PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA

MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH

UNIVERSITY OF "FRÈRES MENTOURI," CONSTANTINE 1

DEPARTMENT OF LETTERS AND THE ENGLISH LANGUAGE

N° ordre:39/D3C/2020
N° serie:05/Ang/2020

Listening to Connected Speech: Students' Problems, Strategies, and the Effectiveness of Direct Instruction

The Case Study of Second Year Students, University of Mohamed Boudiaf- M'sila

Thesis Submitted to the Department of Letters and English Language in Fulfilment for the Requirements of an LMD Doctorate Degree in Applied Linguistics/ Applied Language Studies

Submitted by:

Mr Mohamed LAOUBI

Supervised by:

Prof. Youcef BEGHOUL

Board of Examiners **le 01/10/2020**

Chairwoman: Prof. Farida ABDERRAHIM	University of Frères Mentouri, Constantine 1
Supervisor: Prof. Youcef BEGHOUL	University of Frères Mentouri, Constantine 1
Member: Prof. Riad BELOUAHEM	University of Frères Mentouri, Constantine 1
Member: Prof. Said KESKES	University of Mohamed Lamine Debaghine, Sétif
Member: Prof. Sarah MERROUCHE	University of Larbi Ben M'hidi, Oum El Bouaghi

2019-2020

Dedication

To the soul of my **father** who did not live long to see the end of this work,

To my **mother** who is always there for me,

To my brothers and sisters and all my relatives,

To my beloved nephews,

To all my friends and colleagues,

To everyone who helped and cared,

I dedicate this work.

Acknowledgements

I would like to express my utmost gratitude to my supervisor Prof. Youcef Beghoul for the unstinting support and the continuous encouragement he provided me with throughout all the stages of this work. I thank him for his help and guidance, his confidence in my abilities, and his valuable comments. It has been the greatest privilege and honour to have him as my supervisor.

I would like to extend my gratitude to the members of the board of examiners, Prof. Farida ABDERRAHIM, Prof. Riad BELOUAHEM, Prof. Said KESKES, and Prof. Sara MERROUCHE, for accepting to examine and evaluate this work. With their comments and expertise, valuable enhancements would undoubtedly shape the final version of this work.

I am greatly indebted to the head of the Department of English at Mohamed Boudiaf University- M'sila, Dr Mohammed Saber Djemoui, and to Dr Mourad Touati, the former head of the Department. Any words would not be enough to express my sincere gratitude to them for their warm welcome within the department and for the help which I did not find anywhere else. I would not forget to thank my second year students who took part in this study.

My sincere gratitude goes to my friends and colleagues: Dr Alouache Amel, Dr Boudjemâa Dendenne, Dr Imen Chaalal, Dr Maha Lounis, Mr Abdellah Yassia, Mr Abdennour Ahmed Yahia, for their appreciated encouragement and support.

Abstract

The purpose of this study is to examine the effectiveness of connected speech instruction in improving learners' decoding of spoken English and their reactions to it. It follows the trend of a balanced listening session that addresses bottom-up processes and decoding skills without compromising the need to develop top-down processes. Thus, it attempts to examine the effectiveness of a teaching framework which allows for integrating connected speech instruction into the listening comprehension lessons following the principles of the diagnostic approach *(suggested by Field, 2008a)*. The study included a sample of thirty-eight second year students of English from the University of Mohamed Boudiaf -M'sila, divided into two equal groups; an experimental group and a control group. After a pre-test on connected speech decoding, the experimental group received listening comprehension lessons with an extended post-listening phase to address the connected speech features which were diagnosed as problematic during the listening phase. The control group received similar listening lessons without a specific focus on connected speech. After the treatment, a post-test was administered to both groups and the results were compared. The pre-test results revealed that the connected speech aspects of assimilation, elision, weak forms and liaison pose serious obstacles in the processes of lexical segmentation and word recognition. The analysis of the mishearings showed that the participants may resort rather unsuccessfully to top-down processing in a bid to compensate for their inability to decode connected speech features. The mishearings also revealed a poor application and coordination of some top-down strategies, namely predicting and monitoring. Poorly decoded connected speech at the beginning of utterances was found to constrain the way in which the unfolding text is processed by triggering inferences about syntactic and lexical forms (cognitive effects). Even if they are wrong, the participants tended to stick to these inferences despite the presence of sufficient contradictory evidence in the unfolding text. Quantitative data analysis

confirmed the hypothesis that connected speech-integrated listening instruction is effective in improving learners' decoding of speech. A post-instruction interview with the treatment group revealed positive reactions to the lessons.

List of Abbreviations

BU: Bottom-up

CG: Control Group

CS: Connected Speech

EFL: English as a Foreign Language

EG: Experimental Group

ESL: English as a Second Language

L1: First Language

L2: Second Language

LS: Lexical Segmentation

LTM: Long Term Memory

SF: Strong Form

SS: Strong Strong (sequence)

SW: Strong Weak (Sequence)

TD: Top-down

TOEIC: Test of English for International Communication

WF: Weak Form

WR: Word Recognition

List of Figures

List of Tables

Table of Contents

General Introduction

General Introduction

1. Statement of the Problem

Listening comprehension has long been considered a passive skill in which very little or no effort is made by the listener to comprehend the spoken message (Vandergrift L. 2004). This state has greatly influenced second and foreign language research, theories, teaching and learning, to the extent that listening was considered, at a given time, the *Cinderella Skill* (Nunan, 2002). However, despite that research in the last three decades into the nature of listening has emphasised that it is an active and very demanding skill that needs attention just like the other three skills, it is still being somewhat neglected in second and foreign language teaching. More attention is given to speaking at the expense of listening based on the claim that learners would be enabled to engage in communicative and real life situations (Field, 2008a) such as conversations and meetings. Many questions, however, are raised in the literature about the sufficiency of developing the learners' oral productive skill to engage in communicative situations without considering their ability to understand what they receive (Norris, 1995). By definition, a conversation is a two-direction way of communication in which the interlocutors exchange roles and they are required to *give* and *receive* information. Hence, in any communicative situation, it would be meaningless to speak without being able to receive.

A recently recognised and widely investigated problem related to second and foreign language listening is the learners' failure in recognising words and breaking down sentences produced by native speakers (e.g. Wilson, 2003; Field, 2008b). At the perceptual phase of the listening process, learners find it difficult to cope with the stream of speech and to recognise words and word boundaries, even those that they already know. Research has revealed that learners who lack special expertise in listening to authentic material do have

problems in decoding the messages when they encounter real life situations. It has also emphasised that listening should not be taken for granted, and that language teaching programmes and curricula should pay more attention to it so as to remove the learners' burden in communication.

Being used to types of adapted input inside the classroom only, untrained learners usually fail to recognise chunks of naturally occurring connected speech produced by native speakers (Brown G. , 1990). We have observed through our own experience that, even inside the classroom, university students usually fail to break down sentences and recognise words and reduced forms in speech if the teacher uses a connected native-like pronunciation. Students ask the teacher to repeat what is said, and the problem is usually solved as the teacher repeats but with the use of modified and adapted language, instead. At university, students are given courses about some connected speech aspects on the assumption that these will raise students' awareness of their importance in the accuracy and the intelligibility of the language they produce. In listening, these courses stress mostly declarative knowledge, i.e. knowledge about the language, and give little attention to procedural knowledge i.e., knowledge about the mechanisms which make language work both in production and in comprehension. Research, however, has shown that students have difficulties in listening comprehension regarding the perception of some aspects of English connected speech despite that they are given theoretical courses about them.

By now, it is established by many (e.g. Anderson & Lynch, 1988; Vandergrift L. , 2007; 1999) that developing learner's listening comprehension both as a channel for language learning and as a skill on its own right is a prerequisite in second and foreign language learning. There is also agreement on the need to diagnose learners' problems in listening prior to giving them a practical chance to develop their comprehension abilities with a focus on procedural knowledge. Many studies have been carried out with this aim in

the scope of listening to the English connected speech, the result of which is a total support for the idea of giving the listening skill primary importance in language pedagogy.

2. Aims of the Study

This study focuses on the students' problems in listening to natural speech; it investigates some of the problems that lie behind the students' failure to recognise naturally occurring stretches of speech and the effectiveness of instruction in addressing them. The objective is twofold; firstly, we aim at understanding the difficulties that students at the Department of English Arts and Language, University of Mohamed Boudiaf –M'sila-, encounter when listening to connected speech. We try to diagnose and understand the nature of the students' misperceptions of common weak forms, assimilation, liaison, elision and contractions in English, and to question the strategies related to the top-down or the bottom-up listening modes which may apply during the listening process to cope with such pronunciation aspects in connected speech. Secondly, we address the issue of teaching connected speech to university students and examine the effectiveness of listening-oriented connected speech instruction following the diagnostic approach in improving the students' performance in listening. Previous research has shown that students who lack training and who are unaware of the significance of reduced forms and other aspects of connected speech in listening are more likely to misperceive them while listening to natural speech, which they usually consider "too fast". This is why, we will review some of the approaches followed in teaching listening in general and how connected speech is addressed in specific. The study also makes some recommendations for addressing learners' decoding problems to improve their listening comprehension skill.

3. Research Questions

The main question this study addresses is the efficiency of teaching connected speech with the objective of developing students' performance in listening. Stated differently:

- Is connected speech instruction through the application of the principles of the diagnostic approach to decoding effective in improving learners' decoding of speech?

In addition, other related questions that the study raises are:

- To what extent do undergraduate students at the Department of English Arts and Language, University of Mohamed Boudiaf -M'sila -, fail to perceive connected speech aspects in naturally spoken language?
- What could be the effects of misperceiving the different types of connected speech on the students' decoding of speech, and on their listening comprehension process in general?
- What would be the students' reactions to the instruction that addresses their decoding of connected speech through the diagnostic approach?

4. Research Hypotheses

Based on the research questions raised in this study, there are three hypotheses that we aim at verifying:

- Connected speech aspects would pose problems for the students in decoding naturally spoken English and in making lexical segmentation.

- In addition to the poor bottom-up skills in decoding connected speech features, the top-down processing mode would have negative effects on the participants' recognition and segmentation of connected speech.
- If students received connected speech instruction that focuses on awareness-raising and provides practice following the principles of the diagnostic approach, their ability to decode connected speech phenomena would improve.

In addition, we hypothesise that the participants in the experiment would have positive reactions to the instruction. We assume that the ability to decode connected speech automatically after training would have positive effects on their overall listening comprehension ability.

5. Research Tools

To answer the questions and verify the hypotheses made in this study, a set of research instruments are used. To check the hypotheses, an experiment is conducted on a sample of students at the Department of Arts and the English Language, University of Mohamed Boudiaf -M'sila-, to find out about the effectiveness of listening-oriented direct instruction on connected speech on the students' performance in listening. The sample is divided into two groups; an experimental group and a control group. In a pre-test, students' decoding abilities regarding the aspects of connected speech are evaluated. After the test, the experimental group receives lessons on the common aspects of connected speech diagnosed as problematic for the learners in listening following the principles of the diagnostic approach. The lessons the experimental group receives focus on:

- Introducing the processing modes; the bottom-up and the top-down modes.
- Listening tasks that develop top-down processing and strategies.

- Diagnosing learners' main problems in decoding speech and the features which lead to listening/decoding failure.
- Introducing the main aspects of connected speech and sound modification in spoken English in addition to some basic theoretical information about their nature, the different forms, and some common and systematic patterns.
- Designing listening tasks that aim at raising the students' awareness to the significance of the aspects of English connected speech.
- Focusing on form and other training tasks and small scale tasks whose objective is the improvement of automaticity in processing.

After receiving the instruction, the experimental group together with the control group take a post-test of the same form as the pre-test. The results are compared to see the effects of the experimental conditions on the experimental group.

The answers to the pre-test are analysed to diagnose the students' main difficulties in the perception of some aspects of connected speech, and the effects of the different types of modifications on their decoding process. The test is comprised of a cloze test and a dictation test. The participants are instructed to listen to a script and fill in the blanks or finish utterances on the basis of their listening. From the answers they provide, the misperception patterns are analysed with the purpose of finding the common problematic forms which are likely to hinder the understanding of connected speech. The patterns of mishearings are also analysed to understand the effects of the listening modes on the decoding process.

After the experiment, semi-structured interviews are conducted with the members of the experimental group to analyse their reactions to the instruction, and their knowledge about connected speech before and after the lessons.

6. Structure of the Study

The study consists of five chapters and is divided into a theoretical part and a practical one. The first two chapters constitute the theoretical part of the study, and the results of the experiment and the analyses are presented and discussed in chapters four and five.

Chapter one includes a literature review about the nature of listening comprehension and its status in language learning. The chapter also outlines some theoretical models of listening, and focuses on those designed to describe the language learners' listening processes and the likely problems they encounter. More specifically, the chapter gives more attention to the additional processes specific to a description of non-native listeners' listening skill, and the problems which result from the failure to decode the speech signal.

Chapter two is devoted to introducing the idea of connected speech and sound modification in spoken English. It presents, in the light of the suprasegmental features of spoken English, the reasons that make reduced forms *the norm* instead of the citation forms. The aspects of connected speech are discussed in detail with a particular focus on the common reduced forms in English which are more likely to hinder listening at the perceptual level of the whole process. The chapter also includes a review of the studies that have dealt with the problem of listening to connected speech and the effectiveness of instruction. Two major approaches –the strategy based approach and the skills based approach – are discussed to examine how connected speech aspects have been addressed and investigated.

Chapter three sheds light on the research tools selected for data collection, and the analysis procedures. It includes an illustration of the quasi-experimental design, a description of the connected speech decoding tests and the way they are administered, the treatment conditions of the experimental group and the control group, in addition to an explanation of the data analysis procedures.

Chapter four is concerned with quantitative data analyses; it presents the results of the decoding test, and gives a detailed account of the statistical procedures followed to analyse the results. The results of the two groups in the pre-test and the post-test are presented, and the mean scores are calculated to evaluate the participants' ability to decode the connected speech aspects. Then, the pre-test and the post-test means are compared using statistical procedures to assess the effects, if any, of the experimental conditions on the experimental group.

Chapter five focuses on the qualitative data analysis. In the light of the literature review presented in the theoretical part, this chapter presents the analysis of the participants' problems in listening to the different types of connected speech features, together with the analysis of the common patterns of mishearings. The performances of the two groups in the pre-test are analysed for the purpose of evaluating their ability to decode speech in the presence of the different connected speech aspects. More specifically, the focus of the analysis will be on the effects of the phonological modifications on the processes of word recognition and lexical segmentation. In addition, the patterns of mishearings are also grouped and coded for the purpose of unveiling any effects of top-down and bottom-up listening strategies on the participants' decoding of speech.

Chapter One

Listening Comprehension: Literature Review

Chapter One

Listening Comprehension: Literature Review

Introduction

This chapter provides an overview about the nature of LC, its status in language learning and its sub-processes. It provides a brief historical account of the developing theories about the nature of listening and tackles the issue of how different views have had a strong effect on the roles assigned to the skill in language learning. A thorough understanding of the status of listening and the views about it will provide a basis for analysing the theoretical models that attempted to describe what happens inside the listeners minds as they try to understand spoken language. Despite their inefficiencies in showing how meaning is created, earlier theoretical models had set the foundation for further research into the area of speech comprehension and interpretation and provided an important basis for describing the listener's cognitive processes. Such processes and the way they interact will be examined in the discussion of recent mental models of listening in both native and non-native contexts. For non-native listeners, the task of comprehending speech requires more focal attention and additional processes, especially at the beginning stages of learning. Some of the additional problems that non-native listeners have in comprehending foreign speech will be addressed, with a partial focus on the lower level processes of decoding the speech signal.

1.1. Nature and Definition of Listening

Supported by much research interest and contributions form a range of disciplines, the definition of listening has been subject to considerable changes over the past 50 years. The changes reflect not only the developmental views about listening, resulting from the

accumulation of research findings about it, but also the different perspectives from which the listening process is approached (Rost, 2011). These are two reasons why there are as many definitions of listening as there are different contributions into the phenomenon from a wide range of disciplines, and why there is no one single definition upon which there is a total agreement (Witkin, 1990; Dunkel, 1991; Rost, 2011). The multidisciplinary view of listening was acknowledged by Richards who stated that "current understanding of the nature of listening comprehension (henceforth LC) draws on research in psycholinguistics, semantics, pragmatics, discourse analysis, and cognitive science" (1983, pp. 219-220). For Witkin (1990), the fact that listening has attracted attention from different perspectives is a two edged sword, and this can be considered from two different angles; 1) that the contributions into the listening literature form a wide range of scholarly works provide an enrichment to our understanding of the process and, 2) that, unfortunately, much of the confusion we have about listening in fact stems in the different views.

Although there is an absence of agreement about the definition, what seems to be a point of consensus among scholars is that listening is a highly complex and covert phenomenon that involves a number of skills and processes (Byrnes, 1984; Ur, 1984; Witkin, 1990; Wolvin & Coakley, 1993a; Celce-Murcia, 1995; Brindley, 1998; Vandergrift L. , 1999; Anderson J. R., 2010). The discussion of the multidisciplinary view *per se* is evident of the complexity of listening. In addition, the view of listening as complex may also stem from the fact that it is an abstract skill which is difficult to observe. Unlike the other language skills, listening is a covert skill that involves more than the physical reception of the input. Hauser & Hughes (1988) stated that,

> We can usually tell by observation whether a receiver is paying attention to sound […]; but once the information vibrates the

> tympanic membrane […] we can no longer directly observe what happens as the receiver processes the message. However, we can recognise that whatever happens once the individual physically receives the message stimulus involves cognitively interpreting, storing, and using the information (p. 75).

Hauser & Hughes refer to a key element that is common in most of the definitions and discussions about listening which consider it as a primarily cognitive process. Unlike the traditional definitions such as those which consider listening as the simple process of "the attachment of meaning to aural symbols" (Nichols, 1947, p. 84), Hauser & Hughes acknowledge that perceiving the acoustic signal is only one phase in the process of listening and that understanding requires further cognitive processes. The listener has to actively work on this input in order to create an understanding of the speaker's intended meaning. Rost (1990) clarified this point by defining listening as "an essentially inferential process based on the perception of cues rather than straightforward matching of sound to meaning" (1990, p. 33). He argued that the listener, in addition to perceiving the speech sounds, can make use of the context to infer what will be said through *cognitive effects* –lexical, schematic and syntactic effects (p. 50). The following are some of the commonly cited definitions that stress the cognitive perspective:

- "[listening] is a complex communication behaviour, involving a process of receiving, attending to, and assigning meaning to verbal/or nonverbal stimuli." (Wolvin & Coakley, 1993a, p. 21)
- "Listening can be defined as the process of selecting, organising, and integrating information." (Imhof, 2010, p. 98)

- “Listening is the active and dynamic process of attending, perceiving, interpreting, remembering, and responding to the expressed (verbal and nonverbal) needs, concerns, and information offered by other human beings.” (Purdy, 1997, p. 8)
- “Listening is conceived of as an active process in which listeners select and interpret information which comes from auditory and visual clues in order to define what is going on and what the speakers are trying to express.” (Rubin, 1995, p. 7)
- “It is a complex, active process in which the listener must discriminate between sounds, understand vocabulary and grammatical structures, interpret stress and intonation, retain what was gathered in all of the above, and interpret it within the immediate as well as the larger sociocultural context of the utterance.” (Vandergrift L. , 1999, p. 168)

Instead of adopting a single definition of listening, one good approach is to review common elements in the scholarly definitions to arrive at a broad and comprehensive view about listening and how it is theoretically conceived (Glenn, 1989). In the definitions presented above, the processes of attending, remembering, assigning meaning, interpreting and selecting are, among others, some of the common cognitive terms highlighted by Glenn (1989) in her content analysis of 50 definitions of listening. Glenn’s analysis resulted in the following seven main entries although different terms were also used as synonyms to express similar processes: *perception, attention, interpretation, remembering, response, spoken sounds* and *visual cues* (Glenn, 1989, p. 25).

While *interpreting* and *assigning meaning* were the most frequently cited elements, *attending* was acknowledged as a basic element by nearly half of the authors. Here, a distinction must be done between listening and hearing (Nichols, 1947; Helgesen & Brown, 2007; Imhof, 2010; Prince, 2014) . For Imhof (2010), attending to the aural signal is what distinguishes the conscious process of listening from the physical unconscious act of hearing.

Anyone who hears sounds, noise, music or any other auditory stimulus is not necessarily listening, but the opposite is true. While the terms are often used interchangeably (Purdy, 1997), listening is a more active process that requires a decision form the listener to attend to the acoustic signal so as to make a transition from the passive act of hearing to purposeful conscious listening (Imhof, 2010). Underwood, (1989, p. 1, in Dunkel, 1991, p. 433) highlighted this point in his simplified definition saying that listening is "the activity of paying attention to and trying to get meaning from what we hear".

What is apparent in most of the definitions stated so far is that listening for understanding requires more than attending to and perceiving speech signals. After they successfully perceive sounds and recognise words and phrases, listeners need to move to a higher level of processing to *interpret and create* meaning by making use of a variety of information sources (Field, 2008a; Lynch, 2009). Such information may include linguistic knowledge, background knowledge, sociocultural knowledge, contextual and paralinguistic cues, knowledge about the situation, the roles of speakers, and any other information sources that seem appropriate in interpreting the message. Buck (2001) explained the use of the different sources by defining listening as a process in which:

> […] the listener takes the incoming data, the acoustic signal, and interprets that using a wide variety of information and knowledge, for a particular communicative purpose; it is an inferential process, an ongoing process of constructing and modifying an interpretation of what the text is about, based on whatever information seems relevant at the time (p. 29).

Lynch & Mendelsohn (2010) included one other element that can be considered as an expansion to the ones mentioned earlier; *topic knowledge* can be used to make

expectations about what speakers are likely to say in a particular situation. It can boost understanding by enabling the listener to activate the relevant background knowledge only out of the different types of knowledge stored in his memory. Lynch & Mendelsohn (2010) define listening as follows:

> Listening involves making sense of spoken language, normally accompanied by other sounds and visual input, with the help of our relevant prior knowledge and the context in which we are listening […] it is a bundle of related processes –recognition of the sounds uttered by the speaker, perception of intonation patterns showing information focus, interpretation of the *relevance of what is being said to the current topic* and so on (p. 180; italics mine).

Comprehension is partly dependent on the inferences listeners make based on their background knowledge which can be triggered by knowledge about the topic. For Cook (2008), one cannot get much out of the text if he does not know about its topic. He adds, "The sentences themselves do not change when we know the topic, but the interpretation they have in our minds does" (Cook, 2008, p. 122).

The organisation of the different types of background and world information is usually referred to as S*chematic Knowledge* (Rumelhart & Ortony, 1977; Rumelhart, 1980). *Schemata*, a term borrowed form cognitive psychology, are defined as "the relevant packages of prior knowledge and experience that we have in memory and can call in the process of comprehension" (Lynch, 2006, p. 93). A schema includes regular patterns of events, background information, past experiences, knowledge about the world and other information that are stored in, and retrievable form, long term memory (LTM) (Anderson &

Lynch, 1988; Vandergrift & Goh, 2012). They can be triggered during the listening process to help in inferring meaning and in comprehension (Omalley & Chamot, 1990; Field, 2008a; Lynch, 2009). However, an absence of a particular schema or an activation of an inappropriate one may indeed influence the comprehension process, especially for non-native listeners (Tsui & Fullilove, 1998c; Lynch, 2006).

Two more concepts which are not so widely referred to in the definitions but are recognised as supporting elements in listening are visual stimulus and response. Alien as they might look in a definition of listening which is primarily an oral skill that involves the perception of speech *sounds*, visual clues represent one of the key sources of information that listeners may rely on in understanding what is communicated (Hasan, 2000; Imhof, 2010; Lynch & Mendelsohn, 2010). For Rost (1990), the listener can make inferences about what the speaker intends by relying on, among other, visual cues. He gave a list of variables for three types of visual cues that listeners can rely on: gaze direction (e.g. down, to the right or left of listener, directly at the listener), body position (e.g. touching listener, touching objects, moving head up and down), and facial gestures (e.g. smile, empathetic expression, head nod, angry expression) (Rost, 2011, p. 78). Although there is an argument against conceiving visual cues as prerequisite in listening, as in the examples of successful phone calls (Glenn, 1989), there is evidence form research that some visual cues such as the movement of the head and lips may help in perception and understanding, or influence them if they are incongruent with the auditory stimulus (Skipper, Wassenhove, Nusbaum, & Small, 2007; Imhof, 2010).

It is argued that response is an important element in the process of listening (Ridge, 1993). Through responding to the speaker, listeners can confirm the reception of the message, keep the talk going, negotiate meaning, highlight parts of the message that are not

clear, give signs of appreciation or refusal or any necessary feedback that can be communicated to, and interpreted by, the speaker (Vandergrift L. , 1997a). For Broughton, Brumfit, Flavell, Hill and Pincas (1980) feedback such as body language (nods, glances) and nonverbal noise (uh-huh, mm, Oh) is necessary to keep the talk going and, in cases, to assure the speaker that the message he intended is received so that he can move to the next one. In a list of seven components of effective listening, Purdy (1997) classified response in the sixth position arguing that it is "essential to completing the process of good listening" (p. 9). Yet, he added another type of response that he called *internal* suggesting that it allows the listener to internally integrate and comment on what is comprehended.

Listening skills and listening strategies have also been, at least in the last twenty to thirty years, at the heart of much of the discussions about what constitutes successful listening, especially in non-native learning contexts (Wilson, 2003; Renandya & Farrell, 2010). Whether one or the other should be emphasised in teaching was a subject of much debate (Field, 2003). O'Malley, Chamot and Kupper (1989) added strategy application to their definition of listening stating that:

> Listening comprehension is an active and conscious process in which the listener constructs meaning by using cues from contextual information and form existing knowledge, while relying upon multiple strategic resources to fulfil the task requirements (p. 434).

This definition has clearly emphasised the active dimension of listening by characterising it as a conscious process. By including the use of strategies, whose application implies a conscious mental operation (Peterson, 2001), the authors have highlighted the active role of the listener in trying to create meaning.

1.2. Status of Listening

The status of listening has witnessed a significant change from the past to the present, especially in the context of language learning/teaching. As mentioned before, this change is the result of the developing research and the accumulation of knowledge. Some of the main stages of this development will be discussed in this section.

1.2.1. Traditional Myth: Listening as a Passive Skill

It is not until relatively recently that listening has been recognised as a complex and dynamic process of creating meaning. Before that, however, it was believed that both listening and reading, in contrast to speaking and writing, were passive language skills in which the understanding of messages requires little, or no, effort form the part of the listener/reader (Morley, 2001; Osada, 2004; Vandergrift L. , 2004; Harmer, 2007; Helgesen & Brown, 2007; Field, 2008a; Lynch & Mendelsohn, 2010). The listener's role was limited to the reception of fully packaged messages sent by a given source regardless of the situation, the context, the speakers and the intended meaning, i.e. the message can be directly poured into the ears of the listener whose role is to simply receive the information and "passively register the massage" (Lynch & Mendelsohn, 2010, p. 180).

Until the late 1970s, listening was considered as the ability to receive, memorise and use information (Helgesen & Brown, 2007). The listeners' higher cognitive processes, however, were ignored. Anderson & Lynch (1988) demonstrated the traditional opinions that considered listening as a passive skill through an analogy between the listener and a tape recorder. They consider that the conceptualisation of listening as a passive behaviour is an attempt to ignore the listener's mental capacities and to treat him as a device that performs mechanical tasks. Conceived of this way, listening was stripped from its interpretive

dimension that allows the receiver to build personalised understanding of what speakers say. Just like the tape recorder, the role of a passive listener would be only to receive and store incoming sounds and not go beyond.

One possible explanation for the non-active views about listening is the covert nature of the process in contrast to the other language skills. We can observe speakers speak and even analyse their speech waves in a spectrogram. We can also follow the movements of a writer's hand or a reader's eye movements through a paper. However, as Field (2008a) noted, listening happens "in the hidden reaches of the learner's mind" (1p. 1), and the fact that no similar observations can be made about listeners and the state of their minds during the act of listening has possibly led to these wrong assumptions. Considering listening as passive may also be the logical result of teachers' ignorance of its implicit and complex nature especially.

Some traditional classroom practices in second language[1] (L2) or foreign language (FL) contexts have indirectly supported the claim that listening is passive. This has also been reinforced by some teaching techniques. Reporting on her own experience in teaching listening using the traditional Comprehension Method, in which learners are supposed to listen and *then* answer comprehension questions, White (1998) strongly criticised the approach. She argued that the method does not give students the right to interact, show their misunderstandings of what they receive or respond. Their presence during the listening phase is one of receiving signals only. However, this is not what people usually do either in their first language (L1) or L2 face-to-face encounters that take place in real life. Anderson & Lynch (1988) gave a similar criticism to the approach. They argued that such practices

[1]The controversial issue of second language (L2) and foreign language (FL) and the difference between them is out of the realm of this study. Throughout the thesis, 'L2' will be used, following Mitchell & Myles' definition, to refer to "any languages other than the learner's 'native language' or 'mother tongue' " (2004, p. 6). Hence, any possible use of 'L2' / 'FL' individually will fall within this scope.

would, in addition to unnaturally separating the skills of listening and speaking during instruction, promote a passive view of listening. They added:

> This is likely to lead [learners] to suppose that successful listening is purely receptive activity in which you merely receive and record what you hear, rather than actively attempt to integrate the incoming information and seek clarification when that interpretation-building process meets an obstacle (1988, p. 15).

Not only teaching practices, but also the assumptions about how people learn to listen in their first language have influenced the way the L2 listening skill was perceived. Listening is rarely addressed as a subject matter in first language (L1) the way reading or writing are. Morley (1972, p. VII, cited in Morley, 2001, p. 70) noted that the fact that L1 listening skill develops automatically and that it is seldom given special attention could have been transferred to the views about L2 listening. That is, the assumption goes, L2 listening too will take care of itself without special instruction (Mendelsohn, 1998). Obviously, though, it is only through listening to a language other than the native one that the challenges and the active nature of the listening process become apparent.

For a long time, the passive view of listening has had a strong impact on L2 learning and teaching approaches, especially in the early years of the second half of the 20th century. Listening was neglected as a basic language skill and its role in many teaching approaches was only peripheral. Listening instruction was considered as a means to an end rather than an end in itself, and, compared to the other language skills, it was given the least attention in language instruction.

1.2.2. Past Neglect

As one effect of the traditional myths, little attention was given to the listening skill. Compared to the other language skills, listening was the least recognised language skill. Even in language teaching and learning, much attention was given to the speaking, reading and writing, while listening was taken for granted. Different reasons lie behind this.

1.2.2.1. Listening as a Language/Communication Skill

It is established that oral communication is a two direction flow which requires interlocutors to both give and receive messages (Field, 2008a). Historically, however, listening was given very little importance as a key component in communication compared to speaking (Janusik, 2010). Purdy (1997) demonstrated this saying that "we give honours and awards to great speakers, but how many people do you know who have been recognised for their listening talents?" (1997, p. 1). He maintained that the listener was not recognised as a partner in communication, but only existed for the sake of the speaker. Hauser & Hughes (1988) shared a similar view explaining that research in the field of communication has given much more attention to sending messages. According to them, the covert nature of the listening process is perhaps one reason for the lack of research into it:

> One possible explanation for the paucity of research about the receiving component may be found in the abstract nature of that component. We cannot actually see receiving occurring. We can see someone talking or reading a transcript of what someone said, but we cannot see someone processing information (Hauser & Hughes, 1988, p. 75).

Around the middle of the 20th century, much attention was paid to speaking at the expense of listening. One of the claims was that speaking satisfies the long term needs for most learners (Field, 2008a) and that it allows them to communicate fluently and accurately outside the classroom. This was a rather confusing view because, for Field, what the learner really needs is an ability to keep the talk going during communication and this can never happen if one is unable to understand their interlocutor. The listeners' ability to sustain a conversation outside the classroom and the fact that they can be handicapped in a conversation if they are unable to receive were ignored (Field, 2002).

Not only speaking, but the skills of reading and writing have also been prioritised in language learning to the extent that listening was considered as the *Cinderella Skill* (Nunan, 2002). In research, what is written in each of the other skill areas in terms of quantity exceeds that which is written on listening (Mendelsohn, 1998; Vandergrift L. , 2007). Even reading which shares the receptive aspect with listening has been subject to much more research compared to listening. Three reasons were given for this:

- that written texts are easier to handle and analyse while spoken ones are more difficult and expensive,
- that it is more difficult to manipulate the variables in a listening text and what this implies in the experimental conditions
- that there is a real-time nature of listening and lack of tasks that can predict what pieces of text the listener is processing. For reading, this can be depicted by following the reader's eye movements moment by moment. (Fierreira & Anes 1994, cited in Osada, 2004, pp. 64-62)

Thus, the nature of the listening construct has greatly contributed to its widespread neglect and to the wrong beliefs that overshadowed its importance in language learning.

1.2.2.2. Neglect in Language Teaching

Perhaps one of the main effects of the passive view of listening has been a marginal status and role that the skill was assigned in language learning and teaching (Call, 1985). Around the 1960s, the principles of the then applied language teaching approaches reflected the wrong assumptions about the nature of listening being a non-active behaviour that does not require specific attention to develop. In both the American audio-lingual approach and the British situational approach, which dominated the scope of language learning and teaching by that time, listening was only a means to an end, and learners were seldom taught how to listen. The approaches used listening as a channel for presenting new language elements. However, listening as a skill on its own right was put "on the back burner" (Field, 2008a, p. 1); learners' understanding of speech was taken for granted and was left to develop by its own (Mendelsohn, 1998; Morley, 2001).

An observation of a typical lesson applying the principles of the audio-lingual approach (Larsen-Freeman, 2000) reveals that listening did have a role in the classroom as learners were encouraged to use their ears to memorise and learn. They were instructed to listen to and repeat utterances from the target language. The principle was that listening and repeating supported by praise and reinforcement would help for the acquisition of good language habits and, thus, boost learning (Larsen-Freeman, 2000; Martinez-Flor & Uso-Juan, 2006). This focus on oral skills came in partial reaction to the Reading and to the Grammar Translation approaches for their neglect of oral skills (Celce-Murcia, 2001; Grittiths, 2008). It allowed for both a practice and use of the oral skills that had been ignored in the former approaches. However, all of these approaches have in fact neglected listening, each in its own way.

Whereas the Reading and the Grammar Translation approaches' ignorance of listening was through focusing on reading and writing with a minimal attention paid to the other skills (Grittiths, 2008), the audio-lingual approach's neglect of listening was in actually using it the *way it was used* and in the role it was given in the language classroom. In the latter method, it is not the understanding of language that was sought through listening activities, but rather the recognition and the discrimination of speech sounds (Martinez-Flor & Uso-Juan, 2006). The ultimate goal was the development of a so called speaking competency to communicate fluently, and listening constituted the device that would guarantee this and nothing beyond (Osada, 2004). Teaching learners how to understand native speakers' speech, on the other hand, was mostly ignored. The British situational approach for its part did not exceed the limits of "grammar and pronunciation drills and learners' imitation of dialogues" (Morley, 2001, pp. 70-71). Field (2008a), commenting on the neglect of listening in such teaching approaches, wrote about the limits within which the teacher following them would use listening:

> There was a time when listening in the language classroom was almost entirely subordinate to the presentation of new items of language. Short dialogues on tape provided examples of structures to be learned […] and *this was the only type of listening practice that most learners received* (p. 1, italics mine).

This suggests that the listening skill was marginalised through the role it was given in the teaching methods. In addition, teaching learners how to listen was not a main objective for instruction as the understanding of native speech was expected to develop automatically through exposure to the target language.

Popular beliefs about language proficiency may also explain the tendency to give the listening skill a peripheral role in language learning. It is commonly assumed that proficiency in language is something that should be observable in one's performance (Nunan, 2002). It follows that a learner's mastery of the language is equivalent to their ability to demonstrate the knowledge they have through observable acts like speaking and writing but not listening and reading. Such a belief resulted in the wrong view that listening and reading are, Nunan wrote, "…secondary skills –means to other ends, rather than ends in themselves" (Nunan, 2002).

The audio-lingual approach received much criticism as to its ineffectiveness in developing learners' listening skills. Having the development of speaking as its main objective, and its use of listening as a means to this goal resulted in learners' limited capacity to understand native speakers (Belasco, 1971, cited in Morley, 2001). This focus on productive skills overshadowed the role of listening as a basic skill for successful real-life communication. Too often, learners taught through the conventions of the audio-lingual method were "unable to listen to real discourse without getting lost" (Meyer R. , 1984).

Developments in language teaching approaches have reconsidered the role of listening and how language understanding rather than repetition would guarantee learning (Martinez-Flor & Uso-Juan, 2006). To this end, the listener's role in creating meaning and his active participation during listening had to be first recognised.

1.2.3. Late 21st Century Views: A Revised Perspective

The late 21st century has seen considerable changes concerning the assumptions about the nature of listening and the significance of listeners' cognitive processes. The changes in their turn have influenced language learning and teaching approaches which

revised many of the traditional assumptions and took into consideration research results from different fields. Mendelsohn (1998), commenting on the state of LC teaching and research, wrote: “If we consider the state of teaching (or, more correctly, non-teaching) of listening [in the previous teaching approaches] some 50 years ago, and we look at where we are today, we have certainly made great progress” (1998, p. 81). This state of affairs has followed years of research into the nature and process of listening.

Progressively, listening was stripped form the habit formation and stimulus response roles that it used to play in the audio-lingual approach as soon as comprehension was recognised as basic. Listening came to be considered as a non-passive activity and “special emphasis was given to the mental and cognitive processes involved in the comprehension act.” (Martinez-Flor & Uso-Juan, 2006, p. 31)

One of the main concepts that have been revised was the passive view of listening. Contrary to traditional views, it is now recognised that listening is “anything but a passive activity” (Vandergrift L. , 1999, p. 198). Throughout the 1970s and 1980s, the skill gained an important status as a result of research evidence that contradicted traditional beliefs. Although late, an agreement was reached about reconsidering the nature of listening and recognising its active nature and complexity (Morley, 2001). Accumulating knowledge from a wide range of disciplinary studies dramatically accelerated the replacement of the earlier belief with one that viewed listening as very demanding (Morley, 2001, p. 85).

