Proceedings of the Speech Prosody 2008 Conference*. (pp. 579-582). Campinas, Brazil: Editora RG/CNPq. Retrieved from http://www1.cs.columbia.edu/nlp/papers/2008/fadi_al_08b.pdf

Bloomfield, L. (1933). *Language*. London, UK: Compton Printing Ltd.

Boersma, P., & Weenink, D. (2016). *Praat: doing Phonetics by Computer*. Version 6.0.21. [computer software]. Amsterdam, Netherlands. Retrieved from http://www.fon.hum.uva.nl/praat/

Bolinger, D. (1958). A Theory of Pitch Accent in English. *Word, 14*(2-3), 109-149. doi: 10.1080/00437956.1958.11659660.

Bolinger, D. (1983). Intonation and Gesture. *American Speech, 58*(2), 156-174. doi: 10.2307/455326

Breen, M., Dilley, L. C., Kraemer, J., & Gibson, E. (2012). Inter-transcriber Reliability for Two Systems of Prosodic: ToBI (Tones and Break Indices) and RaP (Rhythm and Pitch). *Corpus Linguistics and Linguistic Theory, 8*(2), 277-312. doi: 10.1515/cllt-2012-0011

Brentari, D., Marotta, G., Margherita, I., & Ott, A. (2013). The Interaction of Pitch Accent and Gesture Production in Italian and English. *Studi e Saggi Linguistici*, *51*(1), 79-97. Retrieved from http://clml.uchicago.edu/~jkeane/sll/Papers/BrentariOttMargherita-The%20interaction%20of%20pitch%20accent%20and%20gesture%20production%20in%20Italian%20and%20English%20(2013).pdf

Brown, G. (1983). Prosodic Structure and the Given/New Distinction. In Cutler, A., & Ladd, R. D. (Eds.), *Prosody: Models and Measurements* (67-77). Berlin, Germany: Springer-Verlag.

Brugos, A., Veilleux, N., Breen, M., & Shattuck-Huffnagel, S. (2008). The Alternatives (Alt) Tier for ToBI: Advantages for Capturing Prosodic Ambiguity. In Barbosa, P. A., Madureira, S., & Reis, C. (Eds.), *Proceedings of the Speech Prosody 2008 Conference*. (pp. 273-273). Campinas, Brazil: Editora RG/CNPq. Retrieved from http://isle.illinois.edu/sprosig/sp2008/papers/id072.pdf

Casasanto, D., & Jasmin, K. (2010). Good and Bad in the Hands of Politicians: Spontaneous Gestures during Positive and Negative Speech. *PLoS ONE, 5*(7), 1-5. doi: 10.1371/journal.pone.0011805

Condon, W., & Ogston, W. (1967). A Segmentation of Behaviour. *Journal of Psychiatric Research, 5*(3), 221-235. http://dx.doi.org/10.1016/0022-3956(67)90004-0

Cruttenden, A. (1997). *Intonation* (2nd ed.). Melbourne, Australia: Cambridge University Press.

Driskell, J. E., & Radtke, P. H. (2003). The Effect of Gesture on Speech Production and Comprehension. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 45*, 445-454. doi: 10.1518/hfes.45.3.445.27258

Esposito, A., & Marinaro, M. (2007). What Pauses Can Tell Us about Speech and Gesture Partnership. In Esposito, A., Bratanic, M., Keller, E., & Marinaro, M. (Eds.), *Fundamentals of Verbal and Nonverbal Communication and the Biometric Issue* (pp. 45-57). Amsterdam, Netherlands: IOS Press.

Freydina, E. (2015). Prosodic Variation in Academic Public Presentations. *Journal of Language & Education*, *1*(2), 13-19. Retrieved from http://old-ojs.hse.ru/index.php/JLE/article/viewFile/22906/21820

Goldin-Meadow, S., Nusbaum, H., Kelly, S. D. & Wagner, S. (2001). Explaining Math: Gesturing Lightens the Load. *Psychological Science, 12* (6), 516-521. doi: 10.1111/1467-9280.00395.

