DISFLUENCY IN SECOND LANGUAGE: A QUANTITATIVE STUDY ON TURKISH LEARNERS OF ENGLISH

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF INFORMATICS
OF
THE MIDDLE EAST TECHNICAL UNIVERSITY

BY

ERKAN VURAL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
IN
THE DEPARTMENT OF COGNITIVE SCIENCE

SEMPTEMBER 2008

Approval of the Graduate School of Informatics

Prof.Dr.Nazife Baykal
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Prof.Dr. Deniz Zeyrek
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Assist.Prof.Dr. Annette Hohenberger
Co-Supervisor

Prof.Dr. Deniz Zeyrek
Supervisor

Examining Committee Members

Assoc.Prof.Dr. Cem Bozşahin (METU, CENG) _____

Prof.Dr. Deniz Zeyrek (METU, COGS) _____

Assist.Prof.Dr Annette Hohenberger (METU, COGS) _____

Assist.Prof.Dr. Bilge Say (METU, COGS) _____

Dr. Ceyhan Temurcu (METU, COGS) _____

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this wok.

Name, Last name: **Erkan Vural**

Signature **:** **_____**

ABSTRACT

DISFLUENCY IN SECOND LANGUAGE: A STUDY OF
TURKISH SPEAKER OF ENGLISH

Vural, Erkan

M.S., Department of Cognitive Science

Supervisor: Prof. Dr. Deniz Zeyrek

Co-Supervisor: Assist. Prof. Dr. Annette Hohenberger

September 2008, 98 pages

The purpose of this study is to investigate disfluency and gesture in the second language under specific conditions such as familiarity vs. non-familiarity, concrete topic type vs. abstract topic type and speaking with native speaker vs. speaker with non-native speaker. The sample of this study was sixteen students from the Department of Basic English in Middle East Technical University (DBE), three instructors from DBE and one instructor from Modern Language Department in Middle East Technical University. Two of the instructors are native and the rest of them are non-native speakers of English.

With an assigned instructor each student spoke on the following topics: making spaghetti, giving directions, spring festival and clashing midterms, and effects of religion on our life. The conversations on each topic were recorded audio-visually. Then the audio-visual data was

annotated in terms of linguistic and gestural elements. In order to explore the relation among disfluency, gesture and controlled variables, quantitative data analysis methods were used.

Levelt's speech production and Krauss's gesture production model were used as a basic framework. Dual Coding theory and Metalinguistic Awareness Theory was used to explain intricate results of the present study.

As a result of the study, it was found that in the concrete topic condition, learners speak more fluently because of time and topic effects. Similarly, in the condition of familiar addressee and native speaker, learners speak more fluently than they do when speaking with a non-familiar or a non- native speaker.

Keywords: Linguistics, Psycholinguistics, Second Language Production, Disfluency, Gestures

ÖZ

IKINCI DILDE KONUŞMADA YAŞANAN AKSAKLIKLAR: İNGİLİZCE KONUSAN TÜRK KONUŞMACILAR ÜZERİNE BİR ÇALIŞMA

Erkan Vural
Master, Bilişsel Bilimler
Tez Yoneticisi: Prof. Dr. Deniz Zeyrek
Ortak Tez Yoneticisi:Doç.Dr. Annette Hohenberger

Eylül 2008, 98 sayfa

Bu çalışmanın amacı ikinci dilde konuşmadaki aksaklıkları ve el ve baş hareketlerini, tanıdıklık-tanıdık olmama, somut konu tipi-soyut konu tipi, anadilinde konuşan konuşmacı ile konuşma-ikinci dilinde konuşan konuşmacı ile konuşma gibi koşullar altında incelemektir. Çalışmanın örneklemi Orta Doğu Teknik Üniversitesi'ndeki Temel İngilizce Bölümünden (TİB) 16 öğrenci ve 3 okutman ve Modern Diller Bölümü'nden bir okutmandır. Okutmanlardan ikisinin ana dili İngilizce, diğerlerinin Türkçe'dir

Her öğrenci makarna yapma, yön tarif etme, bahar şenliği ve çakışan sınavlar ve dinin yaşantımıza etkisi konular üzerinde kendisine atanan okutmanla konuşmuştur. Bu konuşmalar görsel ve işitsel olarak kaydedilmiştir. Daha sonra, kaydedilmiş veriler konuşmadaki

aksaklıklar, el-baş hareketleri ve kontrol edilen değişkenler arasındaki ilişkileri niceliksel olarak araştırmak ve ortaya çıkarmak için yazılı hale getirildi.

Levelt'in konuşma üzerine ve Krauss'un vücüt dili üzerine olan modelleri bu çalışmada temel çerçeve olarak kullanılmıştır.Ortaya çıkan karmaşık sonuçları açıklamak için İkili Kodlama Teorisi ve Dil-üstü Farkındalık Teorisi Kullanılmıştır.

Bu çalışmanın sonucunda, katılımcıların , kısa süreden ve konu özelliğinden dolayı belirlenen somut konuda daha akıcı konuştukları ortaya çıkmıştır. Benzer olarak, katılımcıların tanıdıkları biri ile konuşmaları veya ana dili İngilizce olan biri ile konuştukları durumlarda daha akıcı konuştukları ortaya çıkmıştır.

Anahtar Kelimeler: Dil Bilimi, İkinci Dil Üretimi, Konuşmadaki Aksaklıklar, Vucüt Hareketleri

To My Family

ACKNOWLEDGMENTS

I thank my advisors Prof. Dr. Deniz Zeyrek and Asst. Prof. Dr. Annette Hohenberger for their guidance, expertise and patience through this research. With their support, this study has been an invaluable learning experience for me. It has really been an honor to work with these consummate professionals.

I also would like to present my gratefulness to my family for their great and continues support especially for my brother Okan Vural. Beside, a special thanks to Duygu Ören, there are no words to describe my gratitude for her support during my study.

I would like to express my special thanks to Robert West and Gary Colon from the Department of Basic English in METU. With their contribution and guidance I was able to complete this study.

I also would like to express my special thanks to my friends Hatice Şengür ,Veli Biçer , Kader Biçer , Ersin Çelik , Nazime Çelik, and my colleagues at Usta Yazılım. A.Ş. for their moral support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

DF : Disfluency

DG : Deictic Gesture

FL : Filler

FS : False Start

HS : Hesitation

I : Interlocutor

LG : Lexical Gesture

MA : Metalinguistic Awareness

MG : Motor Gesture

MLI : Misselected Lexical Gestures

NO : Nodding

NE : Negation

PR : Prolongation

RP : Repetition

RPR : Repair

SOT : Slip of the Tongue

SG : Symbolic Gesture

S : Speaker

SP : Silent Pause

t : Time

XML : Extended Markup Language

CHAPTER 1

INTRODUCTION

Daily speech is replete with deviations; uttering wrong words, expressions or ungrammatical sentences. All these deviations are disfluencies considered as one of the main determinants of the complex processes in the language production through stages: conceptualizing, formulating syntax, morphology and phonology.

Fluency of speech in second language is affected by linguistic and non-linguistic factors. Besides, several additional factors emerge, such as the speaker's non-autonomous abilities of speech production because the reflex-like property of language production of first language vanishes when the speaker moves to a second language. Therefore, the complexity of language production in the second language, and linguistic and non linguistic factors affecting fluency are of importance in second language production, too.

Three factors are the most priming deviators of fluency of speech in the second language learning environment (Bortfeld, Leon, Bloom, Schober, & Brennan, 2001). The first one is familiarity with the interlocutor and the second one is speaking with a native English interlocutor. The nature of the topic of the conversations- whether it elicits concrete or abstract words more- is the third one affecting speech production in L2. These three factors have a relation with fluency and speech production stages: conceptualizing, accessing of words, processing syntax, generation of morphology and phonology and, articulation. These stages were developed by Levelt and his colleagues (Levelt, 1989, 1999). The impact of their argument is quite strong (Roelofs, 1992, 1993, 1996; Temple, 2000). Their theory of speech production has been contributed to various fields and language related issues. It is a commonly used framework for many production studies and for the present study as well.

Gestures are other interesting phenomena accompanying both speech and disfluency. Gestures are assisting elements in speech. They hint both the stages of speech production and occurrences of disfluency, i.e. gestures have explanatory power on speech and disfluency production. The relation between gesture production and speech production is exposed by Krauss (2000), who suggested a gesture production model functioning in parallel with speech production model of Levelt. This model was constructed for lexical gestures that are categorized as meaningful gestures and have semantic relations to discourse.

Studies on disfluency require an interdisciplinary approach. In other words, if the variation of disfluency is in focus, contribution from different fields such as psychology and linguistics is required to understand the deeper phenomena in L2. In this perspective, the question of disfluency, addressed in this study, might be rephrased as follows: "How does speech in second language change if there is a deviation? More specifically, how do disfluencies vary with respect to linguistic and non-linguistic factors? Do the speakers use extra communication channels such as gestures? What happens to the gestures of a speaker when speech is suspended? How do gestures vary with respect to linguistic and non-linguistic factors? What does this imply for gesture-speech and gesture-disfluency coordination?"

In the study, a corpus of semi-natural conversational data was created and disfluencies and gestures were annotated and analyzed. The analysis part of the study is divided into three main chapters, first of which addresses the speaker's fluency in terms of linguistic and non-linguistic factors. The second chapter scrutinizes gesture production and its variability under certain conditions. And in the final chapter gestures and disfluency relation and their temporal occurrences is analyzed. Results and their discussion will also be presented at last.

1.1 Purpose of Study

There are two main purposes of the present study. The first one is to understand how disfluency (DF) varies under specific linguistic and non-linguistic conditions, which are abstract vs. concrete topic condition, familiarity vs. unfamiliarity with interlocutor (I) and speaking with a native addressee vs. a non-native addressee in L2. The second goal is to explore how gesture varies under the same linguistic and non-linguistic factors in L2. In addition to these, the relations between gesture and disfluency, was investigated. Their correlation and dual occurrence were statically analyzed.

1.2 Significance of the Study

There are numerous studies on disfluency (Bard, Lickley, & Aylett, 2001; Bortfeld et al., 2001; Cooper & Hale, 2005; DiSS'03, 2003; Eklund & Shriberg, 1998; Howell & Akande, 2005; Lickley, 1996; H. Nicholson et al., 2003; H. B. M. Nicholson, 2004; Shriberg, 1999) and gesture production (Finlayson, Forrest, Lickley, & Beck, 2003; Griffin, 2001; Krauss et al., 2000) . Studies based on first language (L1) mostly focus either on linguistic factors or non-linguistic factors. Very few of them consider both linguistic and non-linguistic factors in the same study. However, none of them have dealt with types of disfluency and gesture in L2 in terms of both linguistic and non-linguistic factors in the same study.

The study is unique since it investigates disfluency rate in terms of both linguistic and non linguistic factors and tries to connect the results with gesture production in the second language. In addition to this, this study tries to explain the phenomena in speech production in terms of an interdisciplinary perspective including contribution from linguistics and psychology.

1.3 Motivation of the Study

Cognitive Science is an interdisciplinary area aiming to shed light on the cognitive processes of human beings. Language is one of the most important and valuable tool to achieve this. Because of that, studies in different areas have been conducted on language to understand its underlying structure.

Disfluency is one of the most challenging problems in speech production and comprehension theories. The syntax and semantics of disfluency put complicated cases into a more complicated form. Therefore, there are two alternative for researchers in the field of Cognitive Science. The first one is ignoring disfluency and constructing theories putting it aside. The second one is handling disfluency cases and shaping theories by considering them.

The theories of the first group of researchers, who ignore disfluency, are weak to explain the phenomena in real cases that human beings confront every day. On the other hand, the second class of theories suggests a mechanism to handle deviations in speech by considering disfluency. The most widely used theories belong to the first group. Nowadays, researchers are working to overcome the complexity of real speech and they are trying to model disfluency occurrence in the speech.

Yet, most of the researchers have been missing or ignoring the fact that disfluencies might additionally be affected by non-linguistic factors. In the future, they will have to include these external parameters as well as the linguistic ones. Then their theoretical systems would be more coherent and accurate to handle all languages.

It is a fact that researchers want to achieve global theories of language rather than a theory handling a single language. The studies on disfluency are one step towards reaching coherent and complete theories that work for all languages. It is vital to determine each specific problem to simulate the human language comprehension and speech production system since they are a priori steps to explain all human cognitive system. There are many studies conducted with this perspective. Although there are theories on production of speech and on behavior and internal structure of disfluency, these theories have not been tested under different conditions, especially in L2. It must be checked whether the assumptions apply for disfluency in L2 as to be in L1. Otherwise, these need to be extended or altered since it is possible that an application of them may fail. Therefore, problems such as disfluency and their effectors must be scrutinized in detail.

CHAPTER 2

PSYCHOLINGUISTIC THEORIES OF SPEECH PRODUCTION

The theory of speech production by Levelt (1989) was preceded by the induction of an important psycholinguistic construct: the lemma.

2.1 Lemma

The term *Lemma, was* firstly coined by Kempen and Hujibers (1983), as a syntactically and semantically specified lexical item which was not phonologically specified. Lemmas were also described by Levelt (1989) as a representation for meaning and syntactic properties of a word. The concept of lemma is rudimentary for most of the speech production theories and the notion of lemma is one of the least controversial assumptions for the language production theories (Pechmann & Zerbst, 2004).

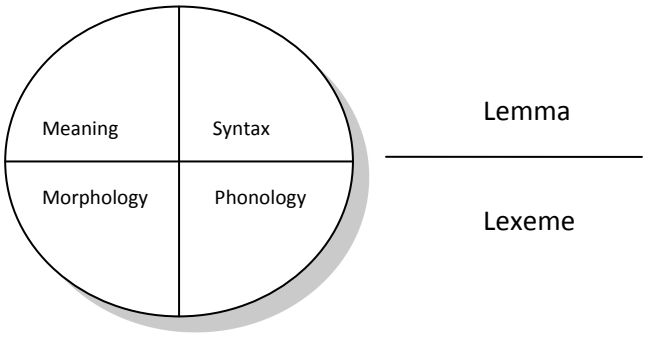


Figure 1 - Internal structure of an item in the mental lexicon (Levelt, 1989, p. 182)

Lemma may be summarized as a bunch of information related to syntactic properties, meanings, syntactic categories and other properties of a word. This bunch contains the required information for constructing phrases and sentences. That is, for a word, it conveys detailed language specific information from the lexicon to the working memory. Lexeme- form part- contains information about phonology, morphology, stress and other articulation information.

Basically, the lemma assumption is that there exist two stages in accessing the mental lexicon during the production of words. These are grammatical and phonological encoding (Dell, Chang, & Griffin, 2001; Levelt, 1989). Figure 1 represents a lexical item's internal structure. This structure is theory dependent and shaped according to theoretical aspects (Levelt, 1989).

2.2 Lemma Retrieval and Lexical Access

Lemma retrieval is the name of the process occurring in grammatical encoding. Lexical access is the more general term for retrieving any information or a concept from the lexicon; it is an issue of long term memory. More specifically, lemma retrieval or lexical access is the process of fetching information from memory in order to construct minor planning which is going to be expressed verbally (Levelt, 1989, 1999; Levelt, Roelofs, & Meyer, 2002). In producing utterances, speakers call their stored knowledge about words, including the words' meaning, syntactic properties, morphological composition and sound structure (Roelofs, 1992). This process is robust and efficient (Roelofs, 1996). Speakers make few mistakes in L1, and in L2 they make relatively few mistakes in spite of the lack of knowledge in L2.

Lexical access is the framework for Levelt's theory of speech production and word production, Word production was just reduced to a simpler and abstract case to avoid the theoretical concern such as connectionism. In the speech production theory, it is known that lexical access and planning continue simultaneously and in a parallel way but each stratum produces outcomes in a serial way (Levelt, 1999; Levelt et al., 2002). There is always a feeding from the conceptual network to formulator networks even though language production is incremental and serial (Levelt, 1989, 1999; Levelt & Meyer, 2000; Levelt et al., 2002; Meyer, Roelofs, & Levelt, 2003).

2.3 Speech Production Theories

Speech production is an incremental process. Specific messages flow through strata. At the conceptual stratum, activation of concepts occurs and the *message* spreads through the formulator stratum which contains grammatical and phonological encoding. In the formulator, syntactic and sound structures and their plans are built (Roelofs, 1992). Lexical access and grammatical encoding are located in the formulation component (Levelt, 1989). After the formulator stratum, the processed message reaches the articulator stratum (Levelt, 1999; Levelt et al., 2002; Roelofs, 1992). All modules or stratums work in a serial way. That is, each outcome of a module is the input of the next module. But modules work in parallel, which means that modules do not wait for the processes of the next modules and continue to process concepts. This form of parallelism is horizontal, hence all speech productions are incremental .In other words, modules work in parallel but they work on different data. On the other hand, connectionist approaches do not posit modularism. They allow for vertical parallelism – activation across layers. Therefore, the model is a serial-modular model and not a fully connectionist model.

Interactive Models of speech production are bi-directional; activation spreads in both ways. This means that these models use the same components for both comprehension and production of language (Levelt, 1989, 1999). On the other hand, Dell (1995; 2001; 1997) has disagreed about this issue by suggesting aphasic people. He argued that aphasics comprehend language but they have trouble in producing sentences. Aphasic people show both good auditory word recognition and phonological encoding but poor production (Dell et al., 1997). Thus, bi-directionality is a controversial issue for speech production theories. For the present study, bi-directionality is not important; speech theories are used in terms of their claims about production.

In the next part, Levelt's Speech Production Theory will be presented in detail.

2.3.1 Speech Production

Levelt, Roelofs and Meyer's Theory is the pioneer speech production theory. Their theory of flow is presented in Figure 2, which proposes a partitioning of the various processes into smaller and abstract parts (Levelt, 1989). In the schema, the boxes represent processing components while the ellipses represent knowledge stores.

It is a modular system; the output of one component becomes the input of the other. Modules work independently and they have few interfaces for communication. Thus, it is a flowing system pushing the output to the next level of processing but all nodes in the flowing process independently.

Levelt classified his theory as the partially connectionist theory. In fact, it is a serial-modular model. It has a spreading activation in implementation which is partially connectionist. The important side of the model is related to parallel processing and it is important for the robust process of speech since the process of thought into speech has been associated with connectionist models (Dell et al., 2001; Levelt, 1989). Theoretically, preservation of parallelism is a difficult issue. Thus, Levelt, Roelofs and Meyer (2002) suggested a new theory (see Figure 3) to simplify complex processes by abstraction and reduction. Demotion to the simpler form reduces the parallelism concern of the theory. Therefore, the Levelt Model is not a connectionist model. He emphasizes that the formal language of connectionism is convenient for talking about parallelism and that there is a lot of parallelism in language production (Levelt, 1989, p. 20). He further says that he will use certain proposals of connectionism or parallel processing particularly in phonological encoding.

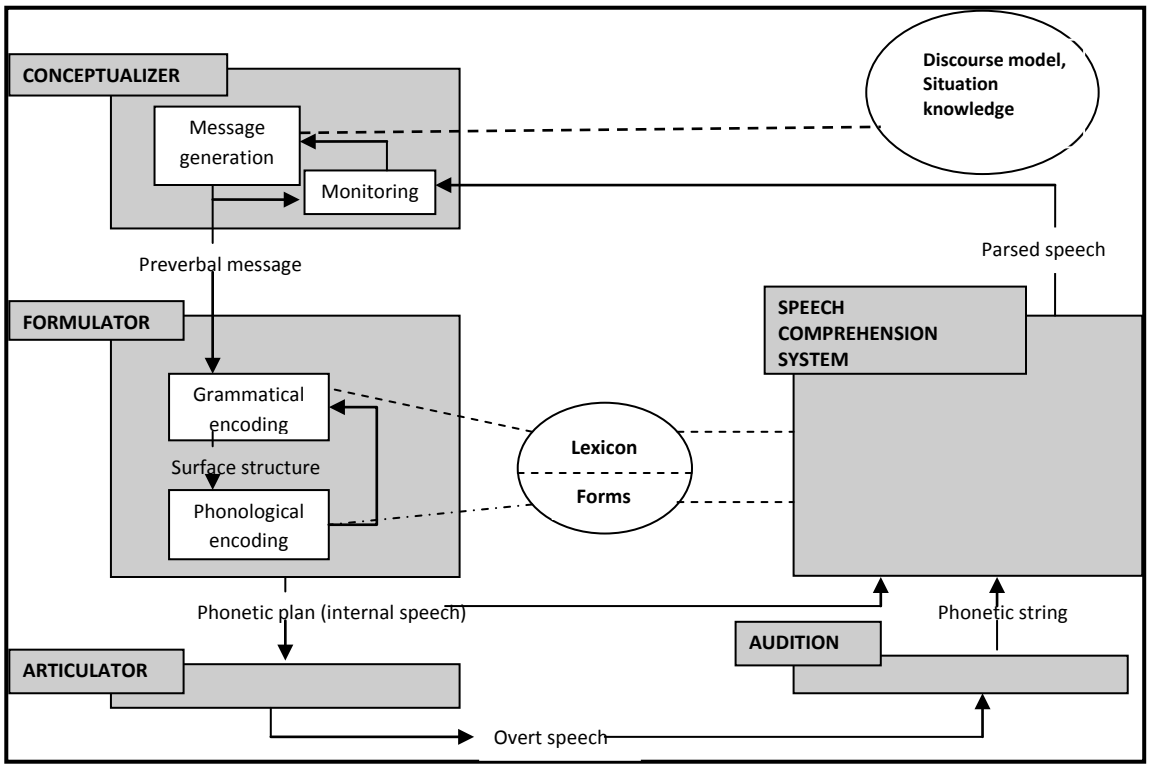


Figure 2 - Speech Production Model (Levelt, 1989, p. 9)

Why is parallelism so important? Levelt (1989, pp. 2,20) emphasizes that processing components are specialized and they do their works in an autonomous fashion. Most of the components accompanying speech production are highly automatic, and operate in a reflex-like way. This automaticity makes it possible for them to work in parallel. It is the result of the modularity of the system. This explains the production of uninterrupted fluent speech and except some deviations robust language production. This is a general approach of Levelt's model and it explains L1 production. L2 production is different to some extent. One can speculate that L2 does not have highly automatic working components. In L2, the speech generation system may work in a semi-parallel way. This means that, if it is assumed that the speaker uses the same semantic knowledge base for every language learnt, it can be speculated that the conceptual processing should work in parallel but the formulation part needs to work in a different mode as processing is possibly interrupted. Word recollection faces with obstacles and it deviates.