For understanding to take place, listeners are now believed to actively work on the input by fully engaging with it rather than passively receive sounds (Harmer, 2007). Lynch & Mendelsohn (2010), assert that current understanding of listening implies that the performance of a listening task requires the same degree of effort that a speaker makes. The ‘active’ label added to listening also implies a crucial additional dimension that takes into

account different possible interpretations which take contextual and non-linguistic variables into account. By using these and other information sources, it is the listener's cognitive operations that guarantee comprehension and learning but not drilling and memorising (Martinez-Flor & Uso-Juan, 2006). That is, understanding goes beyond the mere physical reception of sounds and the recognition of words and phrases. For Broughton et *al*. (1980), listening is active because: 1) the task of decoding requires an active participation form the listener and, 2) the listener has to maintain the line of communication between him and the interlocutor, and this is done through the different types of feedback he keeps sending as an attempt to collaborate during the act of communication.

As a language skill, the development of learners' listening ability has been isolated as an objective in itself in L2/FL contexts. The idea was that, just as one learns to speak, write or read, learning to listen should not be an exception. Around the 1980s, LC instruction and research gained considerable attention and it was only after heated debates that the skill was recognised as a key element in language learning (Vandergrift L. , 1999). The result was a raising consciousness that the listening skill deserves systematic development in language instruction just like any other skill (Vandergrift L. , 2004). Active listening requires that L2 learners should be made aware of their role in creating meaning during communication. It is believed that such awareness has grown, at least, from a theoretical perspective but work is still needed to bridge the gap between theory and classroom practice (Mendelsohn, 1998). By now, it is recommended that learners should be taught how to listen in the target language and deficiencies in their listening skills should be diagnosed and addressed (Field, 2008a).

Around the 1980s, the status of LC in language teaching approaches prospered especially with the recognition of the need to develop learners' LC skill outside the

classroom. L1 acquisition studies greatly influenced research into how L2 can be learnt better and how listening can contribute to that learning.

1.3. Importance of Listening

In addition to the importance of listening as a skill on its own right, it also has a crucial role in language learning and teaching. As a source of language input, it is believed that listening facilitates the process of language learning. In addition, it enables the other language skills to develop.

1.3.1. Listening First Approaches

According to Lynch (2009), a number of teaching methods that are practised today have their roots in the observations made about babies' first experience with language. The observations were mainly two: 1) babies experience a first silent period in which they just listen and do not speak, and 2) at a given point in their L1 learning, they can obviously comprehend more than what they produce (Lynch, 2009).

Such observations influenced and encouraged research in L2 learning/acquisition (Celce-Murcia, 2001; Nunan, 2002). The point was that, as an application of the observations, L2 would be better learnt if focus was first put on listening in the early stages of learning in a similar way as babies learn the language, while speaking should be delayed. The result was the development of what came to be called the Listening First/Comprehension Approaches to language teaching and learning which put a major emphasis on listening as a channel for L2 learning/acquisition and gave it a significant importance (Larsen-Freeman, 2000). The approaches followed the rules of nature; the natural and spontaneous way through which language (the mother tongue) is first acquired by human beings was regarded as the model which should be followed in L2 learning/acquisition as well. Hence, success in

language learning lies but in accepting that listening is the primary channel through which much of the language system is first accessed, acquired and internalised. Krashen & Terrell's (1983) Natural Approach and James Asher's (1969) Total Physical Response were leading approaches that established the importance of listening in the early stages of language learning.

Dupuy (1999) investigated the effects of a technique called Narrow Listening that is based on the principles of the natural approach. The technique requires learners to listen for several times, at the early stages of learning, to comprehensible input and try to comprehend 2-3 minutes topics by native speakers (Krashen, 1996). After surveying participants, significant results were found as to the contribution of the approach to learners' fluency, LC, vocabulary and confidence in the target language (French) (Dupuy, 1999). Similarly, Thiele & Scheibner-Herzig (1983) investigated the effects of teaching learners of English as a foreign language (EFL) following the principles of the Total Physical Response approach and found positive effects on participants. The experimental group (henceforth EG) outperformed the control group (henceforth CG) in oral performance in the post-test despite their low linguistic ability. The researchers reported that "training in listening comprehension combined with an initial delay on oral practice showed a positive effect on the experimental group with respect to their attitudes toward English lessons and with respect to anxiety" (Thiele & Scheibner-Herzig, 1983, p. 277).

Vandergrift (1999) discussed the importance and the benefits of focusing on listening at the early stages of L2 learning suggesting that listening would facilitate its development. He gave four main advantages of such an approach which prioritises listening and delays the focus on speaking:

A cognitive advantage: learners will benefit from having to focus on listening only at least at the early stages of learning to reduce the burden put upon the limited capacities of short term memory. By doing so, there will be more chance for developing comprehension skills. If, otherwise they are forced to simultaneously listen and speak in a language in which they are not yet proficient, there will be a high risk of creating a cognitive overload.

Effective advantage: more effective language learning will take place if learners focus on comprehending language first. Listening to accurate models of the target language will have positive results on the other language skills. This is to be contrasted with a situation where time is spent speaking and listening where learners may be exposed to inaccurate utterances from other learners.

Utility advantage: because listening is the most widely used language skill compared to the other skills, it follows logically that comprehension skills are intrinsic and will be at the heart of the learner's language use.

Psychological advantage: Learners usually have the fear of making mistakes in production and may feel ashamed if they make an error. Forcing learners to speak before they are ready to will run the risk of a potential embarrassment. This may result in a psychological situation that may persist in the future and deprive them from speaking in the target language because of the fear of repeating a mistake. For Vandergrift, "Without the pressure of early oral production there is less potential about producing sounds that are difficult to master, especially for adults and teenagers." (1999, p. 169). Relieving the burden of speaking first, on the other hand, will make room for a relaxed learning environment.

For Anderson & Lynch (1988), an early focus on listening in language teaching would allow for more confidence, motivation and competency for learners in the target

language. These benefits were contrasted with the risks that may arise if, otherwise, learners are required to speak at the early stages of language learning. Reference and criticism in this case were addressed to the previous approaches which had focused on speaking at the early stages of learning. According to Peterson (2001) learners who are forced to speak before they are ready to may resort to L1 habits if they fail to convey what they want in the language in which they are not yet competent. As a result, they may negatively transfer the rules of their L1 into their L2 production. Moreover, a cognitive overload may result from obliging learners to speak, and this may produce "anxiety which further inhibits learning" (Peterson, 2001, p. 88). A listening first method, on the other hand, limits the effects of these risks and relieves the learners form the burden of having to speak in a language in which they are only partially competent (Krashen & Terrell, 1983). From a psychological perspective, it would create a relaxed learning atmosphere (Asher, 1981). Peterson added that teachers of English as a second language (henceforth ESL) / (EFL) "must understand the pivotal role that listening plays in the language learning process in order to utilise listening in ways that facilitate learning" (2001, p. 99).

It should be noted that although the Natural Approach received criticism for overstating the role of listening as the only way L2 can be acquired, still no two would disagree about the role that this receptive skill plays in language learning.

1.3.2. Input for Learning

There is a widespread agreement that listening enables learning and helps for the development of L2 competence (Anderson & Lynch, 1988; Lund, 1990; Brindley, 1998; Vandergrift L. , 1999; Morley, 2001; Rost, 2006; Renandya, 2013). The question that arises here is: how can listening enable learning? The answer is simply that listening provides the

necessary input which allows for learning/acquisition to take place (Rost, 2006). According to Rast (2010), what is meant by input is:

> […] that which is in the environment that the learner can hear or see. It is the linguistic phenomena that are available for being taken in by means of aural systems (hearing) or visual systems (reading and interpreting signs and gestures) (p. 100).

The Concept of input is usually discussed in relation to *intake* which refers to "…that part of the input the learner has actually incorporated into his or her developing knowledge system, that part of the input the learner has perceived and processed, i.e., has learned" (Alanen, 1995, p. 260). It should be noted that not all input to which learners are exposed can become part of their intake. Special conditions are necessary for input to be transformed into intake such as perceptual saliency, frequency and noticing in addition to comprehension (Skehan, 1998). As for frequency, repeated encounters with specific target language forms as in the approach of narrow listening can help in the learning of these forms and in incorporating them into the learner's interlanguage system (Verspoor, Lowie, & De-Bot, 2009). Through listening input, the learner has access to many types of knowledge and learning opportunities for skill development (Rost, 2006).

Processing and understanding input are also crucial in the input-intake process (Long, 1989). Sufficiently processed input allows the learner's language system to adapt, interact and reorganise (Verspoor et al., 2009). In other words, unless it is understood, input provided through listening cannot be accessed and no such interaction may take place.

Even after that a given linguistic form is learnt, be it grammar, syntax, vocabulary or even a language process, it needs to be maintained and consolidated. Again, it is the role of

input to guarantee that new forms can be incorporated into existing knowledge and that already learned target language forms are maintained and kept alive in the interlanguage system (Verspoor et al., 2009). To this end, the listening input as a language modality has been greatly emphasised (Lund, 1990).

Input based practice in language instruction, either through listening or reading, is arguably more effective in developing leaners' linguistic system compared to output-based communicative practice (Renandya, 2013). It was noted that in output-based classrooms "[learners'] communicative skills may increase but their linguistic system remains stagnant and shows feature characteristics of those in the lower levels of proficiency range." (Renandya, 2013, p. 44). Hence, the argument goes for the need for more input to develop learners' linguistic system.

As a pedagogical plan, one goal of listening instruction is to boost learner's ability to transform into intake the input to which they are exposed (Rost, 2006). In this respect, Field (2008a) calls for special attention to listening instruction to increase learner's opportunities for learning both inside and outside the classroom. In comparing listening input and reading input outside the classroom, he argued that the former is more important because of its potential effect on speaking. He maintained that the benefits of listening instruction on language learning may go beyond the actual classroom practice, i.e. it develops learners' independence and opens doors for learning opportunities whenever input is encountered outside the classroom when the teacher is not there. For Harmer (2007), spoken texts also represent a rich source of models that learners can analyse to retrieve and learn the conventions of specific genres in production.

1.3.3. A Basis for Speaking and other Skills

LC is mainly important due to its contribution to the development of the other language skills (Dunkel, 1991; Murphy, 1991; Vandergrift L. , 1999; Cheng, 2004; Vandergrift L. , 2007). Under the relevant conditions, the skills of speaking, reading and writing will emerge spontaneously as a result of LC (Krashen & Terrell, 1983; Celce-Murcia, 2001). Learning to speak a language, for example, is considered primarily as a task of learning to hear it in the first place (Peterson, 2001). For Nunan (2002) listening is fundamental to speaking as, obviously, one cannot speak in the target language before they are able to listen and understand. Listening contributes to the enrichment of "the learner's *spoken* competence with new syntactic, lexical, phonological and pragmatic information" (Field, 2008a, p. 5). It is the skill through which humans first encounter language and which is "the basis of all subsequent communication, the foundation of all life-long reading, writing, speaking and listening activities" (Brown J. I., 1987, p. 5).

Good listeners are generally regarded as good speakers. Purdy (1997) dealt with the benefits of taking a listeners' perspective in performing a speaking task. He argued that a speaker's competence can grow once they are aware of the listeners' needs. Awareness of task requirements would perhaps not have been possible if the speaker had not experienced a similar listening situation which would have aroused consciousness about what helps listeners in that specific situation. This claim is shared by other listening specialists and supported by research findings (Anderson & Lynch, 1988). For instance, there is evidence that speakers who have already taken a listener's perspective can be more effective in giving directions which depend on the listener's needs than those who have not (Anderson, Yule, & Brown, 1984).

1.4. Listening Comprehension Models and Processes

The literature on LC is rich of so many models that describe the cognitive processes which take place in the listener's mind, and what he does with the input he receives. The models can be categorised into very broad lines, with each one showing and focusing on specific components that reflect different authors' perspectives. Variance is to a high degree linked to the areas of processing the models cover, as they represent "for the most part broad descriptions of linguistic and pragmatic competence or narrow descriptions of verbal processes" (Rost, 1990, p. 7). Despite this, there is much overlap between the models built in the fields of, for example, speech communication, information processing, psycholinguistics and cognitive psychology, with the latter having the greatest influence on model building (Martinez-Flor & Uso-Juan, 2006).

The complex nature of listening has had huge implications in model building. Such complexity implies that research is carried out from such a wide range of perspectives that the best endeavours to model what happens during listening constitute a great challenge. This is probably the reason why it is difficult to find one model that is unbiased in its description of the process towards a given theory, process, or any element that constitutes the act of listening. However, as noted by Lynch & Mendelsohn (2010), listening is better understood when models are taken in combination as complementary –rather than mutually exclusive- , with each model clearing the dark areas in another.

Analyses of the different theoretical models can demonstrate how research progressively contributed to shaping and changing an understanding of the listening process. We will present some influential models of listening with a focus on one which was designed to describe L2 listening. As a matter of fact, most of the models demonstrate the listening process in terms of stages that take place in either a linear/non-linear way, with

understanding of the intended meaning being the ultimate goal of processing in all of them. They may also have graphic descriptions that illustrate the process in an economic way (e.g., Goss, 1982, Nagle & Sanders, 1983, Vandergrift & Goh, 2012).

1.4.1. Early Information Processing and Communication Theory Models

While the terms used to describe the different sub-processes may vary across-models, what is currently common between most of them is the division of the listening process into three observable stages presented simply as:

1) Perceiving speech sounds,

2) segmenting the stream of speech and creating basic meaning in the form of propositions, and

3) evaluating and understanding the intended meaning by the listener.

Exceptions to this classification of stages are found in the early models of listening. The early attempts of model building were even simpler, and the last stage was almost absent in their descriptions. It is the stage at which personalised understanding of the message is created by the listener in light of the available information. In the early communication theory (CT) and information processing (IP) models, this role of the listener was peripheral (Rost, 1990; Lynch, 2009). Listening was conceived as the reception of whole messages transmitted by the speaker with almost no effort made by the listener. Meaning was believed to be a receivable entity, and the interpretation of the message seemed to be irrelevant since the input carries a ready-made message for the listener. Figure 1 represents an early communication theory model by Weaver and Shannon (1949).

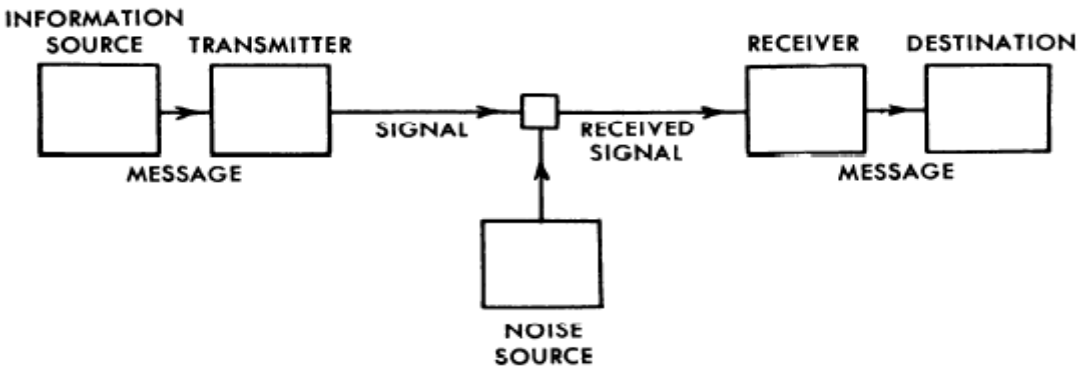

Figure 1: A Communication Theory Model (From Shannon & Weaver, 1949, p7)

The CT model (Figure1) conceptualised communication as the perfect transmission of information from a source to a destination by means of a *transmitter* and a *receiver*. In the case of speech communication, the *receiver* element in the model represents the human ear which is assigned the role of capturing the message and passing it to the brain (Shannon & Weaver, 1949). This looks like a straightforward process of passing a ready-made package of information which requires no effort to be analysed.

1.4.2. Comments on the Information Processing and Communication Theory Models

This early conceptualisation received strong criticism as to its inappropriateness in describing how humans arrive at an understanding of speech, and how the concept of meaning is shaped in both the speaker's and the listener's minds. Peterson, (2001), commented on the IP model saying that it "describes comprehension of a speaker's message as the internal reproduction of that message in the listener's mind, so that successful listening

reproduces the meaning much as the speaker intended" (p. 89). Similar comments were made about the CT Model (Rost, 1990). This explains clearly how the CT and IP models consider meaning as something that should be identical at the two ends of communication; the speaker's and the listener's. What such an attitude fails to capture is that:

1) listeners have objectives while they engage in the act of communication and rely on multiple sources of information, including their own knowledge, to make sense of the input;
2) there is a central role that the listener has to play in creating a personalised understanding of the message;
3) the contextual and the social dimensions of the communicative situation shared by interlocutors affect and shape the meaning of their utterances.

Contrary to the early CT and the IP beliefs, this last idea points to the fact that intended meaning is not something that is inherently carried by the speech signal. According to Anderson & Lynch,

> Understanding is not something that happens because of what a speaker says: the listener has a crucial part to play in the process by activating various types of knowledge, and by applying what he knows to what he hears and trying to understand what the speaker means (1988, p. 6).

In his comments on the IP model, Rost (1990) noted that meaning does not constitute a quantitatively transmittable entity, nor does it have the physical property to move. It is rather a result of a selective effortful process that could be boosted or constrained by multiple internal and external factors such as attention and purpose which may filter the input and

allow but the relevant information only to be processed by the listener. Relevance here may refer to, for example, what the listener decides to attend to in a particular situation and what he evaluates as significant. Redundant or irrelevant pieces and parts of the speech, on the other hand, may be overlooked by the listener and excluded from being processed at higher levels. With all of this in mind, one would easily justify the state of not comprehending a speaker's intended meaning despite that the listener has successfully recognised all of the words, and was able to get the literal meaning of an utterance. This inability to make sense of contextualised meaning is also possible even when language proficiency is very high as in L1 listening situations.

Despite the limitations of the CT model, it is evaluated as having an advantage in at least two main ways. The first one is that it can at least account for one, although limited, part of the listening process. In this sense, Lynch & Mendelsohn (2010) gave the example of taking someone's phone number as a simple listening task. Such a task does not require processing at higher levels. The second one is that models like this formed the basis for investigating the cognitive processes of listening and encouraged research into a view of listening that requires more than the simple passive reception of messages. In Lynch & Anderson's words, "CT stimulated thinking about the ways in which comprehension could *not* be characterised in terms of straightforward reception of a message" (2010, p. 181). Furthermore, with the inclusion of the element *noise* in its model, CT raised attention to the way in which external distractors may come to play and hinder the accurate perception of the input.

1.5. A Two Stage View

Most of the recent discussions of the construct of listening recognise that there are two main stages that characterise the whole process. These are referred to, in general terms,

as the higher level and the lower level stages/processes (Buck, 2001). This conceptualisation of listening is based on the belief that the meaning of an utterance resides not only on the words the speaker uses, but also on the listener's own knowledge and cognitive capacities, knowledge of the world, expectations, the context where the utterance takes place and other paralinguistic factors (Clark & Clark, 1977; Ur, 1984; Rost, 1990; Buck, 2001; Rost, 2011). The listener works as a partner in creating meaning relying on the information resources available that add to the basic meaning of words and phrases, and steer the process of understanding. This is an explicit shift from the traditional modelling of listening that gave a peripheral role to the listener. The latter, in the current view, bears much of the responsibility in *creating,* rather than simply *receiving,* meaning.

The two stages have considerable value in view of their recurrence in the listening literature (Buck, 2001). They are referred to as, to state some examples, "*Apprehension of Linguistic Information*" and "*Linking to Wider Context*" (Carrol, 1972), "*Construction Process*" and "*Utilisation Process"* (Clark & Clark, 1977), "*Listening for Perception*" and "*Listening for Comprehension*" (Ur, 1984), "*Speech Processing*" and "*Speech Interpretation*" (Anderson & Lynch, 1988), '*Decoding'* and "*Meaning Building*" (Field, 2008a). For Buck, variety in the terms used adds more credibility to this view of listening due to the fact that the conceptualisation of the listening process in this case was arrived at under rather individual efforts.

Research in L2 contexts has given support to the division of listening into the two stages. For example, Goh (2000) found that L2 listener's problems fall into a number of categories depending on the stage at which they occur. The results of her study revealed 10 listening problems that can be classified at either the higher or the lower level stages. Analysis of these difficulties demonstrated that a total of 5 and 3 problems were respectively

attributed to the processing failure in the *perceptual processing* and *parsing* phases of listening (equivalent to the lower-level stage) while 2 occurred due to failure in the *utilisation* phase (equivalent to the higher-level stage). The likeliness of facing difficulties with the lower level of processing by novice and lower proficiency learners is a widely accepted view (Field, 2003; Wilson, 2003; Vandergrift L. , 2004; Lynch, 2006). In Goh's study, there was a probable relation between the types of problems encountered and the listening ability of the participant groups. She reported that "a comparison of two groups of learners with different listening abilities showed some similarities in the difficulties experienced, but low ability listeners had more problems with low-level processing" (Goh, 2000, p. 55).

An extensive discussion of the two stages was given by Field in his book *Listening in the Language classroom* (2008a). He devoted two chapters to discuss each of the stages. Using his terminology -*decoding* and *meaning building* – these stages will be further discussed in light of the related literature.

1.5.1. Decoding

According to Field, decoding which represents the lower level stage of listening refers to the process of "translating the speech signal into speech sounds, words and clauses, and finally into a literal meaning" (2008a, p. 125). It is given the name 'decoding' because it entails the reception of the signal and changing its form into a meaning representation by the listener. As we will see in the next sections, a number of sub-processes constitute this stage of processing. For now, we will only discuss the knowledge sources that the listener draws upon to break down the chunks of speech into manageable representations of meaning.

The main factor that guarantees successful processing at this level is the listener's mastery of the linguistic system. Three main types of knowledge constitute this system. The first one is the listener's phonological knowledge. When sound waves reach the ear, knowledge of the sound system provides the listener with cues to depend upon in comprehension (Murphy, 1991), and allows him to perceive the sound units, recognise the stress pattern and the lexical items. The second one is the syntactic knowledge (Marslen-Wilson & Welsh, 1978; Altmann, 1991). By using the syntactic cues provided by words and word order, the constituent structures of the sentence can be identified and the relationship between them can be established (Anderson J. R., 2010). Semantic knowledge, the third type, refers to the "knowledge of the meaning of the words and the meaning of the relations between the words in a sentence. It also refers to the relations between the meanings of the individual sentences making up a discourse." (Flowerdew & Miller, 2005, p. 38). Together, these sources are generally referred to bottom-up (henceforth BU) knowledge (Robert, 2013), and the "application of all three elements of linguistic knowledge helps listeners assign meaning to word-level units and to the relationship between words at the discourse level" (Vandergrift, 2013).

Any gaps in the listener's knowledge of these types will probably lead to difficulties in comprehension (Prince, 2014). In L1 listening, as the language system is appropriately mastered, decoding takes places in an automatic fashion and does not require much attention. In L2/FL listening, the lack of knowledge at any of these levels adds to the problems of the listener (Field, 2003; Lynch, 2006; Rost, 2006; Vandergrift L. , 2007; Chenjun & Li, 2012). Zwaan & Brown explain this difference in L1 and L2 processing:

> During the comprehension of a simple text in their LI, most readers will probably not exhaust their cognitive resources. Thus,

> they will have sufficient activation available for all the cognitive procedures needed to construct a coherent situation model and achieve comprehension. This is presumably different in non-fluent L2 comprehension. In non-fluent L2 comprehension, word- and sentence-level processing will be more resource consuming than in LI (1996, p. 291).

While automaticity in decoding for L1 listeners allows them a spare capacity to concentrate on higher levels of meaning, L2 listeners have to contend with paying attention to the input at the lower level and at the same time process speech at the higher level. This state may greatly constrain the process of L2 listening (Wilson, 2003; Lynch, 2006; Field, 2008a; Lynch, 2009).

The end-up product of decoding is the creation of *propositions*. These are abstract representations of the basic meanings of utterances (Foster, 2013) that are ready to be further analysed at the other stage of processing; meaning building.

1.5.2. Meaning Building

Meaning building refers to the stage at which the listener makes elaborate interpretations of the message to comprehend its meaning. Interpretation refers to the ability to "construct meaning about and draw inferences from messages within a specific context" (Thompson, Leintz, Nervers, & Witkowski, 2010, p. 275). Field defines the higher stage of meaning building as "the process of adding to the bare meaning provided by decoding and relating it to what has been said before" (2008a, p. 125). As stated earlier, decoding the input leads to the creation of propositions that represent the basic literal meaning of the utterances.

When this is done, the resulting data is ready for making elaborate interpretation to arrive at a personalised understanding of the text. According to Prince (2014),

> Word recognition provides the basis for making sense of the input, meaning integration and storing propositional content belong to the subsequent utilisation [/meaning building] stage. Here, the listener makes most use of inferencing skills, encyclopaedic knowledge and, where necessary, compensatory strategies (p. 97).

The main difference between decoding and meaning building processes is that the former are closely linked to the language system and are conducted automatically with the least effort, whereas the latter are more dependent on reasoning, require deeper and longer cognitive processes (Goss, 1982), and draw on multiple internal and external sources of information (Buck, 2001; Lynch, 2009). Prince (2014) refers to such sources as *encyclopaedic knowledge*. To state some, knowledge of the world, background knowledge, the topic of the text and what is understood from it so far, information about the setting and other sources represent the basics of creating meaning with which the listener should be equipped. In addition to these sources of meaning, meaning building is highly dependent on successful processing at the decoding level. It makes use of, and builds on, the raw meaning of the utterances that is created after parsing the speech signal (Anderson J. R., 2010). Another difference between the processes, in the case of L2/FL learning, is that they do not share the same value when it comes to learning. Systematic knowledge is language specific and has to be reconsidered again in learning the target language, whereas the processes that make up meaning building are language independent (Field, 2008a) and need only to be

refined. The latter can be applied in comprehending the target language speech with some adjustments.

Field (2008a) provided examples of the meaning building processes that are important in amplifying literal meaning during listening (figure 2). He contends that the appropriate use of decoding and meaning building is when they are taken in an interactive way. In other words, arriving at an understanding of communicated meaning is the result of the complex interplay between the processes in the higher and the lower levels. To illustrate this, an example of the words that may have multiple meanings was given. At the decoding stage, the listener is able to decipher the stream of sounds and recognise the formal shape of the word. However, the choice in the range of meanings that the word may have leads to a lexical ambiguity that must be solved. The listener may select one meaning from among the possible ones, but he may not be sure whether it is the right one. The role of processing at the higher level is to monitor the understanding by checking the appropriateness of the chosen meaning against the context, the co-text, the topic of the talk or even background knowledge. When sufficient evidence is provided from the relevant information source(s), the listener is able to check the appropriateness of the choice he has made, or revise it if it is incongruent with the suggested evidence.

The example of dealing with ambiguity at word level raises an important question about the time relation between the stages. It suggests that there is a contribution from the information at the higher-level stage in accompanying processing at the lower stage. If this happens throughout the whole process from the time input is perceived until meaning is reached, can we assume that processing at both levels takes place simultaneously rather than successively? Contradictory views exist about the linearity/non-linearity of the listening processes.

- **'Context': using knowledge sources**

  Drawing upon: world knowledge – topic knowledge – cultural knowledge

  Analogy with other similar listening encounters

- **Deriving meaning**

  Storing the literal meaning of an utterance

  Accepting an approximate meaning

  Checking understanding

- **Adding to the meaning**

  Making inferences

  Dealing with pronouns

  Dealing with ambiguity

- **Selecting information**

  Selecting relevant information

  Recognising redundant information

- **Integrating information**

  Carrying forward what has been said so far

  Connecting ideas

  Self-monitoring for consistency

- **Recognising the overall argument structure**

  Noticing connecting words used by the speaker (On the other hand…)

Figure 2: Examples of Important Meaning-building Processes (Field, 2008a)

1.6. L1 Listening Processes and Models

A number of LC models have been suggested to account for how the L1 listening process takes place. Some of them, such as Anderson's three phase model (2010), have been considered as a starting point for L2 listening models, taking into account the differences and the additional processes that take place when listening to a non-native language. Two influential models will be discussed in this section.

1.6.1. Goss's Model of Language Comprehension

Goss (1982) suggested a model of L1 listening in which comprehension is seen as a cognitive process and an act of human information processing. He defines information processing as "the study of how people gather, store, and retrieve information as they attempt to make sense out of their environment" (Goss, 1982, p. 304), and argues that this is possible through listening. The model represents an advanced view of the nature of comprehension compared to the early IP models. This is because it not only takes into consideration what is uttered by the speaker, but also assigns a major role to the listeners' cognitive and reasoning capacities.

Goss breaks down the process into two basic functions in a way that matches the two stages view mentioned earlier, but with some minor differences. The first one is named *auditory perception* and refers to the stage at which systematic knowledge of the language is used. The second function, named *message comprehension,* demonstrates the clear departure from the early IP model in that it depends to a great extent on the listener's intellectual abilities, critical analysis and intelligence. Because of this, notes Goss, it takes a longer time compared the first stage. The two functions are further divided into three main stages: signal processing (SP), literal processing (LP) and reflective processing (RP)

(Figure3). SP is an aspect of the auditory perception function whereas LP and RP together form the basis of

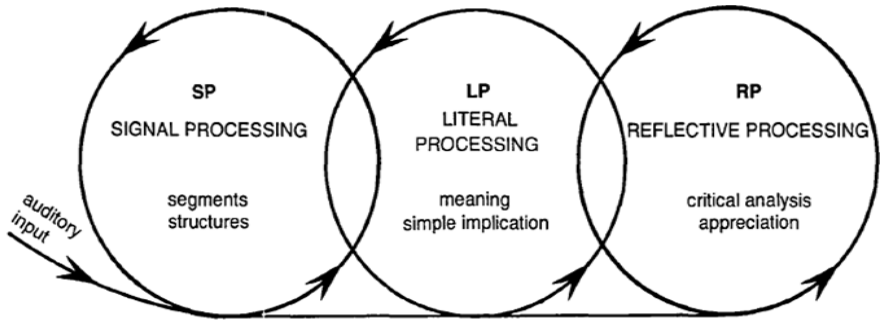

Figure 3: Goss's Information Processing Model of Listening (Goss, 1982, p. 305)

message comprehension. The role of processing at the first stage (SP) is to engage the listener's linguistic competence in the analysis of the phonetic, syntactic and semantic aspects of the signal. When this is done, the speech sounds will be segmented into potentially meaningful parts and, then, it is time for LP to take over. As the name suggests, the LP stage is only referential where the listener will have to assign basic literal meaning to the message. Goss notes that "any inferences that occur during literal meaning processing are simply a function of the listener's basic understanding of the utterance" (Goss, 1982, p. 306). This basic meaning requires to be further refined so as to reach a deeper comprehension of the text. This is the main role of processing at the last stage, RP. Deeper understanding implies a number of cognitive tasks such as inferencing, evaluating, judging speakers and messages, which result in "critical listening and appreciative listening" (Goss, 1982, p. 306). This is why Goss suggests that RP correlates with intelligence and depends on the intellectual abilities of the listener.

1.6.2. Comments on the Model

It is apparent that the model offers a richer view of meaning and of comprehension compared to earlier IP models. It divides meaning into '*basic'* and '*reflective'* and assigns a more rational and reflective role to the listener. In addition, the listener is able to make decisions concerning the extent to which deep or surface analyses are done. Despite all of this, one main reservation about the model is noticeable from its conceptualisation of listening as a sequential process where analysis of the input takes place in a fixed order. There are two different views in the literature about whether different knowledge sources are used hierarchically from the lower to the higher level, or they are all used simultaneously in an interactive way. The latter seems to be a more prevalent view that is supported by logical arguments from research in the listening literature (Marslen-Wilson W. D., 1978; Anderson & Lynch, 1988; Altmann, 1991; Lynch, 1998; Flowerdew & Miller, 2005; Nix, 2016). On the other hand, sequential/linear models have been widely criticised for their descriptions (Graham & Macaro, 2008). This criticism applies to the current model as well.

Goss (1982)discusses non-linearity in the stages as a logical consequence of having to deal with different parts of the utterance at the same time. The real-time nature of listening imposes that all processes, higher and lower, must be in operation, especially when long chunks of speech continuously flow into the listener's ear. By applying this to Goss's model, RP may be in operation at the first part of an utterance –that has already been segmented– at the very time when the last part is being received and segmented through SP. Until now, processing at all stages in taking place simultaneously. But the stages themselves imply a kind of sequence. According to the model, the higher level processes can only be triggered when processing at the lower stage is done and message parts are passed on to it. In other words, reflective processing does not precede signal processing, nor does it operate on the

same signal at the same time when the signal is being received and segmented. Furthermore, higher level sources cannot be called for, say, assisting or solving lower level segmentation problems. This is at least what sequential models, including the current one, suggest and what they have been criticised for (Graham & Macaro, 2008). It is now accepted that there is a degree of interaction between processing in the higher and lower stages where relevant information from both of them is summoned to solve difficulty.

If we take the assumption that the process follows successive stages, as the model suggests, what is it that the listener makes use of at the end of each stage, and that is normally ready to be analysed at the following one? There appears to be some vagueness concerning the outcome of analysis resulting from the first stage. Goss's model does not seem to provide details about the output of the SP stage that can be further analysed in the following stage, LP. It offers an explanation of what the listener does (e.g. making simple inference), but the questions of how this is done or what knowledge the listener uses to make such inferences are probably not sufficiently answered. What is clear, though, is that whatever the outcome is, it does carry meaning. The terms *Meaningful Units*, *Message Parts* and *Structures* are used to refer to the data that can be used to form basic meaning.

Some of the questions raised here are well answered in the model that will be discussed shortly. It represents a typical information processing model (Lynch & Mendelsohn, Listening, 2010) that has been greatly influential and which has been adopted in L2 research, in addition to being primarily designed to describe L1 listening. The model was built by the cognitive psychologist John Anderson in his book "Cognitive Psychology and its Implications" (Anderson J. R., 1985; 2010). It offers a valuable contribution to the understanding of the listening process, and this is due mainly to its empirical foundations. It

also embraces a number of elements that find their validation within research both in linguistics and in cognitive psychology.

1.6.3. Anderson's Three Phase Model

Although Anderson's model (1985; 2010) represents a typical view of IP, it differs in some perspectives from Goss's model in at least two main ways. The first one has to do with the division of the process. In Anderson's model, the creation of the utterance's basic meaning is done *before* that the higher level stage is initiated, whereas, in Goss's model, it is categorised *within* the higher level stage of message comprehension. This means that, for Anderson, segmenting speech and understanding the literal meaning of language structures are considered as aspects of processing at the lower level which can be by no means a detached form to the listener's phonological, syntactic and semantic knowledge; i.e. when engaging these knowledge sources, the listener is actually doing the task of building the basic meaning. This view may possibly bridge the gap between the acts of using linguistic knowledge for segmentation and the creation of literal meaning discussed earlier. In Anderson's view, the latter is regarded as the output result of –rather than being separated from– applying syntactic and semantic knowledge, and both of them take place in one single sub-process called *parsing*. The second one has to do with detailing how information sources interact. Anderson gives a well-detailed explanation of how the information sources, especially at the lower stage, interact in enriching the listener's understanding. Cases where there is conflict between these sources are also discussed in detail, especially in solving ambiguity. Furthermore, in describing what happens in the listener's mind, the arguments given are founded on empirical research in language use and in cognitive science. In L2 listening contexts, for example, there exists some validating evidence to the current model which was adopted in a number of studies (O'malley et al., 1989; Goh, 2000; Vandergrift &

Goh, 2012; Nix, 2016). Three main related stages make up Anderson's model: *perceptual processing*¸ *parsing* and *utilisation* (Anderson, 2010).

1.6.3.1. Perceptual Processing

The first stage of the model is concerned with the perception of speech sounds. These are recognised and separated into segments/units. The most important unit in segmenting sounds is the phoneme. It is the smallest unit of sounds that can make a difference in meaning between words. Listeners recognise the different phonemes based on their distinctive features. For example, the difference in meaning between the words 'park' /pɑːk/ and 'bark' / bɑːk / is the result of the difference in the first consonant sounds /p/ and /b/ which have distinctive features that make them, at least for native speakers, recognisable as separate segments. Features that phonemes share may include, to name only two, the place and manner of articulation. The consonants /p/ and b/ are both bilabial (place of articulation) plosives (manner of articulation). However, each segment of these has a *distinctive feature* that makes it unique in its phonetic description. The /b/ is *voiced* in that it is produced with the vibration of the vocal cords whereas /p/ is not. In English, voicing is a distinctive feature that is recognisable by its users. It should be noted, though, that phonetic features may be the base of discriminating sounds in some languages but not in others. For example, *aspiration* is not a distinctive feature in English. That is, it makes no difference in meaning whether to pronounce the word 'cat' as /k^{h}æt/ or /kæt/. However, aspiration does make a distinction between /p^{h}ə/ (meaning *fruit*) and /pə/ (meaning *moment*) in Hindi (Cook, 2008). An English learner of Hindi may not perceive this difference and, hence, would recognise the two pronunciations as referring to one single word. In addition, in the case of L2/FL listening, the absence of a given phoneme in the listener's native language makes it difficult to be noticed when it is encountered in foreign speech. In this case, "L2 speakers

tend to assimilate the L2 phonemes to their established L1 system" (Flowerdew & Miller, 2005, p. 31) and, hence, find it difficult to discriminate sounds. These are some reasons why basic knowledge of the language sound system is required to decipher continuous speech-sounds and recognise segments.

The way phonemes are produced in connected speech (henceforth CS) raises another obstacle in perception. According to Anderson, one problem in phoneme segmentation has to do with the way sounds in an utterance are articulated. In continuous speech, the organs in the vocal tract may prepare for the articulation of a given sound before the articulation of a previous one is finished. This means that the organs cannot reach the perfect position for pronouncing the sounds which, in this case, will not be pronounced in the same way they are pronounced in isolation (Richards, 1983). This will result in the distortion of the original sounds and, thus, entail "additional difficulties in segmenting phonemes […] it also means that the actual sound produced for one phoneme will be determined by the context of the other phonemes" (Anderson, 2010, p. 52).