Graham, J. A., & Argyle, M. (1975). A Cross-Cultural Study of the Communication of Extra-Verbal Meaning by Gestures. *International Journal of Psychology*, *10*(1), 57-67. doi: 10.1080/00207597508247319

Gut, U., Pillai, S., & Zuraidah Mohd. Don. (2013). The Prosodic Marking of Information Status in Malaysian English. *World Englishes*, *32*(2), 185-197. doi: 10.1111/weng.12018

Gut, U., & Pillai, S. (2014). Prosodic Marking of Information Structure by Malaysian Speakers of English. *Studies in Second Language Acquisition, 36*(2), 283-302. doi: 10.1017/S0272263113000739

Gut, U., & Pillai, S. (2015). The Question Intonation of Malay Speakers in English. In Delais-Roussarie, E., Avanzi, M., & Herment, S. (Eds.), *Prosody and Language in Contact* (pp. 51-70). Dordrecht, Netherlands: Springer-Verlag Berlin Heidelberg.

Hairuliza Idrus & Suzana Meor Abdul Aziz. (2001). Public Speaking: It Can Make a Difference in ESL Students. *The English Teacher, XXX*, 32-44.

Hirschberg, J., & Pierrehumbert, J. (1986, July 10-13). *The Intonational Structuring of Discourse*. Paper presented at the ACL '86: Annual meeting on Association for Computational Linguistics. Stroudsburg, PA: Association for Computational Linguistics. Retrieved from http://www.aclweb.org/anthology/P86-1021

Hirschberg, J., Litman, D., Pierrehumbert, J., & Ward, G. (1987). Intonation and the Intentional Structure of Discourse. In McDermott, J. (Ed.), *IJCAI'87: Proceedings of the 10th International Joint Conference on Artificial Intelligence-Volume 2* (pp. 636-639). San Francisco, CA: Morgan Kaufmann Publishers Inc. Retrieved from http://faculty.wcas.northwestern.edu/~jbp/publications/Hirschberg_etal_IJCAI.pdf

Hirschberg, J. (2002). The Pragmatics of Intonational Meaning. In Bel, B., & Marlien, I. (Eds.), *Proceedings of the Speech Prosody 2002 Conference* (pp. 65-68). Aix-en-Provence, France: Laboratoire Parole et Langage. Retrieved from https://pdfs.semanticscholar.org/25d4/b4965f5faebcb5ea3be6da7ad5930ec6c317.pdf

Hogg, M. A., & Vaughn, G. M. (2005). *Social Psychology* (4th ed.). Gosport, England: Pearson Education.

Holle, H., & Gunter, T. C. (2007). The Role of Iconic Gestures in Speech Disambiguation: ERP Evidence. *Journal of Cognitive Neuroscience, 19*(7), 1175-1192. doi: 10.1162/jocn.2007.19.7.1175

Hostetter, A. B. (2008). *Mind in Motion: The Gesture as Simulated Action Framework.* (Doctoral Dissertation). Available from ProQuest Dissertations and Theses Database. (UMI No. 3327832)

Hostetter, A. B., & Alibali, M. W. (2010). Language Gesture Action! A Test of the Gesture as Simulated Action Framework. *Journal of Memory and Language, 63,* 245-257. doi: 10.1016/j.jml.2010.04.003.

Hostetter, A. B. (2011). When do Gestures Communicate? A Meta-Analysis. *Psychological Bulletin, 137*(2), 297-315. doi: 10.1037/a0022128

Kendon, A. (1972). Some relationships between body motion and speech: An analysis of an example. In Siegman, A., & Pope, B. (Eds.), *Studies in dyadic communication*. (pp. 177-210). New York, NY: Pergamon Press.

Kendon, A. (1995). Gestures as Illocutionary and Discourse Structure Markers in Southern Italian Conversation. *Journal of Pragmatics, 23*(3), 247-279. doi: 10.1016/0378-2166(94)00037-F.

Kendon, A. (2000). Language and gesture: unity or duality? In McNeill, D. (Ed.), *Language and Gesture* (pp. 47-63). Cambridge, England: Cambridge University Press.