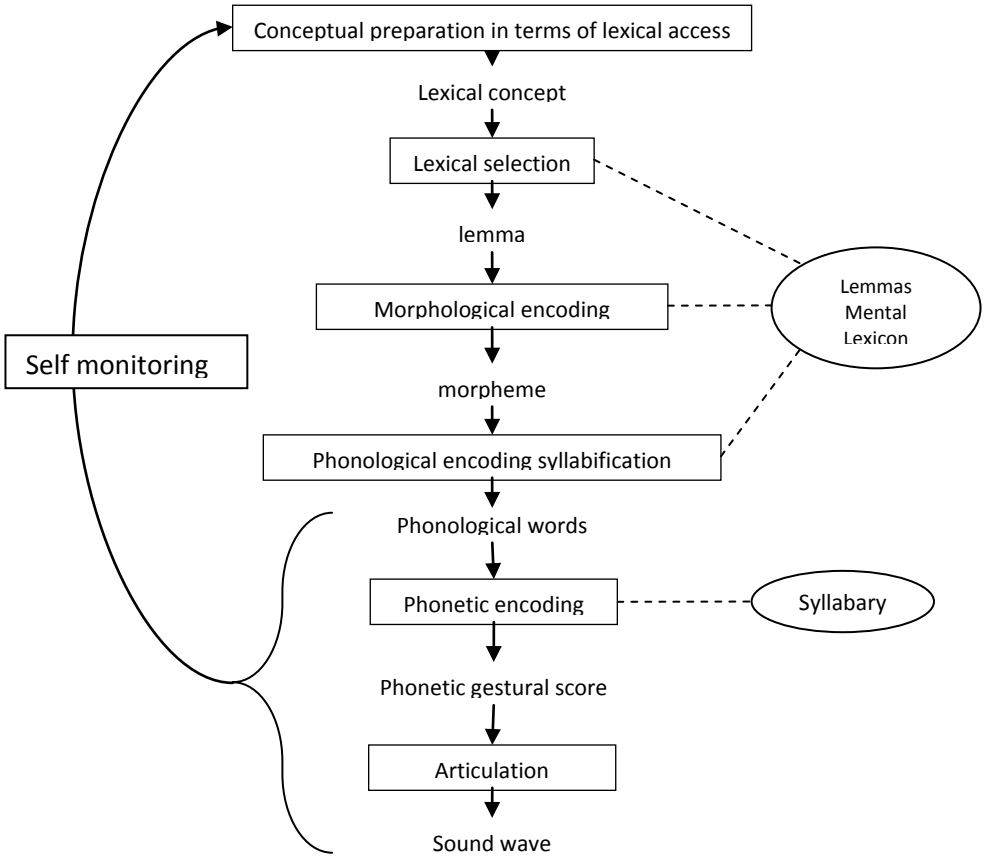


Figure 3 - The Theory Outline of Speech Production Model of Levelt, Roelofs and Meyer (Levelt, 1999)

Besides general arguments and discussion about theory[1], the details and functionality of levels or modules are the most important source of information for the present study. Their embedded structure and functionality has the power to explain some of the L2 issues that are dealt with.

2.3.1.1 Conceptualizer

Uttering sentences accurately and coherently needs an intentional activity. The primary issues of this component are selecting the relevant information, putting information in logical order, ordering them for expression while keeping the track of what is said. Thus, these activities require constant attention of speaker (S). Meanwhile, *S* has to monitor himself while he is speaking. The sum of all of these mental activities is named as *conceptualizing* (Levelt, 1989, p. 9) . All these events occur in the component named *conceptualizer*. The output of this stratum is called *preverbal message* (Levelt, 1989, p. 9). In order to encode a message, the speaker must access and use two kind of knowledge; namely procedural knowledge and declarative knowledge. Depending on the need, the right kind of knowledge should be used for the intended communication.

There exist two distinguished levels of the planning of preverbal messages: macro planning and microplanning. Macroplanning refers to the general idea to be verbalized. Microplanning is the sub parts or chunks of the main goal. Micro planning has to shape minor plans to be more understandable for the addressee. Generally, macroplanning does not change, but according to the situation of discourse, the microplanning may change. The microplanning concerns the further shaping of speech acts (Levelt, 1989, p. 107).

This preverbal message is a conceptual structure that can be accepted as an input by the Formulator (Levelt, 1989, p. 11; Levelt et al., 2002).

1 There is also another important point about the parallel account of lexical access. The parallelism is both required for theoretical and practical reasons. Levelt (1989, pp. 199,200) presented empirical evidence to prove this account. But further details are out of scope of the this study.

2.3.1.2 Formulator

This sub-stage is responsible for processing and translating the conceptual structure- i.e. the preverbal message- into the linguistic structure. This translation process proceeds in two steps; grammatical encoding and phonological encoding. It takes the preverbal message(s) from the conceptualizer for processing it. At the end, it produces a phonetic and articulatory plan for the next component, which is the articulator.

2.3.1.3 Grammatical Encoding

After the preverbal message has been produced, the next stage is grammatical encoding. This is the stage where the grammatical encoding of access the lemma in the lexicon. Minor plans are prepared due to the information contained in the lemma for big plan (Levelt, 1989). It constructs sub-chunks of the bigger plan due to the declarative and syntactic knowledge in the activated lemma. It deploys them in the memory. After all, these greater chunks are built with respect to linguistic structures and the rules pre-stored.

It is important and must be emphasized that grammatical encoding is independent from phonological encoding although they are in the same module and relate to the same lexical item. It is also important for this study that the idea of incremental sentence production is valid at this level (Levelt, 1989, p. 162) . But it is likely that failure in finding the phonological form of the word may provide feedback and trigger reprocessing of the grammatical encoding level and its functions, and most importantly this triggering activity should change the predefined micro planning. Grammatical encoding is the only module affected from two the stages: stages of the conceptualizer and the phonological encoding.

The output of these components is called *surface structure* which is an ordered strings of activated lemmas grouped into phrases and sub phrases of various kinds (Levelt, 1989, p. 11). The results, i.e. the group of lemmas and their combination information, are interim in a buffer called *syntactic buffer* for the next stage (Levelt, 1989, p. 11, 1999; Meyer et al., 2003).

2.3.1.4 Phonological Encoding

The function of the next stage is retrieving or building phonetic plans for activated and structured words, Levelt named this product as *Internal Speech* (Levelt, 1989, p. 12). The syllable structures and several phonological properties are constructed at this level. The stress of the syllables and stress of words i.e. the place and various phonological properties are set here. It pushes the outcome to the next component, the articulator.

Phonological encoding functionality has similarities to grammatical encoding. It reaches lexical collection and fetches the activated lemma's form with respect to the previously activated and stored lexical pointer. That is, by the help of the lexical pointer this layer functions to fetch the form information from the lexicon. It builds minor plans for big plans. If the component faces with an obstacle, it triggers grammatical encoding to change the active minor plans. This triggering function will be presented in next sections.

All prepared plans are stored in the same part of the memory to be used in the next stage which is the articulator.

2.3.1.5 Articulator

The articulator is responsible for the execution of the phonetic plan. It stores phonetic plans in the *articulatory buffer*. The management of the articulatory system is done due to the stored phonetic plans. After the phonetic plans are prepared, muscular plans are executed and the message is verbally expressed (Levelt, 1989, pp. 12, 416 - 420). The product of the articulator is overt speech.

2.3.1.6 Self Monitoring Mechanism

In the Levelt Model, monitoring proceeds via the *speech-comprehension system* because the *S* monitors his overt speech and the speech of interlocutor (I). Due to the situation of the discourse, the module provides information for other modules if required. If there is a problem in overt speech, considering *I* speech, it gives feedback to the conceptualizer about the error. The system has power to access both the form information and the lemma information. It parses the speech

into phonological, morphological, syntactic and semantic parts. For the overt *DF*s, self monitoring is the main trigger.

2.4 The relationship between the Modules and Disfluency

The process of speech production and sources of disfluency phenomena have causal relations. This causality is related to the functionality of the modules. Namely, functionality of each module produces specific types of disfluency. For example, grammatical disharmony may be produced when there is a problem in grammatical encoding. Menyhárt (2003) presented disfluency and their occurrence as depicted in Figure 4. This was intended for L1 but it is also applicable to and valid for L2.

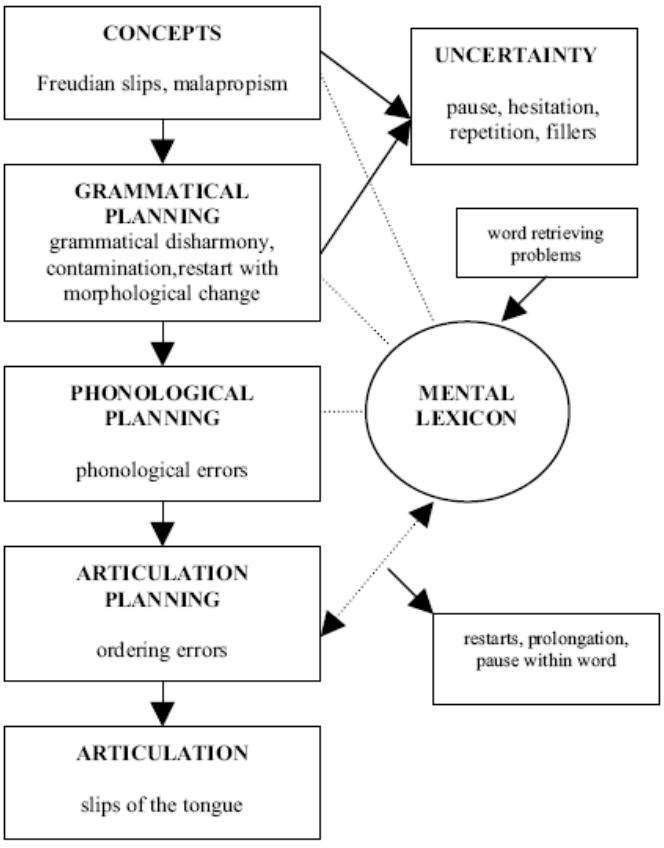


Figure 4 - Sources of Disfluency Phenomena (Menyhárt, 2003)

To sum up, we defined the Speech Production Model of Levelt's stages: conceptualizing, formulating (grammatical encoding and phonological encoding) and finally articulating. We described functionality of each stage and some of their important aspects related to

connectionism and decompositionality. We additionally emphasized deficits of the model. However, Levelt's model and its extensions in the literature provide an effective framework for studies focusing on speech production in both L1 and L2 as modular approaches ease to explain L2 cases violating some of the basic arguments for L1. Otherwise, it would be a great challenge to explain all of them. Extensibility and possible extensions of this model provide a rich explanatory environment for L2 cases.

In the next section, gesture, disfluency and gesture relation and another extension of Levelt's model for gesture production, speech and gesture production parallelism will be presented.

2.5 Disfluency, Gesture and Production Theories of Speech

The planning process for speech production is subconscious and highly automatic. *S* uses implicit linguistic plans from previous experiences. Although plans for speech production are available, daily speech is replete with errors. At first glance, these errors and disfluencies seem to be an unwanted phenomenon though they are a unique source to understand language and its sub-components.

Language production theories have been developed to account for cognitive processes by using speech language errors. Theories on language production are based on the evidence that language production process proceeds through serial strata (Dell, 1995; Dell et al., 2001; Levelt, 1989, 1999; Roelofs, 1996). Especially, *DF* is a unique element helping psycholinguists to differentiate these strata explicitly (Ferreira & Bailey, 2004; Levelt et al., 2002). Besides *DF*, lexical gesture i.e. accompanying elements of speech is another source which distinguishes the stages of speech production (Krauss et al., 2000). Hence lexical gestures are raising clues about *DF* and the overall situation of speech.

Studies have consistently found disfluency as a critical factor influencing speech production and they have highlighted speech errors as an important tool to understand our cognitive system. For example, Levelt and Meyer (2000) used speech errors and picture naming-chronometric approaches as a tool to verify their studies. There are also numerous other studies that have been published on word production and speech errors to understand language phenomena and the internal structure of language faculty (Callaway, 2003; Clark & Wasow, 1998; Cooper & Hale, 2005; Dell, 1995; Dell et al., 2001; Dell et al., 1997; Derwing, Rossiter, Munro, & Thomson, 2004; Eklund & Shriberg, 1998; Ferreira & Bailey, 2004; Kempen & Huijbers, 1983; Kormos,

1999; Krauss et al., 2000; Levelt, 1989, 1999; Levelt et al., 2002; Temple, 2000). All these studies contribute to the theories of speech production.

Although, speech production theories are based on different perspectives, they use same fundamental assumptions: (1) there are two stages in speech production; a semantic/ syntactic layer and a phonological layer, (2) the *lemma* and the *lexeme* are specific information holders in the lexicon (Levelt, 1989). Theories develop in parallel within these two assumptions.

2.5.1 What is Disfluency?

When *S* cannot formulate the whole speech at once, he suspends speech and introduces some elements such as silent pauses and fillers. And when *S* changes his mind or notices his mistake in speech, he could add, delete or replace the words he has already uttered. All these kind of deviations in speech are generally named as *DF*. Of course, they are not bare mistakes in language production. On the contrary, *DF*s indicate a failure of *S* in the design of an utterance before beginning to speak or before an adjustment by replacements (Gurman, Aylett, & Lickley, 2002).

There are various views as to what *DF* is. The encyclopedic meaning of *DF* is "any deviation in speech from ideal delivery" (Ferreira & Bailey, 2004). It is also described as "an interruption in the smooth flow of speech, as by a pause or the repetition of a word or a syllable" (Dictionary.com, 2007). Interestingly, *DF* is seen as an impairment in the language due to some definitions like "impairment of the ability to produce smooth, fluent speech" (Dictionary.com, 2007). Sometimes, it is perceived as a cognitive burden when the speaker is managing speech (Brown & Dell, 1987). Clark and Wasow (1998) defined *DF* as a pragmatic tool of *S* since it is a strategic device for intentionally signaling to *I* that *S* is going to produce a group of utterances.

According to another point of view, disfluency is seen as a pragmatic issue because it is both a message to the interlocutor and an interaction provider between *S* and *I* during the time of lack of fluent speech. It reduces the tension of the conversation, conveys information about what *S* is going to do or it conveys information about the thoughts and mental processes of *S*, i.e. *S* is thinking about something or looking for a word to plan the next utterance (Mehnert, 1998).

DF is one of the primary media used by a number of studies to understand language production. These studies use *DF* both to falsify and to verify the hypothesis of their study (Ferreira &

Bailey, 2004; Levelt et al., 2002) and they provide some striking results about *DF*. Some of the findings are: *DF*s are more common in longer utterances and complex constitutes; they are more probable to occur in early sentences when the later stages are incomplete (Bard et al., 2001; Gurman et al., 2002); increases in *DF* accompanies increases in the use of cognitive functions to communicate (H. Nicholson et al., 2003); increases in *DF* also results in an increase of the use of lexical gestures (Krauss et al., 2000).

There is a controversial issue about *DF*, namely frequency in speech errors. In the literature, the occurrence percentage of *DF* per word is generally presented in different magnitudes. Levelt (1989, 1999) wrote that human beings err, on average, no more than once or twice in 1000 words; on the other hand, Ferreira (2004) stated it is 6 to 10 per 100 words. The numerical data varies from study to study because of several reasons such as topic, gender, familiarity to *I* etc. It is a fact that the rate of disfluency is affected by both linguistic and non-linguistic factors. Although linguistic influences have been well studied, generally non-linguistic influences have been ignored. Thus, one must also check out the non-linguistic influences. Most pioneer studies about non-linguistic influences are on eye contact with *I*, familiarity with *I*, knowledge of the topic, and the role of *S* in the speech (Branigan, Lickley, & McKelvie, 1999). Furthermore age is seen as a non-linguistic influence on the rate of *DF* (Bortfeld et al., 2001; Menyhárt, 2003) because the *DF* types and their occurrences should differ with respect to age (Menyhárt, 2003). The other important non-linguistic influence affecting rate of disfluency may be gender but gender has not been found to have an effect on the occurrence of disfluency at all (Bortfeld et al., 2001; Menyhárt, 2003). In addition to these, some properties of *I* might affect the fluency of *S*. For example, familiarity with *I* will affect the speech of speaker for both L1 and L2, meaning that people speak fluently with familiar people (Bortfeld et al., 2001).

On the linguistic side, the most important linguistic influences affecting number of *DF* have been listed as abstractness and concreteness of the topic, and complexity and frequencies of the vocabulary (Branigan et al., 1999).

In this study, *DF* and the processes causing it, namely non-linguistic effects (familiarity with *I*, speaking with native speaker of English) and linguistic effect (topic) were considered. In the following parts literature on these issues will be presented.

2.5.2 Components of Disfluency

There are some basic elements of *DF*. Basically, *DF* consists of reparandum, suspension point, repair and resumption (Ferreira & Bailey, 2004). This decomposition is related to the self-monitoring mechanism. If an error is caught by the self-monitoring mechanism, its decomposition is always done as in Figure 5.

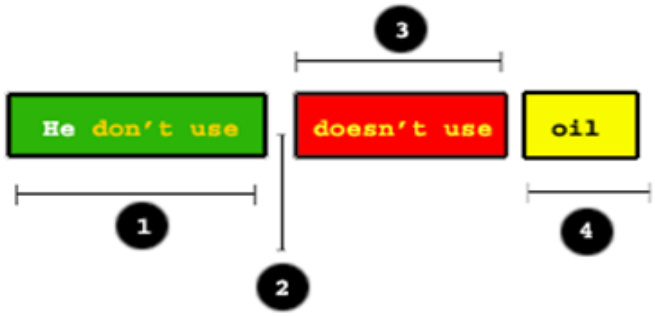


Figure 5 - Basic decomposition of Disfluency.

1 : Reparandum Point, 2:Suspension Point, 3:Repair , 4:Resumption (an example from data)

Reparandum (1) is the first stage in this decomposition structure. This stage contains the original delivery of the disfluency. The second stage is the *suspension point* (2) denoting that *S* recognizes the error he made. The point 2 in Figure 5 refers to the point and time when the recognition of the error occurs. At this point *S* understands that it is not what he intends to say. The next stage is the *repair* (3) stage of the *DF*. In this stage, *S* repairs his mistake by uttering new chunks. That is, reparandum is repaired after the recognition of the error. And the last stage is the *resumption* (4) stage of the *DF*. It is the ongoing part of the speech (Ferreira & Bailey, 2004).

The structure presented above does not cover all types of disfluency, for example fillers, hesitations and silent pauses are left out (Cooper & Hale, 2005). These *DF* types are not recognized by the speaker so they are not repaired. Non-repaired *DF*'s types are composed of one component, only reparandum, instead of three.

These two decomposition structures are true for both L1 and L2 (Cooper & Hale, 2005). Sometimes these two structures are insufficient to explain some of the cases in L2 (see Repair Type of Disfluency for more complex structure). For example, *S* may produce successively repeated utterances as well as unbounded DFs, for instance *S* may repair his errors after uttering a phrase or a sentence. The repair goes beyond boundaries. This type will be explained in section 2.5.5.7.

2.5.3 Differences in Terminology

There are different terminologies and classifications used for the types of *DF*. The same classes or elements are named differently across studies. Before comparing the results and studies in the literature, it is better to understand terminologies and classifications used for the types of *DFs*. For example, hesitation was used as a head title for fillers, silent pauses and repairs. On the other hand, in some other studies hesitation was taken as a different type of disfluency like in the present study.

2.5.4 Overt and Covert Types of Disfluencies

*DF*s could be classified as overt and covert *DF*s in term of the awareness of *S*. Levelt (1989) suggested that repaired kind of errors are overt types of disfluencies. Overt disfluencies always have a repairing process hence they are also named as "edited errors". In the overt type, *S* notices his mistake, i.e. *S* monitors his speech. That is, *S* captures the deviation and then fixes it. On the other hand, Covert *DF* means that *S* could not understand his mistake and continues his speech. Covert *DF*s are also named as "unrecognized errors".

2.5.5 Types of Disfluency

In the literature, basic types of *DF*s are listed as fillers, silent pauses, hesitations, false starts, grammatical errors, slips of the tongue, repetitions, repairs, misseleted lexical items and prolongations. These types of disfluencies were defined for L1. The present study found that while most of them are also valid in L2, some of them such as prolongation are absent. Grammatical errors are not good *DF* types in L2 studies because they are rather developmental

errors of the interlanguage i.e. the stages of development regarding the learner's performance (Selinker, 1972).

In the literature there are few studies on *DF* and L2. The types and results have not been generalized yet. Thus, types of L2 *DF* and the methods of dealing with them will be presented in this study. A new category of *DF* will be defined and its importance will be briefly presented. This new category is named as *Repair*.

In alphabetic order, disfluency types for this study are as follows:

- Fillers
- False Starts
- Hesitations
- Misselected Lexical Items
- Repairs - Fixing Grammatical Error(s) In Context, Misselected Lexical Item Repairs
- Repetitions
- Silent Pauses
- Slip of the Tongue

The next part presents detailed information about these DF types, and their usage in data.

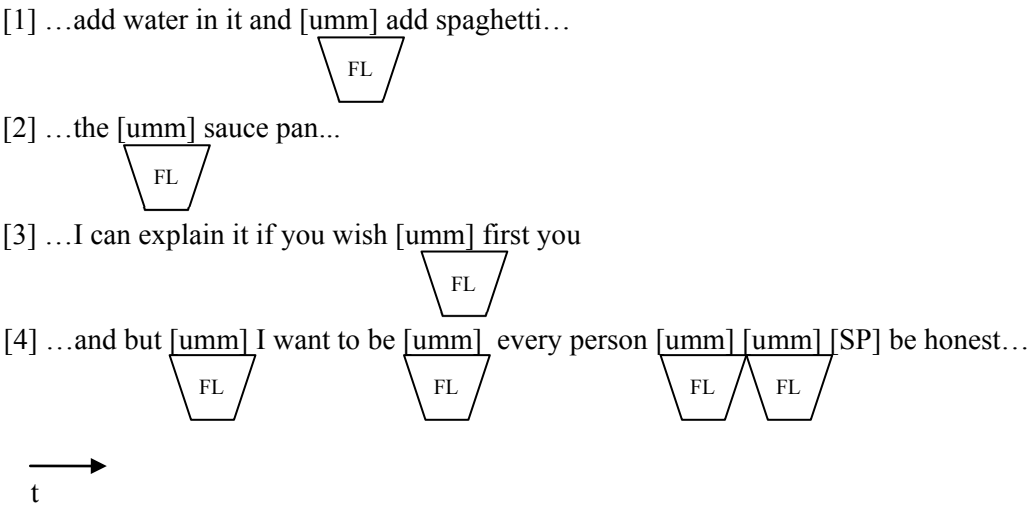
2.5.5.1 Fillers (FL)

When *S* cannot formulate the entire utterance at once, he suspends his speech and introduces the filler before going on (Clark & Wasow, 1998). The Filler (*FL*) type of disfluency may contain one of the meaningless elements such as "*umm*," *uhh"* and *"ımm"*. The occurrence of this type is generally connected to searching a word or planning about the upcoming.

There exists no clear pattern for the occurrence of *FL*. Its place cannot be predicted using any previous information. This is called *uncertainty* (Menyhárt, 2003). Thus, *S* can in principle suspend his speech at almost any point in an utterance (Levelt, 1989). The place of *FL* could be in the phrases between the phrases and even between sentences. Additionally, *FL* could be observed successively in any place of the speech. Thus, the syntactic predictability of this kind is not possible. This is important for those working on artificial intelligence and speech

comprehension systems because *DF* is still a challenge for them (Cooper & Hale, 2005; Stolcke & Shriberg, 1996).

Now *FL* type with examples from the present study will be illustrated. Example [1], [3] are for *FL*s occurring between phrases and sentences. Example [2] is for *FL*s in the phrases, and example [4] is for *FL*s in repetition. Each example is from a different *S* taking part in the study. "umm" is used to stand for "uhh, ımm, umm" and etc.



2.5.5.2 Silent Pauses (SP)

When *S* cannot formulate the rest of the utterance at once and stops and waits for a while, he suspends his speech. In other words, *S* introduces a silent stage before going on (Clark & Wasow, 1998). This is called silent pause (SP).

S can in principle suspend his speech at any point in an utterance because this strategy is mainly used by *S* when having difficulty in lexical selection (Butterworth, 1975). That is, there is not any specific pattern to predict the place of *SPs*. This is similar to FLs in online speech (Levelt, 1989).

This type of disfluency has a controversial side related to what to accept as *SP*. Most researchers working on disfluency accept a pause as *SP* due to their intuitions or impressions but this is not very helpful. Few of them make a clear definition of SP (Kirsner, Dunn, & Hird, 2003). Meyer, Roelofs and Levelt (2003) has a clear definition on when to accept a pause as a *SP*. According to

them, *Ss* produce representations of successive syllables of a word in sequence and they only begin to speak after having planned at least one word completely (Levelt, 1999). Therefore, initial speech latencies should not be longer than short words (Meyer et al., 2003). When *S* does not articulate and pauses more than 0.5 second, as short words may be uttered at 0.5 second, this time span is accepted as a silent pause (Kirsner et al., 2003) . In this study, all pauses more than 0.5 second will be accepted as a *SP*.