Successful processing at the first stage leads to the creation of a phonetic representation of the sounds which is held in the *auditory sensory store* (also called *echoic memory*) before it is sent to working memory for parsing (Vandergrift, 2013). It is argued that speech sounds are held for not more than 10 seconds in echoic memory to be processed (Anderson, 2010). If it is not attended to, due to lack of attention for instance, this sensory information will be lost.

1.6.3.2. Parsing

During the parsing phase, the listener's task is to break down the phonetic representation so as to understand the literal meaning of the utterance. To this end, the

listener has to engage and coordinate a number of knowledge types to parse the utterance. According to Van Patten & Jegerski, "for those working form the linguistic tradition, parsing involves the rapid and automatic assignment of a licit grammatical structure to a sentence" (2010, p. 4).

Words have to be first segmented and recognised. The phonetic representation has to be linked to the listener's mental lexicon that is stored in the long term memory (henceforth LTM), in order to recognise the words encountered in the stream of speech. Difficulty in lexical segmentation (henceforth LS) may rise in defining word borders. In speech, there are no gaps that clearly define where words start and end. By relying on a number of cues like word onset cues, perceptual salience and phonotactic conventions, the correct word is selected from the number of possible candidates in the lexicon (Vandergrift & Goh, 2012).

When words are recognised, meaning is immediately retrieved from each one, and from the way they are combined in the utterance. An all important factor that guarantees successful parsing of utterances is the knowledge of the rules and the structures of the language. Instead of processing blocks of sentences or individual words, listeners use the syntactic and semantic cues of the language to break down utterances into manageable units called *constituent structures*, whose meanings are later chunked together (Tode, 2013). According to Anderson, the main syntactic cues available during parsing are word order and inflectional structures, with the latter being more dominant. This is not to underestimate the value of semantic cues which are proved to be used "immediately to guide syntactic decisions" (Anderson J. R., 2010, p. 375). Most of the time, they are integrated together to interpret the sentence and to depict aspects like the agent and the complement in the phrases. However, in case of ambiguity or inconsistency between the two, one source will be used instead of the other, but the whole process will be, consequently, slowed. In this sense,

Anderson defines one type of ambiguity called *transient ambiguity* which can only be resolved by the end of the sentence. When listeners encounter such ambiguity, they tend to adopt a given interpretation and continue processing the sentence. This interpretation is held in working memory and is subject to revision and reconsideration if it happens that the incoming text unveils contradictory evidence. The use of contextual cues, prior knowledge, and the listener's capacity to hold, manage and monitor interpretations are also important in solving ambiguity (Vandergrift, 2013). Imhof (2010) explains that:

> The resolution of this ambiguity will only be successful if more context information is available and if the listener has the capacity to store the choice of possible interpretations long enough in working memory to reconsider the validity of the initial interpretation (p. 106).

The result of processing at this stage is the creation of propositions. A proposition is a mental representation of basic meaning comprised of "at least one major argument and one or more predications about this argument" (Sato, 1988, p. 375). When basic meaning is constructed, listeners retain a mental representation of the utterance's meaning and the actual words are forgotten.

1.6.3.3. Utilisation

In the utilisation phase, listeners use the basic meaning of utterances and relate it to the types of knowledge stored in their LTM in order to create elaborate understanding of the messages. This is done through *spreading activation* to the concepts in the LTM that are stored in the form of schemas and background information. In the brain, information is represented by cognitive units (called nodes) that are related to each other through links

(Anderson J. R., 1983). Spreading activation happens when concepts from the text that come in the form of meaning representations will activate related nodes (concepts) in LTM. Altarriba & Graves note that "the interaction created by having activation spread across nodes results in overall access of meaning for the concepts involved" (Altarriba & Graves, 2013, p. 578). For example, the activation of the node for the concept 'school' will spread to the related concepts and allow for the retrieval of concepts such as teachers, learners, curriculum or classroom. Similarly, knowing about the topic of the talk will allow listeners to activate related concepts, and foster the process of making inferences and hypotheses about what the speech will be about. In this sense, Anderson describes the stage of utilisation as an inferential process characterised by the generation of two main types of inferences: bridging inferences and elaborate inferences. Inferencing in the first type has to do with making links between the current utterance and the previous ones to understand the relation between them. In the second type, the listener infers implied meaning that is not explicitly stated. This happens, as mentioned earlier, by relating what is understood so far to the listener's background knowledge. The difference between the two is that the first one is done at the text level and in an automatic way while the second requires a careful concentration from the listener on what he already knows. In addition, making elaborate inferences requires the presence of the relevant knowledge based on which inferences can be made. If such knowledge is not available, understanding will suffer (Carrell & Eisterhold, 1983; Tomitch, 1988). It should be noted that contextual cues also can play a key role in making elaborations during the utilisation phase.

1.6.4. Comments on the Model

The model has a major advantage in research on listening. Through the clear division of the listening process into well-defined stages, it makes it possible for researchers to

investigate the listener behaviours at each stage and identify the related problems (Graham & Macaro, 2008). In L2/FL listening, problems encountered at each stage can be diagnosed and categorised as well (Rost, 1990; Goh, 2000). Despite this, the model has some limitations. As Rost (1990) put it, "…while a stage model may appear to account for a range of causes for non-understanding, it does not adequately characterise everyday instances of language understanding," (p. 7) . This is especially that it ignores listener purposes in listening. In addition, Anderson views comprehension as a linear process with utilisation being the final step. In other words, the same criticism to Goss's model seems to apply to the current one (Lynch, 1998; Graham & Macaro, 2008). Finally, the model accounts for one-way listening only in which the listener has the role of receiving without being able to respond (Nix, 2016). Real life listening requires some kind of interaction between interlocutors. The listener can signal understanding or non-understanding, can negotiate meaning, can show appreciation, can ask for clarification and many other ways of responding that guarantee successful two-way interaction. This is one of the main issues addressed by Vandergrift and Goh (2012) in their integrated cognitive-model of L2 listening.

1.7. L2 Listening Processes

According to Vandergrift and Goh (2012), there have been very few theoretical models built for the purpose of synthesising the processing elements in non-native listening contexts. They suggest that a model which comprises the elements of real life L2 listening will allow for a better understanding of the cognitive processes involved in comprehending foreign speech. It will also allow for better classroom practices for improving learners' performance in listening, and addressing their processing problems. This is mainly because listening in a non-native language adds heavy demands upon the listener. In addition to the difficulties that L1 listeners may encounter with comprehension, L2 listeners are faced with

additional difficulties related to their limited linguistic knowledge and some affective factors which constrain their ability to understand target language talk. In this section, we will refer to a recent theoretical model designed specifically to address the major issues related to listening in L2 contexts. It should be noted that many of the processes that enable L1 listening do apply to describing L2 listening with additional demands related to text factors, listener factors and context factors. This is why, in our discussion of L2/FL listening models, more focus will be on the additional issues that characterise L2 listening.

1.7.1. A Cognitive Model of L2 Listening

A synthesis of the cognitive processes and the knowledge sources involved in the process of L2 listening was suggested by Vandergrift and Goh (2012) in their cognitive model of LC (Figure 4). The model is essentially based on Levelt's schematisation of L1 production and comprehension (1993), and is elaborated through the integration of Anderson's cognitive processes of perceptual processing, parsing and utilisation (2010). The cognitive model has the aim of discussing factors and processes specific to L2 listening and presenting them in a coherent framework. Levelt breaks down the listening process into three components: *the acoustic phonetic processor*, *the parser*, and *the conceptualiser*. This model is further extended by Vandergrift & Goh to include the three processes suggested by Anderson discussed earlier; perceptual processing is done in the acoustic phonetic processor, parsing in the parser, and utilisation in the conceptualiser. The conceptualiser is responsible for both the production and the comprehension of speech. Although the main objective of the Levelt's model is to account for speech production, it includes a combined description of both the production and comprehension of speech. The purpose of synthesising the processes of listening and speaking in one single schematic representation is to build a model that is able "to account for what happens as speakers listen to their own speech for

monitoring purposes" (Vandergrift & Goh, 2012, p. 40). For Vandergrift & Goh, the usefulness of the model lies in its inclusion of both the production and perception of speech in its description.

The current cognitive model of listening discusses key processes that are relevant to the description of what happens in the L2/FL listener's mind in the route to comprehend messages. They are synthesised and presented in a framework that shows how they operate together in a coherent way together with the relevant knowledge sources. In the coming sections, we will attempt to deal with each of these processes as a part of Vandergrift & Goh's model, and discuss them in light of a wider context of literature in L2 listening research. Specific to a description of L2 listening are four main cognitive processes that illustrate how language learners' performance could be constrained. Vandergrift and Goh suggest that, in addition to the three phases of perceptual processing, parsing and utilisation, a description of L2 listening must include three other processes: 1) Bottom/top processes, 2) controlled and automatic processing and, 3) metacognition. As we have already provided a detailed account of the lower and the higher levels of comprehension through Anderson's three phase model, we will only briefly discuss them again in relation to the issue of the listening modes, BU and TD.

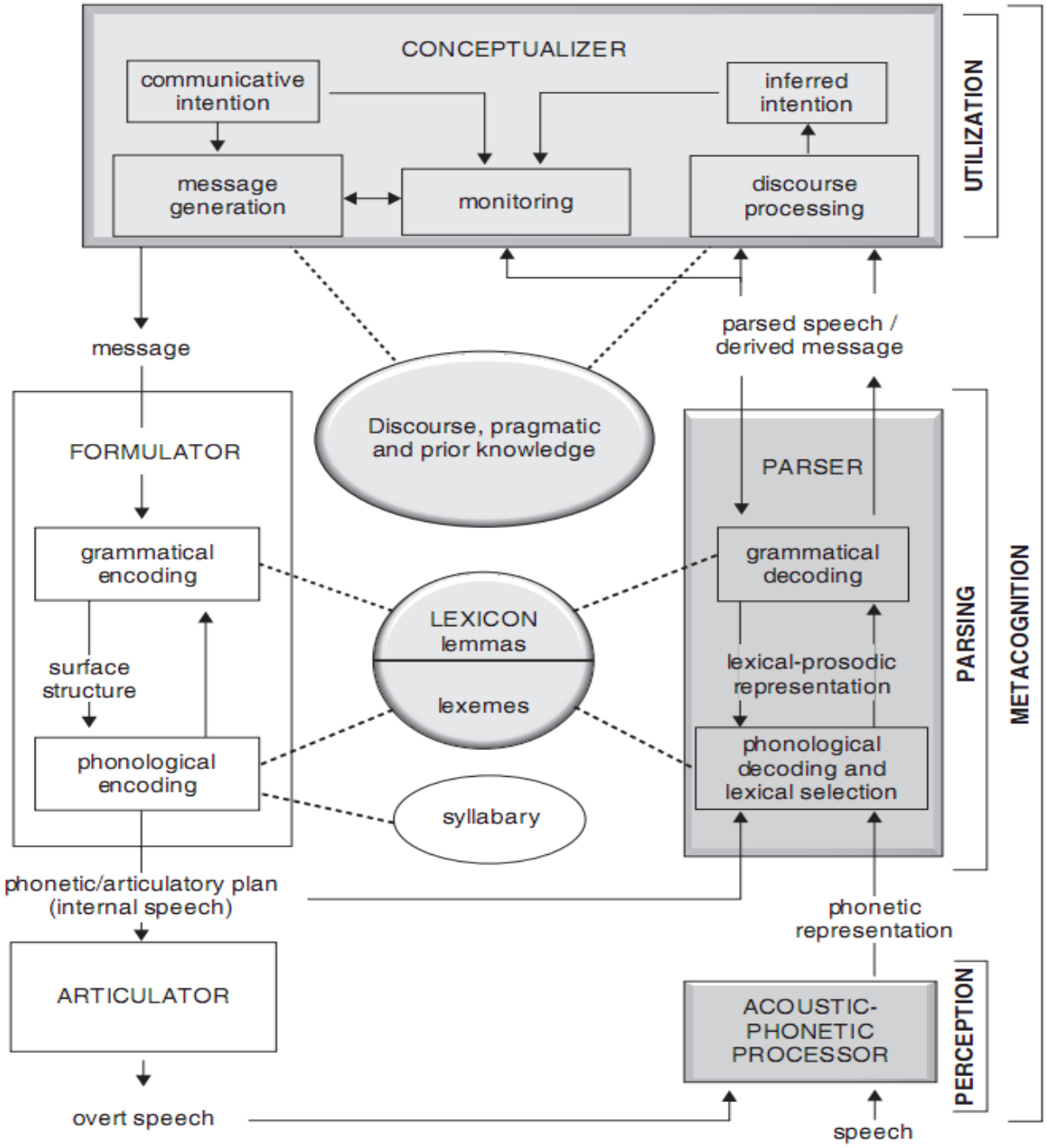

Figure 4: Schematic Representation of the Processing Components Involved in Speech Production and Comprehension (Vandergrift & Goh, 2012, p 39)

1.7.2. Interactive Two-way Listening

A description of listening in real life contexts should take into consideration situations where the listener interacts with a speaker. The main reason for adopting Levelt's model is that it makes it possible to integrate the processes of listening and speaking in a two-way interaction (Vandergrift & Goh, 2012). When we engage in a conversation, there is a need to respond to the speaker in order to express ideas and opinions, answer questions, react to the speakers' ideas or to keep the talk going. These kinds of response constitute a part of the listener's cognitive processes which need to be well coordinated with the receptive part. Vandergrift and Goh explain that "listeners must allocate their limited attentional resources to both comprehension and production in swift succession" (2012, p. 21) .

The implications of this interactive view of communication for non-native listeners are huge. Unlike in one-way situations where the listener has to make sense of the input only, in interactive listening he is obliged to actively participate in the exchange to create a shared understanding between him and the interlocutor. There will certainly be more demands on the L2/FL listeners to manage the interaction and comprehend speech (Farrell & Mallard, 2006). Not only do they have to understand messages, but they are also expected to speak and formulate answers. That latter, especially at the lower proficiency levels, may add more demands on the listener's cognitive ability (Vandergrift L. , 1999; Peterson, 2001). Furthermore, in the presence of such demands, the lack of sufficient knowledge about the target language combined with the limited capacity of working memory may lead to comprehension breakdown and a cognitive load.

Compared to one-way listening, interactive listening provides greater help for interlocutors in terms of cooperation. Listeners can rely on contextual clues (Vandergrift, 2007), while speakers can modify their utterances in a number of ways upon their

interlocutors' requests in order to guarantee a maximum of intelligibility (Richards, 2015). Nation & Macalister (2010) spoke about the benefits of interactive listening:

> An advantage of interactive listening is that the listener can negotiate the meaning of the input with the speaker. That is, they can ask the meaning of words or constructions and they can ask for a repetition of poorly heard material. They can also control the speed of the input by asking the speaker to speak more slowly (p. 91).

In addition to the clarification requests mentioned by Nation & Macalister, listeners can use a number of interactive strategies to negotiate meaning, especially to ask for clarification or to signal understanding/non-understanding (Vandergrift & Goh, 2012). Listening strategies in interactive settings are used to indicate "the ways in which a partner in a conversation may attempt to resolve a comprehension problem by seeking help from the speaker, i.e., to negotiate meaning" (Lynch, 1995, p. 166). They are used by native speakers (Farrell & Mallard, 2006), and they constitute a crucial element in successful native as well as non-native conversations (Pica, 1992; Vandergrift L. , 2007). Commenting on the non-use of back-channelling –sounds or expressions like *'*yes', *"*I see*",* 'mmm', "I got it" to signal understanding – Buck (2001) noted that speakers may feel that there is something wrong and stop speaking if the speaker does not use them. Similarly, Anderson and Lynch (1988) reported that native speakers responding to a call from non-native speakers hanged off quickly when the latter did not engage in negotiation. On the other hand, the non-native speakers who signalled their non-comprehension were given a chance.

Regardless of their proficiency level, foreign learners tend to use interactive strategies to ask for clarification even without specific training (Farrell & Mallard, 2006).

This is a logical result of the need to understand the language in which they are not fully proficient. A study by Rost & Ross (1991) revealed that the types of clarification queries used by L2 learners in interactive listening correlated with their proficiency level. The participants' use of clarifications differed depending on their overall proficiency. Three main types of queries were identified and ordered according to the development of proficiency:

1- *Global queries*: used to show the state of non-comprehension without specifying the problem. Problems that give rise to global queries are probably linked to sound segmentation, lexical or syntax problems.
2- *Local queries*: used to ask for clarifications of specific items at the local and discourse levels.
3- *Inference*: implies asking questions that show elaborate inferencing and anticipation of information based on prior knowledge.

Elaborate inferences are indicative of comprehension and are mainly used by advanced learners. Beginner learners, on the other hand, are more likely to have problems of LS. Their lack of sufficient syntactic and semantic knowledge explains why "they are seemingly forced to allot most of their attention to specific word meanings and parsing the input into basic constituent structure" (Rost & Ross, 1991, p. 262). This may be the reason why they tend to use more global and local queries for signalling misunderstanding. However, this is not without risks; by concentrating on local problems, they might deviate from their main objective in listening (Rost & Ross, 1991) and, eventually, fail to understand communicated meaning.

The ways L2/FL learners prefer to signal their understanding or non-understanding may be bound to cultural differences. In intercultural communication, using clarification

requests may lead to frustration between interlocutors. In a study by Cutrone (2005), it was found that Japanese (EFL) learners used certain strategies differently from native speakers. Although the strategies were shared between them, the way they used them and the frequency of their use lead to negative perceptions and stereotyping.

In another dimension of strategy use, the need to signal non-understanding may be hindered by social and affective factors. In this respect, Carrier (1999) argues that:

> […] one of the barriers to negotiating within conversations is unequal status in the relationship between interlocutors. Consequently, if the status relationship between interlocutors hampers negotiation for meaning within a conversation, and negotiation for meaning aids comprehension, then the implication is that the status relationship between interlocutors has the potential to help or hinder listening comprehension (p. 75).

Social and affective factors may include anxiety, motivation, social distance, fear of taking the risk/losing face, the fear of repeatedly signalling non-comprehension, and other factors. Unwillingness to use interactive strategies because of such factors will deprive L2 listeners an important source of information from interlocutors (Vandergrift & Goh, 2012). They would keep the talk going despite missing important points and pretend that they have understood. Consequently, comprehension breakdown is very likely to take place.

Bygate (1987) discussed FL users' refusal to ask for clarification and negotiate meaning as a reflection of a self-evaluation of their ability to understand foreign speech. Rather than discussing meaning with speakers, they tend to accuse themselves for misunderstanding. The lack of confidence they may experience in listening to foreign speech

may spark off assumptions that whatever they do not understand is not open for negotiation because it is a result of their low listening proficiency. The truth is that, even for native speakers, negotiation of meaning must be a part of conversations for the simple reason that it is difficult to produce messages, as foreign learners may believe, "in a complete and self-sufficient manner, which will be perfectly comprehensible to all hearers" (Bygate, 1987, p. 30).

1.7.3. Listening Modes

As detailed earlier, comprehension is initiated when speech is first received and segmented in the perceptual processing phase and, then, a phonetic representation is sent to the parser to be segmented and assigned literal meaning. The basic meaning created will be related to background knowledge, and textual as well as elaborate inferences will be done during the utilisation phase. What is significant about the Vandergrif & Goh's presentation of the processes is that they act in an interactive recursive manner, rather than in a series of successive stages. The coordination of the two main cognitive processes explain how listening can be looked at as an interactive process; these are BU processing and top-down (henceforth TD) processing. For an understanding of LC, the distinction between these two processes is fundamental (Vandergrift, 2013). This is why, in what follows, separate accounts of the processes and what they involve will be first discusses and, then, the way they may interact in listening will be examined in light of related literature.

1.7.3.1. Bottom-up Processing

The speech signal in a BU view of listening is the main element that triggers the comprehension process. Obviously, LC must be initiated by the processes of receiving sounds and recognising words in the input. Whatever subsequent higher level processes that

take place in the listener's mind, they are all dependent on successful reception and recognition. According to Richards (2015), the speech signal in a BU mode is the "basis for understanding the message" (p. 377). The data received from speech is subject to further analysis, categorisation and interpretation by the listener (Chaudron & Richards, 1986) until meaning is created. This is why a bottom up approach is referred to as *data-driven* processing (Carrelle & Eisterhold, 1983; Field, 1999).

A common element in the definitions of the BU process is the hierarchical and linear direction of processing. Listeners in a BU mode start by analysing the smallest units of speech and move up in a series of hierarchical stages through the different levels of linguistic description until they reach the text level (Nunan, 2002). According to Flowerdew & Miller (2005), a BU view of listening holds that:

> […] listeners build understanding by starting with the smallest units of the acoustic message: individual sounds or phonemes. These are then combined into words, which, in turn together make up phrases, clauses and sentences. Finally, individual sentences combine to create ideas and concepts and relationships between them. According to this model, therefore, the different types of knowledge necessary to the listening process are applied in a serial, hierarchical fashion (pp. 24,25).

In this respect, this view of listening is very close to the early IP model of listening (Flowerdew & Miller, 2005). Field (1999) further breaks the process to include even lower processes. Clues in the acoustic signal are first used to recognise phonemes. The listener's knowledge of the sound system allows for the detection of phonetic features that distinguish the repertoire of phonemes in a language from each other (Rost, 2006). When they are

recognised, phonemes are combined together to build up syllables which are, then, combined to make up words (Field, 1999). The meanings of words are retrieved from the lexicon stored in the LTM (Vandergrift L. , 2011), and further combinations are done at the phrase, clause and sentence levels. Hence, in addition to the listener's knowledge about the sound system which is deemed essential, fundamental in a BU processing is the listener's syntactic and lexical knowledge (Richards, 2015). According to Richards (1990), "lexical competence serves as a mental dictionary to which incoming words are referred for meaning assignment. Grammatical competence can be thought of as a set of strategies that are applied to the analysis of incoming data." (p. 51).

As listeners' attention is directed at receiving signals and changing them into meaningful representations, many scholars consider a BU approach to comprehension as a process of decoding (e.g., Field, 2008a, Nunan, 2002, Richards, 1990). Successful decoding of speech draws on the listener's linguistic knowledge and his ability to recognise the phonological, semantic, and syntactic cues carried in the input. Richards gave examples of some BU decoding processes with reference to such cues. These include: 1) searching for lexical items in the input, 2) breaking down the stream of speech into recognisable words, 3) identifying the focal points by means of phonological cues and, 4) parsing speech to understand propositions using grammatical cues.

While native speakers deploy a variety of linguistic /non-linguistic cues and sources of information, lower level learners are generally inclined to rely heavily on the acoustic signal, making their route to understanding BU (Long, 1990; Peterson, 2001; Brown G. , 1990). The examination of learners' verbal protocols proved that this difference was also found to exist between higher and lower proficiency students, with the latter group having more problems with word recognition (henceforth WR) (Goh, 2000). Of the processes

highlighted by Richards, the first three processes represent a real challenge for listeners, especially non-native ones. Recognising words in the utterance and defining their borders are two of the main difficulties that L2/FL learners face in listening. This is due mainly to the nature of speech in which words are linked together, and they come in a stream with no obvious spaces as in the written language (Levelt, 1993; Vandergrift L. , 2013).

In his discussion of BU processing, Rost (2006) notes that phonological competence is the key which determines how successfully the listener can cope with the problem of decoding. The significance of this competence rises, especially, in the detection of words in the input and defining their beginnings and endings. This is because the lack of reliable cues which allow for the detection of word boundaries in speech makes words rather obscure. L2/FL learners are very likely to face problems in recognising words in speech, even the ones they already know.

Rost maintains that in addition to the segmental features that allow for the identification of sounds, suprasegmental aspects of speech also play a significant role in WR and LS. While these two helpful "complementary phonological processes" (Rost, 2006, p. 58) provide cues to recognition and segmentation for native speakers, language learners' lack of knowledge about them adds difficulty to their listening process. This problem will be further discussed in the following sub-sections.

1.7.3.1.1. Sound discrimination

It is believed that human babies have an innate capacity to acquire the rules of whatever language they are exposed to (Onnis, 2017). As speech is the first language modality they are confronted with, the phonological system that allows for encoding speech signals is in the first place to be acquired. In speech processing research, it is argued that

babies possess *phonetic feature detectors* in their brain, which allow them to make sound contrasts in their mother language (Rost, 2011). As young as one year old, babies are able to distinguish between the sound contrasts of any language more accurately than do adult listeners (Cutler, 2000). This difference has a neurological explanation. The detectors are sensitive to the contrasts that are present in the speech to which babies are exposed, and some of these detectors will account for the native language sound contrasts. When the native-language sound system is efficiently mastered, as one reaches adulthood the brain regions of the detectors which are not used tend to atrophy, and only the ones which are deployed to make L1 auditory contrasts remain functional (Rost, 2011). This means that L1 listeners are able to discriminate between sounds which exist in their native language with ease, while the ones not accounted for by the functional detectors will pose problems. Consequently, non-native listeners may find confusable the sounds that are not contrasted in their native language. Cutler (2000) notes that:

> Adult listeners simply cannot discriminate nonnative contrasts as efficiently as native contrasts. The early specialisation which facilitated learning, and thus enabled the beginning of communication abilities in the native language, has produced disabled second-language learners in later life (p. 3).

Ur (1984) gave examples of sounds that are confusable to learners of different languages. She discovered that, through her experience in teaching English to non-native speakers, her students "did not perceive certain English sounds with any accuracy because these did not exist (at all, or as separate phonemes) in their own language" (1984, p. 11). French learners of English, for instance, find it difficult to recognise /θ/ as distinct from /s/ and /f/. Native speakers of Hebrew may not differentiate between the vowels /ɪ/ and /ɪː/ in English because

these are treated as allophones rather than distinct segments in their language. The same thing holds for the dark/clear /l/ in Russian; native speakers of English do not recognise them as different consonants. Speakers of Cantonese –a variety of the Chinese –cannot discriminate between /l/ and /r/ consonants in English (Brown G. , 1990).

There are notorious implications of a deficient ability to identify sound contrast on the process of WR, and on the listening process as a whole. As discussed earlier, the perception phase of listening is responsible for creating a phonetic representation after the detection of segmental features. Because these representations provide the "access code to the lexicon" (Levelt, 1993, p. 8), the extent to which they are reliable will greatly influence the accuracy of recognising words in the input. With the problem of sound discrimination mentioned above, the recognition processed will be logically constrained.

1.7.3.1.2. Word Recognition

A number of models have been suggested to account for the way spoken words are recognised, and retrieved from the lexicon during listening. One of these is the Cohort model suggested by Marslen-Wilson & Welsh (1978) and Marslen-Wilson & Tyler (1980). The model suggests three main stages in the process of WR: the activation stage, the selection stage and post-selection lexical stage. Activation refers to the process by which the first set of perceived sounds activates words in the lexicon which share the same phonological onset. For example, the recognition of /p/ as a first phoneme of the target word *penultimate* will activated all words (cohorts) which begin with this phoneme like *pen, pencil, put, penguin, park,* and *penury* in addition to the target word. These words are referred to as candidates because there is some kind of competition between them as more acoustic input is perceived. When the second phoneme /e/ is perceived, words whose second phonemes do not match the input like *park* and *put* are automatically eliminated. This way, new evidence form the

input continually decreases the number of candidates until the distinctive point is reached. A distinctive point refers to the point at which the target word is selected from among the activated words based on its uniqueness point that distinguishes it from all other candidates (Levelt, 1993). This stage is called *selection*. Both selection and activation are considered as pre-lexical stages. In the post-selection lexical stage, the syntactic and semantic information of the selected candidate is retrieved to allow for further parsing of the utterance (Marslen-Wilson & Welsh, 1978).

Although the model suggests some interaction between the information from the input (BU), and the semantic and syntactic information (TD), the former has some primacy. TD semantic and syntactic information is applicable only during the post-selection stage. Marslen-Wilson & Tyler explain that "the processing system first determines (on a BU basis) what *could* be there, as the proper foundation for then going on to determine what *is* there" (1980, p. 62, italics original). It should be noted that, despite that the other models of WR – e.g., the Trace Model (McClelland & Elman, 1986) and the Logogen Model (Morton, 1969) – differ from the cohort model in a way or in another, they all share the principles of the activation of candidates and the competition among them (Vitevitch & Luce, 2016).

The number of candidates that can be activated is logically related to the number of words present in the listeners' lexicon. Lower level learners' poor vocabulary will normally reduce the number of candidates activated and, in this case, the absence of the relevant word will lead to comprehension problems. This is why vocabulary knowledge has been identified as intrinsic in the process of listening. Stæhr (2009), for example, provided empirical evidence showing that successful listening to EFL significantly correlated with the participants' rich vocabulary. However, knowledge of vocabulary is impaired by the ability to recognise words in the speech input. Non-native listeners may fail to recognise even the

words they already know if they are unfamiliar with their spoken form. Field (1997, cited in Field, 2004) found that non-native listeners' failure to recognise words leads them to make approximate matches which are far from being accurate. Rather than admitting failure to recognise words, they tend to match the signal to words which have a similar shape.

There is evidence from psycholinguistics research which suggests that the number of candidates that non-native listeners activate is higher than that of native speakers. Consequently, the listening process may be constrained by having to deal with a wide range of candidates at a time. Weber and Cutler (2004) investigated the effects of poor target-language discrimination of phonemic contrasts on increasing the number of activated candidates. They found that, in activating the candidates of a word with a confusable phoneme in its onset, non-native listeners in the study mistakenly activated not only the words that match the target word phoneme, but also the ones with the confusing phoneme they do not discriminate in the target language. For instance, the phoneme at the beginning of the word *think* is confusable for native speakers of French. So, when hearing the first sound of this word, French learners of English may activate the words *think, thin, theme,* and also the words which start with the confusing sound /s/ as *sink, since, sin* and *sing*. Weber & Cutler postulated that, based on the onset, high frequency candidates in both French and English will be activated. It was found that there was a cross-language activation which, in addition to the activation through confusion, led to two main problems in the process of recognition. The first is that the number of activated candidates retrievable from both the L1 and the target language lexicons was increased. The second is that the speed of activation became slower as a result of this increase in the number of competitors for non-native speakers. On the other hand, native speakers in the study showed neither an increased number of candidates, nor a prolonged processing time. Thus, even if the non-native speakers' knowledge of vocabulary is poor, it seems that they activate more candidates due

to their poor ability of making sound contrasts in the target language. Weber & Cutler reported that "lexical competition is greater for non-native than for native listeners" (2004, p. 1), and explained that the reasons are related to the fact that "inaccurate phonetic processing allows spurious candidates from the native language on the one hand and spurious phonemic matches in the L2 on the other" (2004, p. 3).

Problems in LC are likely to be caused by the failure to recognise lexical items in speech. While language teachers tend to assume that word-level problems stem in poor vocabulary, the latter may or may not be the main reason. Field (2008a) suggested five other problems of recognising the words which may be among the non-native learner's lexicon. These include the following:

- the learner does not know the word;
- the learner knows the written form of the word but has not encountered the spoken form;
- the learner confused the word with a phonologically similar one;
- the learner knows the spoken form of the word but does not recognise it in CS generally or in this utterance in particular;
- the learner recognised the spoken form of the word but failed to match it to any meaning;
- the learner recognised the spoken form of the word but matched it to the wrong meaning (Field, 2008a, p. 87).

The discussion so far assumes that WR is made possible by relying on the first segments of a word. That is, word beginnings are assumed to be the key to recognition. Nevertheless, given the fact that speech does not include clear paces that indicate word boundaries as in the written language, the way listeners locate word beginnings and endings

is yet another problem in native as well as in non-native speech comprehension (Batie & Bradley, 1995).

1.7.3.1.3. Lexical Segmentation: Listener Strategies

WR is said to be the foundation of *lexical segmentation* –the identification of word boundaries in CS. This is because one of the goals of WR is that it allows for defining the beginning of the lexical item that immediately follows the recognised word (Rost, 2011). To this end, the listener should first locate the onset of the first word so as to determine its ending. This is not always easy when listening to CS. One of the reasons is that there are no obvious pauses between the words. In addition, there are no reliable cues in speech in which word borders undergo a number of modifications that make the process of LS a laborious task for both native and non-native listeners.

Cutler (2000) suggested that one of the ways listeners follow to segment speech is to use strategies that depend on the speech properties of their language. It is hypothesised that listeners often rely on the metrical characteristics of their languages to locate word beginnings. Cutler & Norris (1988) had subjects listen to a string of nonsense words and asked them to identify monosyllabic imbedded words –e.g., mint, melt, hint – occurring at either of two sequences: two strong syllables (henceforth SS) or a strong and weak syllable (henceforth SW). The results revealed that participants were significantly faster in detecting words in the SW sequences than in the SS ones. Cutler & Norris argued that the difference was imposed by the tendency to locate word boundaries at strong syllables and, hence, participants took a longer time in detecting words at the SS sequences because they had to spend time collecting speech material around the point of segmentation. However, for the SW sequences, no such intention to make the segmentation was attempted, and this made the detection of the word in the string faster. These results led them to argue for the existence

of a segmentation strategy based on rhythmic structure of words. In other words, strong syllables are likely to indicate the beginnings of words, and hence, can be used as cues in segmentation.

This conclusion is motivated by statistical data suggesting that the majority of words in English start with a stressed syllable. Cutler & Carter (1987) analysed three types of corpora to find out the likely dominant rhythmic structure in English words, and to see whether supporting evidence could be found to the LS strategy (suggested by Cutler & Norris, 1988) through word counts of initial strong-syllable lexical items. The first corpus included a collection of 98000 words based on the Shorter Oxford English Dictionary, the second contained 20000 words from the Merriam-Webster Pocket Dictionary (American English), and the third consisted of "34 samples, each consisting of between 5000 and 6000 words, of spontaneous conversation between educated adult native speakers of British English" (Cutler & Carter, 1987, p. 137). The corpora were analysed to find out the distribution of words with initial strong versus weak syllables in addition to their frequencies in real-life encounters. The last corpus served to provide information about both the number of word structures and their frequencies across word classes in actual speech. The results were striking. For the first two corpora, it was found that the words with an initial strong syllable were the most common; monosyllabic words, together with polysyllabic words with either a primary or a secondary stress in their onsets, made up 73% of the words in the first corpus. The remaining 27% of words are those with a weak initial syllable. Similar results were found in the second corpus. When the open class category of words was analysed alone, the number of polysyllabic words with an initial strong syllable was higher than that of words with an initial weak syllable. As for the third corpus –The London-Lund Corpus of English Conversations –it was found that 90% of lexical words (open class category) occurring in

the conversations are those with a strong initial syllable (including monosyllabic words, and polysyllabic words with either an initial primary or secondary stress).

In view of these results, Cutler & Carter postulated that a segmentation strategy would have a high level of efficiency because it accounts for most lexical items in spoken English. There are estimates of only less than 1% of words in natural English speech that may not be accounted for by the strategy (Cutler & McQueen, 1995). Cutler & Carter (1987) concluded that:

> [There is] an adequate basis for the successful implementation of a strategy of speech segmentation whereby strong syllables are assumed to be the onsets of lexical words. The strategy will result in few lexical word onsets being missed, and the false alarm rate will be relatively low (p. 141).

Analysis of slips of the ear mad by native speakers (Cutler & Butterfield, 1992) revealed that listeners mistakenly segment strings of utterances at the onset of strong syllables, and omit segments before weak syllables. This is the result of a metrical segmentation strategy that employs the statistical probability that most English words start with a strong syllable. As grammar words are usually weak monosyllables in CS, and content words more likely to have strong initial syllables, mishearings revealed that listeners erroneously omit boundaries before weak syllables (hence, often interpret them as grammar words), and mistakenly insert boundaries before strong syllables (often interpreted as beginnings of open-class words). Examples of these include: the deletion of word border in "she'll officially" and the insertion of a border in "into opposing camps" which were erroneously recognised as "Sheila Fishley" and "into a posing camp" respectively (Cutler & Butterfield, 1992). These results support the LS strategy hypothesised by Cutler & Norris.

What we are interested in, in this discussion, is the implication of segmentation strategies for non-native listeners. Cross linguistic segmentation studies revealed that speech segmentation is listener-specific. Listeners develop segmentation strategies specific to their native language at the early stages of life. On the other hand, when listening to a non-native language, they do not apply the segmentation strategies of the target language they listen to; instead, they rely involuntarily on the strategies they developed in segmenting their native language (Vandergrift, 2013). When the native and the target languages have different metrical structures, L1 strategies may not be effective to apply in the target language. These strategies are very likely to pose segmentation problems and, eventually, result in reduced comprehension. Lynch (2006) notes that "when trying to recognise L2 speech, learners use the characteristic patterning of their L1 as a mental template for identifying words […] when listening to an L2 we initially transfer the same metrical expectations [of our L1] as we attempt to segment speech" (p. 96).

Cutler (2000) reported the results of studies that dealt with languages having different metrical structures. French and Spanish, for example, have a syllable-based metrical structure while Japanese has a Mora based structure, and English has a stress based one. Consequently, native speakers of French and Japanese deploy different segmentation strategies which fit into the metrical structure their languages, and which are different from the strategies used by native speakers of English. Cutler, Mehler, Norris and Segui (1986) note that "the proper processing model for speakers of English (and languages like it) will differ from the proper model for speakers of French (and the languages like it); a syllabification strategy will be part of the latter but not the former" (p. 397). However, it was found that native speakers of English do not use a syllable-based segmentation strategy when they listen to French (Cutler et al., 1986). The same thing holds for native speakers of Dutch

listening to French (Cutler, 1997). Two important conclusions were reported by Cutler (2000):

> (1) In segmenting speech, speakers of different languages apply different heuristic procedures, efficiently exploiting the specific phonological structures of their various languages. (2) These procedures have become part of the listener's processing system, to an extent that they are also applied when listening to nonnative languages, even though this may lead to inefficiency (p. 1).