Kendon, A. (2004). *Gesture. Visible Action as Utterance*. Cambridge, UK: Cambridge University Press.

Kipp, M. (2016). *Anvil – The Video Annotation Research Tool.* Version 5.1.16. [computer software]. Retrieved from http://www.anvil-software.org/download/index.html

Kita, S. (2000). How Representational Gestures Help Speaking. In McNeill, D. (Ed.), *Language and Gesture*. (pp. 162-185). Cambridge, England: Cambridge University Press.

Lempert, M. (2011). Barack Obama, Being Sharp. Indexical Order in the Pragmatics of Precision-Grip Gesture. *Gesture, 11*(3), 241-270. doi: 10.1075/gest.11.3.01lem

Loehr, D. (2004). *Gesture and Intonation*. (Doctoral dissertation, Georgetown University). Retrieved from http://faculty.georgetown.edu/loehrd/pubs_files/Loehr04.pdf

Loehr, D. (2007). Aspects of Rhythm in Gesture and Speech. *Gesture, 7*(2), 179-214. doi: 10.1075/gest.7.2.04loe

Lucas, S. E. (2012). *The Art of Public Speaking* (11th ed.). Singapore: McGraw-Hill.

Max Planck Institute for Psycholinguistics. (2016). *Elan*. Version 4.9.4. [Computer Software]. Nijmegen, Netherlands. Retrieved from https://tla.mpi.nl/tools/tla-tools/elan/

McClave, E. (1991). *Intonation and Gesture*. (Doctoral Dissertation). Available from ProQuest Dissertations & Theses database. (UMI No. 9219331)

McNeill, D. (1992). *Hand and Mind*. Chicago, IL: Chicago University Press.

McNeill, D. (2005). *Gesture and Thought*. Chicago, IL: Chicago University Press.

Mendoza-Denton, N. & Jannedy, S. (2011). Semiotic Layering through Gesture and Intonation: A Case Study of Complementary and Supplementary Multimodality in Political Speech. *Journal of English Linguistics, 39*(3), 265-299. doi: 10.1177/0075424211405941.

Pierrehumbert, J. (1980). *The Phonology and Phonetics of English Intonation*. (Doctoral Dissertation, Massachusetts Institute of Technology). Retrieved from http://faculty.wcas.northwestern.edu/~jbp/publications/Pierrehumbert_PhD.pdf

Pierrehumbert, J., & Steele, S. A. (1989). Categories of Tonal Alignment in English. *Phonetica, 46,* 181-196. doi: 10.1159/000261842

Pierrehumbert, J., & Hirschberg, J. (1990). The Meaning of Intonational Contours in the Interpretation of Discourse. In Cohen, P., Morgan, J., & Pollack, M. (Eds.), *Intentions in Communication*. (pp. 271-311). Cambridge, MA: Bradford Books, MIT Press.

Rauscher, F. H., Krauss, R. M., & Chen, Y. (1996). Gesture, Speech, and Lexical Access: The Role of Lexical Movements in Speech Production. *Psychological Science, 7*(4), 226-231. doi: 10.1111/j.1467-9280.1996.tb00364.x.

Rosenberg, A., & Hirschberg, J. (2005, September 4-8). *Acoustic/Prosodic and Lexical Correlates of Charismatic Speech*. Paper presented at Interspeech 2005 and the European Conference on Speech Communication and Technology (Eurospeech). Lisbon, Portugal: International Speech Communication Association (ISCA). Retrieved from http://www1.cs.columbia.edu/~amaxwell/pubs/charisma-euro05-final.pdf

Shahrina Md. Nordin & Zullina Hussain Shaari. (2005). Implementing the Toastmasters Tradition in the L2 Classroom. *The English Teacher, XXXIV,* 60-66.