This is a covert disfluency where the self monitoring mechanism is not the aware of the mistake(s). *SP* only includes reparandum. That is, *S* may not recognize the pause if its duration has an acceptable magnitude for both *S* and *I*. There are studies on the tolerance time for silence in conversations. After a period of time, one of *S*s may interrupt the silent stage and change the direction of the conversation.

Similar to *FL* type, this type of the disfluency is frequently encountered in the study. Duration changes from 0.5 second to 8 seconds. Typical examples from the data of this study are presented below. The [SP] stands for Silent Pauses and all data are from different *Ss*.

[5] … spaghetti will be enough for four [SP] people..

[6] … you when you [SP] out of this get out of this building and see the straight..

[7] .. if I am not mistaken [SP] [umm] when you go out there

[8] ..and then [SP] [umm] [SP] industrial engineering

2.5.5.3 False Starts (FS)

Sometimes *S* recollects wrong words from the memory and starts new sentences with a wrong word or words. Then *S* recognizes his mistake and ignores the pre-uttered word(s). He continues to utter new word(s) with respect to the new planning. This phenomenon is called false start (FS).

There are two kinds of *FS* with respect to the place of occurrence. These occur generally before the phrases and sentence beginnings. The constituents of FS could be any type of language elements such as nouns, propositions, verbs, articles or any pairs of these.

The structure of *FS* is similar to the general structure of *DF*. It has reparandum, suspension point and repair part. The false started chunk is the reparandum part. After the false started chunk,

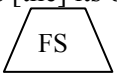
there is a suspension point. Generally, after the reparandum of *FSs*, utterances may be presented so there could be elements at the suspension point, such as *FLs*.

FS is somewhat different then other types of *DF* because repairing can be completely done by different words, chunks, even sentences as in example [10] which is yet another example from the data for this study.

[9] …umm [for] I think some sort of /sɔsə/ sausages you can use meats like...

Filler False Start Resumed Delivery

[10] …I think the whole faculties has [the] its own their own research centers...



Delivery False Start Resumed Delivery

t

FSs show properties of overt type. This is because, *S* understands his mistakes and fixes his errors before continuing his speech(Levelt, 1989).

2.5.5.4 Repetitions (RP)

Repeating an element or group of elements is named as repetition (RP). Any type of linguistic element or any type of phrase might be repeated. Sometimes, it is possible to repeat sentences as well. The occurrence of *RP* is free. It can be observed between sentences, between phrases or even within the phrases. Hence, the prediction of the exact place or syntactic structure is not possible.

There is an important issue for *RP* in speech, some people use repetition as an emphasis. Accepting a repeat as a *RP* or as an emphasis could be understood from the context. Thus, *RP* could be analyzed considering the context of speech. This should be easy for linguistic studies but it is a difficult issue for statistical language modeling of speech disfluencies (Stolcke & Shriberg, 1996).

The structure of *RP* is also different from the general *DF* structure. It consists of an *initial commitment, suspension of speech, hiatus,* and *restart of constituent* (Clark & Wasow, 1998). The difference is the hiatus part. Hiatus is the stage between the suspension and the resumption of speech. *S* may do a variety of things in a hiatus (Clark & Wasow, 1998). Examples from the data are as follows, where the material between curly brackets in the example presented below refers to hiatus;

[11] ...then umm you {[SP]} you put the

[12] ... I {} I don't remember the name …

[13] …what she want {} what she want {} [wants] to do umm…

[14] …I want to {} want to {} want to {} drink

The fourth example is a rare and interesting example for this class of disfluency. The phrase "want to" is repeated twice as the third one also repeats the second one, where the second "want to" is both an *initial commitment* and the *restart of constituent*. Such cases are very few in the data of this study.

RP is a convert type of *DF*. It is a repetition (Levelt, 1989) because it is not clear why it is repaired, why *S* restarts the same constituent rather than going on with it. It may be related to planning or to the upcoming ideas. Thus, there is no a self-monitoring mechanism that helps S to be aware of these type of errors.

2.5.5.5 Hesitations (HS)

Hesitation (HS) consists of uncompleted words in the speech. Its structure is similar to general the structure of disfluency. It consists of an original delivery, a resumption stage referenced by curly brackets and a repair part.

Original Delivery	{...}	Repair

Hesitation is an overt type of disfluency, that is; *S* catches his or her uncompleted group of phones and repairs it. The original delivery is always repaired by the completed word but the middle section, the curly bracket part, has different variations in terms of linguistic elements. Until the speaker understands his mistake, he may utter or continue to speak. This may be seen

in example [16] where the speaker may utter a group of phonemes successively. This is rare but a possible example for L2.

Related examples from this study are presented below.

[15] …some sort of /sɔsə/ sausages you…

[16] …when you umm take /sp/ /spa/ /spa/ umm spaghetti…

[17] …/hol/ I think the whole faculties …

[18] … and then [umm] [SP] one /sp/ spice…

2.5.5.6 Misselected Lexical Item (MLI)

S sometimes uses some words in the wrong place. These words have semantic relations to the real word but this relation could resemble slips of tongue, namely semantic substations. This refers to wrong uses of the words. These kinds of errors are related to the conceptual level of speech. Semantics of the words trigger wrong word selection from the mental lexicon. The example [20] represents the most common case of MLI in our data. S selected Turkish word "kaşık" instead of spoon when speaking with a native I of English. Here are the examples from the data of this study;

[19] ... I am not umm good at where department [state]…

[20]… [spoiler] [kaşık] one spoon of salt..

2.5.5.7 Repair (RPR)

For the present study, the repair type is suggested because of the need encountered in annotations. While annotating the audio-visual data, some disfluency types did not match with those defined in the literature. They were neither included in the basic *DF* types, nor named as a composition of basic types. Thus, a new category has been suggested for the present study.

RPR is composed of the reparandum, the suspension point and the repair part. Some part of the original delivery is repeated and some part of it is repaired. For example, this type may cover repairing the original delivery with some modal words (Example [21]), tense (Example [22]),

new words (Example [23]), prepositions and even changing misselected words with new ones. As it is immediately obvious, this is not a basic type of disfluency at all. It is a hybrid type containing other basic types since decomposition is not possible and not logical.

Examples from this study belonging to *RPR* are given in the following examples;

[21] …or doesn't distract [musn't.distract] peoples life...

RPR

Reparandum Suspension Repair Resumption

[22] …what she want to what she want [wants] to do umm…

RPR

Reparandum Suspension Repair Resumption

[23] …umm they took [they.have] umm…

RPR

Reparandum Suspension Repair Resumption

In *RPR*, *S* monitors himself, recognizes the error and repairs the original delivery. In other words, this type belongs to the overt type of *DFs,* even if sub part stand for a covert *DF*.

2.5.5.8 Prolongation (PR)

If a vowel is pronounced too long or longer than it should be, this situation is named as prolongation (PR). This kind generally occurs in L1. In English, native *S* may be expected to produce this kind of disfluency. Non-native speakers of English in our study did not produce PRs. Thus, this category is not in the scope of the present study.

2.5.5.9 Slip of the Tongue (SOT)

Slip of the tongue is an excellent source of data for understanding natural languages and it has been deeply studied (Dell, 1995; Levelt et al., 2002; Roelofs, 1996). When speaker slips some of the phoneme(s) of a word is transferred to successive word. For example "darn bore" for "barn door" illustrates this phenomenon (Dell, 1995). Slips can be seen in any kind of linguistic unit but it is mostly encountered between words, phonemes and morphemes (Dell, 1995). Slip of the

tongue can occur on both syntactic level and phonological level. Hence, there is no exact stratum for this kind of DF. *SOT* is well known and mostly studied in L1.

Explanatory examples from Dell' studies (1995) are as follows;

[24] "a reading list" → "a leading list" (anticipation)

[25] "black boxes"→ "back boxes" (deletion, omission)

[26] "beef noodle"→"beef needle" (preservation)

[27] "heap of junk"→"hunk of jeep" (exchange)

In our study, some of the learners made phonological and semantic mistakes. The examples [28] and [29] show the phonological slips. All examples were from different subjects and the slips were marked with the brackets.

[28] ... you can add some [species] in it

[29]… [spoiler] kaşık one spoon of salt..

2.5.5.10 Grammatical Error (GE)

Any grammatical error (GE) is accepted as a type of disfluency in L1. But in L2, it is controversial especially for beginner and intermediate learners. *GE* has variations at each level of proficiency since this is an interlanguage issue in L2. In time, some of the *GEs* will decay and some of them will be fossilized. Therefore, *GE* is not a suitable type of disfluency for learners. Advances learners of L2 may be checked in terms of *GE*.

GEs are accepted as competence errors in L2. On the other hand, the present study focuses on performance errors (Selinker, 1972). To sum up, *GE* is not in the scope of this study. Our subjects are learners of English and the focus is on their performance errors, not competence.

2.6 Memory, Dual Coding Theory and Concreteness Effect

There are explanations for certain performance and competence related language phenomena in psychology, especially in the memory related studies. The linguistic factor in the present study; i.e., topic of conversation, also has contributions from the field of psychology. The burden of

concepts in the cognitive system has been scrutinized in terms of concreteness and abstractness. Known results about concreteness and abstractness in the linguistic field was supported by memory theories, too. Paivio (1971; 2001) demonstrated that encoding and retrieving concrete words was faster and more accurate than abstract ones because concrete words were also coded both in visual imagery as well as verbally. That is connection of concrete words with verbal and non-verbal codes have more advantage than abstract ones having connection only with non-verbal codes (see

Table 1). This dualism of coding for concrete concepts are named as *concreteness effect* or *dual coding theory* (Paivio, 1971). According to dual coding theory, if one activates concrete concepts and the related imagery, their connection and interrelated verbal words and concepts are also activated. This bunch of activation eases the retention of the concrete words from memory. On the other hand, abstract words and concepts merely trigger verbal codes. Thus, activation of abstract concepts spreads less than concrete concepts in the brain. Less activation results in hard retention.

Table 1 - Representation of concepts coding in the Mind (Sadoski & Paivio, 2001)

	Imagery	Verbal
Picture	+++	++
Concrete Word	+	+++
Abstract Word		+++

Table 1 represents the coding approach of pictures, concrete words, and abstract words. "+" sign is used to represent the coding ratio in the faculties of the brain –the imaginary and verbal parts of the brain. Pictures are coded in the imaginary and the verbal part of the brain very strongly. Concrete words are also coded in the imagery and the verbal part of the brain. On the other hand, abstract words are merely coded in the verbal part.

2.7 Psycholinguistic Theories of Gesture Production

When people speak, they often spontaneously produce gestures and those gestures are typically used to indicate or present objects and ideas. In other words, speakers use gesture in order to communicate better (Alibali, Kita, & Young, 2000).

Gesture is different than language since gesturing interconnects to personnel issues. It is another communication channel used by human beings. The variations and differences are extremely different as in speech. People use gestures for reducing tension. It is even used to be more fluent and more presentable. Its aspect related to fluency has a theoretical connection to the present study and it will be discussed below.

Occurrence of gesture in speech is another field of study. Why do gestures have a particular place in accompanying the speech (Krauss et al., 2000). Krauss, Chen and Gottesman (2000) have a partial answer for the occurrence of gesture in speech. They present information about the origin and function of gestures.

Kraus and his colleague's model of gesture production accompanying Levelt's Speech production will be presented in detail but before that, gesture categories and their properties will be given. Finally, the relationship of gestures with disfluency and the effect of gesture will be discussed.

2.7.1 Type of Gestures

Krauss and his colleagues (2000) divide gesture into four main categories. These are symbolic gestures (SG), deictic gestures (DG), motor gesture (MG) and lexical gestures (LG). All gesture types have both unique and shared properties. Their differentiation generally depends on the functional properties of the gesture.

2.7.1.1 Symbolic gestures

Gestures including hand configuration and movements having widely-recognized conventionalized meaning are named as symbolic gestures (SG). Interchangeable terms for this type are emblems, autonomous gestures or semiotic gestures. In this study, they are labeled as *SG*s.

*SG*s generally occur in the absence of speech. It is when distance or noise exists or there is an obstacle to communicate vocally (Krauss et al., 2000). It is probable to see this category accompanying speech.

Making a roof shape to mean a house or holding a handle shape to mean a cup is generally accepted as *SG*s. This kind of gesture has a direct relation to uttering words or speech. In the study, all concrete objects gestured by *S*s are classified as *SGs*.

2.7.1.2 Deictic Gestures

This class of gestures is generally related to directional issues. The index finger is the main identifier of this class. When the index finger is used to indicate a person, place, object or location, the gesture should be classified as a deictic gesture (Krauss et al., 2000). This class of gesture has been seen accompanying speech. Using the word "right" in speech and pointing to the right side with the hand or the index finger is a good example for this type of gestures.

2.7.1.3 Motor Gestures

This class of gestures should be discriminated from others by features such as occurring repeatedly, being rhythmic and lacking semantic relation to speech. Generally they are composed of successive and meaningless hand movements. That is, they accompany speech but they have no semantic relation to the content or topic (Krauss et al., 2000).

They are also named as batons and beats in the field. It is known that motor gestures have coordination with speech prosody and stressed syllables (Bull & Connelly, 1985). Their occurrences exhibit rhythmic behaviors.

2.7.1.4 Lexical Gestures

Besides these three types of gesture, there is a fourth category which is not an easy type to define. There is disagreement among researchers on the functions and origins of lexical gestures (Krauss et al., 2000). The lexical gesture would cover gestures related to concepts and actions in the speech. In fact, symbolic gestures and deictic gestures are also sub-classes of lexical gestures. Hence, they are lexically and semantically related to speech. In the study, emblems and direction related gestures are marked as *SG* and *DG* respectively. The rests of the gestures having a semantic relation to speech are marked as *LG*. For example, speaking about walking

and using two fingers to animate of walking is a kind of *LG* .Another example is animating a circle with the hand to specify a property of a sauce pan. This is a *LG* type of gesture.

The last example above is a unique example as one may mark it as a SG. But representing a property of a concrete object is not a SG, it is a LG. This subtle difference must be considered while identifying gestures and their properties.

2.7.2 Functions of Gesture

The traditional view is that gestures have a function in communication. Despite this, there is very limited research for gestures, their functions and origins (Finlayson et al., 2003; Krauss et al., 2000). There are two known important functionalities of gestures: tension reduction and lexical retrieval.

The first main function of gesturing is the tension reduction, which is interesting for the present study because gestures can be used to reduce tension when there is a deviation in the speech (Finlayson et al., 2003; Krauss et al., 2000; Levelt, 1989, p. 35). Gestures let *Ss* take the time; and gestures hold the communication channel open between *S* and *I*. Gestures convey some messages to I about the situation of speech. These messages are always considered as a positive contributor to discourse (Alibali et al., 2000; Krauss et al., 2000).

The second main functionality of gestures is the lexical retrieval (Krauss et al., 2000). When *S*s have the problem in collecting words from the lexicon, they use gestures for both tension reduction and easing the lexical access. The relationship of lexical access and gesture is also an issue. There are studies providing evidence for lexical access functionality. The results show that preventing gesturing of the speaker decreases the fluency of speech (Alibali et al., 2000; Finlayson et al., 2003; Krauss et al., 2000) .

In the next part, the production of speech and its connection with Levelt' speech production model will be presented.

2.7.3 The Production of Speech and Gesture

The three stages; i.e. conceptualizing, formulating and articulation, are important stages for gesture production (Krauss et al., 2000). Because of similar properties of gesture production and

speech production, Krauss and his colleagues proposed a gesture production model cooperating with Levelt Speech production model. This model is serial-modular, too. In addition, like the base model-parallel processing is executed through icrementality. For the relation of speech and gesture, parallelism is assumed as well.

As Levelt, Krauss assumes that at the conceptualizing stage, its functions are in propositional form and only process propositional data. On the other hand, the knowledge in the mental lexicon that constitutes a source concept should be in the form of propositional and non propositional formats. Generally, the idea is that a concept is made up of a set of features that are encoded in propositional and spatial formats. Some features are encoded both in proposition and spatial and some other features are in the intersection set. To process this concept for speech, the non propositional information of concept must be translated into the propositional form. Then they could be processed but gestures are more related to non propositional features. Krauss' (2000) central hypothesis is that *lexical gestures derive from spatial (non propositional) representation of the concept.*

In the model, there is a spatial feature selector that monitors the conceptualizer to select the spatial features. Then the spatial feature selector transforms the information stored in spatial format into a set of abstract properties of movements. This transformed information is also translated by a motor planner into a motor program. This program consists of a set of instructions for executing lexical gestures. The output is lexical gestures, which are composed of gestures and body movements.

How does lexical gesture help lexical access? When concepts are activated in memory to push the conceptualizer, their spatial features are also activated. Thus, both the speech production and the gesture production systems start to function in parallel. When the speech production system has a problem of access or some latency, the phonological layer gets some input from the kinesic monitor because it already knows the features of the activated concepts. This input activates the concepts and their interconnected concepts live longer in the memory. Thus, it may be probable that the more activated is selected first. This whole process does not have detailed information about how the phonological encoder feeds the grammatical encoder and when the feeding from the kinesic motor occurs (Krauss et al., 2000). This whole process works on cross-model priming, i.e., Speech Production model and Gesture production model of Krauss work parallel (see Figure 6).

Krauss' Model of Gesture production has both theoretical and empirical models similar to Levelt's Model of Speech Production; however, both models leave some parts unclear.

Detailed information of the gesture production model is beyond the scope of this study. It is important that gesture facilitates access to lexical items. This is already shown by Krauss (2000). There is a theoretical relation between gesture and disfluency.

To sum up, Krauss' model is the best framework for gesture production since it has a parallelism to Levelt's speech production model. It functions over propositional information at the conceptual level. All related propositional information is converted into gestural movements.

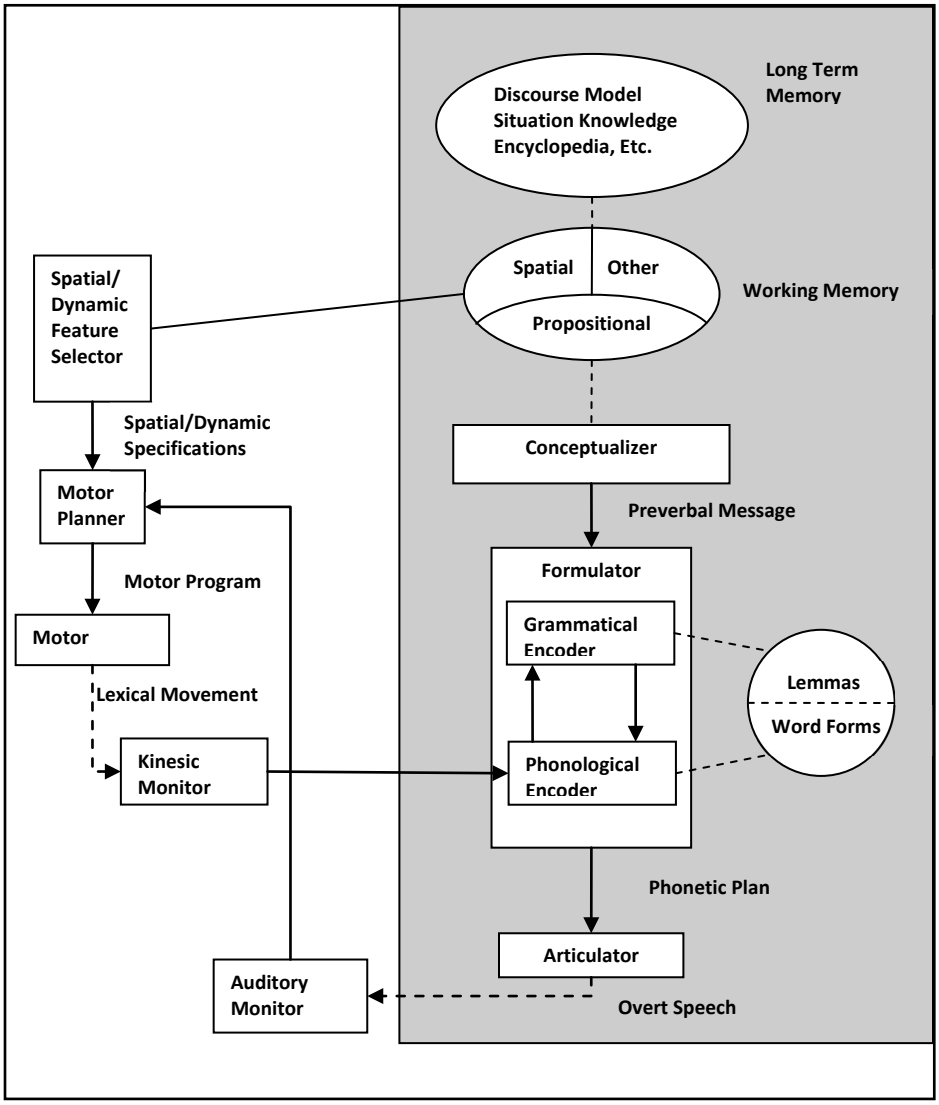


Figure 6 - A cognitive architecture for speech-gesture production process (Krauss, Chen, & Gottesman, 2000)

2.8 Metalinguistic Awareness

Metalinguistic awareness, ability, is a model to explain the interaction between languages (Bialystok, 1990). Metalinguistic awareness (MA) is defined as an awareness of linguistic form and structure in order to consider how they relate and produce underlying meaning of utterances. In other words, *MA* is the ability to analyze a language and its components as an object (Bialystok, 1990).

The model is based on bilingualism and second language acquisition. According to the model, learners may have or transfer knowledge from L1 to L2 but using this knowledge in L2 is mostly inconvenient since learners may lack control on language (Bialystok, 1990). In L2, lack of control on the language might lead to deviations in speech as well structure. It affects language production in some undesired ways. We assume that the longer speech, the more loss of control over the language and confusion of the meta rules of L2 occurs. This phenomenon affects overall process of speech production and produces deviations in speech.

CHAPTER 3

METHODOLOGY

This part includes information about subjects, environmental settings, annotations and annotation principles.

3.1 Subjects

Eighteen subjects took part in the study. Sixteen subject's data were used. We discarded two of them because they were not found qualified by the native speaker .This is because the subject only uttered words and meaningless sentences. Five subjects were female and the rest of them were male. All were aged between 17 and 19 and they were from Department of Basic English (DBE) at Middle East Technical University (METU) at the upper-intermediate level. All subjects were selected from the same age and the same level to understand possible differences in disfluency production (Menyhárt, 2003). Language proficiency and their educational background are nearly the same. They all graduated from high school. They just took the university entrance exam. They had been at METU about six months. They were from various departments. None reported a history of auditory or communication problems. They volunteered to take part in the study.

Three instructors from DBE and one instructor from Modern Language Department of METU took part as *I* in the study. Two of them were native *S* of English. The native *S*s were male. The rest of them were non-native speakers of English. The mother tongue of the two non-native *S*s were Turkish. Non-native *S*s were female. This was important for this study because it is known that gender of *I* is not important (Bortfeld et al., 2001). None reported a history of auditory or communication problems. *I*s volunteered to take part in the study.