In addition to the use of metrical strategies, listeners rely also on phonological cues in order to segment speech. One of these cues is provided by knowledge of the rules that govern which sound combinations are possible within a word and which ones are not. The set of these rules is referred to as phonotactics. In English, /ŋ/ does not occur after a long vowel or a diphthong, nor does it combine with /tʃ/, /dʒ/, /ð/, and /z/, or occupy word initial position (Cruttenden, 2014). In standard Arabic, consonant clusters are not permissible in syllable onsets (Al-jasser, 2008). These phonotactic restrictions may possibly cue word boundaries; in a given language, a sequence of sounds that is not permissible in a given position may possibly indicate a word boundary. Hence, the illegal combination of /kf/ in the sequence of sounds /lʊkfəmɔː/ signals a boundary after /k/: "*look for more*?" McQueen considers such an illegal sequence a segmentation cue, and refers to it as a *phonotactic boundary* (1998, p. 21).

Here also, language specific phonotactic knowledge may help LS (Levelt, 1993; Rost, 2001). This was proved in a study by McQueen (1998) who argued that phonotactic constraints are among the information listeners rely on during segmentation. However, as different languages vary in their phonological rules, the lack of knowledge concerning

permissible or non-permissible combinations will be a disadvantage for non-native listeners. In addition, native-language phonotactic rules may be inappropriately used when listening to a non-native language. For example, the findings of a study by Weber (2000) revealed that native speakers of German were influenced by the phonotactics of their native language in spotting words in English –a language with which they were familiar. Weber & Cutler (2006) found that, while non-native listeners can lean to use the phonotactic rules of the target language, their native language rules may still influence their segmentation of the target language. Similar results were obtained by Al-jasser (2008) who found that EFL learners, with Arabic as their native language, inappropriately transferred the phonotactic rules of their native language when listening to English.

1.7.3.2. Top-down Processing

When listeners start processing from meaning and then consider what they hear in the input, processing is said to be TD; the listener starts from the higher levels of meaning and goes down to the smallest elements of the language (Richards, 2015). The main principle of a TD processing is that, in trying to make sense of a piece of discourse, listeners do not start the task of creating meaning from scratch, nor do they rely exclusively on the input. In other words, the speaker's utterances are not the only source of information that listeners rely on to make sense of utterances. There is generally some kind of shared knowledge stored in the LTM on which listeners have to build in order to arrive at an understanding of communicated meaning. This knowledge is generally referred to as background knowledge, world knowledge or prior knowledge (Chaudron & Richards, 1986; Buck, 2001; Lynch & Mendelsohn, 2010). In addition, the information that the context provides is also utilised by the listener to guide comprehension. The examination of potentially meaningful features in the context can enrich the understanding of the message (Omalley & Chamot, 1990).

Evidence for a TD mode in processing was found when subjects managed to recognise shortened words presented within context while they failed to do so when the words were presented in isolation (Flowerdew & Miller, 2005). This is one reason why inferencing is the key processes in a TD mode (Nation & Newton, 2009). Listeners rely on the knowledge they have about the different aspects of the listening situation such as the topic and the participants in order to predict what speakers will say, and how they will say it. They also make inferences and formulate hypotheses based on the information available (Lynch & Mendelsohn, 2010) even before that the speaker utters his first words. When the input is received, the role of the listener is to check the hypotheses made against the information available in the utterance. The task is one of continuously confirming, modifying, or rejecting the set of hypotheses formulated against the evidence provided by the actual words of the speaker (Nation & Newton, 2009; Lynch & Mendelsohn, 2010). Conceived of this way, TD processing requires that the listener approaches the text in a selective manner, attending more to those parts which are important to verify the hypotheses, rather than processing all linguistic elements in a serial manner as in the BU mode. Chaudron & Richards note that this mode of processing "enables the listener […] to bypass some aspects of BU processing" (1986, p. 113). To recognise the words, the listener does not follow the order of processing sounds, words, phrases, and so on. When the speaker's words are anticipated, the listener may not wait until that the speaker finishes pronouncing the word/words in the utterance in order to recognise it/them. Contextual as well as background knowledge have the potential of reducing the number of possible candidates which can be activated, so that only those which fit into the sentence and the topic of the talk will remain. Buck (2001, p. 3), gave the following example:

"She was so angry! She took up the gun, aimed and ……."

A listener using TD processing does not have to wait until the end of the utterance to know what the woman did, or how the speaker will describe her action. World knowledge and experience can tell us what angry people can do in these situations. So, the listener needs only to check his hypothesis which can be confirmed by receiving the first segment of the last word, be it /ʃ/ or /f/. If redundant information is added before the key word, like *"…aimed and, without thinking about the consequences…*", still the listener will only have to verify his inferences rather than curiously wait to know what happened. The redundant information will be of a lower importance and may not be fully attended to.

TD processing is focused on the interpretation of meaning. The listener infers this meaning by exploiting whatever information he possesses to fill in the missing parts. Richards provided a list of these inferencing processes explaining that they represent the key for using contextual and prior knowledge. These processes include:

- Extracting relevant information which is not stated explicitly.
- Reconstructing relevant information from both linguistic and non-linguistic clues.
- Arriving at an understanding on the basis of what follows on logically, and necessarily, from a given statement.
- Making connections between events, such as cause-effect, problem-solution and problem-explanation.
- Providing extra information that expands on the meaning of a text.
- Understanding what a speaker really means at the level of practical action. (Richards, 2015, p. 383)

One of the strengths of TD processing resides in the capacity to compensate for problems in the lower level of processing. If the perception of sounds or the recognition of

words breaks down because of factors such as noise, lack of attention or weak decoding skills, higher level sources can be utilised to compensate for this deficiency. This is the case of proficient listeners. High proficiency listeners may resort to compensatory strategies to fix lower level decoding problems (Wilson, 2003). Gaps resulting from an inability to decode the signal will be filled by relying on linguistic knowledge, knowledge about the topic, the text so far, the situation, discourse knowledge, experience, or any other source of information. Non-native listeners' lack of relevant knowledge, especially the knowledge about the linguistic system and the target culture, will limit the compensation process and probably make the signal the only source of information.

1.7.3.2.1. Schemata and World Knowledge

According to Rumelhart, the set of knowledge that an individual possesses in his memory, including knowledge about the world, is organised and structured in the form of units called schemata. Schemata refer to the "data structure for representing generic concepts stored in memory. There are schemata representing our knowledge about all concepts: those underlying objects, situations, events, sequences of events, actions and sequences of actions" (Rumelhart, 1980, p. 34). The retrieval of information from memory through the use of schematic knowledge and top down processing is characteristic of effective listeners (Aguilar, 2008). Listeners use their schematic knowledge to guide their interpretation of the text, to make expectations using stereotypical knowledge like knowledge about genre, listeners and speakers, place and time (Brown & Yule, 1983), to construct a framework for predicting and monitoring their understanding (Vandergrift, 2013) and "to aid in the construction of a mental model by providing a framework of expectations for a specific situation" (Tyler, 2001, p. 264). Each schema is composed of a network of interrelated constituents which are activated by the concept related to that schema (Rumelhart & Ortony,

1977). Rumelhart (1980) gave an example of the schema for the *buy* concept. When it is activated, a number of associated variables which are part of the listener's knowledge of the world, such as *purchaser*, *merchandise*, *seller* and *money*, will be automatically activated through the network of connections. To understand speech, listeners rely on their existing schema to make expectations and to infer any related constituents which are not explicitly stated by the speaker. Thus, successful comprehension of speech is believed to be the result of matching what is already known to the unfolding speech. Listeners are said to use the *principle of analogy* (usual qualities of things or behaviours) and the *principle of minimal change* (implies slight deviations from usual qualities) to successfully relate new and already known information (Brown & Yule, 1983, p. 63).

The confirmatory aspect of TD processing is evidenced in the process of schema activation. Schemata can be evoked by contextual cues in a TD mode (Ellis, 2003), by visual cues (Mueller, 1980) or by BU data through textual processing that results from understanding concepts in the propositions (Pan, 2017). In the latter case, it is referred to as BU schema (Carrell & Eisterhold, 1983). When further evidence from the input is received, the listener has to evaluate the activated-schema's goodness of fit against the unfolding text so as to confirm, modify or reject the expectations and the hypotheses made. As Pan explains, if a given schema is confirmed as an appropriate match, it triggers other schemata in a TD fashion:

> [activated schemata] activate the inactivated subschemata in the top-down fashion, in order to compare them with concepts in the successive sentences of the discourse and in an attempt to evaluate its goodness of fit. When one of these schemata has got further positive results about its goodness of fit, it would activate

> even higher level schemata, of which it is constituent, through bottom-up processing. This higher schema would then activate more of its constituent schemata and further their subschemata through top-down processing. These lower-level schemata are compared with other schemata that have been activated from the bottom-up and would eventually either match them or initiate a search for a predicted result (2017, p. 66).

There are two main types of schemata: Content Schemata and Formal Schemata. The first type is defined as the "background knowledge of the content area of the text" (Carrell, 1983, p. 83). Anderson & Lynch (1988) listed three main types of information as part of the listener's schematic world knowledge. These include 1) *general factual knowledge* which refers to general world knowledge shared by people regardless of their nationality or culture, 2) *local factual knowledge* which denotes special knowledge such as information about a specific person, place, or a given event, and 3) *socio-cultural knowledge* which requires the awareness of the cultural and social values of the target language and any related conventions shared by its speakers. To illustrate their use, Ellis (2003) gave the example of a newspaper headline whose interpretation depends on, in addition to the linguistic knowledge, the three types of content schemata. To understand a newspaper headline like "Saddam Slams Doors on Hopes for Peace", one needs to know who Saddam is (type 1), to know what made him "slam the doors" (type 2), and to know the sociocultural meaning of one slamming doors (type 3). In research, content schemata have been operationalized in a number of ways including *topic familiarity*, *religious knowledge*, *cultural knowledge,* or *vocabulary knowledge* (Schmidt-Rinehart, 1994, p. 179). The effects of these types of prior knowledge on both native and non-native comprehension have been studied, and the results indicated varying degrees of influence.

Schmidt-Rinehart (1994) investigated the effects of topic-familiarity on the comprehension of natural authentic speech by learners of Spanish as an FL. Three groups of Spanish learners with varying degrees of proficiency recalled more propositions from a familiar than from a non-familiar passage. A significant difference was found in the number of prepositions recalled in the familiar versus non-familiar topics at all levels of proficiency. However, the group with the highest proficiency level (Intermediate I) obtained the highest scores. Their linguistic knowledge could have played a role in facilitating comprehension. This is because, even in the non-familiar topic, they obtained significantly higher scores compared to the two other elementary groups.

Similar results were found by Long (1990) concerning the effects of linguistic knowledge on comprehension in the absence of an appropriate schema. Participants were able to boost understanding relying on available schemata, and to compensate for their lack of related schema relying on their linguistic knowledge. The results showed that the latter was not as significant as when the relevant schema was available. Despite this, schemata had a negative influence when they were overextended by non-native listeners. There was evidence that "schemata can also have dysfunctional effects on L2 listening comprehension" (Long, 1990, p. 72) when background knowledge of similar events was overused in understanding a non-familiar topic.

On the effects of topic knowledge on the listening process, Tyler (2001) found that when the topic was available, L2 listeners' working memory consumption was lower than when the topic was not. Knowledge about the topic seemed to have had a positive effect on the performance of non-native listeners. In another study by Anderson, Reynolds and Schaller (1977), it was shown that two different titles ("A Wrestler in a Tight Corner" versus "Prison Break") given to the same text affected its interpretation. Two groups of readers

approached the text differently based on the title given to each of them, which triggered different schemata. The interpretations of the message by the two groups varied depending on the title, with no attempt to consider the alternative option that the other group considered.

Students' recall protocols do indicate that schemata are used to make possible expectations, compensate for poor linguistic knowledge, and aid for comprehension. Markham & Latham (1987) investigated the effects of background religious knowledge on comprehending passages about prayer rituals in Islam and Christianity. Participants with relevant background knowledge, either Muslims or Christians, showed higher comprehension of the texts that described their religions' rituals. On the other hand, for the non-familiar religions' rituals, the two groups and another neutral group with no religious background showed recall of fewer main idea units. The results indicate that "religious background knowledge does significantly influence ESL students' listening comprehension" (Markham & Latham, 1987, p. 161). While familiarity allowed for more elaborations on the known rituals, non-familiarity was shown to bring more distortions. In addition, important results were obtained from the immediate post-test interviews with participants. Interviewees indicated that they relied on their prior knowledge about prayer rituals to comprehend the ideas in the passage, to compare them to background information and even to critically evaluate them. When the text exerted obstacles on understanding, schematic knowledge worked as compensatory as well. The following is what one of the interviewees –a Muslim from Saudi Arabia –reported about her process of listening in the study:

- Her use of schemata to comprehend: "I depend upon my information about my religion. I just hear this passage then I remember some ideas…I have the idea then I depend upon my information about my religion" (1987, p. 167)

- Her comments on the difference between the familiar/non-familiar rituals "… [Unlike in the passage about Christian rituals] the first passage I can remember and I have the idea and I have the information from the past…my information […] what he said and what he's going to said about the pray…I knew all about this." (1987, p. 168).
- Her use of known schemata to compensate for poor linguistic knowledge: "I have difficulties in the vocabulary in the first and the second passage, but in the first I can understand because I knew what's going on and what this, but in the second… no" (1987, p. 168).

The second type, *Formal Schemata*, refers to "background knowledge of the rhetorical structures of different types of texts […] differences in genre, differences in the structure of fables, simple stories, scientific texts, newspaper articles, poetry, etc." (Carrell, 1983, pp. 83,84). Anderson, Pichert and Shirey (1979) refer to such knowledge as *Textual Schemata* which they define as the "knowledge of the discourse level conventions of text" (p. 2). Rhetorical conventions are used by speakers and writers to provide a framework for organising ideas. These may include organisation frameworks for expository texts such as *comparison, problem/solution, causation, description* and *collection of descriptions* (Meyer & Freedle, 1984; Carrell, 1984a), as well as narrative texts such as temporal organisation of events in stories. Such formal conventions allow readers and listeners to understand and follow the development of ideas, events or arguments. Carrell (1984b), argued that "knowledge of these conventions aids listeners or readers in comprehending the text as well as in recalling it later" (p. 83).

It was shown that different types of text formal organisation have an effect on information recall with varying degrees; information recall as well as comprehension may

be higher when the text is structured with a more organised type (problems/solution, causation and collection of descriptions) than with loose one (comparison) (Meyer & Freedle, 1984). A study by Carrell (1984b) revealed that comprehension of simple stories by ESL readers may be affected by the formal organisation of episodes. When a story was presented with a simple schema that matched readers' schema for simple stories, the quantity of recall increased. However, when the same story was presented with an interleaved-events schema, both the quantity and temporal recall were strongly affected. This was the result of the new schema that the participants did not possess. Moreover, while native readers were able to sort out the interleaved order of events, non-native readers, who were obliged to pay closer attention to decoding the language, had a lower chance to adapt to the unknown schema (Carrell, 1984b). Carroll noted that the findings of the study apply to a great extent to the comprehension of spoken discourse. Thorndyke (1997) who studied the effects of story structure and content on subjects' recall and comprehension reported similar results; the findings indicated that "rated story comprehensibility and recall were correlated and were found to be a function of the amount of identifiable plot structure in the passages" (1997, p. 104).

The results from these studies indicate that background information structured in the form of schema may assist as well as hinder non-native LC. The lack of appropriate formal/content schemata, as well as the failure to activate available ones, interferes with comprehension (Carrell & Eisterhold, 1983). This issue will be dealt with in more detail in chapter two in relation to our discussion of which listening mode – TD or BU –should be given more importance in teaching LC to non-native listeners.

1.7.3.3. Interactive Listening

A more rational view of the listening process should consider all available information sources at both the higher and the lower levels, and demonstrate how these can be integrated for interpreting messages. Field (2008a) argued against the use of the terms *View* or *Model* in reference to TD and bottom up processes because, according to him, they imply two different theoretical perspectives of listening; one based on the input and the other on context. This way, discourse comprehension requires that the listener uses *either* one *or* the other model, but not both of them. However, this does not reflect the way listeners comprehend speech.

There is a strong argument against a mutually exclusive TD or BU processing view of listening. Meaning interpretation can only be achieved when TD and BU modes combine and operate simultaneously as complementary processes (Chaudron & Richards, 1986; Buck, 2001; Graham & Macaro, 2008; Vandergrift L. , 2013). Top level information and background knowledge will not be brought to bear unless there is some primary decoding of the main concepts and propositions presented in the input (Ellis, 2003; Lynch & Mendelsohn, 2010). On the other hand, processing language at word, phrase, utterance and discourse levels alone cannot account for what a specific speaker intended by their utterances in a certain situational context. Three types of problems may result from exclusively relying on BU processing: 1) word meaning cannot be located in the absence of context/co-text, 2) it will take longer to access words in the lexicon and reduce the number of candidates, and 3) as context is not used, the possible additional word meanings cannot be predicted (Richards, 2015). So, meaning creation is a function of the interaction between TD and BU processes which are mutually dependent on one another. The interaction between the listening modes throughout the three stages of comprehension is demonstrated by forward

and recursive lines in Vandergrift and Goh's (2012) schematic representation of L2 cognitive processes (figure5).

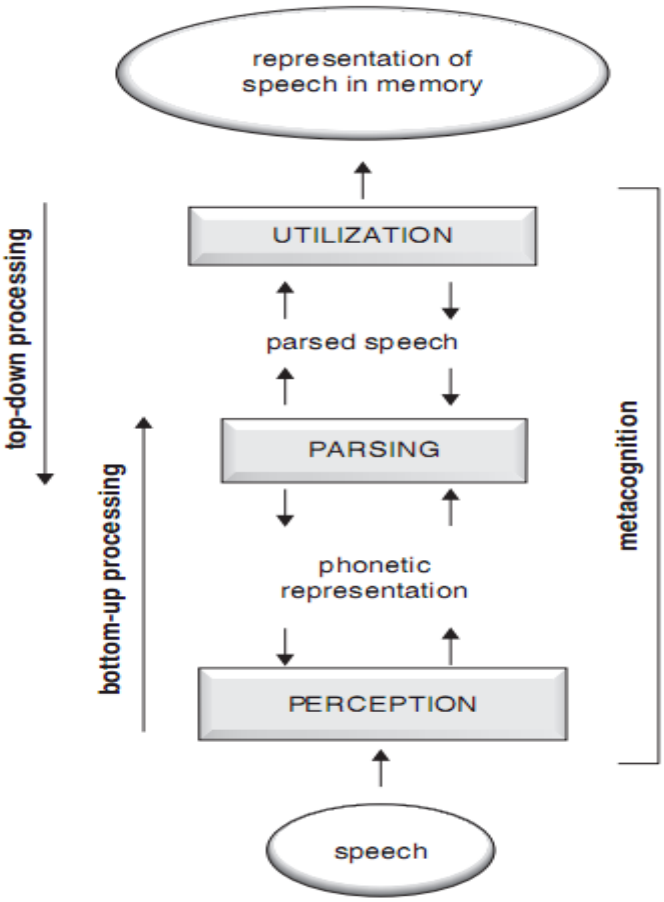

Figure 5: Cognitive Processes in L2 listening and their Interrelationships (Vandergrift & Goh, 2012, p. 17)

Another way of viewing the interaction between top and bottom level processes is through the *interactive compensatory mechanism* suggested by Stanovich (1980). The mechanism suggests that when problems at one level arise, either the linguistic or the non-linguistic, the information sources from the other level can compensate for any processing problems. For example, poor recognition of words due to a noisy environment forces the TD mode to take over to compensate for the deficient input. While the mechanism was basically designed to account for how readers compensate for problems like WR, it is argued that it is also applicable to listening (Graham & Macaro, 2008).

Although interaction is the norm, listeners may be inclined to rely on one mode more than the other under certain circumstances. The extent to which one mode or the other is used may depend on whether or not the listener is familiar with the topic and holds the necessary background knowledge, in addition to their purpose for listening (Vandergrift, 2013). Richards (1990, p. 54) gave the example of a novice cook and an experienced one listening to a description of a receipt. The set of knowledge the first one holds helps him to make expectations about the ingredients and the order of instructions. The task may be one of comparing a known recipe to the one presented, and spotting what is different or what is new. So, a TD processing would be more dominant. On the other hand, the novice, who lacks experience and knowledge, will have to listen in a BU way and pay much attention to every detail of what will be said.

Non-native listeners' language proficiency is another key determiner of the processing mode. Stanovich states that "according to the interactive-compensatory model, the poor reader who has deficient word analysis skills might possibly show a greater reliance on contextual factors" (Stanovich, 1980, p. 63). In this case, resorting to the TD mode is a strategic choice which may be the result of inadequate command of the language rules (Field 1990). It is logical, given the principle of compensation, that beginner language learners are more likely to use higher level sources. However, it has been shown that non-native listeners devote more effort to decoding the signal despite their weaknesses in doing so (Goh, 2000). A trade-off between those two trends is given by Flowerdew and Miller (1995) who suggest that individual listener styles play a key role in the choice between TD and BU modes. In addition, group needs may also constrain learners' processing. According to Flowerdew & Miller (2005):

> At the level of the group, beginners are likely to need to spend more time on developing basic bottom-up skills of decoding. For more advanced learners, however, who have mastered basic phonology and syntax, emphasis on the development of top-down skills of applying schematic knowledge may be more appropriate, although even advanced learners need to work on bottom-up features of fast speech […] (p. 27).

This is not only a pedagogical plan that reflects learners' needs, but also a description of developmental stages though which they pass. As discussed earlier, before the listener activates the appropriate background knowledge, he needs first an ability to automatically decode speech and understand the concepts responsible for that activation. This is why it seems logical that higher and lower level learners process speech differently, and that the latter group is likely to have difficulty to integrate and coordinate the two processes (Vandergrift & Goh, 2012).

1.7.4. Processing Speed

The speed at which the non-native listener can process the target language speech is another element that makes L1 and L2 listening processes different. Non-native listeners usually fail to cope up with the speech rate and, consequently, miss parts of the utterance. One reason is that their processing is *controlled* rather than *automatic.*

1.7.4.1. Automatic Processing

It is believed that part of the cognitive processing in either speech production or comprehension is done automatically. As Levelt noted, "like in all other skills, lower level processes are automatic. They do not use central resources; they come for free" (1993, p. 1).

This is mainly applicable to the BU processing. It is established that, in this processing mode, native speakers do not lend much attention to the basic phonetic and LS processes. This will reduce the cognitive load upon their working memory, resulting in a spare attentional capacity that will be reserved for processing at higher levels (Field, 2008a). In addition, they are able coordinate top and bottom processes to their benefit. When listening in a noisy environment, for example, it is very likely that parts of the speaker's utterances cannot be appropriately perceived. Despite this, L1 listeners are able to compensate for this failure; they can fill in the gaps by guessing the missing parts based on any knowledge source that seems relevant. This way of viewing listening represents the norm against which the study of L2 listening can be placed and compared. For Flowerdew & Miller (2005):

> The processes we use as L2 listeners may be technically somewhat similar to those of L1 situations, but barriers to comprehension and additional processes that L2 listeners must perform can make listening in a second language an arduous task (p. 27).

1.7.4.2. Controlled Processing

In the absence of automatic processing and compensatory strategies, the listener has to pay close attention to all elements at both the higher and the lower levels simultaneously. When comprehension requires special attention, it is said that the listener is using *controlled processing*. This is one of the additional L2 processes noted (above) by Flowerdew and Miller, and which refers to the "temporary sequence of nodes activated under control of, and through attention by, the subject. Because active attention by the subject is required, only one such sequence at a time may be controlled without interference" (Schneider & Shiffrin, 1977, p. 1).

Lower level L2/FL listeners tend to overuse controlled processing due to their limited linguistic competency and lack of experience with the new sound system. Their poor segmentation abilities may lead them to think that successful listening depends on the extent to which attention is paid to the input. By having their attention directed to textual aspects alone, comprehension cannot go beyond the edges of literal meaning (Ellis, 2003). In addition, too much attention and controlled processing may possibly run the risk of exceeding the limited capacity of working memory (Nagle & Sanders, 1983). Zwaan & Brown noted that, "in nonfluent L2 comprehension, word-and sentence-level processing will be more resource consuming than in LI." (Zwaan & Brown, 1996, p. 291). The word *resources* here refers mainly to the human working memory capacity. According to the *capacity constrained comprehension* theory, proposed by Just and Carpenter (1992), humans have a limited working-memory capacity to perform the functions of processing and storage. The number of elements that can be activated and kept in working memory is relatively limited; humans can hold up to seven items in their working memory, plus or minus two (Miller, 1994). This will cause two main constraints for L2/FL listeners who have to rely on controlled processing. The first one is that the different processes and the multiple knowledge sources needed for comprehension cannot be well accommodated and coordinated within the limited capacity of the working memory. The second has to do with the speed of delivery which is controlled by the speaker and over which the listener has no control. This implies that there is a very short time for processing different language elements because new input will always be there to displace an earlier one. To overcome these constraints, it is crucial that non-native listeners develop automaticity in BU processing so as to reduce the demands of real-life listening.

Conclusion

This chapter dealt with the nature of LC and how it has gained an important status. After years of marginalisation and wrong assumptions, it is now widely recognised that listening is an active language skill which forms the basis of language learning and acquisition. We have shed light on early and recent models of LC and showed how important it is to view them as complementary rather than mutually exclusive. Earlier models of listening considered speech comprehension as a process of transmitting ready-made messages that are directly retrievable by the listener. Criticism to these models focused on the peripheral role given to the listener, and their inability to account for the way listening takes place in real life. Recent models added interpretation and meaning creation as important dimensions in describing the process of comprehension. The current view holds that:

- Listeners rely on multiple sources of internal and external knowledge to arrive at an understanding of utterances.
- Comprehension is bound to the extent to which the listener manages to coordinate a number of higher and lower level cognitive sub-processes.

While these processes are more or less identical for native and non-native listeners, the latter do have additional processes and problems related mainly to their poor linguistic knowledge, lack of experience, and to their first language habits. We dealt with some problems non-native listeners have at both the higher and the lower level of processing, and focused on those obstacles which they face in decoding the target language, especially in sound/WR and LS. In the following chapter, we will examine some of the factors that contribute to these problems, and consider how they are pedagogically addressed.

Chapter Two

Implications of the Features of Connected Speech on Listening and Instruction

Chapter Two
Implications of the Features of Connected Speech on Listening and Instruction

Introduction

We discussed in chapter I some major difficulties non-native listeners have in listening, especially at the lower level stage. This chapter takes one step-further in examining learners' problems in decoding. It will focus on the problems of WR and LS created by the characteristics of spoken English, with special attention to those referred to as features of (CS). The chapter starts with an overview of the nature and definition of CS phenomena, and then proceeds to a discussion of the major principles that lie behind their occurrence in spoken English. Because of their implications for communication, both speaker and listener perspectives need to be discussed. In order to highlight the problems CS features are likely to present for language learners during listening, it is important to know the patterns of phonological modifications and the rules that govern their occurrence in naturally spoken English. To this end, the main features and their types will be analysed and illustrated. This will be followed by a brief literature review of the implications CS features have for listeners from a theoretical point of view. Some studies that explored the effects CS features have on the LC process of L2/FL learners will also be reviewed. The final section will be reserved for exploring the effects of CS instruction on learners' LC. Two major teaching trends –the strategy based instruction and the BU instruction –will be discussed and compared for their efficiency in tackling learners' listening problems in light of the related literature.

2.1. Nature of Connected Speech

The literature on CS features is rich of so many definitions and terms used to describe the different phonological modification that take place in speech. As it will be shortly explained, the disparity in the terms used may be misleading to some extent. Hence, before we proceed to any discussion of the problems these CS phenomena may pose for L2/FL learners, it is deemed necessary to understand the nature of CS and the different terms used to describe the different features. In addition, the contexts of occurrence of these features and the principles that lie behind their use in speech should also be highlighted.

2.1.1. Defining Connected Speech

In natural spontaneous speech, words usually undergo a set of phonological modifications that alter their pronunciations in comparison to the way they are pronounced in isolation. Pronouncing a sequence of words in an utterance at a normal rate will have an effect on the quality of sounds both within and across word boundaries (Brown G. , 1990). Sounds may be reduced in quality and/or quantity, replaced by other sounds, or even deleted as a result of the effects of neighbouring sounds. These modifications are usually referred to as *connected speech aspects/ features/ phenomena/ processes/ adjustments.* The term *Citation Form* is used to refer to the ideal pronunciation of a word when it occurs in isolation (Giegerich, 1992). It must be composed of one strong syllable and its vowel quality is not reduced (Ladefoged & Johnson, 2011). On the contrary, the term *Reduced Form* refers to the modified pronunciation of the word when it occurs in natural speech. Crystal (2008) defines CS as:

> A term used in linguistics to refer to spoken language when analysed as a continuous sequence, as in normal utterances and

> conversations. Its significance lies in the contrast implied with studies of linguistic units seen in isolation, such as an individual sound, word or phrase, which were the subject-matter of much traditional linguistic enquiry. It is now realised that important changes happen to these units when they are used in connected speech, as demonstrated by such processes as assimilation and elision, e.g. *and* becoming /n/ in such phrases as *boys and girls* (p. 101).

For a number of principles, speakers adjust the pronunciation of certain segments in adjacent words so that they fit the phonological environment. As Crystal's definition suggests, it is through the *analysis* of spoken language that the phenomena occurring in natural speech can be spotted. This is because, in production, speakers are usually unaware of the modifications happening to words and segments (Cruttenden, 2001). The definition above cites assimilation and elision as two types of these modifications, which Field (2003, p. 331) refers to as "*accommodatory phonological processes*". Other aspects include strong and weak forms (henceforth WFs) of function words, linking, elision, juncture, contracted forms, and high frequency phrases (e.g. waddaya, gonna, wanna). As it will be demonstrated in the following sections, these categories overlap to a great extent and the terms used to describe them also vary. In addition to the term *CS*, the aspects together have been referred to in many different ways.

Brown and Kondo-Brown's (2006) review of the different terms used to describe CS revealed a great disparity. One frequently cited term is S*andhi Variation* or *Sandhi* –a term from Sanskrit which means "putting together" (Henrichsen, 1984, p. 105). Brown and Kondo-Brown note that *Sandhi* and other terms such as *Reductions*, *Reduced Speech*,

Reduced Forms may not be precise; they refer to some particular aspects of CS and, thus, do not cover all the modification processes occurring in CS.

Some authors refer to CS when dealing with special categories of language use such as informal language (Hill & Beebe, 1980; Lass, 1984), fast and informal speech (Mashahiro & Luan, 2012), casual speech (Norris, 1995) or relaxed speech (Weinstein, 2001). Brown and Kondo-Brown (2006) note that this terminology is misleading for the reason that it limits the use of CS to particular types of language. They argue that "connected speech encompasses a variety of phenomena that are found in all registers of a language" (2006, p. 5). Concerning the link between CS and the speech rate, Rost and Wilson (2013) note that CS features can even be found at a speech rate of 180 words per minute, which they consider normal. Another categorisation characterises CS features as examples of speech which is *lazy*, *sloppy*, *careless*, *sub-standard* or *low-class* (Brown & Kondo-Brown, 2006). Usually held by teachers, this view rejects the idea that CS occurs in all styles. Celce-Murcia, Brinton and Godwin (1996) argue against labelling CS as lazy or sloppy, and note that the modification occurring due to adjustment to surroundings "by no means marks a speaker as inarticulate or nonstandard" (p. 159). For Roach (2009), CS phenomena are part of natural authentic language use –in contrast to mechanical speech – as they occur in all instances of spoken language. Pei (1966, p. 238, cited in Henrichsen, 1984, p. 105) maintained that "Sandhi will occur only at normal speed of speech, and will be distorted or obliterated by any slowing-up process". Ladefoged and Johnson (2011) noted that the characterisation of CS features such as WFs and assimilation as lazy are only the result of the artificial notions of those people who have particular beliefs about what constitutes good speech. They went further to suggest that CS processes are more efficient than carefully articulated speech because of their potential for reducing effort, while transmitting the same message.

As a consensus point, CS modifications are believed to occur in all language contexts, formal or informal, with varying degrees. Cruttenden (2001) classified the degree of variation that takes place in word boundaries between two extremes: the casual nature and the formal nature of speech. The more casual speech is the more modifications will be found both within and across words. On the contrary, the more formal speech is, the more likely words will be close to –but not totally resemble- their ideal citation forms. Similarly, Dalton and Seidlhofer (1994) differentiate between what they call *distorted* and *explicit* use of language as a function of the degree of familiarity with the addressee. Brown and Kondo-Brown (2006) concluded their review of the terminology by noting that "[…] the bottom line is that connected speech encompasses a variety of phenomena that are found in all registers of a language, but to varying degrees depending on the register and style involved" (p. 5).

CS is particularly challenging due to the effects of the placement of words together, and the unlimited number of possible word combinations in running speech. As Ladefoged and Johnson (2011) put it, "speech is not really composed of a series of distinct gestures, and anyway, we don't usually speak using isolated words" (p. 107). Despite that studies on sound features provide rules that allow for distinguishing the set phonemes the human speech apparatus can produce, these rules may not be reliable in describing the behaviours and shapes of phonemes occurring in different contexts (Poeppel & Idsardi, 2011). This is because CS phenomena are suprasegmental aspects of pronunciation; they deal with descriptions above the segmental and even word levels (Celce-Murcia et al., 1996). In addition, words in speech come in a continuous stream with no noticeable gaps between them. While there are phonological rules which guarantee within-word possible combinations, such rules do not account for cross boundaries combinations. Sound sequences which are phonotactically legal within words may meet in a number of ways across boundaries. So, just like we need phonotactic rules for describing how segments can

combine in words, we also need rules that bear on the transitions between sounds across borders, which may be *phonotactically illegal*. CS features are descriptions of how sounds accommodate to guarantee a smooth transition among words in speech.

While a mastery of CS features is important in L2/FL reception and production, their significance is higher in listening (Dalton & Seidlhofer, 1994) due to the problems they create in WR and LS. As suprasegmental features, they also have more serious implications compared to the problems created by the limited knowledge about segmental features (Nunan, 2015). In this study, we will use the term CS to refer to the different sound modification features that take place in natural running speech, with special attention to the most frequent ones occurring in spoken English. By the word natural, we specify speech that is produced at normal rates, which is not abnormally slow, and which is produced spontaneously by native language speakers.

2.1.2. Principles of Connected Speech

In this section, we will discuss two main reasons that lie behind the sound modifications that occur in natural speech. We will first discuss the characteristics of the rhythm of English and how it constrains the way words are spoken. Then we will move to the issue of speaker factors and the principle of economy/least effort.

2.1.2.1. Rhythm of English

Of the principles that cause word pronunciation to vary in speech is the property of the rhythmic pattern of spoken English. Rhythm refers to "the way that the alternation of stressed and unstressed syllables within an utterance gives a 'tune' to the utterance" (Nunan, 2015, p. 97). Depending on their rhythm, languages are broadly classified as having a syllable-timed rhythm or a stress-timed rhythm. In syllable-timed languages such as French

and Spanish, rhythm is related to the number of syllables in the utterance. The time spent in pronouncing an utterance in French is bound to the number of syllables it contains because all syllables are given the same value, and no contrast is made between them. On the other hand, in stress-timed languages such as English, it is not the number of syllables that depicts rhythm intervals. As syllables are either stressed or unstressed, the time needed to utter a sentence is almost equal to the number of stressed syllables only. Rhythm in English is a function of the regular time intervals between stressed syllables. Avery, Ehrlich and Jull (1992) demonstrated this regularity by showing how the insertion of a number of normally unstressed words into an utterance may not affect its pronunciation time. In the following examples from Avery et al. (1992, p. 74), stressed syllables occur at equal intervals regardless of the number of unstressed ones inserted among them:

B i r d s	e a t	w o r m s
The bird s	e a t	w o r m s
The bird s	e a t	The worms
The bird s	will eat	The worms
The bird s	will have eaten	The worms

At a regular rhythm, the pronunciation time of the first utterance will almost be akin to that of the other utterances. Only those words which communicate important information such as content words (*birds*, *eat* and *worms*) will need to be made prominent by stressing them. Stressed words and syllables are more prominent in that they are louder, longer and higher in pitch (Celce-Murcia et al., 1996). It follows logically that unstressed ones are less prominent. In order to maintain regularity in time intervals between stressed syllables, unstressed ones are deliberately shortened and squeezed together for the regular interval to be maintained.

According to Celce-Murcia et al. (1992, p. 74), the sound adjustments that words undergo in CS have the function of promoting the rhythm regularity. Syllables or words which are of a lower importance or which the speaker believes they can be recognised based on textual or contextual features will be compressed or squeezed together. Speakers spend a relatively shorter time in articulating unstressed syllables and words. For this to be achieved, the sounds that make up unstressed syllables may be reduced (the /ðiː/→ /ðə/), contracted (will have /wɪl hæv/→/wɪləv/), pronounced intervocalically (will eat /wɪl iːt/→/wɪliːt/), or deleted (eaten the worms /ɪːtənðəwɜːmz/→ /ɪːtə̪wɜːmz/). These are examples of CS features that take place in utterances spoken naturally with a regular rhythm.

Where stressed syllables occur within the sentence is a clear demonstration of the relation between CS features and the rhythm of English. The regularity of word stress that appears when words are articulated in isolation may alternate in speech in respect to rhythmic principles. Cruttenden (2001) notes that, as speakers tend to avoid pronouncing strong syllables successively, word stress may shift in speech so that the distribution of strong versus weak syllables accommodates to the regular rhythm of English. This way, syllables which are strong in the citation form may be weakened in speech (e.g., eight /'eɪt/, thirteen /'θɜːtiːn/ → /'eɪtθɜː'tiːn/) (Cruttenden, 2001).

Bansal and Harrison (1994) share a similar view; they consider that "accent at the level of the sentence is […] much freer than in the word" (pp. 82,83). This is because, they note, the accentuation of a syllable in CS depends on the rhythmic pattern of the sentence and the degree of importance the speaker gives to particular words. This is the case of function words and content words in speech. The difference in pronunciation between content and function words is a clear demonstration of this stress shift, and of the function of CS features in promoting regularity in rhythm (Bansal & Harrison, 1994; Norris, 1995;

Celce-Murcia et al., 1996; Nunan, 2015) Function words usually occur *between* content words to form the structural links that tie them. Their position in relation to content words and the fact that they carry relatively very little meaning makes their reduction a priority; this reduction guarantees that content words will be more prominent and that the rhythm will remain regular.