Siddens, P. J. III. (1998). *An Investigation of the Amount of Emphasis Placed on Content and Delivery by Instructors of Public Speaking.* Retrieved from http://files.eric.ed.gov/fulltext/ED420085.pdf

Slavícková, T. (2013). The Rhetoric of Remembrance: Presidential Memorial Day Speeches. *Discourse Society, 24*(3), 361-379. doi: 10.1177/0957926512471762

Strangert, E. (2005, September 4-8). *Prosody in Public Speech: analyses of a news announcement and a political interview*. Paper presented at Interspeech 2005 and the European Conference on Speech Communication and Technology (Eurospeech). Lisbon, Portugal: International Speech Communication Association (ISCA). Retrieved from https://www.researchgate.net/publication/221481197_Prosody_in_public_speech_analyses_of_a_news_announcement_and_a_Political_interview

Strangert, E., & Gustafson, J. (2008, September 22-26). *What Makes a Good Speaker? Subject Ratings, Acoustic Measurements and Perceptual Evaluations*. Paper presented in the Annual Conference of the International Speech Communication Association 2008 (INTERSPEECH 2008). Brisbane, Australia: International Speech Communication Association (ISCA). Retrieved from https://pdfs.semanticscholar.org/a0b0/fd5976af1a1cd15cd793d3a69c6bfcb791c2.pdf

Sun, Y. C. (2008). The Toastmasters Approach: An Innovative Way to Teach Public Speaking to EFL Learners in Taiwan. *RELC Journal, 39*(1), 113-130. doi: 10.1177/0033688208091143

Thornbury, S. (2002). *How to Teach Vocabulary*. Kuala Lumpur, Malaysia: Longman.

Toastmasters International. (2011a). *Competent Communication. A Practical Guide to Becoming a Better Public Speaker.* Mission Viejo, CA: Toastmasters International.

Toastmasters International. (2011b). *Gestures: Your Body Speaks.* Retrieved from https://www.toastmasters.org/~/media/E202D7AA84E24A758D1BAAE8A77FD496.ashx

Toastmasters International. (2011c). *Your Speaking Voice.* Retrieved from https://www.toastmasters.org/~/media/B7D5C3F93FC3439589BCBF5DBF521132.ashx

Toastmasters International. (2016a). *History*. Retrieved from http://www.toastmasters.org/About/History

Toastmasters International. (2016b). *Club Meeting Roles*. Retrieved from https://www.toastmasters.org/Membership/Club-Meeting-Roles

Tsai, T. J. (2015, September 6-10). *Are You TED Talk Material? Comparing Prosody in Professors and TED Speakers*. Paper presented in Speech Beyond Speech: Towards a Better Understanding of the Most Important Biosignal: Annual Conference of the International Speech Communication Association (INTERSPEECH 2015). Dresden, Germany: International Speech Communication Association (ISCA). Retrieved from http://pages.hmc.edu/ttsai/assets/tedVsSpeaker_Interspeech15.pdf

Whitehead III, G. I. & Smith, S. H. (2002). The Use of Hand Gestures and Smiles in the Inaugural Addresses of Presidents of the United States. *The Journal of Social Psychology, 142*(5), 670-672. doi: 10.1080/00224540209603927

Willyard, J., & Ritter, K. (2005). Election 2004 Concession and Victory Speeches: The Influence of Genre, Context, and Speaker on Addresses by Presidential and Vice-Presidential Candidates. *American Behavioural Scientist, 49*(3), 488-509. doi: 10.1177/0002764205279439

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., Sloetjes, H. (2006, May 24-26). *ELAN: a Professional Framework for Multimodality Research*. Paper presented in the LREC 2006, International Conference on Language Resources and Evaluation. Genoa, Italy. Retrieved from http://www.lrec-conf.org/proceedings/lrec2006/pdf/153_pdf.pdf

Yeşil, R. (2008). Evaluation of Body Language Behaviour in a Class Debate. *Social Behaviour and Personality, 36*(7), 893-902. doi: 10.2224/sbp.2008.36.7.893

Zheng, X., & Pierrehumbert, J. (2010). The Effects of Prosodic Prominence and Serial Position on Duration Perception. *The Journal of the Acoustical Society of America, 128*(2), 851-859. doi: 10.1121/1.3455796