3.2 Topics

Several pre-experiments were performed to choose the most suitable topics for the study. A candidate group of topics was given to random subjects and were asked to speak about the given topics. Each conversation was recorded and analyzed in terms of the number of the abstract and concrete words. Four topics were separated from the rests because they elicited high number of either abstract or concrete words. These topics are below:

- making spaghetti (concrete)
- giving directions from DBE to the research center which is near the Electrical Engineering Department at METU. (concrete)
- spring festival and possibility of midterms clashing with the spring festival next year (abstract)
- effects of religion on our lives. (abstract)

We considered the concrete topics to be one of the well known and familiar dialogs in the second language learning classes at METU for learners. Similarly, spring festival and religion topic are very familiar for students in METU. Concrete group was considered for eliciting words like spoon, building, tree, road and etc. On the other hand, abstract group was considered for words such as feel, idea, thought and etc. Although, the concrete topics are more objective, the abstract topics are more subjective.

Each subject talked about four topics with the assigned *I*. *I* selection was done on the basis of the study design in

Table 2. All participants spoke about all topics in the following order: making spaghetti, giving directions, spring festival possibly clashing with midterms, and effects of religion on our lives.

Table 2 - Study Design for Dependent Variables

	Concrete Topic	Abstract Topic
Familiarity/Non-Familiarity	X	X
Native / Non-Native	X	X

To reduce the difference in the length of time for planning (Mehnert, 1998), none of the topics had been mentioned to the participants before the conversations. Each topic was introduced by *I* and then the subject and *I* started to talk about it. Thus, there exists no preplanning time. Each subject spoke with *I* without planning time. Role of *I* was basically to start the conversation and let the participant talk. He talked very little after their initial opening sentences.

A warm-up stage is important to get students in the mood of the class before the actual activity starts. In the present study, a "warm-up"has not been planned originally. However, when consulting the *I*s who are experienced language teachers and analyzing the data we decided to eliminate the first topic –making spaghetti- as it appeared to function as warm-up. When we did that, we had to exclude one of the abstract topics to equalize the number of the abstract topics and concrete topics. We eliminated the data on "spring festival at METU and possible clashes with mid-terms". This left us with one concrete topic – giving directions – and an abstract topic – the effect of religion on our lives.

3.3 Environment

The experiments took place in DBE rooms. Three straight-backed armchairs were placed in the room. Students and *I* were facing each other about 1m apart. A digital video camera was placed near *I* and student to record both of them. In addition, a voice recorder was used to record the conversation.

The first native speaker spoke with five students. The first native speaker was familiar with two students out of five. The second, *I* spoke with three non-familiar students and one familiar student. The third *I* spoke with five students who are all familiar. The fourth *I* spoke with two non-familiar students (please see

Table 3 for details). Familiarity was reciprocal, i.e. *S* and *I* know each other or not.

Table 3 – Interlocutor and the subject design

	Familiar	Non-Familiar	First Language
Interlocutor-1	2	3	English
Interlocutor-2	1	3	English
Interlocutor-3	5	0	Turkish
Interlocutor-4	0	2	Turkish
Total	8	8	

3.4 Tools

3.4.1 Elan

Elan, a multimedia Annotator, was used to annotate the audio-visual data. It is one of the language tools of Max-Planck-Institute for Psycholinguistics. It is a functional tool which allows the user to add unlimited annotations to the audio or video stream. By this extended property, annotations could be created on multiple layers and they are linked and grouped. Figure 7 exhibits the ELAN user interface environment.

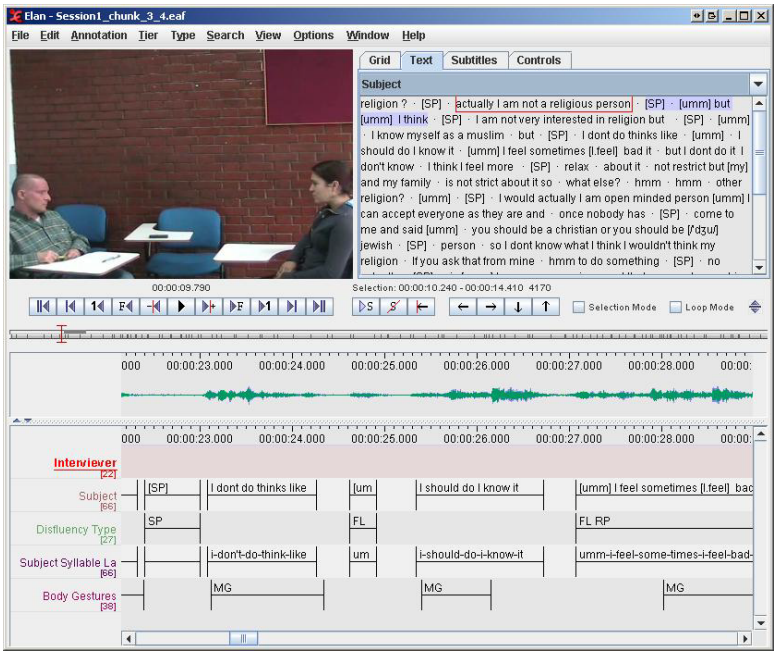


Figure 7 - A view from Elan

The project files of Elan[2] are stored in XML format, which is one of the universal storage formats for computer science. It is an open source and it is both editable and upgradeable.

2 Further information could be gathered from Elan site under "the Language Archiving Technology portal" at http://www.lat-mpi.eu/tools/elan.

3.4.2 Extension to ELAN

By ELAN XML file format, it is easy to automate some works which require too much time, such as creating a syllabification layer of subjects or counting syllables in a layer etc. For all these purposes, custom software was coded for needs. Several new functionalities were used with the help of the custom software. These new functionalities are described below. The First one is "Syllabify layer" [3]. It reads any meaningful layer and creates a syllabified form of the source layer (see Figure 8). The extension does not have a morphological automaton or parser for English. It is a kind of crawler that reads words from the source layer and finds their syllabification form in online-dictionary sites and pushes the result to the target layer.

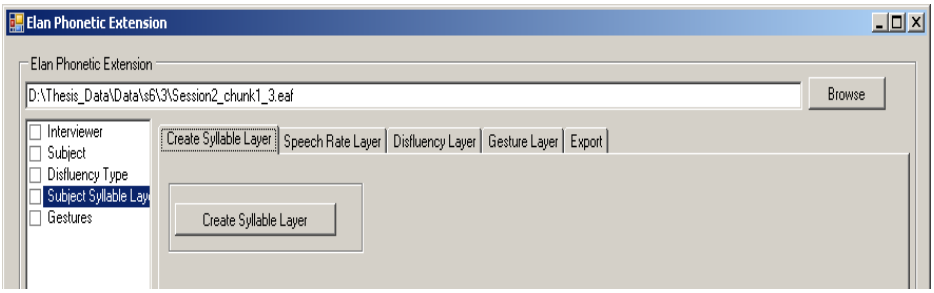


Figure 8 – Custom Software – Syllable Layer Create Module

"Calculate Total Duration" is the second one used for summing all of the speech duration of *S* or *I* in one session. It has one parameter used for including or excluding *SP*s in the calculation of total duration (see Figure 9).

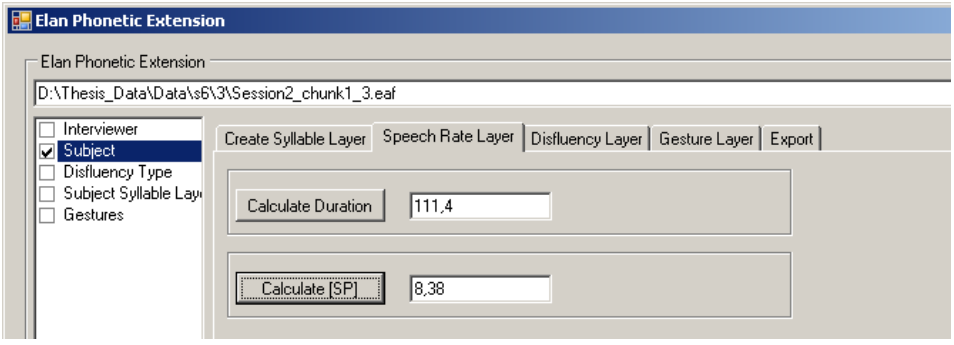


Figure 9 - Calculating Total Duration

3 "Layer" is used to represent a speech or the categorized data of the whole session.

Third, “Calculate Counts of Syllables or Words” is a counter for all of them. It calculates the total number of syllables and words in the selected layer. It has a parameter for including to or excluding FLs from the total counts (see Figure 10).

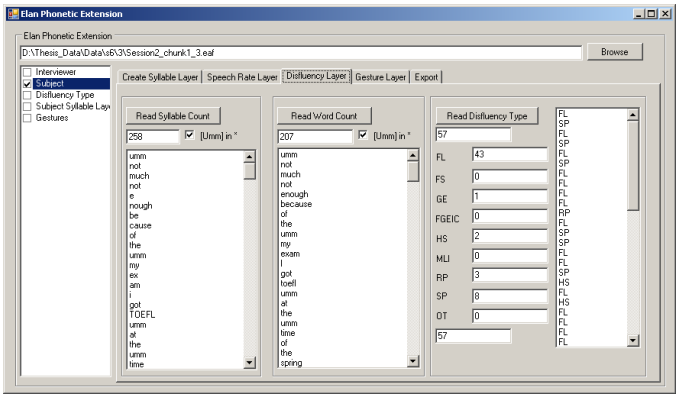


Figure 10 - Calculating Syllable Count, Word Count, Gesture Types Count

Finally “The total count of gestures” and “The total Count of Disfluency” are counters. They count total numbers and also they count the occurrence of gesture types and disfluency types (Figure 11). “Export” function of extension is used to export any layer in the ELAN data file instead of the export function of ELAN (see Figure 12).

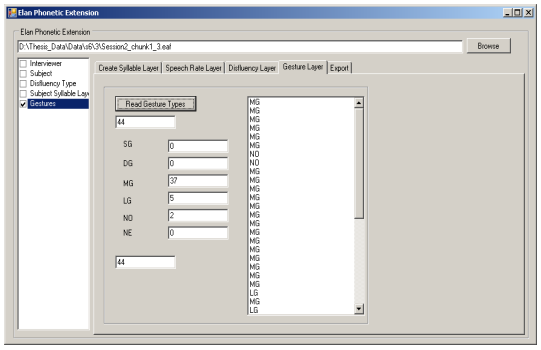


Figure 11 - Gesture occurring in a session

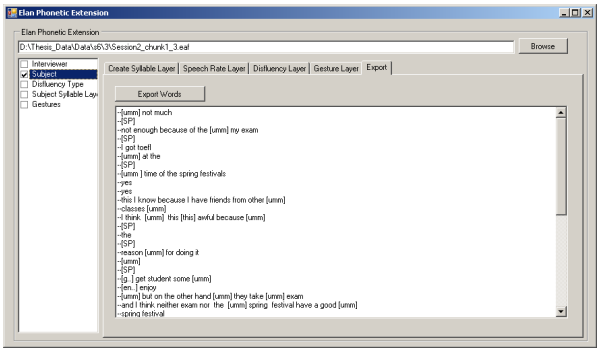


Figure 12 - Export Section

3.5 Annotation

By the end of the experimental part of the study (i.e. conversations) seventy two units of conversations were recorded. They are equal to 210 minute long audio-visual data. The recording was chunked into topics. And each such unit was annotated. For example, the first concrete topic of second subject was one of these units.

In ELAN, the speech of *I* was annotated on one layer without any details. The learners' data was detailed on four layers. These are "Subject Layer", "Disfluency Type Layer", "Subject Syllable Layer" and "Body Gestures Layer". The creation order of layer is not important since Elan has a function that can drop and drag the order of the layers freely.

"Subject Layer" contains sentences and disfluency markings. In this layer, some types of indicators were used. These are [SP], which is the indicator of the silent pause, [umm], which is the indicator of the fillers and brackets [] for emphasizing the reparandum of disfluencies. The detailed information for annotation will be given in the annotation rules (see Figure 13). For a detailed example please see APPENDIX C.

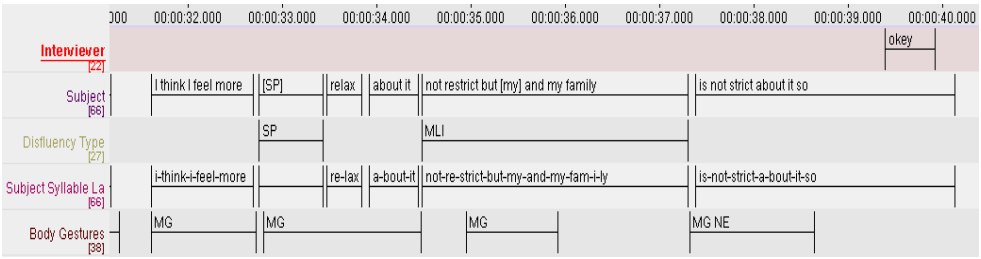


Figure 13 – Sample view from an annotation in the ELAN.

"Disfluency Type Layer" is the identification layer of the disfluencies. Detailed information about the "Disfluency Type Layer" will be given in the next part. "Subject Syllable Layer" is the syllabified form of the "Subject Layer". This layer is the source of information for syllable per second, syllable per word and disfluency per syllable. This layer was used to normalize the data on the basis of the idea that subject's speech rate and length of the speech need to be equalized for the purpose of coherent statistics. This layer was created by the custom software as described in the "Extension to ELAN". "Body Gestures Layer" includes gestures of the subjects. There are six gesture types annotated. The types of gestures include hands and some specific types of nodding and negation gestures. These will be explained in detail in the next sections below.

3.5.1 Annotation Rules

Basically two different types of data were annotated. The first one is the articulated elements and the second part is the gestural elements. In the next two sections the annotation procedure and annotation rules will be described.

3.5.2 Annotation of Utterances

There are lots of difficulties involved in annotating disfluencies because the data is speech data and words and sentences do not have clear boundaries as in written language. Thus, we prepared guide-lines for annotating disfluencies. The subject and disfluency layer was annotated according to the below principles.

Chunks were selected according to the oscillation of the wave spectrum of the speech (see Figure 14), namely a chunk includes all elements between two stopping points of a speaker. This eased the annotation. Otherwise, it should be done on the basis of sentences or phrases but in the online speech the start and the end of the sentences were very difficult to find. It was hard to understand which sentences were left uncompleted and also chunking the speech into phrases or sentences was not a requirement for this study. The study was interested in the frequency of disfluency and gestures. It should not change the result in whatever way they were chunked. In other words, segmentation informal method was used for the annotation purposes. It aims to ease the annotation of the speech into more comprehendible part. The segmentation is not related to disfluency or gesture identification issues.

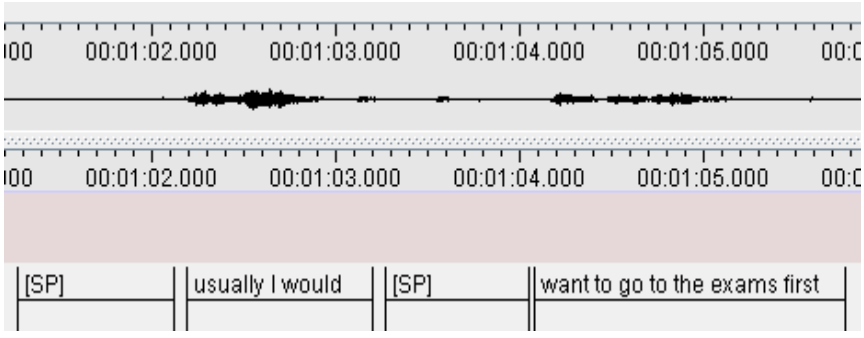


Figure 14 – An example of annotation for harmony of speech.

Silent pauses were annotated separately. The [SP] was annotated on “Disfluency Type Layer”, if the duration of [SP] was greater than 0.5 seconds. The silent stages which occurred while taking the turn of speech were not annotated because they were not disfluencies.

Filled pauses were annotated as other linguistic elements in the annotation chunks (see Figure 15). In the “Disfluency Type Layer” fillers were abbreviated as *FL* and shown as [umm] in the “Subject Layer”

Figure 15 – An annotation example including a filler [Umm] and repetition.

Hesitations and misselected lexical items were annotated as fillers. In the “Disfluency Type Layer” they are abbreviated as *HS* and *MLI* respectively. In the “Subject Layer”, *HS*s were marked with curly brackets including uncompleted words and *MLI*s were marked with brackets including words.

Repairs (RPR) and repetition (RP) were also annotated as *FL*s. They are marked with curly brackets but their reparandum –repeated or repaired part- were marked only. The Rest of the components – original delivery and suspension point – were left unmarked.

3.5.3 Annotation of Gestures

Gesture annotation was another aspect of the study. To discriminate gestures and to follow the continuum of the gestures was not easy. All gestures were identified according to their definition and properties in the literature. They were annotated on the “Body Gesture Layer”. Gestures do not have any annotation rules as disfluencies do. They were annotated in terms of the definition given in section 2.7.1. Examples from the subjects will be explained in the next sections.

3.5.3.1 Deictic Gestures

Figure 16 represents classical deictic gestures (DG). The learner used the pointing gesture to enrich the communication while uttering word “left”. Deictic gestures were abbreviated as DF in the “Body Gesture Layer”.

Figure 16 - A learner gesturing a deictic gesture - "Left"

3.5.3.2 Symbolic Gestures

In Figure 17, the learner is making a symbolic gesture (SG) by drawing a circle in the air while uttering the word circle. Symbolic gestures were abbreviated as SG in the “Body Gesture Layer”.

Figure 17 - An example for Symbolic Gesture

3.5.3.3 Motor Gestures

This type of gesture includes repeated and meaningless hand movements named as motor gestures (MG). All meaningless hand movements which have no relation to the topic were marked as MG in the "Body Gesture Layer".

Figure 18 - a motor gesture example

3.5.3.4 Lexical Gestures

In Figure 19 and Figure 20, the learner is making a lexical gesture for "pour off" at two steps. These gestures are semantically related to topic (making spaghetti). And this type of gesture is the main concern of the present study. All gestures in this category are abbreviated as LG in the "Body Gesture Layer".

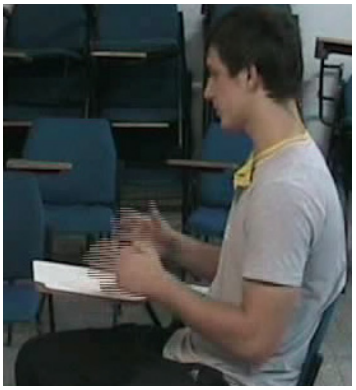


Figure 19- first part of a lexical gesture

Figure 20 - second part of a lexical gesture

3.5.3.5 Nodding and Negation

These types are shaking head for nodding (NO) and negation (NE). The subtle difference is the emphasis. For nodding, the emphasis direction is through front, and for negation; the emphasis direction is to back.

3.6 Assessment of the Data

We wanted to understand whether or not less fluency in speech affects the acceptability of speech by native speaker. After collecting the data, the first and the fourth topic of each subject were examined carefully with a native speaker. Native Speaker was one of I and he didn't see the annotated data; he watched and evaluated each subject's speech by selecting one of the options provided. These proposals were:

- The student communicates his/her ideas very well
- The student communicates his/her ideas not so well
- The student cannot communicate his/her ideas at all

He was also asked to score each speech in terms of fluency, accuracy and coherence out of 10. The results are given in Table 37 and Table 38 in CHAPTER 6 Appendix B. These tables give the assessment of just one native speaker. This information was used to validate subject's data for the present study and also it was used to decide and understand critical issues; such as types of disfluencies for the present study, coherence of topics, warm-up issue and its effects and the acceptability of the speaker speech by native speaker.

3.7 Raw Data

After annotations, the following calculations were made:

- total duration
- total SP duration
- total syllable count

- total word counts
- total disfluency count
- total gesture count
- syllable per second
- disfluency per second
- disfluency per syllable
- frequency of disfluency types
- frequency of gesture types

The count of total syllables and total words and speech rate were calculated. In the calculation we add FLs in the total count of words and syllables and we named these variables with prefix "including FL" and also we calculated another version of these variables which are not including FLs.[4]

3.8 Sampling of Gesture and Disfluency Interaction

To check whether each *DF*s in the data to see whether it intersects with a gesture is a time consuming event. Because of that we created a procedure sampling from our data randomly and we calculated the dual occurrence of disfluency and gesture on the basis of the below sampling method since sub units of the population may represent the properties of a larger unit.

Sampling was required to find the relation between gesture and disfluency. This relation is to calculate the temporal occurrence of the gesture and disfluency. From each conversation, five disfluency occurrences were selected randomly. The randomizing procedure was based on total disfluency number. Each subject's total disfluency was divided by five and the decimal part was ignored. Then this value was taken as a multiplier (abbreviated as *m*). Then formula 1+nm was used to select the disfluency occurrence in the conversation where n is the sequence number from five. Table 4 shows an example to calculate which disfluency occurrence was selected.

[4] We calculated speech rate considering FL and excluding FL. We wanted to normalize the duration of conversations but normalization affects the result because duration was also affected by variables.

After applying this procedure to all, eighty selections were made. All of them were checked and controlled, to see whether one of them was in the intersection by a gesture or not. If it was, then it was scored as 1. At the end of the procedure, the total number was calculated out of eighty. It was converted to percentage. It was applied to both concrete and abstract topics. Then the overall temporal occurrence or the probability ratio of gesture production occurring during disfluency was calculated.

Table 4 - Randomizing Procedure

Total Disfluency	Integer Division of 5 Multiplayer (m)	First Selection 1 +(n)(m) n=1,m=7	Second Selection 1 +(n)(m) n=2,m=7	Third Selection 1 +(n)(m) n=3,m=7	Fourth Selection 1 +(n)(m) n=4,m=7	Fifth Selection 1 +(n)(m) n=5,m=7
39	7	8	15	22	29	36

3.9 Reliability of Annotations

Annotation was done by one person. It was checked by two mechanisms. First one is annotating a sample by different person due to annotation guide-lines and it was checked by original one whether there was a problem in the annotations or not. Second one is the basic one. Annotations were checked by going over the current annotations and discussing problematic parts. Twenty percent of the data was controlled and in this way we tried to eliminate errors from the data by second method.

3.10 Goals of Study

There were two main goals of this study

a) To investigate the effects of the following issues on *DF* and gesture production in L2 (English) in terms of

   i. topic type that is abstract or concrete topic,

   ii. familiarity of *I*

   iii. Property of *I*: native English speaker vs. non-native English speaker

b) To investigate the relation of *DF* and gesture production and to make a generalization about their relation.

3.11 Hypothesis

According to goals of the study, the results and discussions will be based on the following hypotheses:

- The subjects conducting a conversation with a familiar *I* will speak more fluently than the subjects conducting a conversation with a non-familiar *I*.
- The subjects in the concrete topic condition will speak more fluently than those in the abstract topic condition.
- The subjects conducting a conversation with a familiar *I* on the topic eliciting abstract words will speak less fluently than those on the topic eliciting concrete words.
- There will be a positive correlation between the subjects' rate of *DF* and gesture production.

CHAPTER 4

RESULTS

This study was specifically concerned with the differences between the student fluency and gesture production in the abstract and concrete topic conditions under variation of the factors: familiarity vs. non-familiarity and speaking with a native speaker of English or a non native speaker of English. In order to reach this purpose, conversations of sixteen students on the abstract and concrete topic were video-typed and annotated using the ELAN software (see section 3.4.1)

DF and gesture production were analyzed with respect to topic types, and *I* properties – native speaker vs. non-native English speaker. Unfortunately, one big analysis did not provide understandable conclusions since the interaction of variables are too complex and the complexity hide some important facts. Therefore, data were analyzed step by step.

The next section will present overall results of *DF* and gesture, a detailed analysis of length, gesture and *DF*. Gesture and *DF* correlation will be presented. In the analysis, our dependent variables are frequency of *DF*, gesture and conversation duration. On the other hand topic types-abstract concrete-, un/familiarity and speaking with non/native English speaker are our independent variables.