2.1.2.2. Principle of Economy

One of the reasons for the variation of words and sounds in speech is that speakers often aim at articulating with the least energy. According to Rost (2011), economising articulatory effort will "encourage brevity and phonological reduction" (p. 25). This is done by reducing the set of articulatory movements in the speech apparatus to the minimum (Pouplier, 2003). For Dalton & Seidlhofer (1994), this happens in reaction to the need for communicating in the most effective, rather than the clearest, way. As a result, speech is produced with a minimum of both time and energy (Celce-Murcia, Brinton, & Goodwin, 2010).

Given the relatively fast rate of delivery, it is usually difficult to pronounce all words as clearly and accurately as they are pronounced in isolation and, if otherwise done, speech would greatly resemble non-native speakers' language. Most often, the organs in the vocal tract cannot reach the ideal articulatory position under the conditions of fast speech. The time and the communicative pressures put upon the movement of the tongue will not allow for an accurate and easy pronunciation of sounds across word borders, as in the following utterances:

The last game /ðiː lɑːst geɪm/

They sound good /ðeɪ saʊnd gʊd/

I have to go /aɪ hæ**v** t**ʊ** gəʊ/

Law and order /lɔː **æ**nd ɔːdə/

If accurately articulated, the sounds in bold type will be in conflict with the principle of maximum ease as they cost more time and effort. To facilitate pronunciation, speakers tend to unconsciously make smooth transitions from one sound to another and from one word to another, which result in features of CS (Brown & Kondo-Brown, 2006; Skandrea & Burrleigh, 2011). This is why we can hardly find sounds at word borders which are not affected in a way or in another by neighbouring sounds (Kisno, 2012). Hence, the reduced forms (/ðəlɑːs'geɪm/, /ðeɪsaʊŋgʊd/, / /aɪhæftəgəʊ/ and /lɔːʳənɔːdə/) will be more appropriate than the citation forms (listed above).

A number of scholars (Wingfield, Alexander, & Cavigelli, 1994; Cintron-Valentine & Ellis, 2007; Rost, 2011) explain the lack of clarity in articulation by the *Principle of Least Effort* suggested by Zipf (1949, cited in Rost, 2011, p. 25), which illustrates how languages are optimised both synchronically and diachronically. The principle holds that the units which speakers usually select based on their frequency will be automatized by use and eventually reduced (Cintron-Valentine & Ellis, 2007). Reduction here means a limited articulatory clarity in CS which aims at insuring the speaker's communicative goal with the minimum cost possible (Wingfield et al., 1994). This applies well to the example of structure words and bound inflections; due to their frequency in speech, they are mostly shorter and weaker than content words (Cintron-Valentine & Ellis, 2007).

The principle of least effort in articulation appears to be in conflict with the needs of the other side of the communicative situation: the listener. Just like speakers, listeners too apply a principle of least effort to maximise recognition (Rost, 2011). The problem is that the less clearly words are articulated, the more effort will be made by the listener to work

out distorted word shapes, and to recognise reduced forms as realisations of some abstract and ideal citation forms. Lindblom (1990) distinguishes between *hyper speech* and *hypo speech* to refer to clearly articulated and under-articulated speech, respectively. As each one is biased to an extreme, neither of them seems to suite the speaker and the listener together. As a trade-off between the two, Lindblom (1990) asserts that "speakers are expected to vary their output along a continuum of **hyper-** and **hypospeech.** The theory suggests that the lack of invariance that speech signals commonly exhibit […] is a direct consequence of this adaptive organisation" (p. 403).

Volenec (2015) explains that speakers rely on adaptive variability through which they adapt the way they articulate depending on the situation; they alternate between hyper and hypo speech depending on whether or not the situation requires a *phonetic precision*. So, articulation is closer to hypo speech in casual language, while it is closer to hyper speech in situations where clarity is sought. Another variable for resorting to the hypo side resides in the speaker's awareness of the information sources that the listener possesses. So, least effort will be spent whenever the speaker is aware of the availability of any relevant types of information that can be retrieved from context to compensate for lack of clarity (Krug, 1998).

2.2. Aspects of Connected Speech

The phonological modifications that sounds undergo in speech are manifested in different forms depending on the phonological environment where the sounds in question occur. This results in different types which are categorised as the features of CS. Common features include: Assimilation, elision, liaison and WFs.

2.2.1. Assimilation

Assimilation refers to the process by which the citation form of a word changes in continuous speech as a result of the sound changes that take place when sounds co-occur either within or across word borders. The pronunciation of sounds in certain environments changes so that they either share the parametric features of neighbouring sounds which are not part of their distinctive features, or become identical to them. These are referred to as "*Partial Assimilation"* and *"Complete Assimilation"* respectively (Mott, 2011). In "that King" there is a complete change in the place of articulation parameter for /t/ which maintains its plosive manner, but its place changes to share that of the adjacent sound /k/ (velar); the result is an identical sound /ðæk 'kɪŋ/. An example of partial assimilation is found in "that boy" where the /t/ in the first word remains voiceless but changes its place of articulation /ðæp 'bɔɪ/. As in all phonological modifications in CS, assimilation is an example of *obligatory phonological processes*; it is a rule that all native speakers apply and which is motivated by the principle of ease of articulation (Rowe & Levine, 2015).

In all instances of assimilation there is a *conditioning sound* whose articulation causes the features of a neighbouring sound to alter, and an *assimilating sound* which takes the characteristics of the conditioning sound (Celce-Murcia et al., 1996). In the example of '*that boy*', /t/ which is changed to /p/ is assimilating while /b/ is conditioning. It appears that this type of assimilation involves a phonemic change since /t/ is replaced with a different segment. This is, at least, how assimilation processes were traditionally regarded. There has been much debate about the extent of the changes that take place in assimilatory processes (Darcy, Ramus, Christophe, Kinzler, & Dupoux, 2009), and strong arguments were given against the view of phonemic change. We will cite three of these.

Recent definitions of assimilations hold that the outcome of the process is not a new sound, but rather an adjusted one with properties similar to those of an adjacent sound (Celce-Murcia et al., 1996; Roach, 2009; Ladefoged & Johnson, 2010). Roach (2001) regarded as naïve the view which considers assimilation as a replacement of one phoneme with another, and suggested that the generalisation that assimilation is a process which would result in phonemic change does not apply to all instances. Unlike in "good girl", where /d/ becomes more like /g/, examples of lip rounding or spreading in the pronunciation of a given consonant before /iː/ or /uː/ illustrate how assimilation processes do not always involve phonemic change. This is also evident in examples of nasalisation where the velum is kept lowered throughout the articulation of all sounds of a word as an effect of a nasal sound (e.g. /mɔːnɪŋ/) (Roach, 2001, p. 56). Even vowels may undergo assimilation to become nasalised but without a remarkable phonemic change (e.g. /kæ̃nt/) (Dirven & Vespoor, 2004, p. 120).

Even those examples which seem to show a sound replacement are questionable. It is suggested that assimilated sounds do not exactly resemble target sounds. For the example of the alveolar /t/ which is modified before velars /k, g/ (as in "that game" /ðæk 'geɪm/), Brown (1990) notes that the closure of /k/ does not occur in exactly the same place as /k/ when articulated in 'key' /kiː/. The same thing was noted in the articulation of assimilated /ʃ/ in 'miss you' /mɪʃ jʊ/ and /ʒ/ in and 'was young' /wəʒ jʌŋ/ which may not always be the same as /ʃ/ and /ʒ/ in 'sheep' and 'massage' (Lodge, 2009). Lodge (2009) cited evidence from research which demonstrates these differences and raised concerns about the use of an accurate transcription that reflects them. For Brown (1990), using phonetic symbols in transcribing assimilated sounds is misleading.

It is also problematic to assume that all possible instances of assimilation make pronunciation easy. Some differences in assimilatory rules exhibited among languages

suggest that ease of articulation is language specific. Some rules may be found in one language but are not followed under similar conditions in another. For instance, the rule which is that, in French, a voiceless velar /k/ becomes voiced /g/ before a labiodental fricative (e.g. “avec vous”) is not applied in similar examples in English (e.g. “weak voice”) (Roach, 2001). Stranzy (2005) shares a similar view noting that it is not always easy to say whether or not assimilation is due to speech production constraints. He adds that “assimilation often appears to be motivated by ease of articulation, but what seems easy and natural in one language often turns out to be less so in another” (Stranzy, 2005, p. 95).

One alternative to the phonemic-change view is to consider the alternations in sounds as examples of allophonic modifications (Roach, 2009). That is, the /d/ in “good boy” is realised by an allophonic variant /b/ that fits this context. Another view is one which takes into consideration the articulatory gestures and how they interact during the production of words in CS. Much research has been carried out in the area of *coarticulation* which focuses on how the organs in the vocal tract move in the realisation of, among others, assimilatory sounds (Roach, 2001). In this area, we look at the reasons that make phonemes affected in some contexts by observing the combination, instantaneous occurrence, overlap, or sequence of articulatory gestures. Lorenz (2013) contends that coarticulation results in assimilation. Similarly, Yule (1996) considers assimilation as one effect of coarticulation which he defines as “the process of making one sound almost at the same time as the next” (p. 58,59). So, the timing of how the speech organs move is significant in describing how phonemes co-occur and assimilate, especially under the constraints of fast delivery and ease of articulation. Because there is generally no sufficient time for the organs to assume an ideal position in fast speech, they usually tend to accommodate (Field, 2003). Roach (2009) explains that

> […] modern instrumental studies in the broader field of **coarticulation** show that when assimilation happens one can often see some sort of combination of articulatory gestures. In 'good girl', for example, it is not a simple matter of the first word ending *either* in d *or* g, but rather a matter of the extent to which alveolar and/or velar closures are achieved. There may be an alveolar closure immediately preceding and overlapping with a velar closure; there may be simultaneous alveolar and velar closure, or a velar closure followed by slight contact but not closure in the alveolar region. There are many other possibilities (p. 113).

It should be noted that, while both assimilation and coarticulation involve sound modifications, there is a clear difference between them; whereas assimilation is language specific as noted earlier, coarticulation is more independent in that it generally "results from the inertia of the speech mechanism" (Strazny, 2005, p. 95). This seems to be the reason why some assimilatory processes which can be explained by coarticulation may be found in some languages but not in others.

There are two main ways for classifying assimilatory processes. The first has to do with the direction of assimilation. When the conditioning sound comes before the assimilating sound, assimilation is said to be left to right or *progressive* (Cruttenden, 2001) or *lagging* (Collins & Mees, 2013). In terms of the articulatory movement of the speech organs, progressive assimilation happens when the organ maintains its positions of the first sound for the pronunciation of the following one, while it should normally move (Lorenz, 2013). As it is a case of preserving a given articulation, it is also called *perseverative*

assimilation (Lodge, 2009; Lorenz, 2013). In terms of frequency, this type of assimilation is rare in English (Lorenz, 2013). It is usually found in examples of inside word assimilation:

organ /ɔgəŋ/

taken /teɪkŋ/

urban /ɜːbm/

open – /əʊpm/

Note that in these instances, in addition to the change in place of articulation for /n/, the conditioning sounds are also affected. The release of the plosives /g, k, b, p/ is nasal rather than oral, and is almost done simultaneously with the articulation of /ŋ/ and /m/. Another case of progressive assimilation in English is the pronunciation of the inflectional 'ed' in verbs, and the 's' as a morpheme in plural nouns or in the third person singular in the simple present tense.

When the conditioning sound comes after and affects a preceding sound, assimilation is said to be *regressive* (Celce-Murcia et al., 1996) or *leading* (Collins & Mees, 2013). It is also called *right-to-left* (Lodge, 2009) or *anticipatory* assimilation because features of the conditioning sound are anticipated and the organs move into the position of articulating the conditioning sound at the same time of, the or immediately after, the articulation of the assimilating one. Lorenz (2013) explained what happens to the organs in producing this type compared to progressive assimilation:

> While for progressive assimilation the speech organs remain in place too long, for regressive assimilation, their movement comes too early, so that a feature which should only occur later in the

> word can already be heard in an earlier sound. Because of this, regressive assimilation is sometimes also called anticipatory assimilation (p. 86).

Another way of classifying the types of assimilation is according to the particular parametric feature that undergoes the modification. There are three main types: place assimilation, manner assimilation, and voice assimilation. We will deal with each type individually and consider the common rules in each one.

2.2.1.1. Place Assimilation

In place assimilation, the place of articulation of a sound changes depending on that of the conditioning sound or, as Rowe and Levine (2015) explain it, "adjacent sounds are made to agree in their place of articulation" (p. 76). The most common examples of this type in English include the *de-alveolarization* of the alveolars /t, d, n/ when they occur before bilabials /p, b, m/ or velars /k, ɡ/ (Cruttenden, 2001, p. 285). Here are some rules as discussed in Cruttenden (2001), Roach (2009), Celce-Murcia et al. (1996), Brown (1990), and Collins & Mees (2013):

- When a first word ends in /t/ or /d/ and the following starts with a bilabial /p, b, m/, regressive assimilation occurs whereby the /t/ or /d/ maintain their voiceless/voiced features, while they share the bilabial place of articulation:

/t/	before	/p/	→	/p/	white paper	/waɪp 'peɪpə/
/t/	before	/b/	→	/p/	that boy	/ðæp 'bɔɪ/
/t/	before	/m/	→	/p/	court man	/kɔːp mæn/
/d/	before	/p/	→	/b/	hundred pounds	/hʌndrəb 'paʊndz/
/d/	before	/b/	→	/b/	told Brown	/təʊlb 'braʊn/
/d/	before	/m/	→	/b/	vanguard movement	/væŋgɑːb muːvmənt/

- For /n/, while it remains a voiced nasal, its place of articulation becomes bilabial before /p, b, m/:

/n/	before	/p/	→	/m/	action planning	/ækʃm plænɪŋ/
/n/	before	/b/	→	/m/	fan battery	/fæm bætrɪ /
/n/	before	/m/	→	/m/	common mistake	/kɒməm 'mɪsteɪk/

- When they occur before velars /k, ɡ/, /t/ and /d/ change their place of articulation to become velars:

/t/	before	/k/	→	/k/	bright camera	/bɾaɪk kæmərə/
/t/	before	/ɡ/	→	/k/	fat girl	/fæk ɡɜːl/
/d/	before	/k/	→	/ɡ/	wide canal	/waɪɡ kə'næl/
/d/	before	/ɡ/	→	/ɡ/	road game	/rəʊɡ 'geɪm/

- If /n/ occurs in word final position before an initial velar /k, ɡ/, its place of articulation changes to a velar nasal:

/n/	before	/k/	→	/ŋ/	tone control	/təʊŋ kəntrəʊl/
/n/	before	/ɡ/	→	/ŋ/	ten girls	/teŋ ɡɜːlz /

- When a word ends in /t/, /d/ or /n/ and the following one starts with a dental /ð, θ /, the alveolars become dental:

/t/	before	/θ/	→	/t̪/	cut through	/kʌt̪ θruː/
/t/	before	/ð/	→	/t̪/	treat them	/triːt̪ ðem /
/d/	before	/θ/	→	/d̪/	bad thing	/bæd̪ θɪŋ/
/d/	before	/ð/	→	/d̪/	read this	/riːd̪ θɪs/
/n/	before	/θ/	→	/n̪/	ten thousand	/ten̪ θaʊzənd/
/n/	before	/ð/	→	/n̪/	train themselves	/treɪn̪ ðəmselvz/

Although they are alveolar sounds, the fricatives /s/ and /z/ undergo different modifications; before an initial /ʃ/, they move to a post-alveolar position:

/s/	before	/ʃ/	→	/ʃ/	lens shape	/lenʃ ʃeɪp/ or omitted /len'ʃeɪp/
/z/	before	/ʃ/	→	/ʒ/	cheese shop	/tʃiːʒ ʃɒp/ or omitted /tʃiːʃɒp/

These last examples are usually referred to as *palatalization* (e.g. Avery & Ehrlich, 1992; Dalton & Seidlhofer, 1994) despite that "the alternation is between alveolars and palatoalveolars" (Lodge, 2009, p. 147). In the same way, /t/, /d/, /s/ and /z/ may be modified when they precede /j/.

/t/	before	/j/	→	/tʃ/	can't you	/kæntʃ jʊ / /kæntʃə/
/d/	before	/j/	→	/dʒ/	behind you/did you	/bɪhaɪndʒuː/ / /dɪdʒə/
/s/	before	/j/	→	/ʃ/	miss you/ this year	/mɪʃ jʊ/
/z/	before	/j/	→	/dʒ/	was young	/wəʒ jʌŋ/

It should be noted that the outcome of assimilation in these examples appears to be a new sound which not only has a different place of articulation, but also has a different manner of articulation form both the assimilating and conditioning sounds. In the second example, /dʒ/ is a palato-alveolar affricate which shares the manner of articulation of both the alveolar /d/ and the palatal/j/. This is why palatalization is also referred to as *coalescent assimilation*.

In his analysis of authentic speech by news broadcasters, Brown (1990) listed other examples of assimilation not usually discussed. For example, /m/ may be realised as a labiodental nasal /ɱ/ when it precedes a labiodental consonants /f, v/. In another context, /m/ becomes /ŋ/ in in particular examples before velars (e.g. "I'm grateful", "I'm conscious") (Brown, 1990).

2.2.1.2. Manner Assimilation

Compared to the previous type, manner assimilation is less frequent in English. It involves a modification in which the manner of articulation of a given sound changes to conform to that of a conditioning sound which is usually easier to pronounce. In a regressive direction, the fricative feature of /s/ in "that summer" comes early and affects the articulation of /t/ whose manner of articulation requires relatively more energy; i.e. the airflow is not blocked, and it comes in an easier manner in fricatives than in plosives. So, /t/ will be assimilated as: /ðæs 'sʌmə/. Manner assimilation is especially frequent in fast speech where speakers usually tend to make a sequence of sounds easier to articulate. According to Roach (2001), "examples of this type are not easy to find, they involve a change from a 'stronger' consonant […] to a 'weaker' one, and are typical of rapid speech" (p. 55).

The most common types of manner assimilation include:

- The assimilation of the alveolar plosives /t/ and /d/ regressively when they precede alveolar fricatives /s/ and /z/:

/t/	before	/s/	→	/s/	That side	/ðæs 'saɪd /
/t/	before	/z/	→	/z/	Tight zip	/taɪz zɪp/
/d/	before	/s/	→	/s/	Sad song	/sæs sɒŋ/
/d/	before	/z/	→	/z/	Wide zone	/waɪz zu:/

- Before /n/ which is nasal, the plosives /t/ and /d/ become nasal:

/t/	before	/n/	→	/n/	that night	/ðæn 'naɪt/
/d/	before	/n/	→	/n/	good night	/gʊn 'naɪt/

- When /ð/ comes after /n/, the plosives /t/ and /d/, or the lateral /l/, it becomes nasal, dental plosive, or lateral respectively:

/ð/	after	/n/	→	/n̪/	in the army	/ɪn 'n̪ɪjɑːmi/
/ð/	after	/t/	→	/t̪/	get them	/ɡe t̪ t̪ əm /
/ð/	after	/d/	→	/d̪/	read these	/riː d̪ d̪iːz /
/ð/	after	/l/	→	/l/	fail the test	/feɪl lə'test/

Manner assimilation also occurs in word internal positions. This happens when /t/ and /d/ precede one of the syllabic consonants /n/ or /l/ as in the words *kitten, pardon, shuttle, paddle* and *fluently*. Before these sounds, the closure is done in the alveolar ridge, but it is maintained and not released at this position; in the case of /n/ the air escapes through the nose cavity making a *nasal plosion* as the velum is lowered; in the case of /l/, it escapes through both sides of the tongue making a *lateral plosion* (Mott, 2011, p. 150).

2.2.1.3. Voice Assimilation

As the name suggests, voice assimilation happens when the voicing parameter of a sound changes as an effect of an adjacent one so that they become similar; it becomes voiced if the conditioning sound is voiced, or the opposite. Common examples of voice assimilation occur in a regressive manner, especially at word borders, and they usually involve cases of *devoicing* where a voiced sound loses its voicing when it precedes a voiceless one (Mott, 2011). As for frequency, assimilation of voicing is relatively the least common type in English, and it rarely takes place across word borders (Berg, 1998). For Roach (2009), this is because voiced consonants in word initial and final positions have "little or no voicing anyway" (p. 112).

The most common cases of voice assimilation occur when a voiced alveolar, a dental, or a labiodental-fricative precedes its voiceless counterpart:

/z/	before	/s/	→	/s/	Cheese sandwich	/tʃiːs sændwɪdʒ/
/ð/	before	/θ/	→	/z/	With thanks	/wɪθ θæŋks/
/v/	before	/f/	→	/f/	Have to	/hæf tə/

There are cases in which devoicing occurs inside words in a progressive direction when the approximants /l, r, w/ become voiceless after a voiceless stop or a fricative. While the approximants are voiced in *right, win* and *lay*, they are devoiced in: *play*, *twin*, and *fright* (/pl̥eɪ/ /tw̥ɪn/ /fr̥aɪt/) (Roach, 2001, p. 112). Another important example of voice assimilation is the morpheme /s/ in the present tense and in plural nouns (Mott, 2011).

2.2.2. Elision

Elision is the process of CS whereby sounds which are normally pronounced in word citation forms seem to be deleted in running speech. The principle of least effort is manifested in this process, especially in smoothing the articulation of consonant clusters (Field, 2003); because it is usually difficult in English to pronounce sequences of three or more consonants fluently, speakers tend to avoid this complexity by eliding certain consonants on a regular basis. The following example form Roach (2001) illustrates how an utterance may be compressed after eliding sounds:

She looked particularly interesting

/ʃɪ lʊkt pətɪkjələlɪ ɪntərəstɪŋ/	→	/ʃɪ lʊk pətɪklɪ ɪntrstɪŋ/
27 phonemes	→	20 phonemes

The second compressed form is a normal pronunciation of the utterance. The sounds /t/ in *looked*, /ə, l, j/ in *particular*, and /ə/ in *interesting* were dropped. Just like in assimilation, this is a typical pronunciation which by no means denotes careless or lazy speech. On the contrary, speech would sound abnormal if sounds which would normally be elided are deliberately and consistently pronounced in CS (Yule, 1996; Dirven & Vespoor, 2004). Despite this, some variables may render the elision of sounds optional. Knight (2012) cites two factors: 1) speakers may have personal preferences in eliding segments in running speech; 2) some sounds may be deleted depending on the situation of use. Lorenz (2013) makes a link between elision and the fast movements of articulators, i.e. the faster they move, the more likely sounds will disappear.

In terms of coarticulation studies, it is suggested that the process of elision is related to assimilation. It is considered as an "extreme result of assimilation whereby two sounds are articulated so closely in time to each other that a sound or sounds between them are completely obscured" (Pascoe, Stackhouse, & Wells, 2006, p. 170). The main difference lies in the degree of movement the articulators make so that they reach the perfect position. Roach (2001) argued that in the case of elision, consonants may be dropped *despite* the articulatory attempt, but not in its absence. The observation of the tongue activity reveals that it makes its movement toward the appropriate articulatory position, but this movement is only partial. Ladefoged & Johnson (2011) note that in the elision of /t/ in *most people*, the tongue tip reaches the appropriate place in the alveolar ridge, but the sound is so obscured that it cannot be heard due to surrounding context. In this case, /t/ and the following bilabial sound /p/ are articulated almost simultaneously. Celce-Murcia et al. (1996) went further to consider elision as a radical accommodatory process where sounds are not clearly articulated.

Assimilation and elision may also interact. In many cases, the pronunciation of difficult sequences may be simplified either by deleting sounds, or through assimilating them. This is because the conditions for realising the two may be present in the same context, to the extent that one may be confused whether it is a case of assimilation or elision. For example, /d/ may either be assimilated in “emerald forest” /emərəlb fɒrɪst/ as a case of place assimilation, or elided /emərəl 'fɒrɪst/ because the three main rules for eliding it are present (Knight, 2012, p. 215). Another type of interaction is when “assimilation processes also involve elision” (Collins & Mees, 2013, p. 123). This is noticeable in word boundaries as well as in word internal positions where a successive application of assimilation and elision may take place:

Stand beside it /stænd bɪsaɪd/ → /stæm 'bɪsaɪd ɪt/

A learnt behaviour /lɜːnt bɪ'heɪvɪə/ → /lɜːm 'bɪheɪvɪə/

Both assimilation and elision are applied in these examples; in the first one, /d/ is first deleted leaving /n/ in a suitable environment for assimilating (place assimilation). In a similar way, /n/ is assimilated in the second example after the elision of /t/.

Both consonants and vowels may be elided in word internal and final positions. The most commonly elided vowel in English is the weak vowel /ə/, in addition to /ɪ/ in a less frequent manner. Cruttenden (2001) classifies vowel deletion into two main types: *Allophonic variation* and *phonemic elision*. The former happens mainly when the final element of a closing diphthong /eɪ, aɪ, ɔɪ, əʊ, aʊ/ is elided if it is followed by a vowel. This may take place both inside word and across boundaries:

/eɪ/	may appear	/meɪ əpɪə/	/meəpɪə/
/aɪ/	buy a couple	/baɪ ə kʌpl/	/ba ə kʌpl/

/ɔɪ/	the boy intervenes	/ðə bɔɪ ɪntəviːnz/	→	/ðə bɔ ɪntəviːnz/
/əʊ/	go away	/gəʊ əweɪ/	→	/gɜː weɪ/
/eɪ	layer	/leɪə/	→	/leə/

The second type, *phonemic elision*, takes place when /ə/ occurs in initial position before a continuant, and the previous word ends in a consonant:

/ə/	with another	/wɪð ənʌðə/	→	/wɪð 'nʌðə/
/ə/	went alone	/went ələʊn/	→	/went 'ləʊn/

Before an intrusive /r/, the /ə/ may be elided if the following word begins with /ə/. The following are common vowel-elision rules taken form Cruttenden (2001):

/ə/	faster and faster	/fɑːstər ən fɑːstə	→	/fɑːstr ən fɑːstə/
/ə/	after a while	/ɑftər ə waɪl/	→	/ɑftr ə waɪl/
/ə/	gets broader again	/gets brɔːdər əgen/	→	/gets brɔːdr əgen/

In initial syllables, it is elided when it occurs after /t, k, p/ (Roach, 2009):

/ə/	potato	/pəˈteɪtəʊ/	→	/pˈteɪtəʊ/
/ə/	collective	/kəˈlektɪv/	→	/kˈlektɪv/
/ə/	tonight	/təˈnaɪt/	→	/tˈnaɪt/

In unstressed syllables, the weak vowels /ɪ/ and /ə/ may be deleted in word internal positions such as in *aspirin* /æspɪrɪn/→/æsprɪn/ and camera /kæmərə/ →/kæmrə/ (Lorenz, 2013). Elision in this case is optional; the weak vowel in CS is not necessarily dropped (Giegerich, 1992). Note that elision in those last examples affects the syllabic structure of the word. The elision of vowels, which basically function as syllable nuclei, leads to the compression of syllables (/kæmərə/, /æspɪrɪn/: three syllables; /kamrə/, /æsprɪn/: two syllables), and results in consonant clusters /sp/ and /mr/. McMahon (2002, p. 129) gave similar examples where the elision of /ə/ in the first syllables of *connect* and *potato* results in the sequences /kn/ and /pt/ as word initials, which do not conform to the phonotactics of

English. This may also take place at word boundaries; the following examples are from Brown (1990, p. 73):

/tl/ in *back to London* /bæk tə lʌndən/ → /bæk tlʌndən/

/ts/ in *it's the way* /ɪtɪzðəweɪ/ → /tsðəweɪ/

Another common example of schwa elision is when it occurs before sonorant consonants, which may function as syllable nuclei; e.g. *button* /bʌtən/ →/bʌtn/, little /lɪtləl/ →/lɪtl/ (Giegerich, 1992).

Consonant elision is far more frequent than vowel elision. Most commonly elided consonants are the plosive alveolars /t, d/, in addition to /h/. We will start by presenting common instances and rules for the alveolars, and then move to deal with other less frequently elided consonants. As a rule of thumb, alveolar stops are usually elided when they occur in clusters, but different contexts exhibit different effects for each one. For instance, while /d/ may be dropped in word final position when immediately following a vowel (V /d/), /t/ may not:

/d/	in	V/d/+C	who'd been	/huːd bɪn/	/hː 'bɪn/
/d/	in	V/d/C	suspended from	səspendɪd frəm/	/səspendɪ 'frəm/
/d/	in	V/d/C	landed nose	/lændɪd nəʊz/	/lændɪ 'nəʊz/
/t/	in	V/t/C	wait for	/weɪt fɔː/	/weɪt fɔː/

The consonant /t/ is elided when it appears second in a cluster of three consonants (CCC) (Celce-Murcia et al., 1996). Such clusters are of two types: they may be divided across borders (CC+C), or they occur inside words (CCC). Examples of the latter would be:

/t	in	C/t/C	tasteless	/teɪstləs	/teɪsləs/
/t/	in	C/t/C	exactly	/ɪgzæktli/	/ɪgzækli/
/t/	in	C/t/C	listless	/lɪstləs/	/lɪsləs/

The second type includes clusters of two word-final consonants followed by a word initial consonant (CC+C). Cruttenden (2001) provides a well detailed account of plosive alveolars' elision. Cases where /t/ is elided may include combinations of word final voiceless fricative or plosive + /t/, followed by word initial consonant:

/t/	in	C/t/+C	East side	/ɪːst saɪd/	/ɪːs 'saɪd/
/t/	in	C/t/+C	kept quiet	/kept kwaɪət/	/kep 'kwaɪət/
/t/	in	C/t/+C	helped me	/helpt mi/	/help 'mi/
/t/	in	C/t/+C	liked me	/laɪkt mi/	/laɪk 'mi/

Similarly, final /d/ is deleted when it is preceded by a voiced plosive or affricate, and followed by an initial consonant:

/d/	in	C/d/+C	logged behind	/lɒgd bɪhaɪnd/	/lɒg bɪhaɪnd/
/d/	in	C/d/+C	Urged them	/ɜːdʒd ðəm/	/ɜːdʒd ðəm/
/d/	in	C/d/+C	grabbed them	/græbd ðem/	/græb ðem/
/d/	in	C/d/+C	arranged roses	ə'reɪndʒd rəʊzɪz/	ə'reɪndʒ rəʊzɪz/

There are exceptions to the above rules; /t/ is not elided in final consonant + /t/ clusters if the following word starts with a vowel. The same thing holds for final consonant + /d/ clusters, but this does not apply when the following word is the WF of 'and'. When the clusters are followed by the consonants /w, h, j, r/, /t/ and /d/ are linked, or re-syllabified rather than deleted:

/t/	in	/t/+/w/	a locked window	/ə lɒkt wɪndəʊ/	/ə lɒkt wɪndəʊ/
/t/	in	/t/+/h/	right hemisphere	/raɪt hemɪsphɪə/	/raɪt hemɪsphɪə/
/t/	in	/t/+/j/	last year	lɑːst jɪə/	lɑːst jɪə/

/t/	in	/t/+/r/	late race	/leɪt reɪs/	/leɪt reɪs/
/d/	in	/d/+/w/	red wallet	/red wɒlɪt/	/red wɒlɪt/
/d/	in	/d/+/h/	offered help	/ɒfəd help/	/ɒfəd help/
/d/	in	/d/+/j/	bad year	/bæd jɪə/	/bæd jɪə/
/d/	in	/d/+/r/	wide road	/waɪd rəʊd/	/waɪd rəʊd/

The pronunciation of /t/ is often simplified when it occurs in word final position, and the following word commences with /t/ or /d/:

/t/	in	/t/+/t/	I've got to go	/aɪ əv gɒt tə gəʊ/	/aɪ əv gɒ'tə gəʊ/
/t/	in	/t/+/d/	What do you do?	/wɒt dʊ jʊ duː/	/wɒ'dʊ jʊ duː/

Another common case of consonant elision is the dropping of initial /h/ in pronominal WFs; e.g., "I named him Twist*"* /aɪneɪmdɪmtwɪst/. Other consonants may also be deleted in CS, but this is comparably less frequent. Brown (1990) observed some regularities in the omission of /v, ð, l, r, n, k/ in some contexts. The following are some examples:

/v/	in	V+ stressed word	five p.m.	/faɪv piː em/	/faɪ 'piː em/
/ð/	in	the definite article	from the hospital	/frəm ðə hɒspɪtəl/	/frəm ə hɒspɪtəl/
/l/		after /ɔː/	all right	/ɔːl raɪt/	/ɔː raɪt/
/r/		after stressed syllables	terrorist	/terərɪst/	/teːrɪst/
/n/	in	un-/stressed syllables	one wants	/wʌn wɒnts/	/wɐ̃ wɔ̃ts/

2.2.3. Liaison

The opposite of elision is the process of liaison through which sounds are inserted between words to assure an easy and smooth transition between them. This happens especially when there is an articulatory break between syllables or word boundaries that makes the transition from one to the next difficult. An articulatory break is technically referred to as a *hiatus* (meaning *gap* in Latin), while the phonemes that are linked are said to be *in hiatus* (Skandrea & Burrleigh, 2011). Liaison happens when the hiatus is replaced by linking elements which are added between the sounds in hiatus.

The most common case of articulatory gap is when two vowels meet at word borders. This is noticeable in the phrases "in the‿end" and "law‿and order"; if these are articulated with the citation forms, a hiatus deprives an easy articulation and movement from /ə/ to /e/ and form /ɔ:/ and /ə/ in the two phrases, respectively. The role of liaison is to facilitate the transition, and this happens in different ways depending on the context. The degree of linking depends on a number of factors such as the formality level, speech rate, and even the speakers' individual preferences (Celce-Murcia et al., 1996). In addition, different accents also vary in the types of linking allowed.

Linking vowels that are in hiatus takes many forms. A final post-vocalic 'r' which is normally not pronounced in words like 'for', 'player', 'former' is articulated in some contexts. According to Skandrea and Burrleigh (2011), 'r' in such words used to be pronounced in all contexts in Old English, but now it only functions as a means of linking when the word is followed by an initial vowel:

Linking		No linking	
player identity	/pleɪər aɪdentɪti/	player bib	/pleɪə bɪb/
far away	/fɑ:r əweɪ/	far place	/fɑː pleɪs/

As a rule, final ‘r’ is linked in contexts where it follows one of the following vowels: /ə, ɜː ɪə, eə, ʊə/ .It should be noted that this type of linking is found in *non-rhotic accents* only, such as RP, in which the orthographic ‘r’ is usually written but not pronounced at most word endings (Lorenz, 2013). On the other hand, in *rhotic accents* such as American English, this type of linking is not found since the final ‘r’ is pronounced in all environments, (Collins & Mees, 2013).

In other cases, /r/ is added as a linking element even if it does not make part of the orthographic shape of a word. When two vowels at word borders are in hiatus, the linking happens with the insertion of an *intrusive /r/*. Unlike the linking /r/ discussed above, the intrusive /r/ does not have any historical justifications. It is usually inserted before an initial vowel when a previous word ends in a vowel:

media event	miːdɪə ɪvent	→	/miːdɪər ɪvent/
law and order	lɔː ən ɔdə	→	/lɔːrənɔːdə/

As for the transcription, Skandrea and Burrleigh (2011) suggest a normal phoneme to transcribe a linking /r/ (e.g. /fɔːregz/,), and a small one to transcribe the intrusive /r/ (e.g. /miːdɪəʳ ɪvent/).

The use of the intrusive /r/ has been criticised by language purists, especially when used in word internal positions (Skandrea & Burrleigh, 2011). Cruttenden (2001) notes that speakers who share this view tend to avoid linking with an intrusive /r/; the hiatus in this case is filled with a glottal stop instead. This tendency to use a glottal stop has also been generalised to the cases where a linking /r/ is normally articulated. Hence, a final post-vocalic /r/ may not be heard even if it is followed by an initial vowel:

	Linking /r/		Linking with a glottal stop
war and peace	/wɔːrəm piːs/	→	/wɔːʔ əmpiːs/

Linking also takes place when /j/ or /w/ is inserted in word borders. When a word ends in /ɪː, ɪ, eɪ, aɪ, ɔɪ/ and the following word is initiated with a vowel, a linking /j/ is inserted. Similarly, when a word ends in /uː, aʊ, əʊ/ and the following one starts with a vowel, a linking /w/ is inserted.

Two eggs	/tuː egz/	→	/tuː ˈʷegz/
Tomorrow afternoon	/təmɒrəʊ ɑːftənuːn/	→	/təmɒrəʊ ˈʷɑːftənuːn/
Three eggs	/θriː egs/	→	/θriː ˈʲegs/
My uncle	/maɪ ʌnkl/	→	/maɪ ˈʲʌnkl/

Here also, the linking /j/ and /w/ are not to be regarded as phonemes, and the transcription also reflects how they differ. Cruttenden (2001. p. 290) provided examples of minimal pairs to illustrate this difference:

Phonemic /j/		Linking /j/	
my years	/maɪ ˈjɪəz/	my ears	/maɪ ˈʲɪəz/
I yearn	/aɪ ˈjɜːn/	I earn	/aɪ ˈʲɜːn/

Phonemic /w/		Linking /w/	
too wide	/tuː ˈwaɪd/	two eyed	/tuː ˈʷeɪd/

It is noticeable that the articulation of /j/ in “my years” and of /w/ in “too wide” is somehow elongated compared to “my ears” and “two eyed”, respectively. Cruttenden (2014) notes that there is an inserted linking /ʲ/, but it is slight and “[cannot] be equated with phonemic /j/” (p. 317). Heselwood (2006) went further to suggest that these glides are “low level articulatory transitional phenomena” (p. 80). Furthermore, acoustic parameters were found to cue the distinction between the transitional glides (/ʷ/, /ʲ/) and the phonemes (/w/, /j/) (Hunt, 2009). Similar to the linking /r/, the glides /j/ and /w/ may also be replaced with a glottal stop.