4.1 Average Results

Table 5 exhibits the arithmetic mean of the results. These results are frequency results of some of the most important counts for the present study.

All results must be considered row by row independently and magnitudes are in their unit. According to the results subjects spoke more than the interlocutor in both topics. The abstract topic had greater magnitude in syllable, word, *DF* and gesture count.

Table 5 - Average Values of Duration, Word and Syllable Count, Gesture and Disfluency Count

	Concrete	Abstract
Interviewer Duration	47.10	37.68
Subject Duration	64.02	88.43
Total Disfluency Count	32.31	42.25
Total Gesture Count	24.44	30.13
Total Uttered Word	148.25	198.56
Total Uttered Syllable	185.38	255.13

4.2 Disfluency Overall

All the data related to *DF* satisfied our basic assumptions of analysis which are normal distribution of data and homogeneity of variance.

It was expected that some *DF* types may be more dominant than others and some of them might be very limited and few and some of them might not be observed in the data at all. The results show that two categories of the *DF* were dominant in the data. These are *FL* (Percentage $_{Abstract}$ =45, Percentage $_{Concrete}$ = 45) and *SP* (Percentage $_{Abstract}$ =41, Percentage $_{Concrete}$ = 41) in both abstract and concrete topics. They have same percentage.

Figure 21 represents the overall *DF* distribution for all the subjects in the concrete topic condition. Figure 22 exhibits similar results for the abstract topic condition. Only difference in abstract topic is the percentages of the non-dominant types: *RP*, *MLI, HS*, *RPR* and *FS*. Variation of non-dominant types is affected by topic type and other factors.

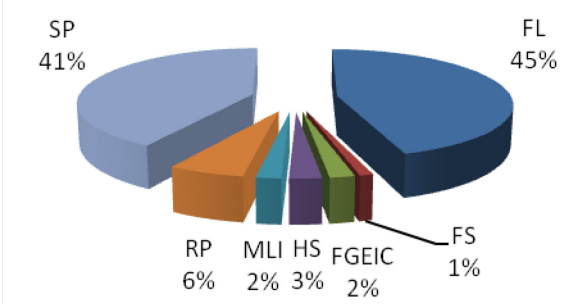


Figure 21 - The Overall Disfluency Distribution of the Concrete Topic (Topic: Giving Direction)

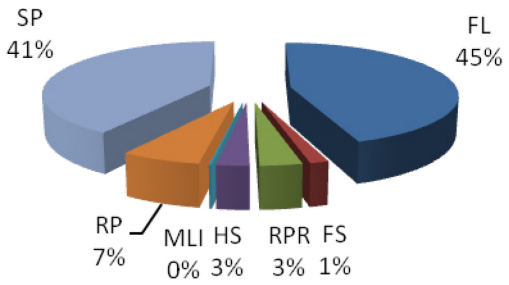


Figure 22 - The Overall Disfluency Distribution of the Abstract Topic (Topic: Effects of Religion)

Results show that *PR*s were not observed in the present study and *MLI* vanished in the abstract topic. *RPR* type was observed only in the abstract topic. In the concrete topic, this hybrid type was not observed.

4.3 Gesture Overall

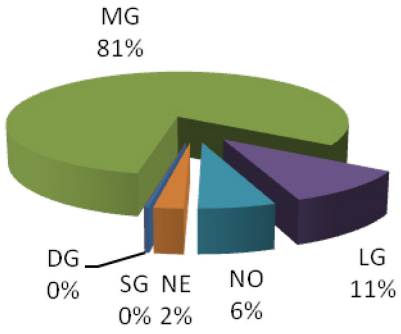


Figure 23 - Abstract Topic Gesture Distribution

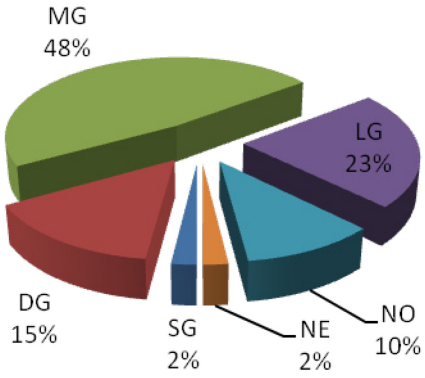


Figure 24 - Concrete Topic Gesture Distribution

The overall gesture distribution is exhibited in Figure 23 and Figure 24. Figures show that the gesture distribution is different in the two the topics. Most common seen type was *MG* which is successive and rhythmic meaningless hand movements. Second common was *LG* and its occurrence varies with the topic. The *DG* type has different behavior in the concrete case. It was

used by the speakers in the concrete topic – giving direction– and its usage vanishes in the abstract case since there is no need to use them.

4.4 Length of Subject Talk

The results did not straight forwardly support all the hypotheses because there is a complex interaction among the dependent and the independent variables. Firstly, the duration of conversation varies with respect to the topic. The repeated ANOVA shows that duration of concrete conversation is shorter than that of the abstract topic. This implies that fewer numbers of syllables and words was produced in the concrete topic. Thus, the duration of conversation was significantly affected by the topic type, $F (1, 12) = 9.78$, $p < .05$. Table 6 shows that duration of concrete conversation is less than the abstract one ($Mean_{Concrete}$=67.22, $Mean_{Abstract}$ =94.75).

Table 6 - The Descriptive Information of Topic Type in Repeated ANOVA

Topic Type	Mean	Std. Error
Concrete	67.22	5.54
Abstract	94.75	10.56

Secondly, the results of repeated ANOVA shows that the interaction of topic type and familiarity with *I* affects the length of conversation, $F (1, 12) =8.24$, $p < .05$. Type of topic and familiarity interaction varies significantly. That is, the subject speaking with a non-familiar *I* on concrete topic spends less time. It is the same for *S*s speaking with a familiar *I*. On the other hand, the variation of the duration between concrete and abstract topic is not significant in the speech with a familiar *I*. It may not be generalized that *S* spends less or more time on the abstract topic when speaking with familiar *I* (see Table 7).

Table 7 - Descriptive Information of the interaction of Familiarity and Topic Type

Familiarity	Topic Type	Mean	Std. Error
Non Familiar Interviewer	Concrete	70.70	8.26
	Abstract	123.54	15.74
Familiar Interviewer	Concrete	63.75	7.38
	Abstract	65.96	14.08

Thirdly, it was found that familiarity also affects the duration of subject's talk whether or not other independent variables vary. *S*s, speaking with a familiar *I*, spoke less than *S*s speaking with a non familiar *I*, $F(1, 12) = 5.30$, $p < .05$ (see Table 8).

Table 8 - Descriptive Information of Familiarity on duration between subjects

Familiarity	Mean	Std. Error
Non Familiar Interviewer	97.11	10.72
Familiar Interviewer	64.85	9.59

The overall results of length are graphically represented in Figure 28 and Figure 29 in Appendix A.

4.5 Correlations

The relation of duration and gesture added another dimension to the present study. Comparing data from repeated subject design[5], the duration was affected from controlled variables (topic type, familiarity and native English *S*) but it was not expected to be highly dependent on them. And it was not expected to have high correlation with other depended variables (*DF* and gestures).

The results show that length and total number of *DF* are highly correlated, $r = .921$ where $R^2 = .848$, $p < .01$. Thus, 85 % of the variation of duration can be affected from total disfluency in

5 Analyzing of same subject in two topic conditions.

the concrete topic - giving direction. That is, there is a positive high correlation between duration and disfluency (see Table 18 at Appendix A).

Similarly, in the abstract topic condition, duration of conversation and total disfluency correlation is high, r = 0.854 where R^2=.729, p <.01.Thus, 73 % of the variation of length can be affected from total *DF* (see Table 19 in Appendix A). Comparing the performance of *S* in both the concrete and the abstract topic condition, it is clear that there is a high correlation between duration and *DF*.

Although speech rate, total syllable count, and word count were calculated, none of them worked in the main statistics since they are related to duration and the duration was a dependent variable which had been not predicted. To be sure about this issue, the repeated subject design was converted to between subject designs[6]. Groups were created randomly. In between subject design, duration was used as covariate[7].

Like the repeated design, results show that there is a high correlation between duration and *DF* in two distinct randomly selected groups, r= .892 where R^2=.795 p<.01 (see Table 20 at Appendix A).

Results also show that there is no correlation between topic type and *DF* if we assign the duration as covariate, r= .17. But it is clear that there was a correlation.

To sum up, consolidation of all the results implies that there is a relation between duration and familiarity type, topic type and interviewer type. The duration is definitely affected by familiarity, topic type and interviewer type. We also found that there is a highly positive correlation between DF and duration (see

Table 21 at Appendix A).

4.6 Main Analysis

6 Separating the data into two groups and comparing the duration variation of these two groups in terms of independent variables.

7 A covariate is a variable that is possibly predictive of the outcome under study. A covariate may be of direct interest or be a confounding variable or effect modifier (Dictionary.com, 2007).

The aim of this study is to scrutinize the *DF* behavior of learners of English under certain conditions. Data from sixteen subjects show that subjects spoke more fluently in the concrete topic than the abstract topic without considering any controlling variables, F (1, 15) = 4.886, p=.43 < .5 (see Table 22 at Appendix A).

Table 9 indicates that variations are very different (*Mean* $_{Concrete}$=32.31, *Mean* $_{Abstract}$ =42.25).

Table 9 - Descriptive information of Disfluency Rate between Concrete and Abstract Topic

Topic Type	Mean	Std. Error
Concrete	32.31	3.18
Abstract	42.25	5.15

For detailed information and a more coherent explanation, the type of the topic, familiarity and the type of the interlocutor were considered in the analysis. The results show that (

Table 23 in Appendix A) topic type has an effect on the DF occurrences, F (1, 12) = 9.947, p=.008 <.05. It is highly significant that the concrete topic has less number of *DF* than the abstract one (*Mean* $_{Concrete}$=34.64, *Mean* $_{Abstract}$ =47.18) (see Table 10).

Table 10 - Descriptive Information of Topic type in the main analysis.

Topic Type	Mean	Std. Error
Concrete	34.64	3.39
Abstract	47.18	3.99

Table 11 - Descriptive information of Topic Type and Familiarity interaction

Familiarity	Topic Type	Mean
Non-Familiar	Concrete	37.08
	Abstract	61.91
Familiar	Concrete	32.20
	Abstract	32.43

The interaction of topic type and familiarity also has an effect on *DF* variations, F (1, 12) = 9.580, p=.009<.05. Further descriptive information emphasizes that (*Mean* $_{Non\ Familiar\ Int.\ +\ Concrete\ Topic}$=32.20, *Mean* $_{Non\text{-}Familiar\ Int.\ +\ Abstract\ Topic}$ =32.43) subjects speaking with a non familiar *I* on an abstract topic produce more *DF* in the speech (see Table 11). The case is similar if *S*s speak with

a familiar *I* on an abstract topic condition but the variation is not as big as in the case of the non-familiar *I* ($Mean_{Familiar\ Interviewer\ +\ Concrete\ Topic}$ =32.20, $Mean_{Familiar\ Interviewer\ +\ Abstract\ Topic}$ =32.43). There is contrasting effect of two groups of *S*s who spoke with a familiar *I* those speaking with a non-familiar *I* independent of whether he is native English or not. This contrast between two groups is shown in Table 24 in **Error! Reference source not found.** where it is shown that there is a highly significant effect of familiarity in the variation of *DF* productions between groups F(1,12)= 3.865 , p=.017<.05. This effect is exhibited in Table 12 ($Mean_{Non\text{-}Familiar}$ =49.50, $Mean_{Familiar}$ =32.32). This emphasizes fact that the familiarity issue has a big effect on *DF* production.

Table 12 - Descriptive Information for Disfluency Rate and Familiarity of Interviewer

Familiarity	Mean	Std. Error
No	49.50	4.64
Yes	32.32	4.15

In addition to the familiarity effect, interviewer type is marginally significant, F (1, 12) = 3.865, p=.073~>.05. Table 13 shows that *S*s talk with a native English *I* more fluently ($Mean_{Native}$ =49.50, $Mean_{Non\text{-}Native}$ =32.32).

Table 13 - Descriptive Information for Disfluency and Interviewer Type

Interviewer Type	Mean	Std. Error
Native	34,792	4,017
Non-Native	47,025	4,753

Although the results were in accord with hypotheses, it is disputable that duration, syllable count and word count were varied. This variation raises unpredictable effects on the results of number of *DF* under controlled conditions. It is a suspicious case for the study. Because of this result, the *DF* per syllable and *DF* per word were calculated and the main analysis was reapplied to them to see whether there is any difference affecting results.

The results show that none of the controlled variables or their interactions had an effect on the variations in *DF* per syllable (see Table 25 in Appendix A). Descriptive information suggested that there was no difference in the variation of *DF* per syllable. All previously found effects vanished. This was an unwanted situation. Because it was expected that considering syllable count or word count would produce more coherent information than previous calculations. This

issue was tested by converting the data into two distinct groups where subjects were selected randomly and the intersection of the data was empty.

The results show that none of the independent variables have a significant effect on *DF* except duration. It is known that duration has a correlation with *DF*, that is why we used it as covariate F (1, 8) =13.458, p=.006<0.05. Only the interaction of *I* type and familiarity exhibits a limited marginal significance, but these were not accepted as a strong effect because the analysis gave marginal significant results about their intersection (see Table 26 in Appendix A).

These results show that this was not only the effect of duration but an effect of independent variables (topic types, familiarity and type of *I*) on the length. This information was not sufficient to make a conclusion about duration and its behavior.

In sum, the present study cannot defend that in the concrete situation, subjects produce less disfluency because concrete topics are generally shorter. Conveying ideas in concrete topics is simpler, require less duration, and *S* does not need many words and time to express himself. On the other hand, the abstract topic is different. *S* talks about abstract topics longer than concrete ones because there is no exact way to express personal ideas and the necessary vocabulary is limitless. Most importantly, abstract topics are subjective.

To sum up, we cannot say that *DF* is merely affected by a single factor such as speech on abstract concept, speech with a non-native speaker or speech with a non-familiar speaker. All these variables also effect the duration of the topic. And duration has causal effects on *DF*. Thus, it may be inferred that concrete speeches are more fluent because they are shorter and affected by the concrete properties of topic as well as speech with a familiar speaker, and speech with a native speaker. Familiarity, native-speaker status, abstraction have a unified effect on *DF* and they determine the length of speech.

4.7 Gesture

It is known in the literature that gesture has a relation with *DF* (Krauss et al., 2000). Duration has relation, with *DF* as well so gesture has connection with duration. This inference was complex but was proven by present results. Results reveal that in both abstract topic and concrete topic the production of gesture has correlation with *DF* and duration.

The results show that total number of gestures both in concrete topic and abstract topic had significant correlations, $r_{\text{Concrete Topic}} = .831$ where $R^2_{\text{Concerete Topic}} = .690$, $p < .01$ and $r_{\text{Abstract Topic}} = .913$ where $R^2_{\text{Abstract Topic}} = .833$, $p < .01$. These two results show that gesture production has a positive correlation with duration in both abstract and concrete topics (see Table 27 and Table 28 in Appendix A). It is natural to expect that speakers will utter more words in longer talks and speaker speech rate is necessary stable; on the other hand, DFs which constitute 5% of speech to 30% of speech (Mean = 23%) affect the deviation of speech. The results show that 70% variation of gesture production should be affected by duration in the concrete topic and 83% variation of gesture production should be affected by duration in abstract topics.

Also in both topics, gestures are highly correlated with DF production. In Appendix A Table 29 and Table 30 shows the high rate of the correlations results; $r_{\text{Concrete Topic}} = .742$ where $R^2_{\text{Concerete Topic}} = .550$, $p < .01$ and $r_{\text{Abstract Topic}} = .696$ where $R^2_{\text{Abstract Topic}} = .484$, $p < .01$. These are high correlation values but the percents are lower with respect to percents of duration. *DF* production affects 55% of gesture production in concrete and 48% of gesture production in abstract topic. We conclude that the correlation between gesture production and duration is more significant than the one between gesture production and disfluency production.

In the present study, further analysis was performed. The variation of gesture productions in terms of topic type, familiarity, type of native speaker and duration was analyzed by using uni-variate ANOVA. None of the variables affected gesture production except duration and *DF* (see Table 33 in Appendix A). Bare analysis of total gesture production and raw data should not say anything to us. It was focused on the specific type of gestures to understand phenomena underlying conditions and conditions' interactions.

Multivariate ANOVA was applied after carefully selecting most strong pairs of gesture types, which are Deictic Gesture and Lexical Gesture. The results show that *Duration* ($F(2, 7) = 31$, $p < .05$), *I Type* ($F(2, 7) = 18.45$, $p = < .05$.), *I Type* * *Familiarity* ($F(2, 7) = 4.79$, $p < .05$), have significant effects on gesture productions. In addition to these, Familiarity ($F(2, 7) = 3.35$, $p = .095 > .05$) and topic type ($F(2, 7) = 4.33$, $p = .06 > .05$.) was marginally significant affecting the production of gestures.

For further explanation and interaction of all variable effects in gesture production, see Table 34 in Appendix A. We list the basic inferences below:

- Deictic Gesture is produced more in the concrete topic. Deictic Gesture is a kind of Lexical Gesture and it has a semantic relation to discourse so giving directions have a strong relation to deictic type of gesture.
- Deictic Gesture and Lexical Gesture are produced more in the case of Speaking with a Native addressee. The meaningful body gestures are used more when speaking with a native speaker.
- Deictic Gesture is produced more in the case of Speaking with a Familiar addressee. The speaker prefers to use deictic gesture more with a familiar interviewer.
- Lexical Gesture is produced more in the case of speaking with a non-familiar addressee. The speaker produces more gestures while referring to the properties of the concrete object with the non familiar interlocutor.

More detailed and interactive results can be listed as;

- Subjects speaking in the Concrete Topic situation with Native Interviewer (no matter he is familiar or not) produces same amount of Lexical Gesture and Deictic Gesture.
- Subjects speaking in concrete topic condition with a non-native and a familiar interlocutor produces fewer Lexical and Deictic type of gestures.
- Subjects speaking in Concrete Topic condition with a Native Interviewer produce more gestures than in when Interlocutor is non-native.
- Subjects speaking in the Abstract Topic condition with a native Interviewer produce more LG than when the interlocutor is non-native.
- Subjects speaking in the abstract topic condition with non-familiar addressee produce more Lexical Gesture than the case of familiar interlocutor.

All these results were explanatory but the *DF* relation and gesture need to be generalized and their relationship must be statistically modeled. The likelihood of the intersection of disfluency and gesture is one of the key concepts for the present study. If the study clarifies the temporal occurrence of gestures and the relationship of gestures to disfluencies, it will be more accurate and coherent. Thus, random sampling was done to calculate the probability of dual occurrence of gesture and DF (For details of sampling, see *Sampling of Gesture and Disfluency Interaction* Section). The results showed that in the concrete topic situation, 67.75 % of the DF occurrence was accompanied by gestures and in the abstract topic condition this percentage decreases to

48.75 %. Although the total DF is less in the concrete topic situation, subjects produced more DF coupled with gesture.

Table 14 - Gesture and Disfluency Probability of occurrence together

Concrete	Abstract	% Concrete	% Abstract
55	39	68,75	48,75

4.8 Further Analysis

Recall that all the results on DF and DF variation repeated so far are base on two topic conditions, "giving directions" and "effects of religion on our lives". However, the study aims to reach a generalization about DF and DF variation in L2 rather than making comments on variation between two these topics. It was important to understand whether our findings were limited to these two topics. Therefore, we conducted further statistics on the pairs – "giving directions" and "spring festival and clashing midterms" (the eliminated abstract topic). Fortunately, the results of these conversations, which are between subject patterns, are in line with the main analysis.

The correlation results of "giving directions" had been already reported. On the other hand, we needed to analyze correlation of "spring festival and clashing midterms". The results show that length and total number of *DF* are correlated as in the main pair correlation, r= .755 where R^2=.570, p<.01. Thus, 57 % of the variation of duration can be affected from total disfluency in the abstract topic – spring festival and clashing midterms. That is, there is a positive high correlation between duration and disfluency (see Table 36 at Appendix A).

The results show that (Table 35 in Appendix A) topic type has a marginal effect on the DF occurrences, F (1, 16) = 3.830, p=.076 >.05. It is marginally significant that the concrete topic has fewer number of DF than the abstract one (Mean Concrete=28.97, Mean Abstract =37.19) (see Table 15).

Table 15 - Descriptive Information of Topic type in the further analysis.

Topic Type	Mean	Std. Error

Concrete	**28.97**	**3.52**
Abstract	**37.19**	**3.38**

Table 16 - Descriptive information of Topic Type and Familiarity interaction

Familiarity	Topic Type	Mean
Non-Familiar	**Concrete**	**35.00**
	Abstract	**40.40**
Familiar	**Concrete**	**24.40**
	Abstract	**32.67**

The interaction of topic type and familiarity also has an effect on *DF* variations, F (1, 16) = 4.633, p=.054>.05. Further descriptive information emphasizes that ($Mean_{\text{Non Familiar Int. + Concrete Topic}}$=35.00, $Mean_{\text{Non-Familiar Int. + Abstract Topic}}$=40.40) subjects speaking with a non familiar *I* on an abstract topic produce more *DF* in the speech (see Table 16). The case is similar if *S*s speak with a familiar *I* on an abstract topic condition but the variation is not as big as in the case of the non-familiar *I* ($Mean_{\text{Familiar Interviewer + Concrete Topic}}$=24.40, $Mean_{\text{Familiar Interviewer + Abstract Topic}}$=32.67). The contrast between familiar and non familiar case emphasizes fact that the familiarity issue has a big effect on *DF* production as in the main analysis. Besides considering other factors, familiarity is clearly seen in the Table 17 ($Mean_{\text{Familiar Interviewer}}$=31.11, $Mean_{\text{Non-Familiar Interviewer}}$=35.05).

Table 17 - Further Analysis of Familiarity

Familiarity	Mean	Std. Error
Non-Familiar	**35.05**	**3.86**
Familiar	**31.11**	**3.62**

CHAPTER 5

DISCUSSION

In this chapter the discussion of the results and conclusions will be presented. At the end of this chapter, possible further studies will be outlined and the recommendations inspired by the obstacles faced within the study will be stated.

5.1 Discussion of Results

The two main purposes of this study was to understand

(a) the disfluency and gesture variation, and

(b) the connection of disfluency and gesture production

in L2 under the following conditions:

i. differences in topic types: concrete topics or abstract topics

ii. familiarity with the interlocutor

iii. speaking with a native or a non-native English Speaker.

This study was conducted in order to investigate the issues mentioned above under controlled conditions. Empirical data was collected and annotated by the means of the software ELAN. Learners spoke about two topics with an interlocutor and their audio-visual data was annotated in terms of language and gestural elements. The frequencies of gesture and disfluency types were converted into numerical values. Various types of statistical approaches were used to

analyze the data. To testify the accuracy of the results, data was retested by converting it from "within subject design" - the term used to compare same subjects in more than one condition- to "between group design" – the term used for comparing more than one distinct group. Although the results of the present study were rich enough to generalize some of the hypotheses, it has also some limitations that will be explained in this section.

The results revealed that gesture and disfluency vary under controlled conditions and it is found that there is a statistical relation for the connection of gesture and disfluency occurrence in L2 speech. Discussion of these results will also be detailed in this section. All the results must be considered in terms of L2 production.

5.1.1 Duration, Word Counts and Syllable Counts

As explained before, speech duration of the interlocutors and the subjects in the conversations was different. The subjects spoke more than the interlocutors because the interlocutors were asked to speak less and to interfere only when the conversation has a problem. The real actors were the speakers. This chapter includes an overall discussion of the raw data. The details will be presented in further sections.

Subjects in the concrete topic condition (giving direction) uttered fewer words and syllables, and produced fewer disfluency and gesture than in the abstract topics condition (effects of religion on our life). In addition, the concrete topic condition was quite restricted and straight-forward. The subject in the concrete topic condition did not need deep knowledge or personal ideas. Thus, concrete topic conversations lasted shorter and had fewer words and syllables than the abstract topic conversations.

The concrete topic condition elicited concrete and more frequently used words. Speakers can explain a concrete topic by using daily used simple words, such as road, walk, tree etc. On the other hand, speakers have to use emotional words and find relatively rarely used words in the abstract topic condition. These words, such as worship, religion etc., have limited usage in subject's L2 lexicon. However, concreteness of the words was not quantified.

To sum up, duration, number of the words and syllables were less in the concrete topic condition and more frequent and concrete words were used in this condition. On the other hand, the abstract topic condition lasted longer and included more abstract words.

5.1.2 Disfluency Overall

Fillers and silent pauses were more frequently used than the other disfluency types. In terms of L2, the results suit Levelt's Theory of Speech Production well. Levelt (1989, 1999; 2000; 2002) strongly emphasized that the modules of the model work in a parallel manner because of the reflexive-like production properties, but the internal processes of modules work incrementally in language production. These properties led Levelt to classify his model as a serial-modular model, which can, however, be couched in the formal language of connectionism, as far as parallel configuration is concerned. All these assumptions and suggestions are for L1. On the other hand, in L2, where parallelism and compatibility of modules may be paralyzed, the modules could not work as perfectly as in L1. The main problematic stages in the model for L2 are lexical access and grammatical encoding stages, which are the main stages that trigger the production of fillers and silent pauses. Menyhárt (2003) suggests that speakers may have a problem under uncertain conditions and this may cause filler and silent pause type of disfluencies because concept activation and grammatical encoding are the places where uncertain situations occur generally. All these are also reasonable for L2 production. In particular, vocabulary access and grammatical encoding are difficult for the learners. Results about fillers and silent pauses pointed out the vocabulary related problems of L2 learners.

Repetition was found as the most common disfluency for L1 (Bortfeld et al., 2001) but the results of this study shows that fillers and silent pauses were more common in L2. This may be a result of the L1 being Turkish but there are not any studies addressing whether repetitions are common disfluency types in Turkish. The suggestion of this study is nevertheless viable because the influence of the first language on the second language is an important factor in the acquisition process. On the other hand, repetitions may not be common disfluency types across all languages. To sum up, there is not enough information about repetition, filler and silent pauses in Turkish and in other languages. All we can say is that Turkish speakers of English produce silent pauses and fillers more than the other types of DF and there are two dominant types of DF for Turkish learners of English. The predominance of fillers and silent pauses is not surprising given that L2 learners find it hard to access words in their relatively small L2 lexicon and also experience problems with grammatical encoding. Therefore, the results fit into the particular problem of L2 learners.

The new DF type suggested by this study, namely repair, was only seen in the abstract topic situation. Thus, it may be related to the complexity of the topic as the abstract topic situation is

more complex both syntactically and semantically. This complexity of abstract topics makes the speakers to utter more complex syntactic structures than in the concrete topic conditions. Therefore, the complexity also affects the occurrence of the DF syntax and structure. If we consider the disfluency types encountered in both topic conditions, the percent of disfluency types in the abstract topic condition was higher than in the concrete one. The overt disfluency types are more complex and affect the cognitive system directly since self-monitoring is in process and triggers the reproduction of reparandum. Repair, which was seen only in the abstract topic condition, can be an interesting field of research but in this study, there is not further information on this issue since this study was not designed to understand the variations of this specific type of disfluency.

The results also showed that there is not any example for prolongation type of disfluency. Again, it may be related to the L1 effect but this idea would be speculative as there are no studies on the prolongation type of disfluency in Turkish. This is a further interesting research area as one may expect to find a relationship between the frequency of occurrence of long vowels and prolongation.

5.1.3 Gestures Overall

It was found that gesture distribution is related to topic type related patterns. It is additionally sensitive to the properties of the interlocutor, whether being a native speaker or a non-native speaker of English. The topic type affects the type of gesture while the property of the interlocutor affects the frequency of the gestures.

The distribution of gesture types encountered in this study suits the theoretical perspective of Krauss well. The motor gesture production was fewer in the concrete topic condition and higher in the abstract topic condition; because, in the concrete topic condition lexical and deictic gestures were more frequently produced than the motor gestures. In the abstract topic condition, this ratio - motor gesture and lexical gesture relation- has changed. That is, the number of motor gestures has increased and the number of lexical gestures has decreased. The total number of gestures in the abstract topic condition has also increased. It is clear that lexical gestures are related to discourse and motor gestures are meaningless gestural elements which have no connection with the topic. Therefore, it is expected that concrete topic situation elicits more lexical gestures and the abstract topic situation elicits more motor gestures.

The occurrence of gestures in the abstract topic condition is more frequent than in the concrete topic condition. In the abstract topic condition, the speakers may have more problems in accessing the lexicon and organizing their thoughts, which may, as a result, have caused them to gesture more frequently. They were stuck more frequently and tried to communicate with other channels while trying to suppress the tension of speech. However, the semantic relation of the gestures to the topic was less in the abstract topic condition than in the concrete topic condition. Further, the overall gesture increase in the abstract topic condition may be a result of the longer durations of the abstract topic conversations since there is a positive correlation between gesture production and duration. This correlation will be explained in the following section.

5.1.4 Correlations Overall

Correlations in this study were very complex as there were three-way interactions among disfluency, duration and gesture. They were positively correlated with each other. A high positive correlation was observed in each controlled condition.

Independent of other factors, as the duration of the conversation increased, the number of gestures and disfluency increased as well. These three factors are correlated but the increase in the ratio is not linear. Introspectively, there was a refraction point in the speech of the speaker, where the distribution of gesture and disfluency changed. Until that point, speakers produced more gestures and disfluencies but after that point the production of these was stabilized. However, this study was not designed to test whether they would diminish totally. On the other hand, it can be stated that, that the particular point is the reference point where the main idea has just been transferred to the interlocutor. After that point, supportive ideas are conveyed. In other words, after the communication of the main goal to the interlocutor, the burden of the cognitive system decreases; and it is highly probable to expect that this causes the decrease of both gesture and disfluency production.

The intricate interactions between the duration of the conversations and the number of disfluencies and gestures (dependent variables) as well as their relations with the independent variables are very challenging issues for this study. Although there is statistical information about their relations, there is no explicit information about the direction of these interactions, namely, the interconnection of dependent variables.

At a first glance, a significant correlation is found between disfluency and gesture under various conditions (familiarity vs. non familiarity, abstractness vs. concreteness and native speaker vs. non-native speaker). However, when duration has been taken into account, all effects vanished. Two possible scenarios could explain this intricate result.

First, there are no such effects between the disfluencies, gestures, topic types, familiarity, and speaking with a native speaker, when duration is factored out. The minimal result is that people talk longer about abstract topics and shorter about concrete topics. It would assume that duration covariates with the number of disfluencies positively, meaning that the longer we talk, the more occurrence of disfluency. On the other hand, there are good reasons to believe that this is not the entire story.

Second, there is an intricate inter-relation between duration, disfluency, gesture, familiarity, speaking with native speaker and the topic types, which introduce various feedback loops or indirect relations in the environment. Following this line of thought, duration does not appear to be solely a factor that needs to be controlled and factored out but it reflects the influence on the disfluency and gesture indirectly in that it is itself affected by a dependent variable. Thus, there would be two possible directions of causality.

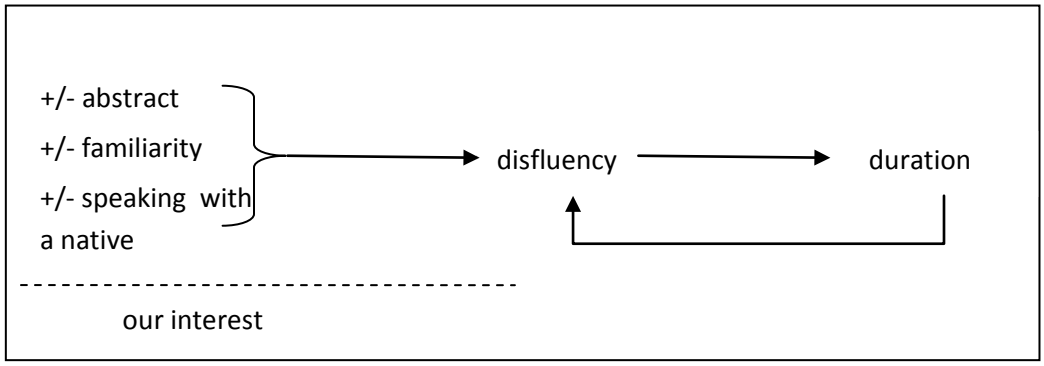


Figure 25 - First causal scenario for disfluency

In Figure 25 the first direction of causality is represented, where it is assumed that all independent variables affect disfluency and then disfluency affects duration, and meanwhile duration has an inter-connection with disfluency. Thus, abstractness of the topic determines fluency; there will be more disfluency in the abstract topic, and hence, overall length will increase as well. This direction of causality is the same for non-familiarity and speaking with non-native speaker. The detailed results regarding why they affect each other will be discussed in the next section. The important point is that, disfluency causes duration but it is itself caused

by abstractness, familiarity and speaking with native speakers. Thus, the factoring out of duration will diminish the effects of causal variables behind it.

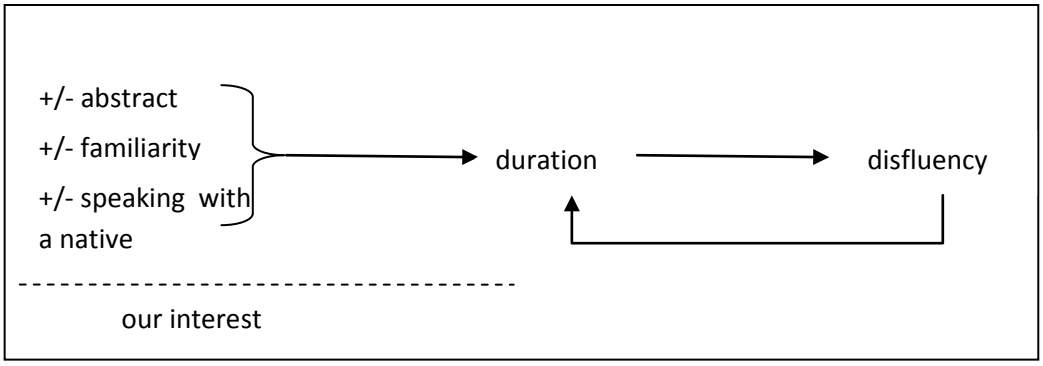


Figure 26 - Second causal scenario for disfluency

Second possible causal scenario is shown in Figure 26. The only difference is that disfluency is not directly affected by controlled variables. Like in the first scenario, it is not logical to factor out duration in this scenario, otherwise all effects of causal variables behind it would be lost.

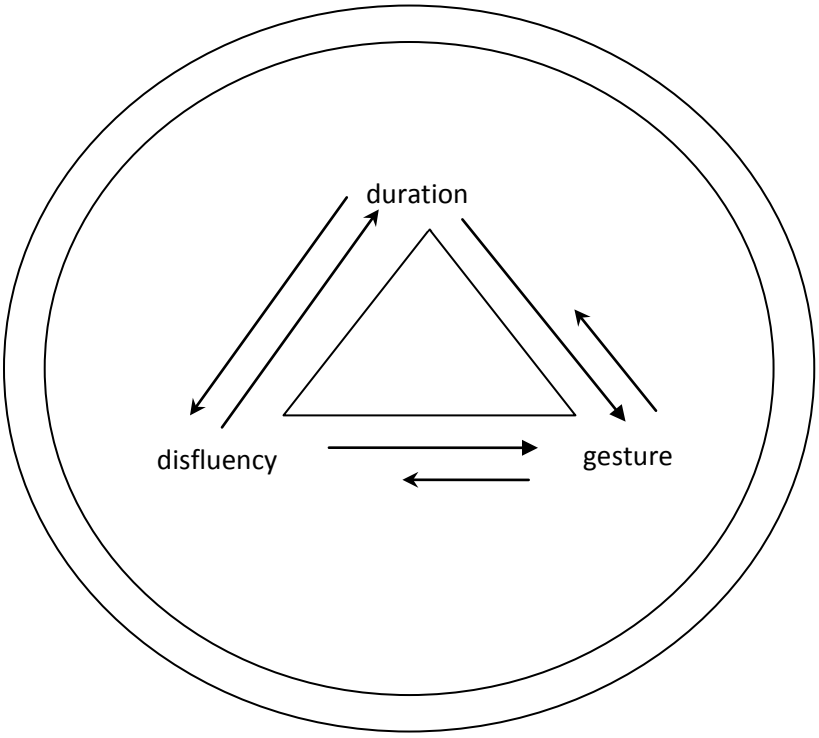


Figure 27 - overall causal relation

5.1.5 Discussion of Analysis

In Figure 27, the possible inter-connection is shown (size of the arrow represents the size of the effects of variables to each other). The intricate relation is very complex. Besides the intricate relation, duration, gesture and disfluency are affected from environmental factors (abstractness, familiarity and speaking with native speaker are represented as a circle in the figure).

5.2 Discussion of Main Analysis

Duration was the most interesting and challenging part of the study. It was not expected that duration would be affected by the controlled variable and would correlate with both disfluency and gesture. There may be four reasons resulting in spending less time for the concrete topic conversations and more time for the abstract topic conversations.

The first reason is a memory issue. We know that concrete words are processed faster and more accurately than abstract words in a variety of cognitive tasks since they are dually coded – verbally and imaginary (Sadoski & Paivio, 2001) . Their imaginary properties and coding make them strong, easily retainable and prone to decay. In addition, the connections of concrete words (concept) in the semantic network are plentiful. These connections are related to their entities and "is-a" kind of links. For example, a table: it is a kind of furniture and it is generally wooden, and has four legs. Chair is also wooden and it has four legs, too. Most importantly, when both of them are used together, there is a stronger relationship between them. When one of them is activated, the other one is activated, too.

The second one is the frequency of words. The more frequent words are retained faster and more accurately since the frequency of a word, i.e. the usage ratio in the speech, affects the lexical access of that word (Levelt & Jescheniak, 1994). There is no direct relation of frequency and concreteness. Although some abstract words have higher frequency in speech, the overall classification of concrete and abstract words shows that the frequency of concrete words is higher than the frequency of abstract words for learners in L2. On that account, similar to the former one, concrete topics and abstract topics would differ in length because they are composed of more frequently or less frequently used words.

The third reason is related to the content of the topic. Talks on concrete topics are more restricted than talks on abstract topics; in other words, the speech in the concrete topic condition

is based on facts, and it is difficult to enrich the conversation in contrary to the abstract topic condition. In addition, talks on concrete topics are more objective (such as giving directions). However, talks on abstract topics are more subjective (such as religion since everyone has something personal to say about). Limited vocabulary use in the concrete topic condition limits the duration and increases the fluency. Fluency is high because accessing the limited vocabulary is faster and more accurate. Sentences have straight-forward meaning in talks on concrete topics.

The fourth reason is related to familiarity. It was found that duration of speaker's speech is less when speaking with a familiar interlocutor. Duration was not only affected by topic type; speaking with a native or the nonnative familiarity appeared to be a dominant factor on the duration of the conversations. This is related to the psychological condition of the speaker; in other words, duration of speech is affected by the anxiety level of the speaker (Levelt, 1989). Tension is an old and well known concept affecting speech and it was studied well in psychology (Bonner, 1943). As the tension of the speaker increases, the time needed to speak increases as well. That is, it affects the communication channel and distracts the constant attention of the speaker.

To sum up, we speak with familiar people in a more relaxed way and thus consume a shorter time. Independent of other factors, in the concrete topic condition, talks take less time and a speaker can express himself in shorter talks compared to the abstract topic condition.

The main concern of the study is the variation of disfluency rate under specific conditions. It was found that in concrete topics, disfluency rate was less compared to abstract topics. However, topic type is not the only aspect affecting disfluency rate. Disfluency was less in the concrete topic condition because both the topic types were concrete and the duration of conversation were shorter. It is shown that both duration and disfluency were affected by topic type. This study also showed that disfluency and duration are positively correlated. Considering all of these findings, it can be said that speakers' speeches on concrete topics are more fluent and shorter than on abstract topics but it has to be emphasized that fluency is not related only to the property of topic but also to the duration of talk in L2. The results are not surprising for L2. Following Bialystok (1990) we may assumed the coordination of knowledge (in this case knowledge of lexicon) gets harder when the time spent increases. The time spent increases because there are not any exact ways to explain an abstract topic which is subjective. Therefore, both longer and less fluent conversations take place in the abstract topic condition than in the concrete topic condition.

Like length, disfluency is also affected by the familiarity with the interlocutor, topic type, total duration and properties of the interlocutor - being native or non native speaker of English. There are three way interactions between topic type, duration and disfluency as stated previously in detail. It is known that the burdens of concrete words on the cognitive system are less and discourse on concrete topics is generally short; thus L2 speakers are more fluent in concrete topics than in abstract topics. Moreover, frequency of concrete words are higher than the frequency of abstract words since frequency affects lexical access, as well (Levelt & Jescheniak, 1994). For example, in giving directions, speakers uttered words like tree, road, to walk, building, and library. On the other hand, speakers tried to utter words like pray, worship, head scarf etc. in the abstract topic (effects of religion on our live). For learners, the first group is probably more accessible than the second one.

Disfluency is affected by the familiarity with the interlocutor in at least two ways. The first one is related to the psychological aspect considering the tension or anxiety levels of the speaker (Branigan et al., 1999). Anxiety affects performance when speaking to an unfamiliar person, because in speech with a familiar interlocutor, the speaker is more relaxed. On the other hand, the learner has more problems when speaking to a non familiar addressee. His tension arouses and the fluency of his speech reduces. The second factor is related to trying to be more cooperative in speech. The speaker selects words more carefully when speaking with an unfamiliar interlocutor. The performance of the speaker decreases since it affects the planning process and results in slow performance and higher rate of disfluency. Therefore, speakers are more fluent when speaking with a familiar addressee (Bard et al., 2001).

Native speech decreases the rate of disfluency. In other words, fluency of native speakers affects the speech of the non-native speakers. The learner is more relaxed with a native speaker because a native speaker could understand the intention of the non native speaker more easily. The native speaker is an experienced speaker and could help the non-native speaker using his native speaker language. This affects the planning processes. Simpler and successive sentences are uttered. In addition, it also reduces the tension. Therefore, speakers are more fluent when speaking with a native speaker (Branigan et al., 1999).

A speaker might be more creative (produce more gestures) when speaking with an unfamiliar interlocutor, since he would try to cover his speech deficiencies by gesturing more. Gesture is an interesting part of this study; it is found that familiarity with the interlocutor and the native property of the addressee affect gesture production. The results of this study also show that learners prefer to gesture more while speaking with a non-native addressee, an unfamiliar native

addressee or an unfamiliar non-native addressee in L2. This is valid for L1 also (Bard et al., 2001; Branigan et al., 1999).

The results provide a statistical model for dual occurrence of gesture and disfluency. This statistical approach is well suited with Krauss's Theory of Gesture production and Levelt's Speech production Model. In the concrete topic condition, 67.5% of randomly selected deviations intersected with the gestures. That is, the productions of gesture and disfluency at the same time are more common though both gesture and disfluency productions are less in concrete topics. On the other hand, the percentage decreases to 48% in the abstract topic condition, because motor gestures occur more frequently in the abstract topic condition and lexical gestures occur less in the abstract topic condition. That is, the percentage represents lexical gesture occurrence in the condition and it is expected that in the abstract condition learners produce fewer lexical gestures.

5.3 Discussion of Further Analysis

This section is interested in further statistics to prove what was discussed in the main analysis. The new statistics was performed on the new pair - "giving directions" and "spring festival/clashing midterms". The results of the correlations concerning with "giving directions" have been already discussed in the previous sections. Thus, this section only includes the discussion of the correlation results of the eliminated abstract topic – "spring festival and clashing midterms", and overall analysis of this new pair.

According to results of the further analysis, disfluency and duration were positively correlated in the eliminated abstract topic condition. In other words, as the duration of the conversation increased, the number of disfluency increased similar to the main analysis. Therefore, the same conclusions about the correlations in main analysis are also valid for this topic. For example, duration is correlated with disfluency and both duration and disfluency are dependent variables. Also, eliminating effects of the duration diminished all the other effects. Therefore, there is also an intricate correlation in this pair. So, the suggested scenarios in the overall correlation section are applicable for this case.

We will now discuss analysis of new pair. The duration was a challenging issue for this pair and the topic type and familiarity also affected the results. The results pointed same conclusions as the proposed in the main analysis. Therefore, we can safely assume that four reasons were

responsible for spending less time for the concrete topic conversation and more time for the abstract topic conversation. The first reason is the memory issue. The difference between coding of the concrete and abstract words results in different recalling time of words (Dual Coding Theory, Sadoski & Paivio, 2001) and recalling time affects durations of the conversations. The second reason is the frequency of the concrete and abstract words. The lexical access is affected from frequency (Levelt & Jescheniak, 1994) and affects the time of conversation. In other words, frequent words are accessed fast and more reliable. For example, in the eliminated abstract topic, students used words related to feelings (which are abstract and limited than those in the concrete topic condition). The third reason is the content of the topic. The eliminated abstract topic is subjective as the other abstract one, and it was affected from personal experiences, affecting planning of the speech .On the other hand, in the concrete topic condition, vocabulary is limited and well defined for everybody and it is objective. Therefore, it affects the fluency as well. The fourth reason is related to familiarity. Like the main analysis, speakers spent less time while speaking with a familiar speaker in this pair result. This result proposes that speakers are more fluent when speaking with a familiar addressee (Bard et al., 2001; Branigan et al., 1999). In the main analysis, this issue was explained by anxiety issue (Bonner, 1943; Levelt, 1989). The interconnection of duration and familiarity is same and valid as in the main analysis.

As a result of the correlations and the analysis, it is also concluded that in the concrete topic, disfluency number was fewer compared to abstract topic. There are two reasons for this; the topic affect and the concrete topic being shorter than abstract topic. Therefore, it has to be emphasized again that the fluency is not related only to the property of topic but also to the duration of the talk in L2. Besides these, familiarity affects the number of the disfluency as well. It is related to anxiety of the speaker and it affected speaker's willing to produce cooperative speech.

Although the native speech has effect on disfluency as in the main pair, results showed that there was no significant effect of the native speech in this pair. However, it might be effect of the topic, i.e., spring festival. In other words, maybe in the spring festival case, the effect of native speaker on learner diminished because he might be unfamiliar with the spring festival. Thus, they could not help the learners and not understand the intentions of learners as they did in the religion topic. But it might be a speculative and an open question. This situation must be studied with different topics to conclude native speech effects on the conversations in a more general way.

To sum up, all results of the further statistics were parallel with the main analysis. This parallelism hints the generability aspect of the study in L2.In other words, it would be stated that this additional analysis shows, indeed, that the findings from the main analysis can be generalized to some degree and do not seem to be specific to exactly the one pair of topics in the main analysis.

CHAPTER 6

CONCLUSIONS

In this study, significant information about disfluency and gesture is found, though further research on this topic is needed to clarify and generalize the findings. The results of this study are summarized below. All conclusions were made for L2.

- Speakers are more fluent when speaking on a concrete topic then when speaking on an abstract topic, since concrete topics are shorter than abstract topics.
- Speakers are more fluent when speaking with a familiar addressee.
- Speakers are more fluent when speaking with a native English addressee.
- Speakers produce more lexical gestures in concrete topics although they are gesturing less in concrete topics.
- Speakers gesture less when speaking with a familiar addressee.
- Speakers talk longer about abstract topics, and they have more problems in controlling L2 to express their ideas.