2.2.4. Weak Forms

Despite that they are usually treated separately, WFs are considered as "another aspect of connected speech" (Dalton & Seidlhofer, 1994, p. 113). They are different from other accommodatory aspects in that only a limited number of words are regularly weakened in CS. These are mainly structure words (also called function words/ functors) which do not have a dictionary meaning the way content words have, and which function as grammatical links that hold content words together (Collins & Mees, 2013). Unlike content words, functors have two main pronunciation forms: A strong form (henceforth SF) (also called full form), which is usually listed in dictionary entries, and a WF that occurs in naturally spoken language. An example is the preposition 'from' which has a SF /frɒm/ that is used when the word occurs in isolation or in other special environments, and a WF /frəm/ that is used in CS (Brown & Kondo-Brown, 2006).

The *normal* pronunciation of function words in CS is the WF; speech would sound unnatural if the SF is deliberately and constantly used (Kisno, 2012). As noted earlier in the discussion of Rhythm, this type of weakening in function words allows for maintaining a regular rhythm of spoken English. According to Dretzke (1998), the regular intervals between strong syllables will be distorted if SFs are regularly used. He went further to suggest that the use of full forms where the weak ones are more appropriate would interfere with comprehensibility. The weakening of grammar words has to do also with the amount of information they carry compared to content words (Kisno, 2012). Because they carry relatively little meaning, they must be weakened so that words which communicate important information will be more prominent for the listener (Mott, 2011).

For Jenkins (2000), however, this argument is not satisfactory. He contends that speech would remain intelligible if speakers fail to use WFs. He also notes that, in

international meetings for example, speakers tend to use full forms rather than WFs for the reason of intelligibility. This argument is acceptable to some extent. If we consider syllable timed languages, we find that listeners are able to concentrate on focal points despite that all syllables receive the same degree of stress, and no such weakening in syllables occurs. But Jenkins is probably concerned with the use of English as an international language, in contexts where participants may not have the required competency to understand highly modified speech. Consequently, as speakers aim at ensuring a maximum comprehension of their ideas by participants, their awareness of the needs of the audience forces them to avoid WFs. Despite this, citation forms will always make speech unnatural and, whenever used, speakers may compromise the principles of efficiency and least effort (Bloomer, Griffiths, & Merrison, 2005).

The manner of reduction for WFs in CS depends on the context where the word occurs. When weakened, the WFs appear as realisations in which citation forms lose either their sound quality, quantity, or both (Roach, 2009; Dretzke, 1998). Weakening also results in some general characteristics: reduced forms are pronounced quickly, their length is reduced, and they are relatively lower in volume. As for qualitative reduction, it affects the vowel sounds, especially in monosyllabic words. This often happens with the "obscuration of vowels towards /ɔ, ɪ, ʊ/" (Cruttenden, 2001, p. 252), or a replacement of the vowel with a more central one (Figure 6), usually the schwa sound /ə/ (Field, 2003). For instance, the long vowel in 'for' /fɔː/ is reduced to a schwa sound as in *tea for two* /tiː fə tuː/.

Function words are reduced quantitatively when sounds, either consonants or vowels, are deleted. The /h/ in 'him' is usually dropped in CS as in "I found him there" /aɪ faʊnd ɪm ðeə/. An example of vowel elision is the dropping of /ə/ in 'and' as in "fish and chips" /fɪʃntʃɪps/.

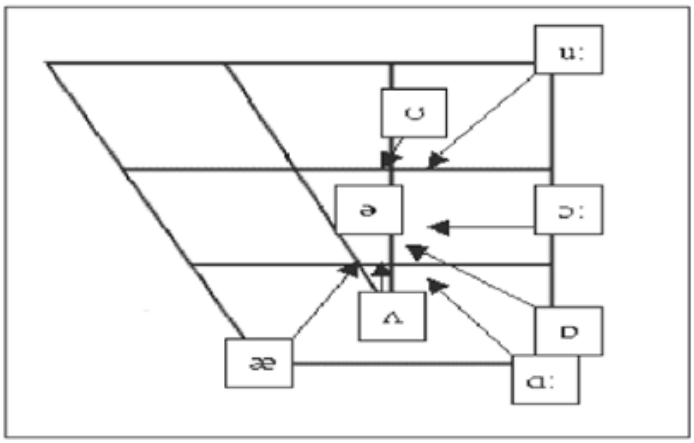

Figure 6: The Change of Position of Vowel Production for the Articulation of Weak Forms (from Kisno, 2012, p. 117)

It should be noted that, while function words have special rules of reduction in CS, they may also be subject to modification through the other aspects of CS, especially the linking to content words, the elision of consonants and vowels, and the assimilation of consonants (Dretzke, 1998). The fact that they are among the most common words in speech increases the likeliness that they get influenced in different environments. Dirven and Vaspoor (2004, p. 122) used the term *Complex Processes* to refer to the "series of processes whereby careful and relaxed pronunciation can be related" in the articulation of function words. In the phrase "girls and boys", there is first a reduction of the vowel in 'and' and then then a dropping of the consonant /d/ followed by the assimilation of the nasal /n/; the outcome would be /gɜːlzəmbɔɪz/ (Ibid).

Although the use of the weak version is the norm, there are situations in which the SF is obligatory or, in fewer cases, optional. The full form is used when the word is pronounced in isolation. In sentence final position, many function words are pronounced in their SF. The following table summarises the exceptions to the normal pronunciation of grammar words listed by Yavas (2006), Cruttenden (2001), Roach (2009), Mott (2011), Jenkins (2000), Ladefoged and Johnson (2011) and Kisno (2012):

Contexts of Strong Form use	Examples			
	Function word	**Strong form**	**Weak form**	**In context (strong form)**
When the function word is quoted	have	/hæv/	/əv/	The auxiliary '***have"***' has two pronunciations.
When it is emphasised	was	/wɔːz/	/wəz/	A: You were absent yesterday. B: No, ***was*** here.
In sentence final position	at	/æt/	/ət/	What are you looking *at*?
When it is contrasted with another word	from to	/frɒm/ /tuː/	/frəm/ /tə/	I said ***from*** Rome, not ***to*** Rome
When auxiliary verbs/ modals are negative	can	/kæn/	/kən/	I ***can't*** bear that anymore!
Before fillers (pauses, hesitations…)	and	/ænd/	/ənd/-/ən/	She looked angry ***and***, errr and very sad.
When 'that' is used as a demonstrative	that	/θæt/	/ðət/	Please sign ***that*** paper.
When 'must' is used in the sense of concluding something	must	/mʌst/	/məs/	His car is parked outside. He ***must*** be at home.

Table 1: Contexts where the Strong Form is often Used

One situation in which the use of the SF is optional is when prepositions precede a pronoun in its WF; e.g. in "*I have to go to you"* the second preposition 'to' is pronounced either /tə/ or /tuː/ (Cruttenden, 2001, p. 251).

Grammatical Category	Word	FFull form	Weak form	Example of weak form
Verbs	am	æm	m	That's what I'm trying to say.
	are	ɑː	ə	Where are you from?
	is	ɪz	əz/z/s	Where's he from? / Where is he from?
	was	wɒz	wəz	That's where he was born.
	were	wɜː	wə	That's where my children were born.
	do	duː	də	Where do you live?
	does	dʌz	dəz	Where does he live?
	have	hæv	əv/v	He will have left by now./ they've gone
	has	hæz	həz/əz/z/s	The baby was swallowed by a stone. He's gone
	had	hæd	həd/əd/d	He had already gone./He'd already done
	can	kæn	kən	I'm not sure if I can lend it to you.
	could	kʊd	kəd	Well, what could I say?
	would	wʊd	wəd/əd	Well, what would you have done?
	should	ʃud	ʃəd/ʃd	Well, what should I have said?
Personal pronouns	you	juː	jə	How do you do?
	your	jɔː	jə	What does your boss think?
	he	hiː	hɪ/ɪ	Where does he work?
	him	hɪm	ɪm	I'll give it to him later.
	she	ʃiː	ʃɪ	She's leaving tomorrow.
	her	hɜː	hə/ə	I'll give it to her later.
	us	ʌs	əs	They'll give it to us later.
	them	ðem	ðəm	I'll give it to them later.
Prepositions	to	tuː	tə	He's already gone to work.
	at	æt	ət	He's at work, I think.
	of	ɒf	əv	That's the last of the wine!
	for	fɔː	fə	He's away for two weeks.
	from	frɒm	frəm	She comes from Scotland.
Conjunctions	and	ænd	ən/ənd	She's tall and fair.
	but	bʌt	bət	She's here, but Juan isn't.
	than	ðæn	ðən	She's older than you.
Articles	a	eɪ	ə	He's a doctor.
	an	æn	ən	She's an architect.
	the	ðiː	ðə	She's the person I told you about.
Indefinite adjectives	any	eniː	əniː/niː	Have we got any biscuits?
	some	sʌm	səm	There's some tea in the pot.
	such	sʌtʃ	sətʃ	It's not such a big deal, really.

Table 2: Weak Forms of Function Words (Kelly, 2000, p. 74)

Table 2 presents a list of common function words with their SF and WF pronunciations. Cruttenden (2001; 2014) provided a larger list of 50 function words that includes, in addition to the items listed above, other less frequent yet common words in speech. Some grammar words which have one WF only may be reduced in special manners in fast speech. For instance, the reduction of the grammar words *I, by, my,* and *nor,* results in a new alternative for each one, which has a schwa as a main vowel: /ə/, mə/, /nə/ (Cruttenden, 2014).

2.3. Implications for Non-native Listeners

The set of modifications discussed in the previous section may have negative effects on non-native listeners in decoding speech. Due to their physical characteristics and/or the modifications they result in, CS features would make the processes of LS and WR that we discussed in the first chapter laborious tasks. This is in addition to the listener factors which add difficulty to the task of speech decoding. The features may also affect the process of language learning as a whole. We will discuss some of the implications of CS features in general and, then, report some of studies that dealt with the effects of the features on non-native listeners.

2.3.1. Acoustic Characteristics

While one of the CS aspects' *raison d'être* is to facilitate the speaker's task by allowing for a relaxed articulation and maximum ease, they are evidently challenging for listeners, especially non-natives. Their presence in speech interferes with the comprehensibility of the message, and with the listener's ability to perceive and recognise words. One reason for this is related to the acoustic/physical characteristics of words and sounds after the modifications. Modifications reduce both the qualitative and quantitative

qualities of words; they become shorter in pronunciation time and, where sounds are changed, the new sounds are weaker compared to the original ones (Larsen-Freeman, 1976). An example of this is the schwa /ə/ sound which replaces vowels in WFs. As the weakest vowel sound in English, the schwa is usually difficult to perceive. For Field (2008a), because of the qualitative and quantitative characteristics of unstressed syllables, there is "small wonder […] that they pose perceptual problems for the foreign listener" (1p. 329).

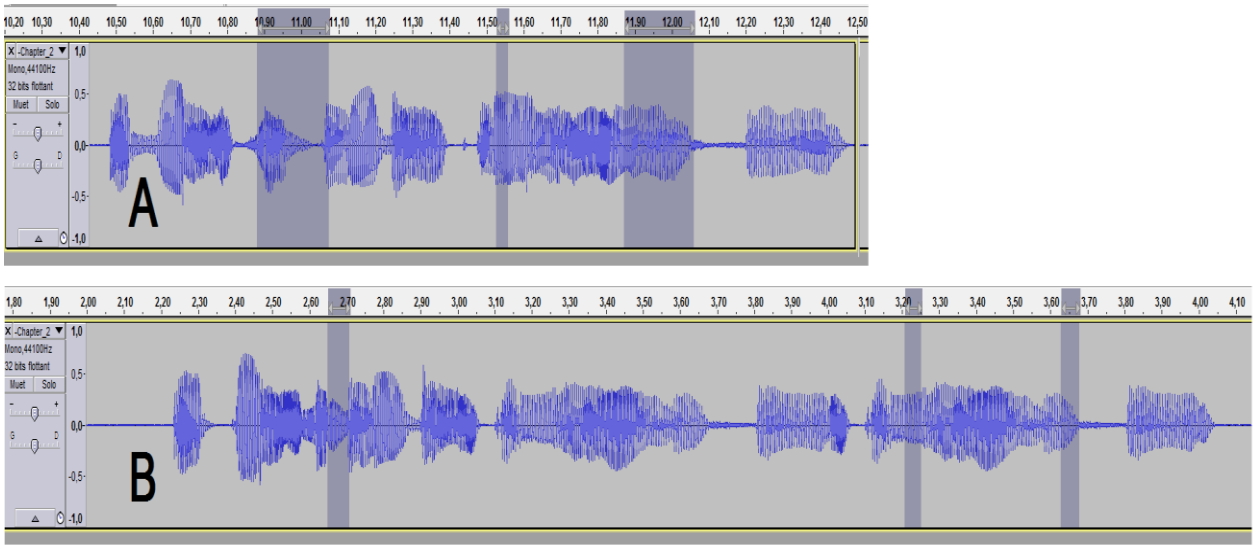

Figure 7: Speech Waveforms Generated for an Utterance Pronounced in a Natural/ Unnatural Way

Figure 7 presents waveforms generated by an audio software for the sentence "*It might have been back a mile or so*" (audio file from Weinstein (2001)) in two conditions: The first (A) represents the unnatural pronunciation in which the words are pronounced in their citation forms, while the second (B) shows the natural relaxed pronunciation. The three highlighted areas in each waveform stand for "have been", 'a' and 'or' from the utterance. There is an apparent difference between the two conditions which is noticeable in the pronunciation time allocated to the pronunciation of not only the function words, but also the content words. In addition, there are more salient gaps in the second condition, while the lengths of the waveforms demonstrate the difference in quality between the two conditions.

These physical characteristics of natural speech potentially affect the perceptual saliency of speech, and cause BU decoding problems for non-native listeners.

2.3.2. Perceptual Saliency

One major effect of phonological modifications is that they reduce the perceptual saliency of the input. In chapter one, we discussed the notions of input, intake, and the relations between them. One main factor for input to become intake is the perceptual saliency of the language items in the input; the more salient they are, the more easily they are perceived and transformed into intake (Henrichsen, 1984). However, due to their physical characteristics, it is argued that CS phenomena reduce the perceptual saliency of the input. Hakuta (1976) found that more perceptually salient grammar forms were acquired earlier than less salient ones. Reduced saliency was found to delay the acquisition of some forms like 'gonna' compared to more salient ones (Hakuta, 1976). Henrichsen (1984) and Ito (2001; 2006) found that reduced forms stand as a filter to the input-intake process and cause comprehension problems, especially for language learners.

Another potential problem related to the saliency of reduced forms is that they stand as a barrier for the learning of new language items. There are many examples which demonstrate how CS phenomena cause mishearings. For instance, "put the" may be perceived as 'putta', and "they are for" as 'therefore'. For language learners whose knowledge about the target language is not adequate, such mishearings provide unreliable data about the target language system. Consequently, the learner's interlanguage may remain stagnant as erroneous data will deprive their approximate system from developing (Odlin, 1978). Henrichsen (1984) spoke about the condition in which Sandhi may stand as a filter that disrupts grammatical information from being integrated into the learner's approximate system. He states that:

> Without back-up knowledge of the system of English, the learner may misperceive this input. For example, *d'you* may be misperceived as merely *you*. These misperceptions may provide erroneous support for the learner's developing hypotheses about how English operates. His or her interlanguage may be incorrect at these points, but continued exposure to more natural input may merely reinforce the same misperception (p. 120,121).

This condition may also be generalised to the learning of lexical items. Since content words are affected by assimilation and elision phenomena, the way they are perceived could be hampered by the reductions occurring at their borders.

CS phenomena not only lead to wrong hypotheses to be made about the nature of the grammatical system, but also deprive the non-native listener form the necessary grammatical cues to parse speech. In this sense, function words and bound grammatical morphemes are believed to cause comprehension problems. According to Cintron-Valentine and Ellis (2007), because structure words and grammatical morphemes provide significant syntactic cues about the structure of utterances, the inability to perceive them means that the information they carry will not be available for the listener to make successful syntactic parsing. In the example of /t/ elision in "*I walked back*", the past tense clue is omitted, forcing the listener to rely on the context or on their linguistic knowledge (Cruttenden, 2001). However, the latter is exactly what language learners lack (Cintron-Valentine & Ellis, 2007). Similarly, the misperception of structure words which tie content words together makes the relation between parts of the utterance loose unless the learner compensates for this with linguistic/non-linguistic information.

It can be thought that structure words are sufficiently salient by matter of their frequency. Through repeated encounters, learners are able to notice the recurrence of, say articles, which makes them amenable to acquisition. However, CS features may have stronger effects on saliency than frequency. Hakuta (1976, p. 338) reported that "telling against their salience is the operation of liaison in English which causes articles to be frequently slurred and hard to tell apart, *that's a* becoming Thassa, and *put the* becoming Putta or Pudda". This also leads to reduced attention which by its turn filters the input-intake process (Ito, 2006) and forces the learner to spend much effort to understand the message (Osada, 2004).

2.3.3. Unfamiliarity and Lack of Training

The presence of CS in speech interferes with the ability to decode spoken language particularly for learners who are not familiar with naturally spoken English (Kisno, 2012). Those who have been taught with a focus on reading and on written texts often have problems with parsing the spoken language (Underwood & Wallace, 2012; Rylander, Clark, & Derrah, 2013). For Ringbom (2007), this is because they are familiarised with the word written form but not the spoken form. In addition, they are usually exposed to fine-tuned language inside the classroom, which is loaded with carefully and slowly articulated words. The risk that this entails is that, due to CS phenomena, spoken language does not necessarily include the ideal word forms which are found in written language (Elliott & Wilson, 2013). Given this, learners make wrong expectations about the pronunciation of the target language, which will be broken at the first encounter with naturally spoken English outside the classroom; they are often shocked when they listen to CS (Brown G. , 1990).

According to Rost (1990), "listeners who anticipate hearing ideal pronunciation of words will have considerable difficulty decoding CS since all phonemes change their

perceptual features in different phonetic environments" (p. 38). The phonemes which they expect to find may be elided in CS (Roach, 2009). The words encountered, together with their meanings, may be part of the learner's lexicon, but their *new* shapes may not. Thus, words which are easily recognised in isolation will become difficult to decode in CS (Rost & Wilson, 2013), and access to the meanings in the lexicon will be blocked even for the already known words (McQueen, 2007). This is one reason for the widespread idea that native speakers speak fast. Buck (2001) argues that it is not the speech rate that causes the difficulties, but rather the lack of preparation for CS phenomena. For Ito (2006), it is the lack of training about these modification phenomena that makes learners unprepared to listen to real English by native speakers. In technical terms, these learners lack the required knowledge about the sound system, which leads to listening difficulties (Mashahiro & Luan, 2012). They could even be completely unaware of the existence of CS phenomena (Dalton & Seidlhofer, 1994). Celce-Murcia et al. (1996) show that lack of awareness about subtle phonological information causes communication breakdown. Among other information, the listener who does not know that auxiliary verbs are always strong in the negative form will possibly understand that he is allowed to park his car when a policeman says /jukæm'pɑːk hɪə/; through a complex phonological process, the /t/ in 'can't' is elided and /n/ undergoes place assimilation as the following word is initiated with a bilabial consonant.

2.3.4. Lexical Segmentation Problems

While most foreign learners complain that the speech rate is what causes most of their decoding failures, it is arguably not the case. Even in low speech rates, learners are still unable to recognise words and segment speech. The reason lies in the fact that spoken language lacks the luxury of the pauses that enable learners to determine word beginnings

and endings. The speech waves in Figure 7 demonstrate how in a normal speech rate words are articulated in a continuum as if no divisions exist between them. It is the CS phenomena, more than the speech rate, which are responsible for the lack of pauses and the blurring of word boundaries, and which cause learners' LS problems (Brown G. , 1990). In a study by Brequist (1994, cited in Lynch (1998)), three conditions were compared for their effect on comprehension. It was found that comprehension suffered more under the condition in which pauses were missing, more than when the speech rate was slow or normal.

Field (2008b) found that non-native listeners adopt a matching strategy to determine word boundaries in the acoustic signal. This is based on linking sounds or clusters of sounds from the input to the set of words stored in the lexicon. Given the tentative nature of speech, there may be different options for dividing the stream of speech, but usually only one option is appropriate according to the context/co-text. The common example of "I scream" and "ice cream" shows that special information must be used by the listener (e.g., context and small phonetic details) to determine where the boundary falls. It was found that the risk with the matching strategy is that learners are more hesitant compared to native speakers to revise the appropriateness of the gaps they inserted. In other words, native speakers make revisions based on different information sources, whereas non-native learners were found to be more reluctant and very slow to do so due to the lack of confidence, greater cognitive demands and insufficient spare attentional capacity. Another possible reason is that the non-native listeners might be unaware of the existence of alternative segmentation possibilities. Despite that the words in Field's test were among the vocabulary they should possess, they hesitated to abandon wrong segmentation hypotheses.

Field (2008a; Field, 2008b) extended the tentative nature of speech to include cases in which CS features lead to segmentation weaknesses. He specified three main features:

WFs, contractions, and formulaic phrases. For function words, the fact that they are usually short monosyllabic morphemes makes them appear as parts of longer words. For instance, the reduction of the pronoun /hɜː/ results in /ə/ which may be perceived as a part of /əsist/ in "assist her" →/əsɪstə/ "a sister". Similarly, "the standard the hotel achieves" may be segmented as "stand at the hotel". Segmentation failure in the last example is particularly bound to the extent to which the learner is familiar with WFs. This is because /ət/ is in the first place a WF of which beginners may not be aware. Weaker learners may potentially hear "stand the hotel" as the final /d/ in 'standard' becomes dental as a result of assimilation (which makes it similar to /ð/), or even "standard hotel". For Lynch (2009), the frequency of WFs of function words in English "adds [more problems] to the L2 listener's difficulty in segmenting speech" (Lynch, 2009).

Dalton and Seidlhofer (1994) discussed the problem of segmentation in relation to one type of linking referred to as *Catenation*; it results in a consonant being displaced to appear as part of a neighbouring word if the latter starts with a vowel. This displacement is a source of confusion in determining word borders. The authors gave the example of "get out*"* which may be segmented as "ge-tout", and commented that "this is often responsible for learners' confusion regarding the identification of word boundaries" (Dalton & Seidlhofer, 1994, p. 123). Because of such linking, the learner will probably think that the sequence of sounds includes a new word which is not part of his mental lexicon. Or, he may just adopt a matching strategy, the result of which is a LS which does not conform to the speaker's words.

Elision is also problematic, especially at word boundaries as it often produces illegal sound clusters, which are sequences which do not conform to the phonotactics of English. Consequently, learners with partial knowledge about elision may opt for inserting a word

boundary in the illegal cluster with respect to the phonotactic rule. This means that knowledge about phonotactics alone is not sufficient in enabling the listener to segment speech. It is true, for instance, that /pt/ and /kn/ are not possible clusters in initial positions. Given that these clusters may be the result of vowel elision in /pəteɪtəʊ/ and /kənekt/, the listener may segment "right potato" and "we connect to the body" as /raɪp teɪtəʊ/ and /wiːk nek tə ðə bɒdi/.

2.3.5. Word Recognition Failure

Phonological modifications of words often constrain learners' ability to recognise words in running speech, especially those who expect to hear the citation forms of words. According to Peterson (2001), "although [intermediate learners] have internalised the phonemic system of the language fairly well, they may have little understanding of the complexities of phonological rules that govern fast speech, elisions, assimilations and so forth" (p. 94). What these learners often do is that they scan the input for possible matches between the sounds and the words they have by relying on phonetic features. This is however of a limited help given that phonemes do not maintain their regularities in CS (Levelt, 1993). One possible implication is that the changes in phonemes decrease the lexical contrasts between words and "potentially disrupt lexical recognition" (Darcy, Ramus, Christophe, Kinzler, & Dupoux, 2009). One single modification in a sound can neutralise the lexical contrast by making the word similar to another one. In some contexts, the changes are so extensive that the outcome of modification appears as a new lexical item which does not fit the context, neither meaningfully nor syntactically. The following examples are taken form Davenport and Hannahs (2010) Knight (2012) and Knowles (2014):

hand bag	ham bag	elision and assimilation
tinned salmon	tin salmon	elision

creamed potatoes	cream potatoes	assimilation
Sutton Coldfield	Sutton Coalfield	elision
schedule times	scheduled times	elision
clothes	close	elision

Knight (2012) notes that such modifications do have the *potential* to cause misunderstanding despite that context and other resources may provide reliable cues. Knowles (2014) acknowledges that the modifications at word borders leave some subtle features which allow the listener to recover from the effects of modifications such as assimilation and elision. Despite this, even native speakers may be confused with such modifications which may also have effects on writing (Knowels, 2014).

Several examples of neutralisations can be found in function words whose WFs may be identical. For example, the WFs of 'has' and 'is' are neutralised in some environments:

/ɪtsˈbɪn/ (It's Bin who did that)

/ɪtsbɪn/ (It has been a long time)

Even those which are not identical can be neutralised through other CS features. The words in the sequence of sounds /jʊwəfaɪndeksrəmʌni/ can either be recognised as "you *are* fined / you *were* fined extra money" as a result of the linking /ʷ/. The WF of 'have' can also be identical to that of 'are' in some contexts, as in "*you have found extra money"* /jʊwəfaʊndeksrəmʌni/. Similarly, 'her', 'of' and 'or' can be realised as /ə/, 'an', and the conjunction 'and' as /ən/, 'are' and 'or' as /ər/, 'as' and 'has' as /əz/, 'the' and 'there' as /ðə/ (Dretzke, 1998). Failure in recognising these words and the contrasts between them affects the process of syntactic parsing and, consequently, results in reduced comprehension.

The problem of modification is more severe when word initial phonemes are affected. We discussed the problem of WR in chapter I in relation to the processes of activation and lexical search that are initiated with the perception of word initial sound clusters. Following the Cohort model of WR, we saw how the activation of potential candidates is affected when segmental information is not accurate. Marslen-Wilson and Zwitserlood (1989) found that initial sounds have a special status in that they can block the process of lexical access if there is a mismatch in the initial phoneme. They went further to suggest that any phonemic mismatch will have severe consequences on WR.

Just like the consequences limited knowledge about phonetic features can have on WR, modification processes such as assimilation, linking and elision can also have disrupting effects. As word beginnings are changed and sounds are inserted or deleted, lexical search could be triggered based on the *new* sound/sounds rather than the original one and, consequently, the original word will not be part of the activated candidates. This is analogous to a process where one looks for a given entry in a dictionary section of a letter with which the entry does not start. For instance, non-native listeners who do not differentiate between the phonemic /j, w/ and the linking /$^{j, w}$/ will possibly fail recognise words that undergo linking processes. In the phrase "over the ears", the /ə/ and the /ɪ/ are in hiatus and a linking /j/ is inserted between them as a glide to fill in the articulatory gap. Learners who are not aware of this type of link may consider the linking /j/ as phonemic, and misrecognise the words as "*over the years*".

McQueen (2007) suggests that it is not only the phonological knowledge that is needed to recognise words that undergo modification, but also language independent information about subtle phonetic information that is left by the modifications. Other types of information also include non-linguistic and contextual features. Despite that the Cohort

model has been criticised in that contextual information may also be deployed to compensate for the misperception of segmental information (especially at word beginnings), the model is likely to be reliable for describing lower-level learners' listening process which usually follows an analytic way in segmenting speech, and relies often exclusively on the signal. For this category of learners, missing the initial segment of a word will eventually block lexical access or lead to the activation of wrong candidates based on the wrong segment. They could also think that the words they could not recognise are unknown to them.

2.3.6. Effects of CS: Evidence from Research

In language learning research, several studies have looked into the effects of CS features on learner's ability to comprehend natural speech. Henrichsen's study (1984) was probably the first to establish a link between learners' reduced comprehension and the presence of CS phenomena. A group of native speakers, and two groups of higher and lower proficiency ESL learners were tested on their ability to comprehend speech under two conditions; the first in the presence of CS features and the second in their absence. It was hypothesised that participants whose linguistic ability is high would obtain higher scores compared to those with a low ability, and that the effects of CS would be greater for lower-proficiency learners. The results confirmed the hypotheses; while no significant difference was found in the performance of native speakers in the two conditions, CS phenomena had an effect on the two groups of learners in that they obtained significantly lower scores in the presence condition compared to the absence condition. However, the higher proficiency group's scores were significantly higher than those of the lower proficiency group.

The knowledge about the language system could have allowed the higher proficiency group to compensate for the lack of perceptual saliency. However, it is important to note that knowledge about the language system was not sufficient to compensate for all problems

faced by the upper group; despite that their performance was higher than that of the lower ESL group, it remained significantly low compared to the performance of the native speakers in the presence condition. This implies that, even when the linguistic knowledge is high, it is not sufficient to compensate for poor BU decoding skills. Although the experiment's results are not informative as to what exactly allowed the higher proficiency group to outperform the lower proficiency group, it is unlikely to be knowledge about CS features. According to Henrichsen, the ESL participants (international students who had only one semester studying in an English speaking country) were more exposed to foreign talk, and they had more contact with international students than with the natives from the relatively small local area in which the university is situated. It is possible, then, that their knowledge about this natural way of English pronunciation and Sandhi is not mature. If this was the case, their higher scores were not due to their knowledge about the natural way of how English is pronounced, but rather from their knowledge about other areas of the language system which they used to compensate for their inability to perceive Sandhi variations. Henrichsen reported that the results "support the idea that perception is not dependent on the signal alone and that at least one other factor –knowledge of the language –plays an essential part in the input-intake process" (Henrichsen, 1984, p. 120).

Ito (2001; 2006) replicated Henrichsen's study by adding sentence complexity as a variable, in addition to analysing the effects of the different types of modifications. CS phenomena were categorised into two types which were referred to as lexical reduced forms (they can be memorised as one lexical item such as "do not" → 'don't'), and phonological (non-lexical) reduced forms. The results indicated that the non-native speakers scored significantly higher in the absence condition compared to the presence condition, while no significant difference was found in the performance of the native speakers. In addition, the level of proficiency did lead to a significant difference between the upper and the lower

proficiency groups. Concerning the types of reductions, phonological reductions were found to have a more significant effect on listening than lexical reductions. The overall results showed that reduced forms present a barrier to listening. One limitation in this study is that all types of the reductions tested were contractions, regardless of the category to which they belonged (lexical or non-lexical). Reduced comprehensibility in the study cannot be taken as an effect of CS phenomena in general because the other types of reductions were not tested as to their effects on listening.

Zahdi, Sahragrad and Nasirizadeh (2007) addressed the effects of a wider range of phonological features on listening. The total of 125 EFL university learners with differing proficiency levels –elementary, intermediate and advanced– were tested on their ability to decode CS, with a particular focus on the features of assimilation, elision, liaison, juncture, palatalization and transition. The results showed no interaction between the decoding of CS and the proficiency level. Gender, for its part, was not found to be a significant discriminator in the participants' performance. On the other hand, all groups "were generally weak at perceiving phonological features presented in L2 listening materials" (Zahdi, 2007, p. 115). As for the categories of reductions, assimilation and elision exerted more salient effects on the testees compared to the other features, especially transition and junctures which were the least problematic. Consequently, the study gives more support to the argument that CS features stand as a barrier to listening for language learners.

Gao (2014) found that most of the problems Chinese learners of English faced were attributed to their lack of phonological knowledge in addition to their lack of awareness concerning CS features. The problems of LS and poor identification of already known words were among the most problematic areas as indicated by the three instruments used: A questionnaire, a retrospective report and a listening test. Joyce (2014) developed a listening

test that aimed at evaluating L2 learners' ability to comprehend CS. The test which included 13 sentences featuring different aspects of CS was administered to 548 Japanese university majors ranging from beginner to upper intermediate levels. Unexpectedly, the results demonstrated that not only beginners, but also intermediate learners had serious problems in understanding naturally spoken English. Decoding problems included even the most common words when they were spoken naturally.

Some recent studies tackled specific features and processes. In a recent study, Wong et al. (2017) investigated a number of phonological skills to find out which ones predicted and aided more accurate reduced form perception. In addition to confirming the correlation between reduced forms perception and the general LC ability, the results suggested that learners' ability to recognise parts of words and their received vocabulary were more significant skills in CS perception. The former allows the listener to act in a TD manner to recognise the word before it is fully articulated by utilising the parts of the received signal for accessing the mental lexicon. This aids the process of recovering from the effects of modifications, especially those that occur at word endings. It has a solid ground in listening since modifications tend to affect word endings more than word beginnings (Field, 2003). Because there is an interaction between the signal and the learner's stored knowledge, the skill requires both TD and BU processing. Similarly, receptive vocabulary is also a significant skill as it provides useful contextual cues as to which words are likely to be used.

2.4. Teaching Connected Speech

Two important questions to raise at this stage are: 1) Can CS be taught to address learners' listening problems? 2) How CS features could be best addressed in actual classroom practice? These questions will be addressed in the following sub-sections in light of the related literature.

2.4.1. Effectiveness of Instruction

There is a large body of research that investigated the effects of systematic training on CS in improving learners' decoding skills, and their overall LC ability. We will review a number of studies in both EFL/ESL contexts that focused on the different aspects of CS.

Brown & Hilferty (1986) conducted a classroom research to see the effects of CS instruction on learners' decoding and LC. An EG received ten-minute daily lessons on reduced forms while a CG was given word discrimination lessons and pronunciation drills. After four weeks of instruction, three types of tests were administered to see the possible effects of reduced forms instruction. The first was a dictation test that aimed at evaluating the effects of the ten-minute lessons in improving learners' ability to perceive these forms. The other two tests were focused on learners' listening and linguistic abilities, and they included: 1) An Integrative Grammar Test, and 2) a multiple choice-listening test. The results revealed that the instruction the EG received had a significant effect on learners' performance in both the reduced forms dictation test and the Integrative Grammar Test. The learners who received lessons on reduced forms outperformed those who did not; the post-test results of the two groups had a significant statistical difference. For the LC test, the weak performance was attributed to the period of the experiment which was deemed insufficient to have a positive effect on the overall listening skill.

Similar results were found by Matsuzawa (2006). The participants were 16 Japanese business people who received short lessons of about 30 minutes on CS features over a period of four weeks. The lessons included a brief explanation of TD and BU processing modes, the rhythm of English, direct explanation/discussion of rules and dictation tasks. At the outset of the study, a cloze task revealed that CS features present serious obstacles for the participants; only about 50% of the test items were answered correctly. Most problematic

features included WFs, flapping, linking, contractions, palatalization, and elision. After the instruction, the participants' performance improved significantly, indicating that the short lessons had a positive effect on their ability to perceive reduced forms. Due to their strong effect on listening, Matsuzawa urged for a special attention to WFs and contractions. This study was recently replicated on by Cormier et al. (2013) with an additional delayed post-test to see the long term effects of the instruction. The post-test and the delayed post-test scores of the Chinese business people who participated in the study were significantly higher than those of the pre-test.

What is interesting about the study of Cormier et al. is the categorisation of the participants according to whether their work implied a certain amount of exposure to natural English, which may have an effect on their achievement. The degree of exposure to the language that varied according to the work environment was not found to have a significant effect on their performance on the pre-test. This may imply that mere exposure to the language without particular attention to the way words are modified cannot be of much help to the non-native listener. It also suggests that any successful communication in the work context when CS is used is the result of relying on contextual or linguistic knowledge to compensate for poor decoding skills.

The use of songs in teaching reduced forms was found to be effective in improving international learners' perception of CS. We report the results of two studies that investigated this. Carreira (2008) used pop songs in weekly lessons of 90 minutes to teach aspects of CS to a group of international learners of English. The participants were pretested and post-tested using two measures: the listening sections of the Test of English for International Communication (henceforth TOEIC), and dictation cloze tests. The testees obtained significantly higher scores in the dictation cloze tests in the post-test, indicating that the use

of songs was effective. However, no significant differences were found in the TOEIC listening test. This was attributed to the fact that the TOEIC test requires not only a BU skill but also a TD ability to use background and contextual knowledge. The experiment was followed by a survey that included open-ended and closed-ended questions to probe the participants' reactions to the instruction. The findings revealed different views among higher and lower proficiency learners. While the former had positive views about the lessons, the latter found CS lessons not useful or very easy for them. These results match those found by Goh (2000) which revealed that listening problems varied depending on the proficiency level, and that lower level learners have more problems with decoding than the upper level ones. Although the language of songs may not have a natural rhythm like ordinary speech, using them to teach reduced forms was effective.

In the second study, Ting and Kuo (2012) found significant gains on the part of an EG –EFL learners– which received explicit instruction on some CS phenomena through songs. Of the six categories of CS features, the participants' scores were significantly higher in the perception of C-V linking, contraction, palatalization, and elision. The two other categories of /h/ deletion and flapping were found to be more difficult, and were classified as *learnable* rather than *teachable*. For the authors, the difference lies in the assumption that "teachable components can be achieved by concrete instruction during a shorter time. But it takes longer exposure to improve components which are regarded as learnable" (Ting & Kuo, 2012, p.88). Despite this, the six week period of instruction was sufficient to have a positive effect on the overall scores of the participants. Table 3 summarises the findings of other studies on the perception of CS.

Study	Aspects included	Method	Training Period	Training Focus	Testing instruments	Results
Wang 2005	Elision, assimilation, contraction, linking	EG (N=37), CG (N=35). EFL Taiwanese freshman students Intact classes	7 weeks (every 2 weeks), 30-minute lessons (3.5 hours)	Self-developed units based on video clips form a movie. Comprehension check, dictation, CS processes description, short dictation	- Sentence dictation test, LC test (GEPT)	- EG significantly increased their awareness of reduced forms. EG did not outperform the CG in the LC post-test.
Crawford 2006	A number of reductions	EG (N=23), CG (N=26). EFL Japanese freshman students. Intact classes	7 weeks, 15-minute lessons (2 hours)	- Self-developed sentences. - Description, examples, cloze dictation.	- Cloze dictation	- For most of the reductions, gains were over 30% in the post-test
Kuo, 2009	C-V linking, /h/ deletion, C-C linking, V-V linking	EG (N=33), CG (N=32). EFL Taiwanese six graders Intact classes	14 weeks 40 minute lessons, twice a week (18 hours)	- Self-developed worksheets. - Explanation, modeling, repetition, practice, song teaching, production, recording	- Sentence reading, listen and circle, mark linking cloze dictation	- EG significantly improved their speech production and developed phonological awareness

Lee & Kuo, 2010	C-V linking, C-C reduction and palatelization	EG1 (explicit instruction), EG2(communicative instruction), CG, N=30-32 each. EFL Taiwanese nine graders Intact classes	3 weeks daily 15-minute lessons (4 hours)	- Self-developed. - Explicit CS group: description, cloze dictation. - Communicative group: 15 activities with CS processes in topic sentence.	- Cloze dictation, LC test	- The communicative group performed significantly better than the other groups on the cloze text. Explicit instruction group did not perform significantly better than the CG on the cloze text; no significant improvement was found on the LC among the three groups.
Chenjun & Li (2012)	Chunks, WFs, resyllabification, assimilation, elision, cliticisation (in addition to other decoding skills)	EG (N=60) CG (N=57) Both were non-English major EFL freshmen	Six weeks	- Training framework suggested by Goh (2000) - Explicit Instruction on basic decoding skills - Explanation of rules - Skill training	- Pre-test - Post-test - Delayed post-test - College English Test (CET) that included multiple choice questions and a compound dictation	- Improved comprehension and control of the listening process by L2 learners. - Better general comprehension.