CHAPTER 7

RECOMMENDATION FOR FURTHER STUDY

- Disfluency variations need to be scrutinized in terms of different languages, especially Turkish.
- The same research may be repeated by putting a time restriction on both talk time and gestures.
- Prolongation can be studied in terms of specific properties of a language.
- The parallelism of syntax and disfluency in different languages can be investigated.
- A study with more abstract and concrete topics need to be conducted since in this study, only one concrete and one abstract topic have been studied.

REFERENCES

Alibali, M. W., Kita, S., & Young, A. J. (2000). Gesture and the process of speech production: We think, therefore we gesture. *Language and Cognitive Processes, 15*(6), 593-613.

Bard, E. G., Lickley, R. J., & Aylett, M. P. (2001). Is Disfluency Just Difficulty? *ISCA Tutorial and Research Workshop (ITRW) on Disfluency in Spontaneous Speech*.

Bialystok, E. (1990). *Communication Strategies: A Psychological Analysis of Second-Language Use*: Blackwell Publishers.

Bonner, M. R. (1943). Changes in the speech pattern under emotional tension. *American Journal of Psychology, 56*, 262-273.

Bortfeld, H., Leon, S. D., Bloom, J. E., Schober, M. F., & Brennan, S. E. (2001). Disfluency Rates in Conversation: Effects of Age, Relationship, Topic, Role, and Gender. *Language and Speech, 44*(2), 123-147.

Branigan, H., Lickley, R., & McKelvie, D. (1999). Non-linguistic influences on rates of disfluency in spontaneous speech. *Proceedings of the 14th International Conference of Phonetic Sciences*.

Brown, P. M., & Dell, G. S. (1987). Adapting Production to Comprehension: The Explicit Mention of Instruments. *Cognitive Psychology, 19*(4), 441-472.

Bull, P., & Connelly, G. (1985). Body movement and emphasis in speech. *Journal of Nonverbal Behavior, 9*(3), 169-187.

Butterworth, B. (1975). Hesitation and semantic planning in speech. *Journal of Psycholinguistic Research, 4*(1), 75-87.

Callaway, C. (2003). Do we need deep generation of disfluent dialogue. *AAAI Spring Symposium on Natural Language Generation in Spoken and Written Dialogue*.

Clark, H. H., & Wasow, T. (1998). Repeating Words in Spontaneous Speech. *Cognitive Psychology, 37*(3), 201-242.

Cooper, A. A., & Hale, J. T. (2005). Promotion of Disfluency in Syntactic Parallelism. *Disfluency in Spontaneous Speech*.

Dell, G. S. (1995). Speaking and misspeaking. *Language, 1*, 183–208.

Dell, G. S., Chang, F., & Griffin, Z. M. (2001). Connectionist Models of Language Production: Lexical Access and Grammatical Encoding. *Connectionist Psycholinguistics*.

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review, 104*(4), 801-838.

Derwing, T. M., Rossiter, M. J., Munro, M. J., & Thomson, R. I. (2004). Second Language Fluency: Judgments on Different Tasks. *Language Learning, 54*(4), 655-679.

Dictionary.com. (2007). disfluency. (n.d.). *Unabridged (v 1.1)*, 2007, from http://dictionary.reference.com/browse/disfluency

DiSS'03. (2003). Disfluency in Spontaneous Speech Workshop.

Eklund, R., & Shriberg, E. (1998). Crosslinguistic Disfluency Modelling: A Comparative Analysis of Swedish and American English Human--Human and Human--Machine Dialogues. *Fifth International Conference on Spoken Language Processing*.

Ferreira, F., & Bailey, K. G. D. (2004). Disfluencies and human language comprehension. *Trends in Cognitive Sciences, 8*(5), 231-237.

Finlayson, S., Forrest, V., Lickley, R., & Beck, J. M. (2003). Effects of the Restriction of Hand Gestures on Disfluency. *ISCA Tutorial and Research Workshop on Disfluency in Spontaneous Speech*.

Griffin, Z. M. (2001). Gaze durations during speech reflect word selection and phonological encoding. *Cognition, 82*(1), 1-14.

Gurman, E., Aylett, M., & Lickley, R. (2002). Towards a Psycholinguistics of Dialogue: Defining Reaction Time and Error Rate in a Dialogue Corpus.

Howell, P., & Akande, O. (2005). Simulations of the Types of Disfluency Produced in Spontaneous Utterances by Fluent Speakers, and the Change in Disfluency Type Seen as Speakers Who Stutter Get Older. *Disfluency in Spontaneous Speech*.

Kempen, G., & Huijbers, P. (1983). The lexicalization process in sentence production and naming: Indirect election of words. *Cognition, 14*(2), 185-209.

Kirsner, K., Dunn, J., & Hird, K. (2003). Fluency: Time for a Paradigm Shift. *ISCA Tutorial and Research Workshop on Disfluency in Spontaneous Speech*.

Kormos, J. (1999). Monitoring and self-repair in L2. *Language Learning, 49*(2), 303-342.

Krauss, R. M., Chen, Y., & Gottesman, R. F. (2000). Lexical gestures and lexical access: A process model. *Language and gesture*, 261–283.

Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*: MIT Press.

Levelt, W. J. M. (1999). Models of word production. *Trends in Cognitive Sciences, 3*(6), 223-232.

Levelt, W. J. M., & Jescheniak, J. D. (1994). Word Frequency Effects in Speech Production: Retrieval of Syntactic Information and of Phonological Form. *Learning and Memory, 20*(4), 824-843.

Levelt, W. J. M., & Meyer, A. S. (2000). Word for word: Multiple lexical access in speech production. *The European Journal of Cognitive Psychology, 12*(4), 433-452.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (2002). A Theory Of Lexical Access In Speech Production. *Psycholinguistics: Critical Concepts in Psychology, 22*(99), 1-75.

Lickley, R. J. (1996). Juncture cues to disfluency. *Spoken Language, 1996. ICSLP 96. Proceedings., Fourth International Conference on, 4*.

Mehnert, U. (1998). The Effects of Differrent Lengths of Time for Planning on Second Language Performance. *Studies in Second Language Acquisition, 20*(01), 83-108.

Menyhárt, K. (2003). Age-dependent types and frequency of disfluencies. *DiSS'03, 10*(1.1), 6.99.

Meyer, A. S., Roelofs, A., & Levelt, W. J. M. (2003). Word length effects in object naming: The role of a response criterion. *Journal of Memory and Language, 48*(1), 131-147.

Nicholson, H., Bard, E. G., Lickley, R., Anderson, A. H., Mullin, J., Kenicer, D., et al. (2003). The Intentionality of Disfluency: Findings from Feedback and Timing. *ISCA Tutorial and Research Workshop on Disfluency in Spontaneous Speech*.

Nicholson, H. B. M. (2004). Disfluency: The Cost of Attending to Listener Feedback? *Psychology and Language Sciences*.

Paivio, A. (1971). *Imagery and verbal processes*: Holt, Rinehart and Winston, New York.

Pechmann, T., & Zerbst, D. (2004). Syntactic constraints on lexical selection in language production. *Multidisciplinary approaches to language production*, 279-301.

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition, 42*(1-3), 107-142.

Roelofs, A. (1993). Testing a non-decompositional theory of lemma retrieval in speaking: retrieval of verbs. *Cognition, 47*(1), 59-87.

Roelofs, A. (1996). Computational models of lemma retrieval. *Computational psycholinguistics*, 308–327.

Sadoski, M., & Paivio, A. (2001). *Imagery and Text: A Dual Coding Theory of Reading and Writing*: Lawrence Erlbaum Associates.

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching, 10*(3), 209-231.

Shriberg, E. (1999). Phonetic consequences of speech disfluency. *Proceedings of the International Congress of Phonetic Sciences (ICPhS-99), 1*, 619–622.

Stolcke, A., & Shriberg, E. (1996). Statistical language modeling for speech disfluencies. *Acoustics, Speech, and Signal Processing, 1996. ICASSP-96. Conference Proceedings., 1996 IEEE International Conference on, 1*.

Temple, L. (2000). Second language learner speech production. *Studia Linguistica, 54*(2), 288-297.

APPENDICES

APPENDIX A

Figure 28 - Length Variations in terms of Topic and Familiarity when Interlocutor is native

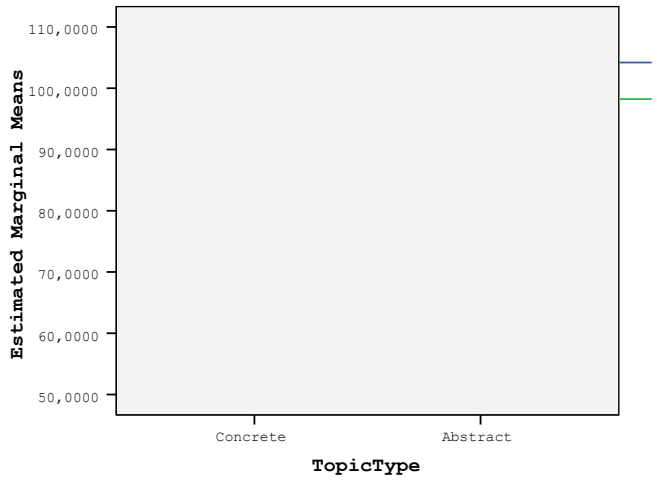


Figure 29 - Length Variations in terms of Topic and Familiarity when Interlocutor is non-native

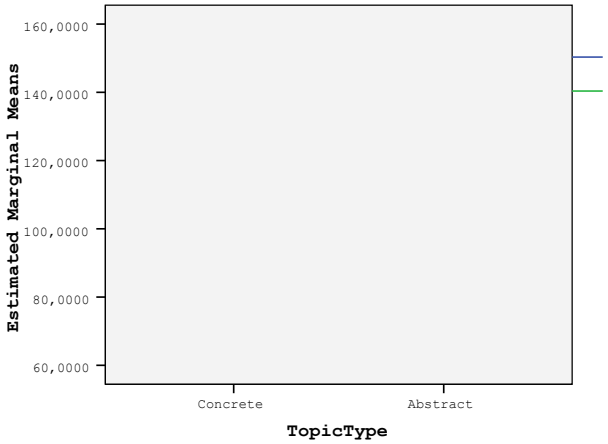


Table 18 - Correlation of Length and Disfluency in Giving Direction Topic

		Subject Duration	Total Disfluency
Subject Duration	Pearson Correlation	1	,921(**)
	Sig. (2-tailed)		,000
	N	16	16
Total Disfluency	Pearson Correlation	,921(**)	1
	Sig. (2-tailed)	,000	
	N	16	16

** Correlation is significant at the 0.01 level (2-tailed).

Table 19 - Correlations of Length and Disfluency in Effects of Religion to Our lives Topic

		Total Disfluency	Subject Duration
Total Disfluency	Pearson Correlation	1	,854(**)
	Sig. (2-tailed)		,000
	N	16	16
Subject Duration	Pearson Correlation	,854(**)	1
	Sig. (2-tailed)	,000	
	N	16	16

** Correlation is significant at the 0.01 level (2-tailed).

Table 20 - Correlation result of Length and Disfluency in Two Group Design

		Subject Duration	Total Disfluency
Subject Duration	Pearson Correlation	1	,892(**)
	Sig. (2-tailed)		,000
	N	16	16
Total Disfluency	Pearson Correlation	,892(**)	1
	Sig. (2-tailed)	,000	
	N	16	16

** Correlation is significant at the 0.01 level (2-tailed).

Table 21 - Correlation of Topic Type and Disfluency by assigned covariate: Subject Duration

Control Variables			Topic Type	Total Disfluency
Subject Duration	Topic Type	Correlation	1,000	,170
		Significance (2-tailed)	.	,545
		Df	0	13
	Total Disfluency	Correlation	,170	1,000
		Significance (2-tailed)	,545	.
		Df	13	0

Table 22 - Within-Subjects Effects of Disfluency and Topic Type without considering any controlled variable.

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Topic Type	Sphericity Assumed	790,031	1	790,031	4,886	,043
	Greenhouse-Geisser	790,031	1,000	790,031	4,886	,043
	Huynh-Feldt	790,031	1,000	790,031	4,886	,043
	Lower-bound	790,031	1,000	790,031	4,886	,043
Error(Topic Type)	Sphericity Assumed	2425,469	15	161,698		
	Greenhouse-Geisser	2425,469	15,000	161,698		
	Huynh-Feldt	2425,469	15,000	161,698		
	Lower-bound	2425,469	15,000	161,698		

Table 23 - Tests of Within-Subjects Effects: Disfluency variation between Topics in Repeated ANOVA by to controlled variables and their interactions.

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Topic Type	Sphericity Assumed	1047,230	1	1047,230	9,947	,008
	Greenhouse-Geisser	1047,230	1,000	1047,230	9,947	,008
	Huynh-Feldt	1047,230	1,000	1047,230	9,947	,008
	Lower-bound	1047,230	1,000	1047,230	9,947	,008
Topic Type * Interviewer Type	Sphericity Assumed	45,646	1	45,646	,434	,523
	Greenhouse-Geisser	45,646	1,000	45,646	,434	,523
	Huynh-Feldt	45,646	1,000	45,646	,434	,523
	Lower-bound	45,646	1,000	45,646	,434	,523
Topic Type * Familiarity	Sphericity Assumed	1008,600	1	1008,600	9,580	,009
	Greenhouse-Geisser	1008,600	1,000	1008,600	9,580	,009
	Huynh-Feldt	1008,600	1,000	1008,600	9,580	,009
	Lower-bound	1008,600	1,000	1008,600	9,580	,009
Topic Type * Interviewer Type * Familiarity	Sphericity Assumed	244,017	1	244,017	2,318	,154
	Greenhouse-Geisser	244,017	1,000	244,017	2,318	,154
	Huynh-Feldt	244,017	1,000	244,017	2,318	,154
	Lower-bound	244,017	1,000	244,017	2,318	,154
Error(Topic Type)	Sphericity Assumed	1263,400	12	105,283		
	Greenhouse-Geisser	1263,400	12,000	105,283		
	Huynh-Feldt	1263,400	12,000	105,283		
	Lower-bound	1263,400	12,000	105,283		

Table 24 - Tests of Between-Subjects Effects: Repeated ANOVA – Dependent Variable : Total Disfluency ; Independent Variable: Familiarity and Intervier Type

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	44626,446	1	44626,446	172,863	,000
Interviewer Type	997,696	1	997,696	3,865	,073
Familiarity	1968,446	1	1968,446	7,625	,017
Interviewer Type * Familiarity	1553,807	1	1553,807	6,019	,030
Error	3097,933	12	258,161		

Table 25 - Tests of Within-Subjects Effects: Repeated ANOVA - Dependent Variable : Disfluency per Sylable ; Independent Variable :Familiarity and Interview Type

Source		Type III Sum of Squares	Df	Mean Square	F	Sig.
Topic Type	Sphericity Assumed	7,54E-005	1	7,54E-005	,064	,805
	Greenhouse-Geisser	7,54E-005	1,000	7,54E-005	,064	,805
	Huynh-Feldt	7,54E-005	1,000	7,54E-005	,064	,805
	Lower-bound	7,54E-005	1,000	7,54E-005	,064	,805
Topic Type * Interviewer Type	Sphericity Assumed	3,99E-005	1	3,99E-005	,034	,857
	Greenhouse-Geisser	3,99E-005	1,000	3,99E-005	,034	,857
	Huynh-Feldt	3,99E-005	1,000	3,99E-005	,034	,857
	Lower-bound	3,99E-005	1,000	3,99E-005	,034	,857
Topic Type * Familiarity	Sphericity Assumed	8,95E-005	1	8,95E-005	,075	,788
	Greenhouse-Geisser	8,95E-005	1,000	8,95E-005	,075	,788
	Huynh-Feldt	8,95E-005	1,000	8,95E-005	,075	,788
	Lower-bound	8,95E-005	1,000	8,95E-005	,075	,788
Topic Type * Interviewer Type * Familiarity	Sphericity Assumed	7,36E-005	1	7,36E-005	,062	,808
	Greenhouse-Geisser	7,36E-005	1,000	7,36E-005	,062	,808
	Huynh-Feldt	7,36E-005	1,000	7,36E-005	,062	,808
	Lower-bound	7,36E-005	1,000	7,36E-005	,062	,808
Error(Topic Type)	Sphericity Assumed	,014	12	,001		
	Greenhouse-Geisser	,014	12,000	,001		
	Huynh-Feldt	,014	12,000	,001		
	Lower-bound	,014	12,000	,001		

Table 26 - Test of Between-Subject Effects: Univariant ANOVA in Between Subject Design - Dependent Variable : Disfluency ; Independent : Topic Type, Familiarity, Type of Interviewer ; Covariate : Subject Duration

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	6059,269(a)	7	865,610	10,092	,002
Intercept	529,256	1	529,256	6,171	,038
Subject Duration	1154,332	1	1154,332	13,458	,006
Interviewer Type	186,994	1	186,994	2,180	,178
Familiarity	93,871	1	93,871	1,094	,326
Topic Type	40,978	1	40,978	,478	,509
Interviewer Type * Familiarity	342,974	1	342,974	3,999	,081
Interviewer Type * Topic Type	44,616	1	44,616	,520	,491
Familiarity * Topic Type	,918	1	,918	,011	,920
Interviewer Type * Familiarity * Topic Type	,000	0	.	.	.
Error	686,168	8	85,771		
Total	33233,000	16			
Corrected Total	6745,437	15			

a R Squared = ,898 (Adjusted R Squared = ,809)

Table 27 - Correlations: Concrete Topic; Duration and Total Gestures

		Duration	Total Gestures
Duration	Pearson Correlation	1	,831(**)
	Sig. (2-tailed)		,000
	N	16	16
Total Gestures	Pearson Correlation	,831(**)	1
	Sig. (2-tailed)	,000	
	N	16	16

** Correlation is significant at the 0.01 level (2-tailed).

Table 28 - Correlations: Abstract Topic; Duration and Total Gestures

		Duration	Total Gestures
Duration	Pearson Correlation	1	,913(**)
	Sig. (2-tailed)		,000
	N	16	16
Total Gesture	Pearson Correlation	,913(**)	1
	Sig. (2-tailed)	,000	
	N	16	16

** Correlation is significant at the 0.01 level (2-tailed).

Table 29 - Correlations: Concrete Topic; Total Disfluency and Total Gestures

		Total Disfluency	Total Gestures
Total Disfluency	Pearson Correlation	1	,742(**)
	Sig. (2-tailed)		,001
	N	16	16
Total Gestures	Pearson Correlation	,742(**)	1
	Sig. (2-tailed)	,001	
	N	16	16

** Correlation is significant at the 0.01 level (2-tailed).

Table 30 - Correlations: Abstract Topic; Total Disfluency and Total Gestures

		Total Disfluency	Total Gestures
Total Disfluency	Pearson Correlation	1	,696(**)
	Sig. (2-tailed)		,003
	N	16	16
Total Gestures	Pearson Correlation	,696(**)	1
	Sig. (2-tailed)	,003	
	N	16	16

** Correlation is significant at the 0.01 level (2-tailed).

Table 31 - Univariate Analysis in Between Subject Design - Dependent :Total Gestures; Independent :Topic Type, Familiarity , Type of Interviewer

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1821,833(a)	6	303,639	,905	,531
Intercept	10644,054	1	10644,054	31,740	,000
Interviewer Type	7,514	1	7,514	,022	,884
Familiarity	136,013	1	136,013	,406	,540
Topic Type	733,426	1	733,426	2,187	,173
Interviewer Type * Familiarity	,762	1	,762	,002	,963
Interviewer Type * Topic Type	6,857	1	6,857	,020	,889
Familiarity * Topic Type	181,551	1	181,551	,541	,481
Interviewer Type * Familiarity * Topic Type	,000	0	.	.	.
Error	3018,167	9	335,352		
Total	17836,000	16			
Corrected Total	4840,000	15			

a R Squared = ,376 (Adjusted R Squared = -,039)

Table 32 - Multivariate Test - Dependent Variables : LG,DG ; Independent Variables : Topic Type, Familiarity, Interviewer Type ; Covariate : Duration

Effect		Value	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	,721	2,000	7,000	,011	,721
	Wilks' Lambda	,279	2,000	7,000	,011	,721
Duration	Pillai's Trace	,899	2,000	7,000	,000	,899
	Wilks' Lambda	,101	2,000	7,000	,000	,899
Topic Type	Pillai's Trace	,553	2,000	7,000	,060	,553
	Wilks' Lambda	,447	2,000	7,000	,060	,553
Familiarity	Pillai's Trace	,489	2,000	7,000	,095	,489
	Wilks' Lambda	,511	2,000	7,000	,095	,489
Interviewer Type	Pillai's Trace	,841	2,000	7,000	,002	,841
	Wilks' Lambda	,159	2,000	7,000	,002	,841
Topic Type * Familiarity	Pillai's Trace	,025	2,000	7,000	,915	,025
	Wilks' Lambda	,975	2,000	7,000	,915	,025
Topic Type * Interviewer Type	Pillai's Trace	,426	2,000	7,000	,144	,426
	Wilks' Lambda	,574	2,000	7,000	,144	,426
Familiarity * Interviewer Type	Pillai's Trace	,578	2,000	7,000	,049	,578
	Wilks' Lambda	,422	2,000	7,000	,049	,578
Topic Type * Familiarity * Interviewer Type	Pillai's Trace	,000	,000	,000	.	.
	Wilks' Lambda	1,000	,000	7,500	.	.

Table 33 - Univariate Analysis : Dependent Variable : Total Gestures; Independent Variables: Topic Type, Familiarity, Interviewer Type ;

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	3535,915(a)	7	505,131	3,099	,068	,731
Intercept	156,644	1	156,644	,961	,356	,107
Total Disfluency	1714,082	1	1714,082	10,515	,012	,568
Topic Type	3,490	1	3,490	,021	,887	,003
Interviewer Type	300,343	1	300,343	1,842	,212	,187
Familiarity	215,214	1	215,214	1,320	,284	,142
Topic Type * Interviewer Type	12,631	1	12,631	,077	,788	,010
Topic Type * Familiarity	3,318	1	3,318	,020	,890	,003
Interviewer Type * Familiarity	362,002	1	362,002	2,221	,175	,217
Topic Type * Interviewer Type * Familiarity	,000	0	.	.	.	,000
Error	1304,085	8	163,011			
Total	17836,000	16				
Corrected Total	4840,000	15				

a R Squared = ,731 (Adjusted R Squared = ,495)

Table 34 Descriptive Statistics of Multivariate Analysis of DG and LG

	Topic Type	Familiarity	Interviewer Type	Mean	Std. Deviation	Number
DG	Concrete	No	Native	3,33	3,215	3
			Total	3,33	3,215	3
		Yes	Native	3,00	,000	2
			Non-Native	2,00	1,732	3
			Non-Native	2,00	1,732	3
	Abstract	No	Native	,00	,000	3
			Non-Native	,00	,000	2
		Yes	Native	,00	.	1
			Non-Native	,00	,000	2
			Total	,00	,000	3
			Non-Native	,00	,000	4
	Total	No	Native	1,67	2,733	6
			Non-Native	,00	,000	2
		Yes	Native	2,00	1,732	3
			Non-Native	1,20	1,643	5
		Total	Native	1,78	2,333	9
			Non-Native	,86	1,464	7
LG	Concrete	No	Native	5,00	2,000	3
		Yes	Native	5,00	,000	2
			Non-Native	1,33	2,309	3
			Total	2,80	2,588	5
		Total	Native	5,00	1,414	5
			Non-Native	1,33	2,309	3
	Abstract	No	Native	7,33	5,859	3
			Non-Native	3,00	1,414	2
			Total	5,60	4,827	5
		Yes	Native	2,00	.	1
			Non-Native	3,50	4,950	2
			Total	3,00	3,606	3
		Total	Native	6,00	5,477	4
			Non-Native	3,25	2,986	4
			Total	4,63	4,340	8
	Total	No	Native	6,17	4,119	6
			Non-Native	3,00	1,414	2
			Total	5,38	3,815	8
		Yes	Native	4,00	1,732	3
			Non-Native	2,20	3,194	5
			Total	2,88	2,748	8
		Total	Native	5,44	3,539	9
			Non-Native	2,43	2,699	7
			Total	4,13	3,462	16

Table 35 - Tests of Between-Subjects Effects

Source	Type I Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1606,387(a)	4	401,597	4,618	,020
Intercept	17358,062	1	17358,062	199,612	,000
TopicType	333,063	1	333,063	3,830	,076
Adressee_Type	,907	1	,907	,010	,920
familarity	402,899	1	402,899	4,633	,054
SUB_DUR	869,518	1	869,518	9,999	,009
Error	956,550	11	86,959		
Total	19921,000	16			
Corrected Total	2562,937	15			

a R Squared = ,627 (Adjusted R Squared = ,491)

Table 36 - Correlations in the further statistic for “spring festival and clashing midterms”

		T3_TotalDis	T3_SUB_DUR
T3_TotalDis	Pearson Correlation	1	,755(**)
	Sig. (2-tailed)		,001
	N	16	16
T3_SUB_DUR	Pearson Correlation	,755(**)	1
	Sig. (2-tailed)	,001	
	N	16	16

** Correlation is significant at the 0.01 level (2-tailed).

APPENDIX B

Table 37 - Evaluation of Subjects' Performance

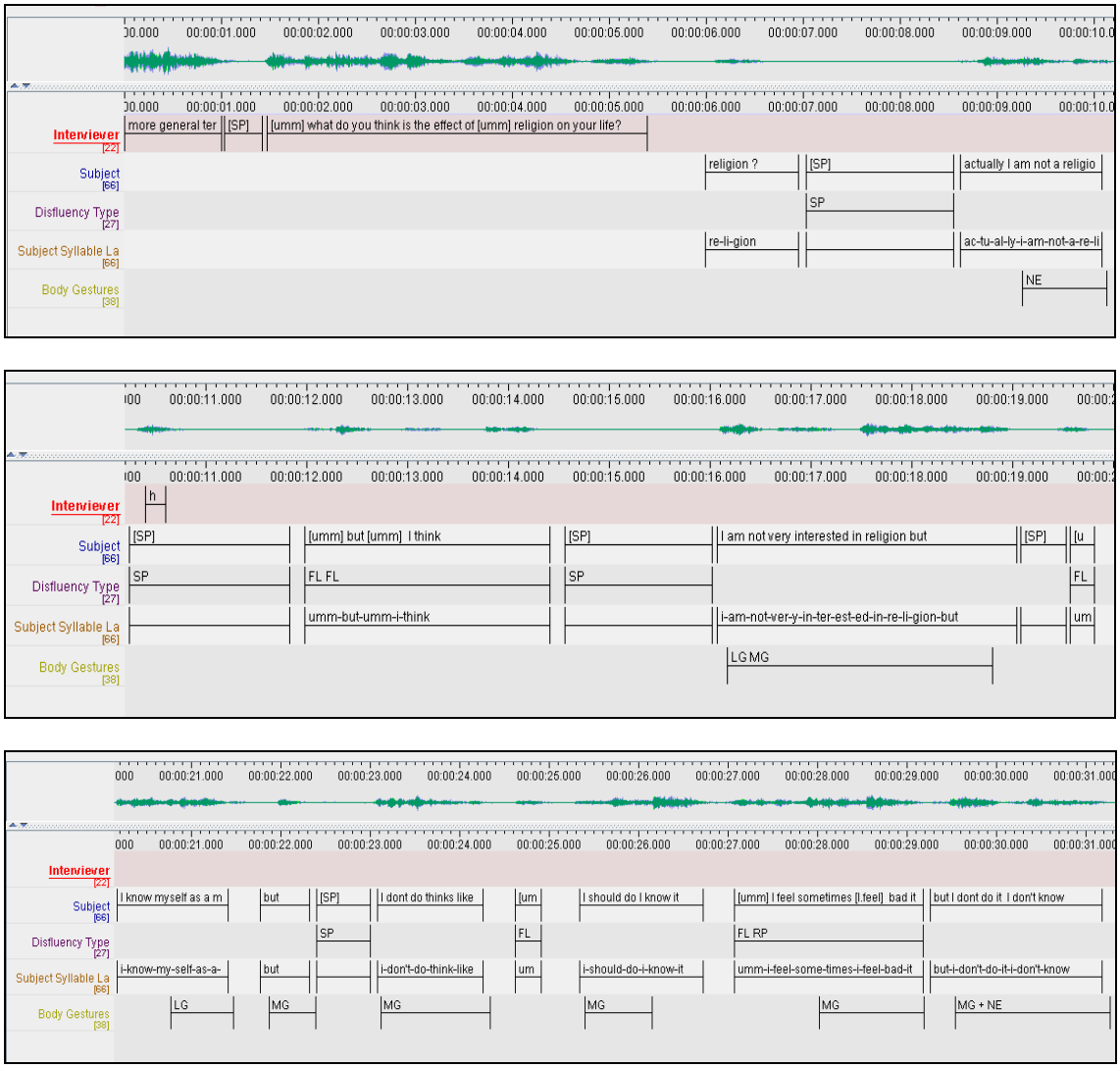
Subject No	Topic	Fluency	Accuracy	Coherence
1	Concrete	10	9	10
	Abstract	10	9	10
2	Concrete	4	8	5
	Abstract	6	9	9
3	Concrete	10	10	10
	Abstract	10	9	10
4	Concrete	6	5	6
	Abstract	7	8	8
5	Concrete	2	9	5
	Abstract	8	9	9
6	Concrete	7	8	8
	Abstract	4	5	6
7	Concrete	2	3	2
	Abstract	4	7	6
8	Concrete	7	6	6
	Abstract	8	5	5
9	Concrete	10	8	9
	Abstract	9	8	8
10	Concrete	9	8	9
	Abstract	9	9	9
11	Concrete	10	10	10
	Abstract	10	10	10
12	Concrete	8	7	8
	Abstract	8	8	8
15	Concrete	8	7	8
	Abstract	6	8	7
16	Concrete			
	Abstract			

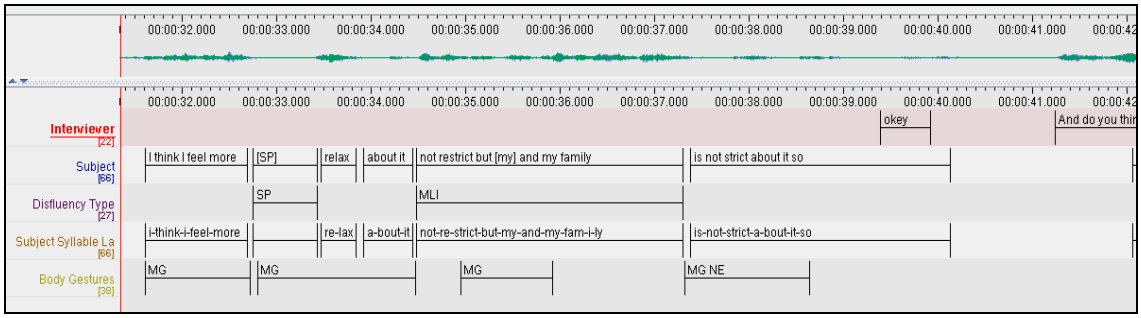
Table 38 - Evaluation of Performance overall

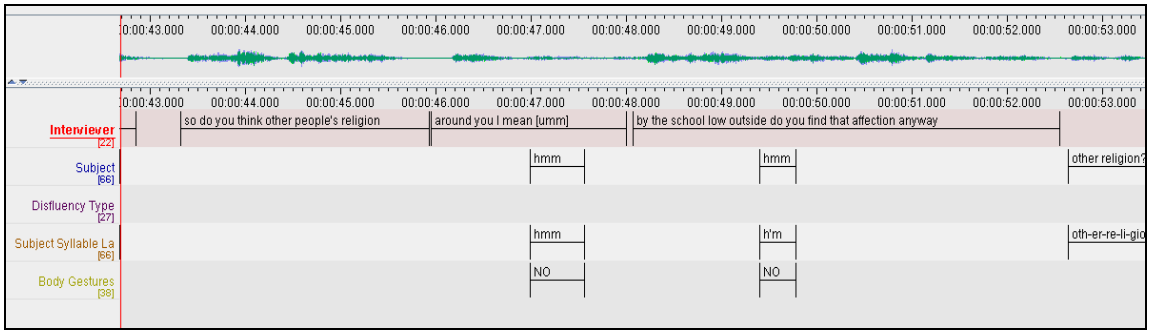
	1		2		3		4		5		6		7		8		9		10	11			12		15	
Topic	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2
The student communicates his/her ideas very well	x	x		x	x	x				X	x						x	x	x	x	x	x			x	
The student communicates his/her ideas not so well			x				x	x				x		x	x	x							x	x		x
The student cannot communicates his/her ideas at all									x				x													

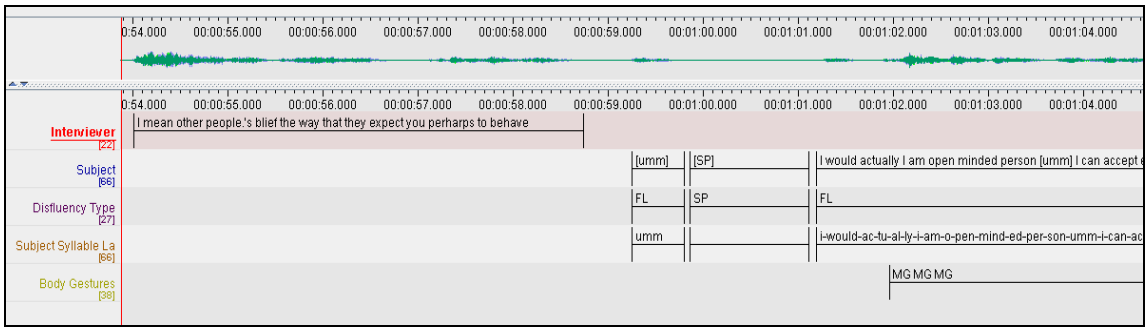
APPENDIX C

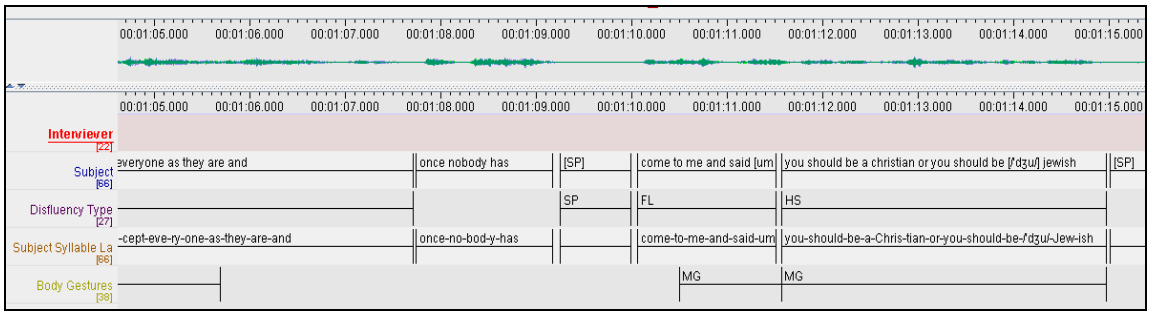
A sample Annotation example from study data

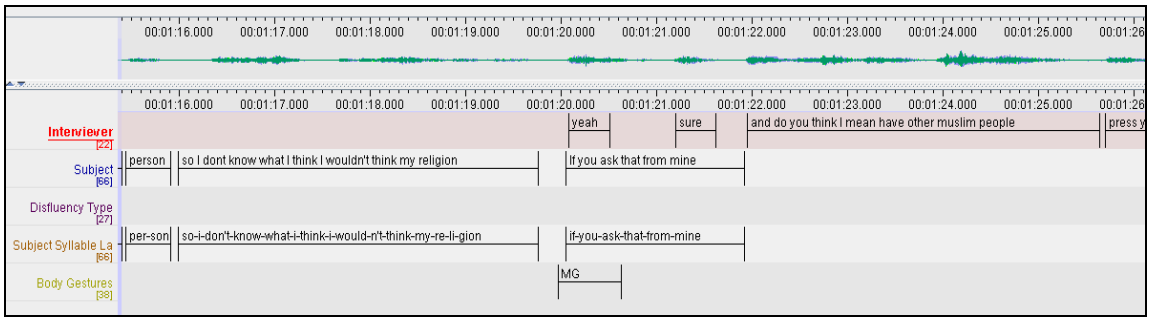


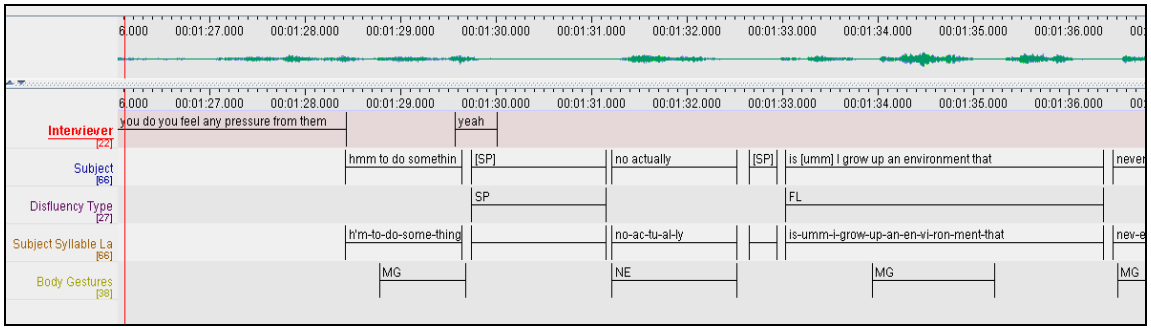

Interviewer
Subject
Disfluency Type
Subject Syllable La
Body Gestures
okey
And do you thin
I think I feel more
[SP]
relax
about it
not restrict but [my] and my family
is not strict about it so
SP
MLI
i-think-i-feel-more
re-lax
a-bout-it
not-re-strict-but-my-and-my-fam-i-ly
is-not-strict-a-bout-it-so
MG
MG
MG
MG NE

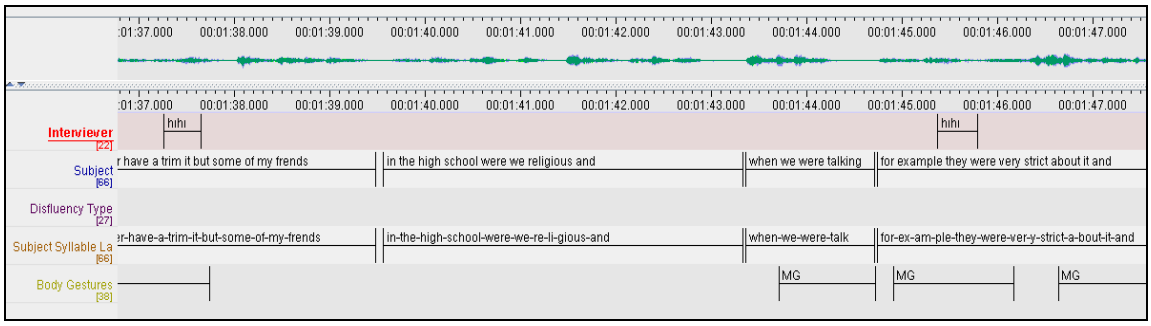

Interviewer
Subject
Disfluency Type
Subject Syllable La
Body Gestures
so do you think other people's religion
around you I mean [umm]
by the school low outside do you find that affection anyway
hmm
hmm
other religion?
hmm
h'm
NO
NO

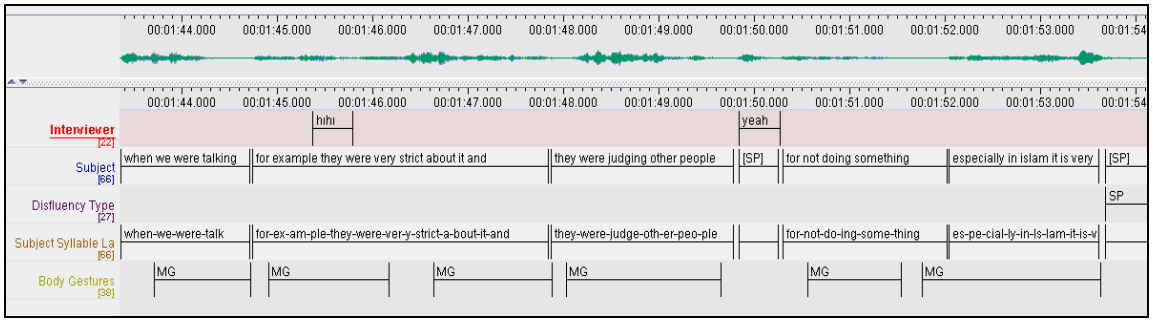

Interviewer
Subject
Disfluency Type
Subject Syllable La
Body Gestures
I mean other people.'s blief the way that they expect you perharps to behave
[umm]
[SP]
FL
SP
FL
umm
MG MG MG

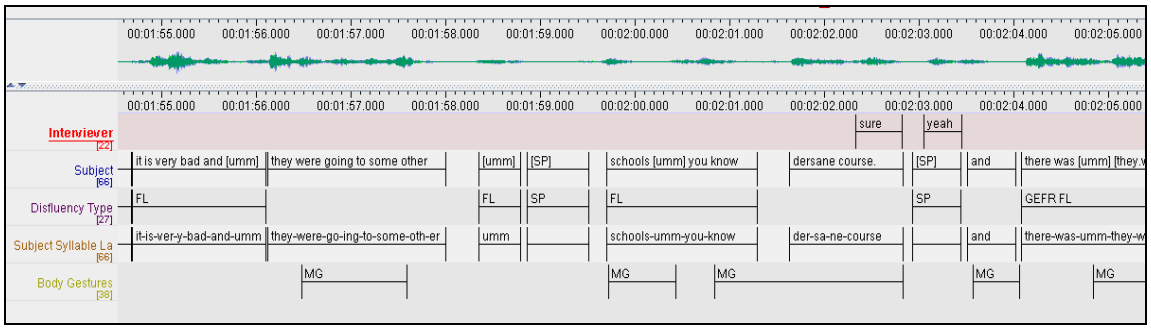

Interviewer
Subject
Disfluency Type
Subject Syllable La
Body Gestures
everyone as they are and
once nobody has
[SP]
come to me and said [um
you should be a christian or you should be [ʃ'dʒuʃ] jewish
[SP]
SP
FL
HS
-cept-eve-ry-one-as-they-are-and
once-no-bod-y-has
come-to-me-and-said-um
you-should-be-a-Chris-tian-or-you-should-be-'dʒuʃ-Jew-ish
MG
MG

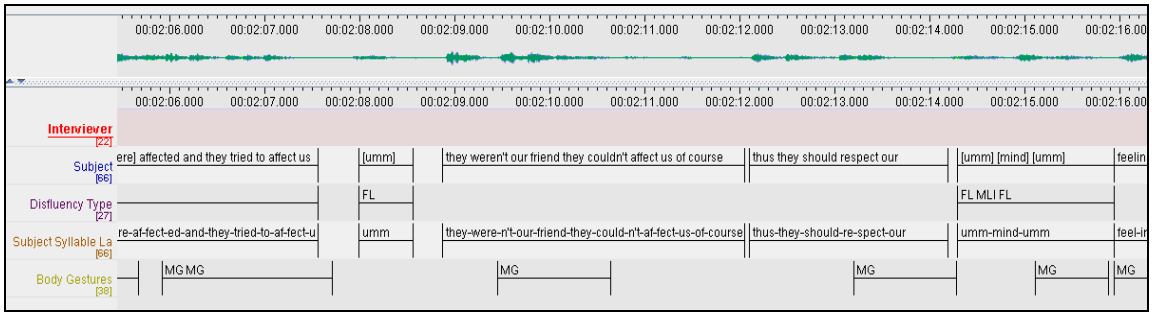

Interviewer
Subject
Disfluency Type
Subject Syllable La
Body Gestures
yeah
sure
and do you think I mean have other muslim people
person
so I dont know what I think I wouldn't think my religion
If you ask that from mine
per-son
so-i-don't-know-what-i-think-i-would-n't-think-my-re-li-gion
if-you-ask-that-from-mine
MG

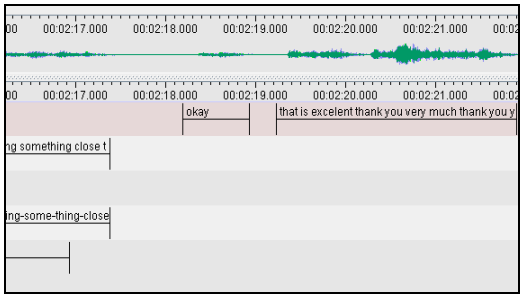

Interviewer [22]
you do you feel any pressure from them
yeah
Subject [66]
hmm to do somethin
[SP]
no actually
[SP]
is [umm] I grow up an environment that
Disfluency Type [27]
SP
FL
Subject Syllable La [66]
h'm-to-do-some-thing
no-ac-tu-al-ly
is-umm-i-grow-up-an-en-vi-ron-ment-that
Body Gestures [38]
MG
NE
MG
MG


Interviewer [22]
hihi
hihi
Subject [66]
have a trim it but some of my frends
in the high school were we religious and
when we were talking
for example they were very strict about it and
Disfluency Type [27]
Subject Syllable La [66]
in-the-high-school-were-we-re-li-gious-and
when-we-were-talk
for-ex-am-ple-they-were-ver-y-strict-a-bout-it-and
Body Gestures [38]
MG
MG
MG


Interviewer [22]
hihi
yeah
Subject [66]
when we were talking
for example they were very strict about it and
they were judging other people
[SP]
for not doing something
especially in islam it is very
[SP]
Disfluency Type [27]
SP
Subject Syllable La [66]
when-we-were-talk
for-ex-am-ple-they-were-ver-y-strict-a-bout-it-and
they-were-judge-oth-er-peo-ple
for-not-do-ing-some-thing
Body Gestures [38]
MG
MG
MG
MG
MG
MG


Interviewer [22]
sure
yeah
Subject [66]
it is very bad and [umm]
they were going to some other
[umm]
[SP]
schools [umm] you know
dersane course.
[SP]
and
Disfluency Type [27]
FL
FL
SP
FL
SP
GEFR FL
Subject Syllable La [66]
it-is-ver-y-bad-and-umm
they-were-go-ing-to-some-oth-er
umm
schools-umm-you-know
der-sa-ne-course
and
Body Gestures [38]
MG
MG
MG
MG
MG


Interviewer [22]
ere] affected and they tried to affect us
[umm]
they weren't our friend they couldn't affect us of course
thus they should respect our
[umm] [mind] [umm]
Subject [66]
Disfluency Type [27]
FL
FL MLI FL
Subject Syllable La [66]
re-af-fect-ed-and-they-tried-to-af-fect-u
umm
they-were-n't-our-friend-they-could-n't-af-fect-us-of-course
thus-they-should-re-spect-our
umm-mind-umm
Body Gestures [38]
MG MG
MG
MG
MG
MG


00:02:17.000
00:02:18.000
00:02:19.000
00:02:20.000
00:02:21.000
okay
that is excelent thank you very much thank you y
ng something close t
ing-some-thing-close