Table 3: Some ESL/EFL Studies on the Effectiveness of CS Instruction – (Adapted from Alameen (2014))

In sum, the available studies provide reliable evidence to the effectiveness of CS instruction although they differ in the teaching approaches followed, and their results show varying degrees of effect. Except for Kuo's study (2009), the training sessions for all other studies ranged between 10 and 30 minutes in time with a primary focus on decoding skills. None of them, however, suggested a framework where CS aspects could be addressed as a part of a larger LC training programme that develops, in addition to the BU decoding skills, the learner's ability to bring schematic and contextual knowledge into play during listening. As Carreira's study suggests (2008), training in decoding skills alone is not sufficient in allowing a better listening performance. The remaining of this section will be reserved for discussing some theoretical and practical issues related to CS speech instruction.

2.4.2. Enhancing Learners' Listening: Focus on Strategies

Developing learners' listening comprehension through a focus on developing learners' strategies is a major trend in language teaching. This section outlines the major principles behind this trend, and highlights its limitations as discussed in the literature.

2.4.2.1. Overview of Learner Strategies

One approach that concentrated on improving learners' listening has emphasised learner strategies. This area has received much interest in EFL and ESL research and instruction. Learning strategies are defined simply as "the conscious thoughts and actions that learners take in order to achieve a learning goal" (Chamot, 2004, p. 14). Research into learner strategies has focused on, among other issues, the types of strategies used by learners, how learners differ in their application of strategies, what strategies make some learners more successful compared to others, the relation between strategy application and the general proficiency level, and the teachability of strategies. We will briefly present some of

the principles of learning strategies in general, and then discuss how listening instruction adopted this strategic approach.

The main principle behind a focus on learner strategies is to teach learners how they should approach the target language, and take responsibility of their own learning. Weaver and Cohen (1994) noted the shift in foreign language teaching through which the learner has become the centre of the learning process, while the teacher's role is only to guide this process. The application of this principle requires that learners should know how to take responsibility of their learning and be aware of the strategies that meet their needs:

> Our point of view is that learning will be facilitated if students become aware of the range of possible strategies they could use successfully throughout the language learning process. With learning strategy instruction, students can "learn how to learn" a foreign language when they are provided with the necessary tools to self-diagnose their learning difficulties (Cohan, 1994, p. 286).

Through the focus on strategies, the teacher allows the learner to reflect on his own learning, to depict his own problems, and to evaluate and monitor his own comprehension/learning processes.

There are taxonomies for the types of strategies learners use. For Cohen & Weaver (1998), learner strategies include both *language learning strategies* and *language use strategies* which are "selected by the learner either to improve the learning of a second language, the *use* of it, or both" (p. 3). Oxford (1990) provided the following taxonomy which includes four main general types of strategies: *cognitive strategies* which allow for the manipulation and transformation of the language; *metacognitive strategies* which include "actions which go beyond purely cognitive devices, and which provide a way for learners to

coordinate their own learning process" (p. 136); *affective strategies*, the indirect strategies that help learners regulate their emotions, reduce anxiety, and control attitudes; and *social strategies* that allow for interactive learning especially in groups.

It should be noted that, although some strategies may be recognised as 'effective' or 'good', this is not something inherent in them (Cohen & Weaver, 1998). It is argued that different learners may use the same strategies but with varying degrees of success (Littlejohn, 2008). That is, the way the strategies are used makes them effective for some learners, but not for others. In addition, it is argued that the use of more strategies does not necessarily guarantee success.

Research into learner strategies has been in favour of an instructional approach that teaches learners how to use strategies for their benefit. One significant line of research was focused on the strategies used by successful learners and the possible transfer of these strategies to less successful ones. These strategies have been elicited through, to mention only some procedures, think-aloud protocols, questionnaires, classroom observations, and learner diaries (Cohen & Scott, 1996). Another one has concentrated on the sequence of strategy application in the foreign language classroom. Weaver & Cohen's review (1994) of the sequences suggested by Oxford (1990), Pearson and Dole (1987) and Chamot & O'Malley (1994) revealed similar components:

- Presenting, explaining, and modelling strategies
- Giving learners a chance for practice and consolidation
- Encouraging the transfer of successful strategies to new tasks
- Evaluating the effectiveness of the strategies selected (e.g., for particular tasks)

2.4.2.2. Listening Strategy Instruction

Available research into learning strategies in listening has resulted in the "*Strategy-based Approach*" (Lynch & Mendelsohn, Listening, 2010). Mendelsohn (1995, p. 134; cited in Mendelsohn, 1998, p. 87) defines this approach as:

> […] a methodology that is rooted in strategy instruction. It sees the objective as being to teach students how to listen. This is done, first, by making learners aware of how the language functions and second, by making them aware of the strategies that they use – i.e., developing "metastrategic awareness." Then, the task of the teacher becomes to instruct the learners in the use of additional strategies that will assist them in tackling the listening task.

Attention is paid to the way learners use strategies to approach different listening tasks, to overcome their listening problems, and to learn how to listen. Table 4 summarises the findings of some studies that attempted to look at the effectiveness of strategy instruction on learners' LC, the results of which reveal varying degrees of significance. The table also shows the types of listening strategies that are believed to aid the process of comprehension. A detailed classification of listening strategies, suggested by Lynch and Mendelsohn (2010) is shown in Table 5. It divides listening strategies into three main categories: Cognitive, Metacognitive and social/affective strategies. A similar, but more detailed classification was provided by Vandergrift (1997b).

Study	Strategies	Participants	Method and Procedure	Results
O'malley, Chamot, Stewner-Manzanares and Kupper (1985)	– selective attention – note-taking – co-operation	70 High school ESL learners	Pre-training interviews Class observation guides	Varying results depending on tasks
		75 High school students	Strategy Training: Metacognitive group Cognitive group CG (50 minutes a day/8 days)	Strategy training likely to be effective for integrative tasks
Henner-Stanchina (1986-1987)	– Predicting/generating hypothesis using schematic knowledge – Inferencing – Integrating new information	ESL learners	End-of-term evaluation Questionnaire	Positive reactions to the training (direct and efficient way of attaining goals)
McGruddy (1995, cited in Cohen & Macaro, 2007)	– Selective attention – Inferring	EG (N=10) CG1 (N=10) CG2 (N=12)	-Standardised test -Researcher designed test -Strategy-use questionnaire (14 weeks of instruction)	Significant results, especially in the selective attention strategy, demonstrated by the standardised test scores

Thompson and Rubin (1996)	**-Metacognitive strategies:** planning, defining goals, monitoring, evaluating. **-Cognitive strategies:** predicting content, listening to the known (cognates, words, phrases), listening to tone and intonation, resourcing, special genre-related strategies (drama, interviews and news)	Intact Classes: EG (N=24) CG (N= 12)	**Measures:** -video and audio comprehension tests -Students' reports/ comments **Training:** -strategy training using videos	-Significant improvement in LC after the training (especially in the video measures) -Positive reactions form participants (e.g. reports, comments)
Guan (2014)	**Top-down Strategies:** inferencing, elaboration **Metacognitive Strategies** Directed/selective attention, **Complex strategies:** note taking, voice/paralinguistic inferencing	52 beginner-level ESL learners: -EG (N=30) -CG (N=22)	**Measures:** interviews, listening test, background surveys and classroom observations Explicit instruction on strategies	Positive Effects: -Reported improvement in listening -Statistically significant improvement

Table 4: Studies on the Effectiveness of Strategy Instruction in Listening Comprehension

Cognitive	Meta-cognitive	Social/affective
Predicting/ Inferencing • from the text • form the voice • from the body language • between discourse parts **Elaboration** • from personal experience • form world knowledge • from academic learning • form imagination **Contextualisation** **Imagery** **Summarisation** • mental • physical (notes) **Translation** **Repetition** **Transfer from other language (s)** **Deduction** **Fixation** • stopping to think about spelling • stopping to think about meaning • stopping to memorise	**Planning** • advance organisation • self-management **Comprehension monitoring** • confirming comprehension • identifying words not understood **Directed attention** • concentrating • persevering despite problems **Selective attention** • listening for familiar words • listening for the overall message • noticing the information structure • noticing repetition and reformulation • listening to specific parts **Evaluation** • checking interpretation against predictions • checking interpretation against knowledge • checking interpretation against context	**Questioning (two-way tasks)** • asking for clarification • asking for repetition • using comprehension check **Cooperation** • working with other learners **Anxiety reduction** • encouraging yourself • comparing yourself with others • focusing on success **Relaxation** • using physical techniques • using visualisation

Table 5: Listening Comprehension Strategies – From Lynch & Mendelsohn (2010, p. 187)

The main issues that dominated this approach to listening was how listening difficulties can be addressed by applying compensatory strategies (Cohen & Macaro, 2007),

and how learners can plan for, monitor and evaluate the process of comprehension (Vandergrift L. , 1999). In other words, both cognitive and metacognitive strategies have been focal points. This is of particular interest to learners whose linguistic knowledge is not sufficient to allow them to listen fluently to spoken language. It has been suggested that, through the application of TD strategies that rely on background knowledge, these learners are able to bypass the signal whenever language-related problems arise. Knowledge of the topic or the context can fill in the gaps in the learner's knowledge through TD processing which fosters the strategies of prediction and inferencing. It can also help the learners organise thoughts, retrieve the appropriate knowledge, and prepare themselves for the listening task (Vandergrift L. , 2012). Decoding problems arising from the weak knowledge of the sound system, the interference of L1 segmentation strategies or poor WR processes will be overcome by adopting a TD mode that starts from meaning and then moves to the input for confirmation.

Oxford (1990) divided the compensation strategies of guessing into two types: Using linguistic cues (e.g., affixes, word order, stress pattern, and first language knowledge), and using other cues which are related to, but go beyond, language (e.g., forms of address, tone of voice, facial expressions, body language, background noise, knowledge about participants, behaviours of people, and the general background knowledge). For Cohen (1991), successful learners resort to the following cues as a strategic choice in their search for meaning:

(1) the knowledge that they have of the world;

(2) their knowledge of the given topic;

(3) their knowledge of expected utterances for the given context;

(4) cues from prior utterances within discourse;

(5) cues from the stress that individual words have received;

(6) knowledge of the speaker; his/her tone of voice and body language. (pp. 111,112)

At different levels of proficiency, such cues are believed to "help learners to overcome knowledge limitations in all four skills" (Oxford, 1990, p. 90) .

In research studies as well as in in teaching frameworks, the way TD strategies have been operationalized was mainly through the use of advance organisers. These are activities designed before listening tasks to stimulate learners' thinking about the topic so as to allow them to call background information into play. They take different forms such as videos, pictures, short sentences, brief discussions, or anything that can activate text-related schemata.

There are studies which investigated the use of advance organisers as an aid to comprehension. This has been recognised as an attempt to see whether teachers can activate learners' TD strategies through pre-listening tasks (Cohen & Macaro, 2007). Chung (2002) found that the use of two types of advance organisers –vocabulary pre-teaching and question previewing –had significant effects on the performance of 188 Taiwanese students in listening. Those who had access to two types of advance organisers performed significantly higher than those who had access to only one type or to none. Herron, Hanley and Cole (1995) found that when a description was accompanied with a visual support as an advance organiser, the comprehension of learners significantly improved compared to the condition in which a description was given alone. They maintain that "extensive listening is facilitated by the richness of context that visual organisers provide" (Herron et al., p. 387). Similar results were also found for the significance of written words/sentences with pictures (Wilberschied & Berman, 2004) aural vocabulary and/or character presentation (Chung & Huang, 1998), and written textual versus contextual visual aid (Dixon, 1991).

The above results indicate the significance of encouraging learners to adopt a strategy of using their schematic knowledge both inside and outside the classroom alongside the information provided by the text. If learners' awareness is raised to this, and sufficient practice is provided about how to use this knowledge efficiently, the comprehension process will be improved.

Using advance organisers is closely related to metacognitive strategies which constitute another important dimension of the strategy-based approach to listening. In addition to their potential for encouraging the strategy of guessing, organisers raise awareness to the need for planning before listening. As a metacognitive strategy, planning involves "developing an awareness of what needs to be done to accomplish a listening task, developing an appropriate action plan and/or appropriate contingency plans to overcome difficulties that may interfere with successful completion of the task" (Vandergrift L. , 1997b, p. 392). It also includes advance organisation, selective/directed attention, and self-management (Vandergrift, 1997b).

Other metacognitive strategies include monitoring and evaluation. Monitoring refers to the strategies of "checking, verifying, or correcting one's comprehension or performance in the course of a listening task" (Vandergrift, 1997b, p. 392). It is used to revise and to evaluate ones' understanding by comparing different sources of information for a consistent match. For instance, monitoring involves that previously held propositions are verified, revised or rejected in case the unfolding text unveils contradictory information. Evaluation refers to "checking the outcomes of one's comprehension against an internal measure of completeness and accuracy" (Vandergrift, 1997, p. 392). It allows the learner a chance to think of the strategies he applied, how successful he was, and what problems interfered with comprehension.

While listening strategies exist in one's L1 use, their use may not be transferred automatically to target language use. That is why the objective of strategy-instruction is to insure their transfer. Vandergrift (1999) suggested a framework for explicit instruction on metacognitive strategies to raise learners' awareness of them. By relying on the traditional framework of pre-during-post listening, some tasks were suggested for raising awareness of the strategies of planning, monitoring and evaluating. As an instructional tool, pre-listening tasks have much to provide for the purpose of raising metacognitive awareness. In addition to helping learners to recall related content and formal schemata, they encourage them to listen purposefully and focus on important information and/or selected details by setting well defined objectives for listening. With all of this in a learner's mind, useful predictions can be made as anticipation to the possible content and organisation of the text. In terms of metacognitive strategies, this phase of listening is about planning, and the listeners preparing for the listening task are actually thinking strategically. The framework also includes guiding checklists suggested to make learners ready for listening tasks and to urge them to evaluate the set of strategies they select before and after listening. With group discussions of the strategies used by different learners, effective strategies can also be shared; this may help less successful learners to evaluate their own strategies, try new ones and incorporate them if they are effective.

The strength of a strategic approach lies in making the learner aware of the processes of listening and the strategies that overcome limitations in linguistic knowledge. While both TD and BU strategies have constituted an area for strategy elicitation, emphasis was greater on the former. According to Graham and Macaro (2008), one reason for this is attributed to the listening models such as Anderson's which fostered attention on TD strategies and their facilitating nature, such as inferencing (bridging inferences and elaborate inferences). This has spread the notion that a successful non-native listener is one who "skips over textual

difficulties by deploying compensatory schematic knowledge" (Graham & Macaro, 2008, p. 749). On the other hand, the basic linguistic knowledge and the skills which constitute the listening processes, especially the lower level ones, do not seem to be an area of concern for this approach. Because of this, strong criticism has been addressed to the approach.

2.4.2.3. Limitations of the Approach

One major limitation of the strategy-based approach to listening is that, by focusing on TD strategies and compensatory strategies, it turns away attention from the basic decoding skills required for making LS and WR. It is true that native speakers deploy such strategies whenever they encounter problems such as poor speech signal, special speaker accent, or background noise. In these cases, they are compelled to exploit sources other than the signal in order to understand messages. However, the strategies to which they resort must not be considered as an alternative to the basic decoding ability which constitutes part of the linguistic knowledge that any listener should possess. It is argued that if learners' listening skills are to be developed, attention should be placed on lower level skills in the first place.

Wilson (2003) notes that, unlike TD processing and listening strategies, "bottom-up approaches that focus on word recognition […] have been comparatively undervalued" (p. 335). Similarly, Norris (1995) noted that available listening materials almost ignored the phonological characteristics of speech. He criticised the practice in teaching listening to L2/FL listeners where learners are required to develop TD processing skills before that they are able to appropriately decode speech. From the part of listeners, having to follow this order is a very demanding task; it is just like "putting the cart before the horse" (Norris, 1995, p. 47). This is in addition to being a source of frustration and confusion to the learners. To put things in the right order, Norris suggested that learners should first focus on micro-skills to develop an ability that makes BU processing an easy task.

In view of the cognitive ability of the non-native listener, using strategies during listening adds more demands on his working memory. This is mainly because strategies are *conscious* operations which have to be performed simultaneously with other listening processes. The risk is that such additional attention reduces the limited working memory resources. Ridgway (2000) asserts that the act of listening requires the full engagement of the listener with the text in a way that does not leave place for the application of conscious strategies. He goes on to contend that strategy instruction turns focus away from the need to teach language skills. If this is true, it is necessary to question the time spent on teaching strategies to cope with lower level processing skills while ignoring the fact that they are supposed to be automatic. Renandya and Ferrell (2010) believe that, from a teacher's perspective, strategy instruction is unrealistic and burdensome in that it necessitates many requirements from the part of the teacher. They list five main requirements:

- [Teachers] need to know the theories and principles behind strategy training
- They need to know how to select the strategies that the students need to learn
- They need to know the order in which these strategies should be presented and practised
- They need to know how to integrate these strategies into the curriculum
- They need to know how much time should be allocated for strategy training (Renandya & Ferrell, 2010, p. 55)

With strategy instruction becoming a standard in language teaching, some arguably wrong assumptions have prevailed. One is that decoding problems that arise from weak perceptual ability are of lower importance and, hence, do not need special attention. (Renandya & Farrell, 2010). The belief that listening strategies will take over this poor decoding ability is the second assumption. Field (2003b) argued that the opposite is true,

noting that many higher level problems do originate in learners' inability to decode the signal. To verify predictions, what the learner needs is an ability to *perceive* the input accurately to check for consistencies/inconsistencies. Asking a beginner learner to verify hypotheses and make inferences out of a text which he cannot decode is beyond any logic. Even if he successfully activates relevant schematic knowledge, he still needs an ability to decode the text in order to make appropriate matches. According to Brown (1990),

> even if you do manage to develop a rich set of predictions you still need to be able to monitor the incoming acoustic signal so that you know which of your predictions is being confirmed and which is not. You need to be able to use what segmental cues there are and to recognise how they are likely to be distributed in the acoustic continuum […] you also need to be able to see the information provided by the regular saliencies of speech, which draws your attention to the bits of language which the speaker is treating as crucial to the message (pp. 11,12).

In the absence of a reliable decoding ability, even non-matching schemata will come into play and influence the interpretation of the message. Reporting on the results of studies that dealt with the effects of prior knowledge, Cohen and Macaro (2007) concluded that one condition for background knowledge to be appropriately used is that it should appropriately match the schemata present in the text. That is, incongruent background schemata can have distorting effects. As a second condition, learners' lexical knowledge should be beyond threshold level. This is, however, not always possible, especially for beginner learners.

The first of the above conditions has empirical grounds. To answer the question of which processing mode, BU or TD, is more important in listening, Tsui and Fullilove

(1998c) conducted a study based on the analysis L2 candidates' performances in large scale listening tests. The study included 150,000 multiple-choice items in which learners' answers to comprehension questions were analysed according to whether these answers were influenced by the activation of two schema types: 1) a non-matching schema suggested by the initial part of the text. A learner's correct answer will imply rejecting an initial wrong hypothesis activated by the non-matching schema, and revising it based on new evidence in the text using BU processing; 2) a matching schema whose activation is 'congruent' with the text and a learner's activation of it will help in comprehension. The results indicated a link between the non-matching schema and the wrong answers given by candidates. A high percentage of candidates were distracted by the choices activated by inappropriate non-matching schemata. On the other hand, those who obtained correct answers, and did not have such interference from the non-matching schemata, had better performance. The results suggested that "bottom-up processing was more important than TD processing in discriminating the listening performance of L2 learners in the test items" (Tsui & Fullilove, 1998, p. 432).

2.4.3. Bottom-up Primacy: Enhancing Decoding Skills

A different approach from the strategy based instruction is one which encourages the development of learners' BU decoding skills. This approach has emerged in part as a reaction to the overemphasis on higher level processes and TD strategies. In this respect, Wilson (2003) wrote that "we need to respond with practical classroom activities that shift the balance towards 'bottom-up primacy' " (p. 341). Indeed, a number of teaching frameworks have been proposed with the objective of enabling learners to improve an ability to deal with the physical characteristics of speech, including CS phenomena.

Increasing learners' automatic processing of the lower level aspects of speech is a main objective for the BU approach. Because of the demands of real time listening, it is argued that part of the listener's processes should take place automatically (Vandergrift & Goh, 2012; Vandergrift L. , 2013). Field (2008a) states that:

> One of the characteristics of an expert listener is the ability to process spoken input in a highly automatic way –one which does not make heavy demands upon the listener's attention. That kind of automaticity can be achieved only by extensive experience of actually using the skill (p. 32).

This is an obvious rejection of the assumption that strategies allow for basic skills to take care of themselves. It is through special attention and practice that skills can be developed and automatized. For Field, this example of the expert L1 listener who does not lend much attention to lower level processes needs to be transferred to a view of how L2/FL listening should develop.

There are advantages for the learner from developing lower level skills. When the learners are able to perceive CS with a certain level of accuracy, the problems they encounter with LS and WR will be reduced. If, otherwise, their perceptual skills are not efficient, they may get stuck at the lower level of perception and parsing, with few chances to concentrate on other aspects of the message (Goh, 2000). In addition, learners are believed to gain more confidence during listening if they possess the required knowledge and ability to perceive speech with ease. More importantly, when this decoding capacity becomes automatic through practice, it will allow the learner a spare attentional capacity to concentrate on higher levels of meaning and message interpretation (Ellis, 2003). After all, what learners need is a

skill which directly works on the input and facilitates its reception, not the use of contextual guesswork (Wilson, 2003).

In view of the studies on CS and the frameworks suggested for teaching these phenomena, the application of a BU approach appears to have two main principles. Because learners may not be familiar with the types of modifications that take place in natural speech, the first thing to consider in teaching is to raise their awareness of them (Norris, 1995). This takes place either with direct explanation of rules and modification types (Chenjun & Li, 2012) or through listening tasks that encourage learners to discover the problems created by the signal by themselves (Hulstijn, 2001; Wilson, 2003). By doing this, the learners will have declarative knowledge about CS features and some general rules about them. The second principle is to proceduralize this knowledge. In other words, learners need to acquire the knowledge *how*, or the "mechanisms which make language work in production and in comprehension" (Towell & Hawkins, 1994, p. 172). For Field (2008a) this is an important goal for decoding practice because it moves the learner "from processing that is controlled and demanding to processing where the recognition of words and chunks comes easily" (1p. 116). This takes place by providing sufficient time for practice. Extensive practice will eventually result in increased efficiency and more automaticity in decoding.

2.4.4. Practical Pedagogical Suggestions for Connected Speech Instruction

The Literature on the BU instruction is rich of many teaching frameworks. We will review some of these with the aim of exploring the range of tasks and techniques suggested in this respect. Peterson (2001) urges that BU processing should be practised at all levels of proficiency, with the condition that CS features are presented to learners after that they have mastered the segmental features of speech. Tasks suggested aim at helping intermediate-level learners notice the features by themselves through differentiating between content

words and function words, listening to short sentences/words, dictations, or even pronunciation tasks (e.g., pronouncing consonant clusters to discover which sounds are normally dropped).

Wilson (2003) believes that the first step to address learners' weaknesses is to enable them to *notice* their problems, and what caused them. He proposed a learner centred framework called *Discovery Listening* which encourages learners to focus their attention on words and sounds by reconstructing a text after listening to it. The learners' main task is to make group discussions divided into three phases:

1. Learners focus on what their precise listening problems are after reconstructing the text and comparing their answers to the original
2. They think of the causes of miscomprehension
3. They assess the importance of these causes (for learning).

When problems are self-diagnosed, the learners have what is called *final listening* to compare their level of understanding before and after locating the problems.

The validity of Discovery Listening lies in the fact that it addresses learners' needs directly. The elements to be learnt are selected by the learners themselves, rather than being imposed by a syllabus. Technically, this is referred to as "*Focus on Form"* as opposed to "*Focus on Forms"* (Harmer, 2007). The main advantage for focus on form is that the desire for learning comes after a breakdown in comprehension/ communication when the primary focus is on meaning (Doughty & Williams, 1998). Comprehension/ communication breakdown forces the learner to focus attention on what causes the problem, and gives a well-established reason for learning.

Norris (1995) is convinced that teachers' knowledge and awareness of phonological modifications is imperative in BU instruction. He suggests a similar framework to Discovery Listening which focuses on learners developing self-awareness of their decoding problems, and provides chances for practice both inside and outside the classroom. The framework includes a series of tasks ordered in a way that allows learners to build awareness of CS, to practise listening to it, and to focus their attention on the problems it creates in listening. The types of tasks include:

1- listening to reduced and citation forms to differentiate between them (awareness)
2- listening to sentences and writing them with the full forms of words
3- comparing answers to the original scripts
4- analysing examples of modifications by comparing spoken and written forms
5- filling in blanks with full forms based on spoken texts (Cloze task)

At advanced levels, learners can perform tasks in which they have to interact with each other by asking/answering questions which require the recognition of the CS features. To create a relaxed learning atmosphere, group work games and competitions which motivate learners to practise listening to reduced forms were suggested.

In addition to the BU tasks, the framework also focuses on learners' self-evaluation of their decoding problems by urging them to keep listening journals to report on their experience with listening to English. Exposure to spoken English outside the classroom allows for more practice, and may encourage learners to build awareness of their listening skills. For the teacher, this type of learners' reported experience provides insights into their problems and processes, and allows for an informed feedback.

For Renandya and Farrell (2010), learners can be helped with the features of spoken English such as fast speech, the blurring of word boundaries and other phonological modifications through listening to the target language in an enjoyable environment. They suggest Extensive Listening as an instructional method where learners listen to comprehensible input. They suggest a set of "simple and enjoyable listening activities" (p. 7) which provides the learners with much listening-practice time. The tasks include dictation, repeated listening, reading and listening simultaneously and narrow listening which are evaluated as being "at least as good as, or may even be superior to, strategy-based teaching" (Renandya & Ferrell, 2010, p. 7). For the authors, devoting much time to listening is an integral part of a continuous process to proceduralize learners' declarative knowledge about features of spoken language.

Some methodologists urge for a balanced approach that comprises work on higher level skills and TD strategies in addition to a focus on lower-level BU skills. For instance, Hulstijn (2001) believes that the two trends should be regarded as complementary rather than mutually exclusive. To develop automaticity in WR, which is deemed the most important process in bottom up listening (Rost & Wilson, 2013), he proposes that learners should listen to *i* minus-1 input which includes lexical items that are all familiar to the learner. After exploring the elements and dealing with comprehension questions at one lesson, subsequent lessons should include texts which allow for repeated encounters with the elements explored. The teacher's role is to direct learners' attention to the features of speech. This way, learners can notice how the pronunciation forms of familiar words change in different environments, and how spoken and written versions differ. According to Vandergrift (2007), "such an approach can also call attention to other phenomena in connected speech such as reduced forms, assimilation, elision, resyllabification and cliticization" (p. 179). Other tasks include listening and reading scripts at the same time to detect deviations, playing parts of the text

and checking recognition, listening to the text for several times, and listening and counting the number of words in a chunk.

Field (2008a) also supports the integration of BU skills and TD strategies. However, he stresses that once decoding skills are developed, compensatory strategies will lose their reason to exist, and their use will normally be limited to the cases when problems arise. The *signal-based approach* (2003b) and *the diagnostic approach* (2008a) which Field proposed have the aim of addressing the weaknesses of the Comprehension Approach, especially the point that the latter emphasises the product rather than the process of listening. Field (ibid) notes that learners are evaluated on the basis of the correct answers they provide, without questioning the reason why they arrived at right or wrong answers. If the teacher does not explore the reasons behind this, he cannot help learners with their problems. The approaches, then, are based on diagnosing learners' problems and providing appropriate remedy.

According to Field (2008a), teachers can integrate the higher and lower level skills and strategies by adopting the pre-during-post listening framework. Diagnosing learners' problems takes place during the listening phase by analysing their answers. The justifications learners provide to their answers to comprehension questions and/or the group discussions they have are the main diagnosis tools at the teacher's hands. In addition, teachers should use their knowledge about CS features to anticipate the types of problems that may cause decoding obstacles and eventually lead to comprehension breakdown. Through time, special portfolios can be kept to note down the common types of decoding problems learners usually face. When problems are located and diagnosed, an extended post-listening phase will be appropriate to provide the necessary remedy. The following are types of remedial micro-listening exercises suggested by Field (2003; 2008a) which focus on particular skills and features:

1- listening with a focus on stressed syllables as keys for segmentation (especially if the learner's mother tongue does not have a stress-timed rhythm)
2- 5-minute dictation sessions: dictating sentences where learners write full forms
3- repeating parts of the input that cause recognition/segmentation problems
4- gap filling tasks; writing citation forms
5- *final play;* listening to the text after dealing with comprehension questions and reading the transcript with a focus on difficult parts
6- Listening to the text and focusing on the pronunciation of WF words
7- *revising segmentation decisions*: Dictating parts of ambiguous sentences , and urging learners to check segmentation by unfolding subsequent parts
8- dictating neutralised contractions/WFs and encouraging the use of syntactic and grammatical cues (e.g., it's been raining/ it's raining).
9- directing learners' attention to words frequently modified, such as formulaic phrases (e.g., didn't, /dɪnt/; more and more /mɔːmɔː/; do you know what I mean /nɑːpmiːm/ (field, 2003, p. 331).
10- Sentence transcription task

Conclusion

While CS features play an integral role in maintaining the rhythmic pattern of spoken English, and in facilitating the articulatory effort of the speaker, they present serious obstacles to decoding spoken language and interfere with the process of message comprehension. Due to their physical characteristics, CS features are not only difficult to perceive from the part of learners, but they also make the processes of WR and LS laborious tasks. This is in addition to being a barrier to language learning in general as they may filter the input-intake process. Due to this, it is deemed necessary that special attention should be

paid to these aspects from instructors to improve learners' listening skills and to relieve some of the burden they have with spoken English. We have shed light on two different teaching trends; the first one focuses on TD compensatory strategies and higher level processes and the second one directly focuses on BU decoding skills. To address learners' weaknesses in listening, it is argued that an approach that integrates the two trends is necessary.

Chapter Three

Methodology

Chapter Three
Methodology

Introduction

The objective of this study is to look at the problems learners face in decoding CS and to find out whether specific instruction on CS would have a significant effect on the participants' decoding skills. This chapter is meant for describing the research design followed in our investigation. To this end, the experimental design of the study and the participants are first presented. Then, a description of the tests used and the way they were administered is provided. In a following section, a detailed account of the treatment condition is given with a description of the lesson framework followed, and the main types of tasks used. The final sections are reserved for a description of the post-instruction interviews, followed by an explanation of the qualitative and the quantitative data analyses procedures.

3.1. Restatement of the Objectives

The objective of this study is to investigate the learners' LC ability by focusing on the lower level decoding skills. The literature review revealed that CS phenomena pose serious problems for untrained learners in decoding naturally spoken English, and that there is a need for addressing these problems through instruction. The studies that took this perspective, however, did not provide sufficient information about how such a focus on developing decoding skills can be applied and adopted in a balanced way that comprises most of the aspects of the listening process. Most of the related studies focused on short lessons ranging between 5 and 30 minutes, and provided practice tasks, direct explanations of rules or open lists of reduced forms. The way these CS features can be integrated into LC

lessons needs to be investigated. This study takes a more balanced perspective by integrating CS instruction into teaching LC that targets decoding skills, without compromising the need for developing top-level skills.

For teaching to be effective, it is deemed necessary to target those aspects which really hinder comprehension. This is why the study, as a first objective, looks into the obstacles created by the CS features by analysing learners' performance in a CS listening test. Both quantitative and qualitative data are obtained by using a mixed approach. After analysing the participants' problems, we provide instruction on these sound phenomena in order to find out the effectiveness of a remedial approach to improve the learners' decoding skills. To this end, participants are pre-tested and post-tested. As the learners' reactions to the instruction are also deemed necessary, post-intervention interviews are carried out to explore the learners' evaluations of both the instruction and their own LC skill after the instruction.

3.2. Quasi-Experimental Design

In order to test the effectiveness of the instruction on the learners' performances, a quasi-experimental design is followed. The difference between this design and the true experimental design lies in the ability to have full control and manipulation over the experiment's conditions (Broota, 1989). In the absence of ideal conditions which may not be available, it is worth considering a quasi-experimental design (Campbell & Stanley, 1963). For this study, randomisation in assigning the participants to groups is not within our control.

This study includes both an EG and a CG. At the outset of the investigation, a pre-test is administered to the two groups, and their results are compared. The learners' answers

are subject to a qualitative analysis which aims at depicting the likely effects of the CS features on the recognition and segmentation of spoken language. The learners' mishearings are also analysed to find out which processing mode the participants rely on most. Then, through a period of 8 weeks, the EG receives specific instruction on CS, while the CG does not. During the instruction, classroom observations are also taken into account to analyse the performance of the participants and their reactions to the different types of the tasks designed. After the instruction, a post-test is administered to the two groups. Their scores are compared to verify whether the lessons the EG receives have any significant effect on their performance compared to the pre-test, and also in comparison to the post-test scores of the CG. Another measure of analysing the lessons' significance is through semi-structured interviews carried out after the instruction.

3.2.1.Participants and Settings

The study took place at the Department of English Language and Literature, University of Mohamed Boudiaf -M'sila. Participants in this study were 38 second year students of English. They were members of either a CG (N=19) or an EG (N=19). As they were members of intact classes, randomness in assigning them was not feasible. The experiment took place during the second semester of the academic year, and the students had already been grouped. It is worth mentioning that grouping the learners was not done under any measure of academic achievement, but rather through mere administrative procedures. Other variables were taken into consideration such as the regional backgrounds or the students' own choice in some cases. The choice of having volunteers as participants was avoided due to the relatively long period of the instruction and its likely effect on the commitment and attendance of the participants. On the other hand, having intact classes at least guaranteed that the lessons would be easily integrated into the lessons the EG received in the Oral Expression Module. All the lessons were delivered in the language *practitiatory*

–in contrast to *laboratory* – as the main objective was to provide practice rather than to test the participants (Labed, 2001).

The reason for selecting second year learners was that they had already dealt with segmental features of speech in their first year. They were supposed to possess the basic knowledge about the sound system of English which should have made subsequent work on suprasegmentals easier and logically ordered. In addition, according to the designed syllabus, it was in the second year that they were supposed to deal with the suprasegmental aspects of speech in the module of Phonetics.

3.3. Procedures

In what follows, we will provide a description of the data-collection procedures, the way the tests were administered, the treatment that the two groups received, and the way the interviews were conducted.

3.3.1. The Pre-test

3.3.1.1. Description of the Pre-test

The objective of the pre-test was two-fold. In addition to being a measure of the level of CS comprehension for the experimental and the control groups before the treatment, it also provided qualitative data about the problematic features and the reasons for their occurrence. The test was composed of two main types of tasks. A listening cloze task comprised of 34 sentences was administered to measure the learners' ability to perceive function words. Each sentence included one function word that was blanked out in addition to the content word that precedes or follows it. The sentences were well selected so that each function word occurred at least twice in the test but in different phonological environments.

It should be mentioned that, although the function words were the main items omitted in the sentences, the test was also a measure for the perception of the other CS features as these words undergo other types of reduction through elision, assimilation, liaison, and complex reduction processes (Table 6). Because the task was designed to test the participants' ability to *decode* the signal, the sentences selected were relatively short, and they did not have any topic relation. The reason for this was mainly to avoid that test-takers use TD processing to compensate for/guess the missing items (although syntactic and semantic information at the sentence level could have cued this).

The second part of the test included a dictation task. Twenty one sentences were selected to test the learners' ability to decode speech in the presence of the following CS features: Assimilation, elision, juncture and liaison. Here also, the sentences were relatively short, they did not have any topic relation, and each one included at least one CS feature. The two tasks used in this test are widely used in teaching as well as in testing CS (Brown & Kondo-Brown, 2006). It should be noted that all the sentences selected in these tasks together with their audio versions were mainly adopted from commercial materials; these were listening and pronunciation training materials which included, among others, tasks on the CS features. The main sources were Harmer and Lethaby (2005), Bowler and Cunningham (2003), and Roach (2009).

Sample Test Items	Pronunciation		Modification Features Involved
	Citation Form	Reduced Form	
We went **to her** room	/tʊ hɜː rʊm/	/tʊʷərʊm	*Complex process*: elision + linking with /ʷ/+vowel reduction
We **wanted her to** stay	/wɒntɪd hɜː tʊ/	/wɒntɪdətə/	vowel reduction + elision
You must come	/jʊ mʌstkʌm/	/jʊməskʌm/	
So **that you** can talk to the person you are with and you really get to know her	/ðæt jʊ/	ðətʃjʊ/	Vowel reduction + coalescent assimilation
If they had searched more carefully, they might **have found** the jewels.	/hæv faʊnd/	/ əf faʊnd/	Consonant Elision+ Vowel reduction + consonant assimilation
There is some soup in the fridge	/ðeər ɪz sʌm/	/ðərɪs səm/	Vowel reduction + regressive consonant assimilation
…rather **than buying** a digital file	/ðæn baɪɪŋ	/ðəmbaɪɪŋ/	Reduction + regressive assimilation
…**send them** anything	/send ðem/	/sen̪ðəm/ (dental n)	Complex process: Consonant Elision+ assimilation
…sounds absolutely terrible **to us** today	/tʊ ʌs/	/tʊʷəs/	Linking with /ʷ/+vowel reduction
Everyone knows me as Hag, apart from a couple of annoying people, **as simple** as that	/æz sɪmpl/	/əs sɪmpl/	Vowel reduction+ regressive assimilation
Aren't there some letters **for her** to open	/fɔː hɜː/	/fər ə/	Vowel reduction + elision+ Linking with /r/

Much teaching and learning **that goes** on today	/ðæt gəʊz/	/ðək gəʊz/	Reduction + assimilation
He wants to come **and see us at** home	/ænd siː ʌs/	/ənsiːʲəs/	Elision+Linking with /ʲ/+ vowel reduction
When are you taking him **to see her**	/tʊ siː hɜː/	/təsiːʲə/	*Complex process*: consonant elision + linking with /ʲ/+vowel reduction

Table 6: Samples of Multiple Phonological Modifications in the Cloze Task Items

3.3.1.2. Pre-test Administering

The pre-test was administered to both groups in a language laboratory. The laboratory was equipped with 20 computers, each with good quality headphones. There was a console which allowed that each sentence could be played to all testees at the same time. Before administering the test, the headphones were tested for their sound quality and volume. Prior to the test, the dictation and the cloze test sentences were cut and grouped into single files using the audio software *Audacity*, version 2.0.5 (Audacity-Team, 2013).

For the cloze test, each participant was given a sheet which included the corresponding sentences (with blanks). They were told that the objective was to test their listening skill, and they were asked to write the numbers of the computers where they sat, not their names. The instruction was explained first; the participants were instructed to listen to the sentences and fill in the blanks in the corresponding sheets on the basis of what they listened to (Table 7). It was explained clearly that they were supposed to write the full forms of the words (e.g., they should write *she is*, not she's). For each sentence, the testees listened first to its number, which signalled that it would be played. Each sentence was played once, and the testees were given sufficient time to fill in the blanks. When there was more than one blank in a single sentence, they were allowed to listen twice. As for the dictation test,

almost the same procedure was followed. The testees were given sheets with numbers from 1 to 21 corresponding to the sentences in the test. They were instructed to listen to each sentence twice and write it down on the paper. The whole test took about 35-40 minutes for each group.

Example items in the test-sheets	What testees heard	What testees should have written
1. The computer which we bought ----------.	The computer which we bought *was stolen*. /wəs stəʊlən/. (*reduction + assimilation*)	The computer which we bought **was stolen.**
2. He wants ----------home.	He wants to come and see us at home. /təkʌmənsiːjəsət/ (*reduction + elision+ liaison*)	He wants **to come and see us at** home.

Table 7: Sample Items form the Cloze Task in the Pre-test

3.3.2. Treatment

3.3.2.1. Experimental Group

The lessons the EG received were delivered in well-equipped language practitiatory supported with high audio and video technology. The room includes 20 computers with headphones connected to a console that the teacher can control. The audio/video files can be played by the teacher so that all the participants can listen and watch in their computers simultaneously. The console also allows for controlling the microphones; it allows the teacher to speak while the learners listen. It also allows one learner to speak at a time or many learners speak to each other (in small groups/pairs). These features were very useful as they met the requirements of the tasks during the listening and the post-listening phases

of the lessons. The room also included an interactive board which allowed for projecting the tasks with an electronic version to the whole class. As for the sessions, the class met twice a week for a total time of 3 hours.

In the first session after the test, we gave the members of the EG an explanation of the objectives of the instruction. They were told that the aim was to develop their listening skills through practice tasks. Then, a brief explanation of the TD and the BU listening modes was given. This was done through examples and figures using Power-point slides (e.g., Figure 8). The second session was designed with the purpose of motivating the members of the group. Because we were new to them, since we substituted their usual teacher, it was important to make an ice-breaker to insure the learners that the instruction would meet their needs and to make them feel at ease. To this end, the first lesson included a funny short story with listening tasks from Harmer (2005). Through the different phases of the lesson, we observed that the learners were active and motivated to participate. They even expressed their appreciation of the lesson after the session had finished.

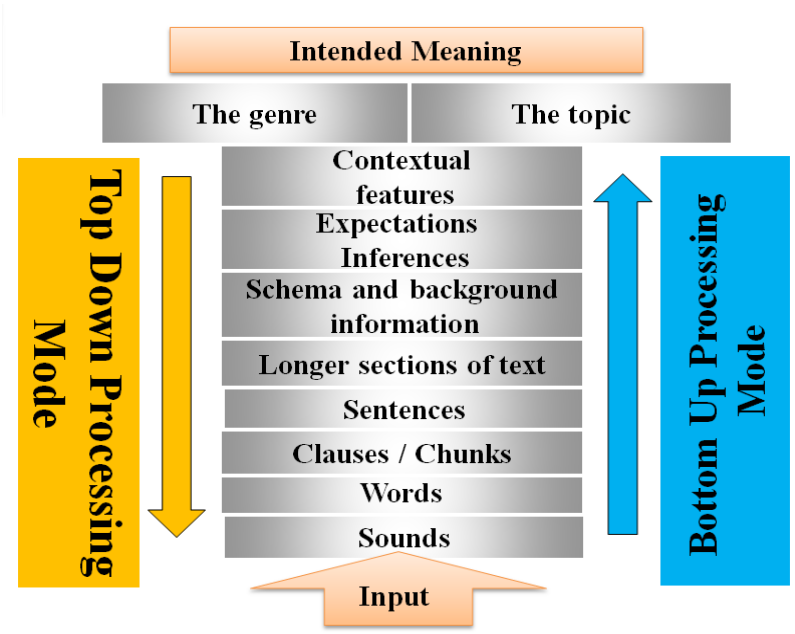

Figure 8: Difference between the Top-down and the Bottom-up Listening Modes (Buck, 2001; Rost, 2001; 2011; Lynch, 2009)

All subsequent sessions were designed following the pre-during-post framework. The framework was used for its suitability to meet the requirements of the instruction, especially that it allowed for integrating the decoding instruction into the LC lessons the learners had already been taking. It should be noted that the learners were used to this framework before the instruction. However, the tasks that targeted their decoding skills in the post-listening phase were new to the learners (as they had been used to different speaking tasks).

The tasks in the pre-listening phase had the objective of activating the learners' background knowledge about the topic of the text to which they would listen. Each lesson included one type of advance organisers in the form of pictures, discussion questions, vocabulary teaching, listening to short descriptions and others. Some tasks involved making predictions about certain information or about people, which the learners had to confirm/reject during the listening phase. These tasks encouraged them to deploy the information they had about the topics to activate their TD processing. The listening-phase, which will be explained in detail in the following sub-section, included tasks which focused on listening for gist or listening for detail. It also had the objective of spotting the learners' comprehension problems which originated in the failure to recognise CS features. The last phase, which was the main focus of the instruction, included small-scale listening tasks that aimed at addressing the problematic features identified, raising the learners' awareness to them and providing practice.

Before starting with the treatment, the participants' answers in the pre-test were analysed and scored. The analysis was done for two main objectives; 1) to diagnose their common decoding problems to get an idea of the possible reasons for their occurrence, and 2) to specify which of the features and their sub-categories caused more decoding problems.

Such information provided the basis for a well-informed remedial programme which addressed only the features which had been found to pose barriers to decoding. This is one of the main issues this study focused on. As the literature review suggested, previous studies have concentrated on providing short lessons on open lists of CS features, or direct explanations of rules without taking into consideration the learners' needs. By doing so, the features which actually hinder listening may not be well covered. In addition, time could be wasted in teaching a range of non-problematic features at the expense of those which require more attention and practice. By having a background of the learners' problems and needs, special tasks and listening scripts can be selected to develop their listening skills.

The lessons that constituted the listening-phase of the lessons were taken from Harmer and Elsworth (1989), Harmer (2004), Harmer and Lethaby (2005), and Dunkel and Lim (2006). They included a variety of texts (e.g., radio interviews, narrative texts, announcements, short lectures, and short stories), some of which were authentic. When the listening texts included specific features which had been found troublesome for the learners (in the pre-test), special tasks were designed before the listening sessions to address them. In other words, the problems these features may cause had been anticipated. For instance, almost all types of function words were found to cause mishearings. So, the frequency of words like *had, would, was* and *were,* especially in narrative texts, was very likely to cause recognition and LS difficulties. So, special tasks and questions needed to be designed to address them beforehand. However, anticipating problematic-features was not the only basis for the lesson design and the choice of the tasks. The diagnosis of the learners' decoding obstacles was also done in the during-listening phase. Field (2008a) argued that listening sessions should be regarded as mainly diagnostic, where the teacher should look into what caused miscomprehension. This implies considering the *process* rather than the *product* of listening. As an application of this principle to the instruction, special tasks and

comprehension questions were selected/designed for the during-listening phase, the answer of which required the accurate perception of the CS phenomena. The learners' answers and justifications to these questions can cue the reasons for miscomprehension.

Another technique implied re-playing portions of the text that were miscomprehended (as revealed by the wrong answers to the comprehension questions) and discussing /reconstructing them with the whole group afterwards. This was supposed to be more reliable in revealing the specific parts that were misperceived and the reasons behind them as CS features were part of the barriers to learners' comprehension.

Small-scale tasks were given to address the problematic features in the post-listening phase. A number of tasks were designed especially to raise learners' awareness to the features. Others encouraged them to practise listening to the features in order to allow for more automaticity in decoding. For WFs, a list of common reduced forms was provided to give the learners a general idea about them and to serve as a reference whenever a certain task dealt with a given form. Many of the materials for these tasks were taken from the listening texts that the EG had in each lesson. Using the audio software *Audacity*, the sentences from the listening texts were extracted; they served as the main items for the tasks. This way, the learners can be motivated to do the tasks as the materials that constituted them were taken from the very parts which they dealt with, and which caused their decoding failure. Table 8 summarises the main types of tasks which constituted the post-listening phase. To provide more practice, some tasks were borrowed from commercially available books such as Weinstein (2001), Bowler & Cunningham (2003), Hancock (2003) Cunningham & Bowler (2003), Hewings (2004), and Gilbert (2005).

Task	Instruction	Objectives/focus	Examples
Dictation	-Write the sentences that you hear using the full forms/citation forms	-To raise learners' awareness -To consolidate and proceduralize the perception of the aspects (automaticity)	*Do I* have a letter? (linking) He told me *that he had* been ill. (WFs)
Dictation of ambiguous sentences	-Listen to the parts of the sentences dictated and write the correct words	-to provide practice in speech segmentation /raise awareness to the possibility of making more than one segmentation and revising it based on new evidence	*I scream* all day/ *Ice cream* is my favourite… (more than one candidate)
Cloze tasks	-Listen and fill in the blanks in the sentences		
Multiple-choice questions 1	-Listen and choose the transcription that corresponds to the way the underlined words are pronounced	-To draw learners' attention to specific pronunciations/forms -Automaticity in processing	-She <u>*was* suffocating</u> from lack of oxygen (WF/assimilation) -<u>*Did you go* and see her?</u> French (assimilation)

Task	Instruction	Objectives/focus	Examples
Multiple-choice questions 2	Listen to the questions and choose the right answer	To enable learners to differentiate between homonymous function words -To encourage learners to use syntactic /grammatical information to decide between candidates	It's been here (= yes, it has) It's a good idea (homonymous contracted forms) (= yes, it is) -*She is* scared / *She has* seen the letter.
Listen and Repeat	-Listen to the sentences/passage focusing on the pronunciation of X/Y -Listen and write the sentences. Then, read them as they were pronounced.	-Awareness raising -Practice & automaticity	*There are some* new ones *he's* brought
Noticing tasks	-Listen and notice how the words change /are reduced -Consider the examples and draw the rules	-To show different reductions of function words -To highlight systematic patterns of reduction (rules)	-I *have to* go/ I *have* a meeting (assimilation) They were /wər/ absent / they were /wə/ here (WFs)

Table 8: Types of Tasks Used in the Post-listening Phase

3.3.2.2. Control Group

The participants in the CG received no specific instruction on CS. They continued to have their listening lessons in the same language practitiatory as the EG following the pre-during-post framework. For them, the post-listening phase included tasks that developed their speaking skills. However, it should be mentioned that both the EG and the CG dealt with CS features at the same time of the instruction in the module of Phonetics. This took place in an ordinary room, and included direct explanations of rules and types which targeted pronunciation skills. As our main interest in the instruction was to improve *listening* skills and to increase procedural knowledge through extended practice, the lessons of the Phonetics module were unlikely to have a significant effect on the performance of the participants in the listening tests.

3.3.3. The Post-Test

3.3.3.1. Description of the Post-test

To guarantee a similar level of difficulty, the post-test was similar to the pre-test in both the form and the types of features tested. It included a listening cloze task and a dictation task. The cloze task contained 36 sentences to test the participants' ability to perceive CS in addition to function words. The same criteria for selection were followed. Each function word occurred at least twice but in different phonological environments. In the answer sheets, these were blanked out systematically with the word that occurred before or after each one. As for the dictation task, 18 sentences were selected to test the learners' ability to recognise other aspects of CS; these were assimilation, elision, linking, juncture and WFs. For both tasks, the sentences selected were short and had no topic relation.

3.3.3.2. Post-test Administering

The same procedure as the pre-test was followed. The EG and the CG were tested in the language laboratory in which they were given two sheets; one included sentences with blanked parts for the cloze task, and one with numbers for the dictation task. The participants were asked to listen and fill in the blanks for the first task, and write what they heard in the second. After playing each sentence, they were given sufficient time to write down their answers.

3.3.4. Post-instruction Interviews

After the treatment, the members of the EG were interviewed about their perspectives regarding their experience in developing their decoding skills through CS instruction. To this end, semi-structured interviews were conducted. The main reason for choosing this tool was that it allowed us some freedom and flexibility in collecting richer data and to compare the participants' responses (Zacharias, 2012). Unlike the structured interviews, semi-structured interviews start with a number of pre-determined questions that turn around a selected topic, but new questions may arise as the discussion goes on (Lodico, Spaulding, & Voegtle, 2010). That is, in the course of the discussion certain unexpected information may be provided by the interviewee; the researcher might evaluate such information as significant, and more details may be sought. So, new questions which had not been included could be added. This was very reliable in collecting information about the participants, about the lessons from the point of view of the learners and about their reactions to it. The interviews were conducted in the language laboratory during the afternoon and the interviewees were allowed to use their mother tongue if they wanted. Each interview lasted for about 10 minutes.

The interview included the following main themes. First, we tried to get an idea about the participants' knowledge concerning CS phenomena before the instruction. We were interested in what they knew exactly about CS features, if they had any information at all. In case they did, we wanted to know the source of that knowledge, i.e., whether it was a personal effort, some lessons they had received, or some other sources. In the case of the lessons, we wanted to know whether the focus was on listening or speaking (pronunciation), and in what way they had benefited from them. Another main objective was to know how they reacted to the lessons in the instruction, and how they evaluated them. The interviewees were asked to speak about their listening problems in relation to CS; they were asked to specify the features or the types which they thought were most problematic, and which ones needed more focus in learning/teaching. They were free to express their opinions and to determine any advantages or shortcomings they had noticed about the lessons. In addition, they were asked to evaluate their listening skill before and after the instruction, and whether they felt any difference after the lessons. As a last point, they were asked if they had any intention to continue working on CS features and why.

3.4. Data Analysis Procedures

To answer the research questions, both qualitative and quantitative data were gathered and analysed. To analyse the data form the pre-test, the answers of the CG and the EG were all typed into a Windows Excel template that was designed for this purpose. The sentences were organised in two main ways. In one option, the answers of each individual participant could be displayed under each other. This helped in the scoring process of individual testees. Another option was used to display all the participants' answers to a particular sentence, which could appear under each other. This way, the answers to particular items could be easily compared and analysed at the same time, and qualitative data could be

gathered. The following sub-sections provide an account of the analysis procedures for the different types of data.

3.4.1. Quantitative Data Analysis

The scoring process of the two tasks in the pre-test was done separately and, then, the total scores were calculated. For the cloze task, the total score was 80; one mark for each item. This is an easy and recommended scoring procedure for this type of tasks (Buck, 2001). For unfilled blanks, it was assumed that the test-taker was unable to decode the words and, hence, no score was given. While the blanks included both function words and content word, only the function words were evaluated. Answers which included minor spelling mistakes that did not change the word forms were not rejected. For the dictation test, a similar scoring process was followed. Only the words which underwent CS modifications were considered in the scoring process.

When the total scores were obtained, a series of quantitative data analysis procedures were conducted using the Statistical Package for Social Sciences (SPSS). The data from the two groups were first entered and compared. In order to test the hypotheses that CS features led to decoding problems for the participants at the outset of the study, descriptive statistics of the pre-test scores of the two groups were generated. Then, the mean scores of the two groups were compared using an independent samples *t*-test to make sure there was no statistically significant difference between the two groups before the treatment. This was done after the data were analysed for their normality of distribution. The same procedure was followed in analysing the results of the post-test. In order to test the second hypothesis in the study and to find out if there was a statistically significant difference between the mean scores of the two groups after the experimental condition, an independent samples *t*-test was conducted. In addition, the scores of each group in the pre-test and the post-test were

compared to find out if there was a significant improvement in the gain scores of the two groups as an effect of the treatment. This was done using a paired samples *t*-test.

3.4.1. Qualitative Data Analysis

The qualitative data analysis of the pre-test included the analysis of the participants' mishearings. The answers they provided for the two tasks were analysed for two main objectives:

- to gather information about the common mishearings, their error patterns, and the possible causes for them in relation to the processing modes.
- to examine the effects that phonological modifications had on the test-takers' ability to segment speech and to recognise words. We were mainly concerned with the cases in which the testees were unable to recover from the effects of the modifications.

As for the post-treatment interviews, the participants' answers were first analysed and coded. Then, the prevalent themes were grouped and categorised.

Conclusion

This chapter has provided details about the methodology followed in testing the hypotheses of the study. The research tools we selected for approaching the questions of the study would allow for gathering both qualitative and quantitative data. The Analysis of the qualitative data would help in depicting the common problems that the CS features pose for the learners, and in categorising the patterns of misperception so as to diagnose these problems. It would also provide insights into the participant's own opinions about, and reactions to, the instruction they receive. The quantitative data would serve in measuring the

overall abilities of the participants in decoding the CS features and, through inferential statistics, their performances would be easily compared to find out the effects of the instruction on the treatment group's listening abilities. The following chapter will provide the results of the data analysis and the discussion of the findings.

Chapter Four

Learners' Decoding Ability and the Effects of Connected Speech Integrated Listening Instruction

Chapter Four
Learners' Decoding Ability and the Effects of Connected Speech Integrated Listening Instruction

Introduction

The purpose of this chapter is to present the results of the EG and the CG regarding their performance in the tests in the form of descriptive and inferential quantitative data. The aim is two-fold. Firstly, the pre-test results will be used to depict the participants' ability to decode the CS aspects. The total score of each participant together with the mean scores of the two groups will be presented and analysed. These results will also serve the objective of comparing the levels of the two groups before the treatment to find out if any inherent differences exist between them. To this end, descriptive and inferential statistical methods will be deployed. Secondly, the post-test results will be presented, analysed and compared to find out if any significant improvement took place after the treatment the EG received compared to the CG. Similarly, descriptive and inferential statistical methods will also be used to depict the levels of the two groups, and to compare their scores to one another and to those obtained in the pre-test.

4.1. Restatement of the Research Questions and Hypotheses

Let us not forget that the main focus in this study is on the learners' listening problems that may originate in the aspects of speech, particularly CS phenomena. The objective is to explore the problems these phenomena pose in listening and to investigate the effectiveness of instruction in addressing them. The literature review has revealed two different perspectives as to how learners' decoding skills should be developed; one focuses on developing higher level skills and TD strategies, and the other on the basic decoding

skills. Following the view that these trends may be regarded as complementary, we attempt to investigate whether the learners' ability to decode CS could be developed through the principles of a diagnostic approach (Field, 2003, 2008a) that allows for integrating CS instruction into LC lessons by focusing on both higher and lower level skills. The research questions are the following:

- Do second year learners in this study have problems in listening to CS phenomena?
- Do CS phenomena pose problems for the participants in this study in recognising words and segmenting speech?
- What are the effects of the different phenomena on the learners' decoding of speech?
- Is CS instruction through the application of the principles of a diagnostic approach to decoding effective in improving the learners' decoding of speech?
- What would be the students' reactions to the CS integrated LC lessons?

Accordingly, the questions correspond to the research tools selected and discussed in the previous chapter. As for the first question, it examines the possible obstacles that CS features may create regarding the BU decoding skills of WR and LS discussed in chapter one. Because of their potential for changing sounds and word forms, these phenomena may render lexical access and LS problematic for untrained learners. Quantitative analysis of the pre-test scores of the EG and the CG serves the objective of answering this question; the total scores of the participants were calculated together with the means. As for the second and third questions, the pre-test data were also subject to qualitative analyses to find out what effects the phonological modifications may have on the students' perception of words in speech. The patterns of mishearings were grouped in accordance with the feature involved

and interpreted in terms of the processing modes followed. The answering of the fourth question has a relation with the experiment and the treatment that the EG received. It required the comparison of the mean scores obtained by the two groups in the pre-test and the post-test. Finally, the learners' reactions to the experiment were explored through the analysis of the semi-structured interviews.

Based on the questions of the study, three main hypotheses were formulated so as to be verified after the data analysis. The three main hypotheses are:

- CS aspects would pose problems for students in decoding naturally spoken English and in making LS.
- In addition to the poor BU skills in decoding CS features, the TD processing mode would have negative effects on the participants' recognition and segmentation of CS.
- If students received CS instruction that focuses on awareness raising and provides practice, following the principles of the diagnostic approach, their ability to decode CS phenomena would improve.

As for the students' evaluation of the lessons, we hypothesised that the learners would have positive reactions to it.

4.2. Results of the Pre-Test

In this section, the scores obtained by the two groups in the pre-test will be presented together with the descriptive statistics.

4.2.1. Results of the Experimental Group

The test scores of the EG and the CG were calculated for each participant. Table 9 presents the scores obtained by the members of the EG in the pre-test. As the table shows, the scores of the cloze test and the dictation test were first treated separately and then calculated to get the total score which reflects the participant's ability to recognise the

different types of reduced forms. It should be recalled that the total scores of the cloze and the dictation tests were 80 and 21, respectively. Although both tasks included the different types of reduced forms, it is noticeable from the results that the test-takers obtained relatively lower scores in the dictation test compared to the cloze test.

Students	Pre-Test		Total Score
	Cloze Text	Dictation Test	
S1	8	3	**11**
S2	11	5	**16**
S3	40	9	**49**
S4	28	7	**35**
S5	14	3	**17**
S6	13	5	**18**
S7	23	2	**25**
S8	30	3	**33**
S9	29	7	**36**
S10	16	5	**21**
S11	25	14	**39**
S12	18	3	**21**
S13	8	5	**13**
S14	23	13	**36**
S15	36	11	**47**
S16	25	4	**29**
S17	24	3	**27**
S18	19	8	**27**
S19	11	8	**19**

Table 9: The Experimental Group's Pre-test Raw Scores

A visual inspection of the scores obtained by the EG gives more insights into the participants' performance in the pre-test. Figure 9 shows that the scores varied significantly among the participants. Despite this, almost all of the participants' scores were less than the average score, and some of them were very low compared to it. The figure shows that no participant obtained a score above the average. All the other members' scores were relatively low: 2 between 40 and 50, 5 scores were between 30 and 40, 6 between 20 and 30, and 6 between 10 and 20.

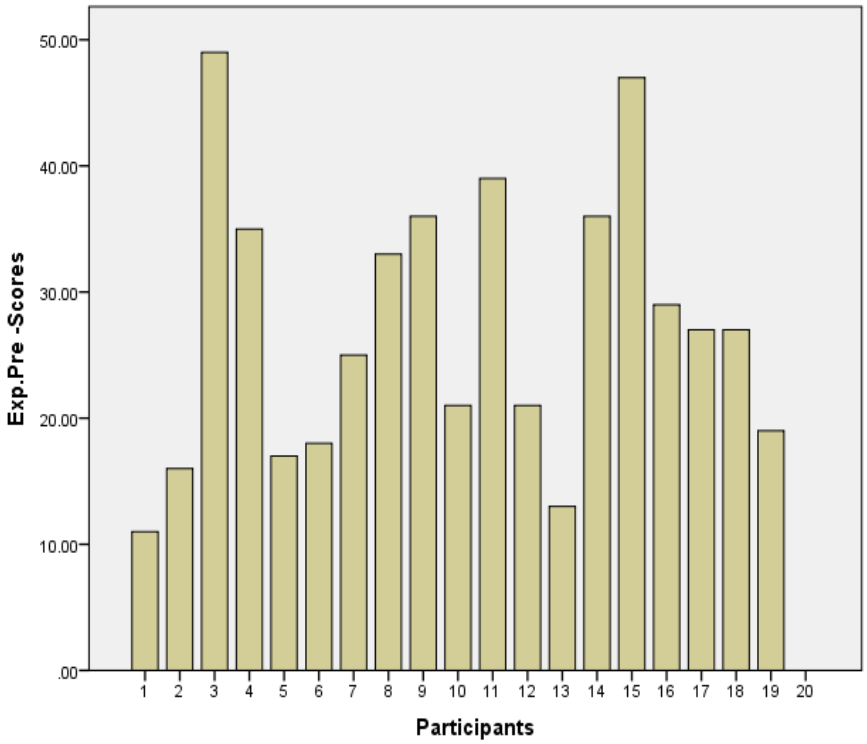

Figure 9: The Experimental Group's Pre-test Scores

The descriptive statistics of the EG's pre-test data are reported in table 10. The mean score was M=27.3158 (N=19, SD=11.02045). In comparison to the test's overall scale, the obtained mean is very low. The table also shows the maximum and the minimum scores obtained; 49 and 11. That is, by considering the test's overall scale, all of the scores obtained were below average.

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Exp.Pre	19	11.00	49.00	27.3158	11.02045
Valid N (list wise)	19				

Table 10: Descriptive Statistics of the Experimental Group's Pre-test Scores

4.2.2. **Results of the Control Group**

Table 11 presents the scores obtained by the members of the CG in the pre-test. They represent their ability to decode the different types of CS features in the pre-test. Following the same procedure as in the EG, the results of the cloze test and the dictation test were first treated separately and then the total scores were calculated. Here again, as the table shows, the members of the CG obtained relatively lower scores in the dictation test compared to the cloze test.

Students	Pre-Test		Total Score
	Cloze Text	Dictation Test	
S1	30	7	**37**
S2	47	8	**55**
S3	23	3	**26**
S4	7	2	**9**
S5	32	7	**39**
S6	13	3	**16**
S7	19	6	**25**
S8	19	4	**23**
S9	32	5	**37**
S10	28	6	**34**
S11	14	4	**18**
S12	12	5	**17**
S13	15	0	**15**
S14	9	3	**12**
S15	23	4	**27**
S16	24	8	**32**
S17	23	5	**28**
S18	21	4	**25**
S19	16	3	**19**

Table 11: The Control Group's Pre-test Raw Scores

A visual inspection of the scores obtained by the CG (Figure 10) reveals similar remarks about the performance of this group in the pre-test compared to the EG. Only one of the participants obtained a score above, but very close to, the average. All the other members' scores were relatively low: 5 scores were noted between 30 and 40, 6 between 20 and 30, 6 between 10 and 20, and only one score below 10.

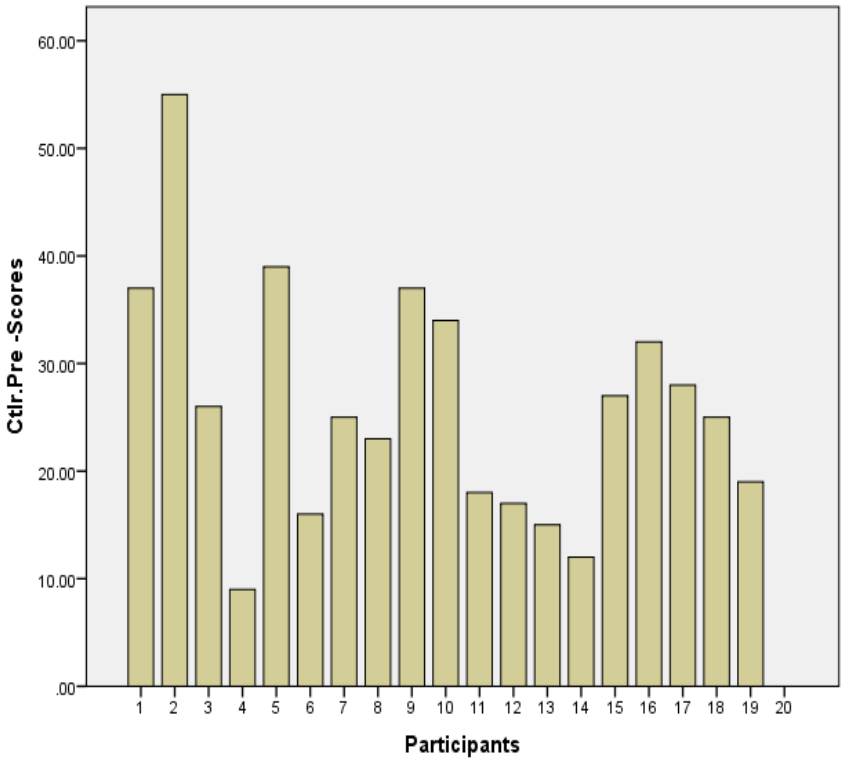

Figure 10: The Control Group's Pre-test Scores

Table 12 reports the descriptive statistics of the CG's pre-test data. The group received a mean score of *M*=26.0000 (*N*=19, *SD*=11.22497). The mean is comparatively very low in view of the total test scale. The maximum score obtained was 55, whereas the minimum score was 09. In other words, except for one participant, all of the group members obtained a score below average.

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Ctlr.Pre	19	9.00	55.00	26.0000	11.22497
Valid N (list wise)	19				

Table 12:Descriptive Statistics of the Control Group's Pre-test Scores

4.3. Comparing Pre-test Scores

Before the treatment, it was necessary to insure that the EG and the CG did not differ regarding their ability to decode the CS features. In statistical terms, the pre-test mean scores had to be compared to find out whether:

- the difference in the mean scores was large enough to the extent that the two groups could be considered as coming from different populations. Thus, the EG and the CG differed significantly in their ability to decode the CS aspects before the treatment; or,
- the difference in the mean scores was not large enough to consider that the groups came from different populations. Hence, there was no significant difference between the two groups regarding their ability to decode the CS aspects before the treatment.

The descriptive statistics revealed that the EG obtained a numerically higher mean score compared to the CG. However, this numerical comparison of the mean scores was not sufficient to conclude that the two groups differed in their ability to decode the CS features in the pre-test. According to Dörnyei (2007), descriptive statistics "do not allow drawing any general conclusions that would go beyond the sample" (p. 209). This is why inferential statistics that allow for finding out the existence of significant group differences (Larson-Hall, 2010) had to be followed.

Since the two sets of the pre-test scores came from two independent groups (the EG and the CG), the appropriate statistical test to follow in order to compare between them was the independent samples t-test (Larson-Hall, 2010). According to Larson-Hall:

> The t-test determines if the differences between groups are small enough to attribute them to the random variation in scores that wold happen each time we take a new sample of the same

> population, or whether the differences are large enough that the two groups can be said to belong to two different populations (Larson-Hall, 2010, p. 136).

For Larson-Hall (2010), the t-test is appropriate in finding out group differences in which there are two variables; one *categorical variable* which is the independent variable, and one *continuous variable* (or interval variable (Dörnyei, 2007)) which is the dependent variable. The independent variable in our study consisted of two main categories represented by group membership; the participants were members of either the CG or the EG. The pre-test scores of the participants represented the dependent variable which might vary depending on the group membership.

The inferential statistical procedures and all other data analysis were performed using the Statistical Package for Social Sciences (SPSS) version 20; it is statistical software frequently used in the field of Applied Linguistics (Dörnyei, 2007; Larson-Hall, 2010). After obtaining the scores of the two groups, the data were inputted into the software and the coding frames of the variables were defined. This implied creating the data file, naming the variables, keying, and defining the values and the labels for the independent variable (we used the values: 1=ExpGroup, and 2=CtlrGroup) (Dörnyei, 2007).

4.3.1. **Test Assumptions for the Pre-test Data**

There are two main assumptions in conducting a t-test, the violation of which may falsify the results obtained. This means that, for instance, group differences which really exist may not be found if these assumptions are not met (Larson-Hall, 2010). These assumptions will be analysed in the following sub-sections.

4.3.1.1. Normality of the Pre-test Data Distribution

The first of the assumptions is that the data should be normally distributed to resemble a bell-shaped curve. This suggests that the scores should be clustered around the

mean, rather than skewed in the tails. For the t-test, a Mann–Whitney U test should be run in case the data violate the normality assumption (Larson-Hall, 2010). The normality of the data can be checked in two mains ways:

1- by a visual inspection of the data presented through histograms and Q-Q plots; or

2- by running statistical tests or examining statistical parameters.

Larson-Hall (2010) urged that both ways should be followed for examining the data for normality.

To test the normality of the distribution of our pre-test data, we first generated histograms and Q-Q plots. In addition to giving a general idea about the normality of the data distribution, a visual inspection of such figures may help in spotting any outliers[1] in the data, or in checking any abnormality in the data that might result from mistakes in keying (Dörnyei, 2007).

The histogram in figure 11 shows the obtained pre-test data of the EG. The curve stands for a normal distribution to which the data obtained should be compared. The figure shows that our main data approximately resemble a bell-shaped curve.

Figure 12 is a Q-Q plot generated for the same data. The line in the figure represents the normal value to which the actual data obtained (the dots) should normally adhere in case the distribution was normal. The figure shows that there is a slight deviation of the data from the normal value.

[1] The term O*utlier* refers to the: "data which is markedly different from the rest of the data" (Larson-Hall, 2010). It can possibly affect the mean of a given group. Generally, the presence of outliers is dealt with by deleting the score representing an outlier, or replacing it with a score closer to the maximum/minimum score that it approaches (Dörnyei, 2007).

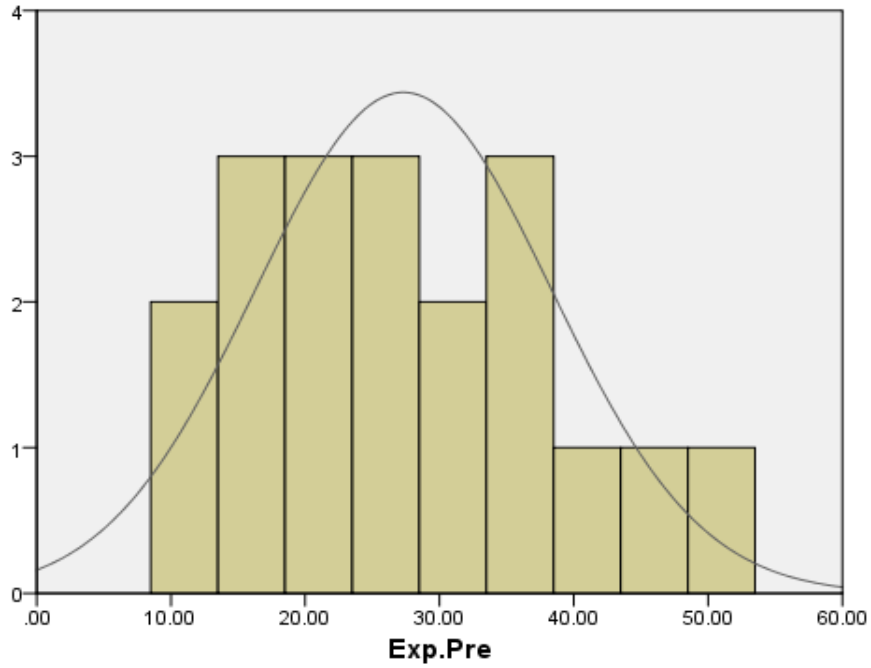

Figure 11: A Histogram of the Experimental Group's Pre-test Data

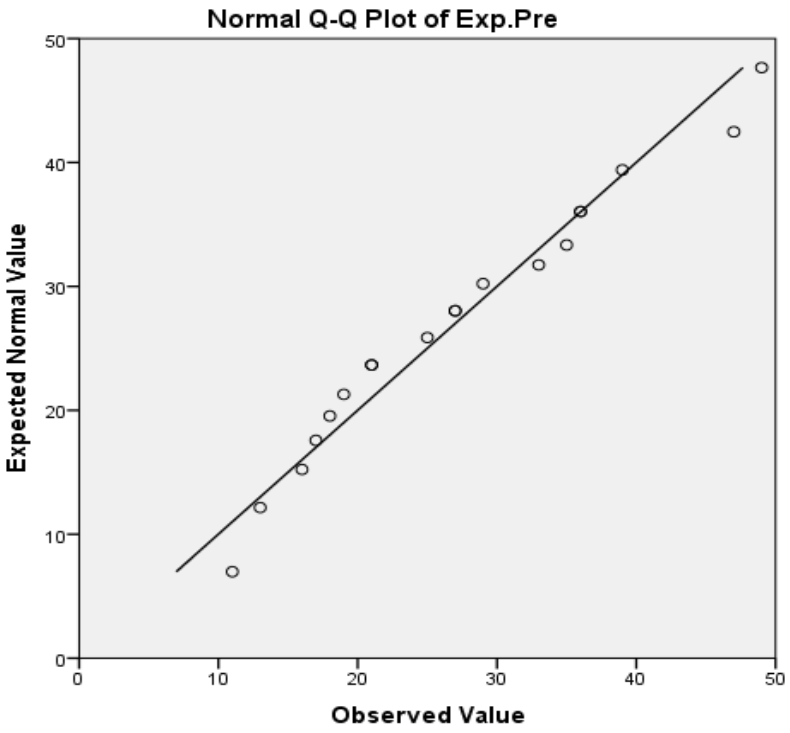

Figure 12: Q-Q Plot of the Experimental Group's Pre-test Data

The histogram in Figure 13 presents the obtained pre-test data of the CG. Here again, the data tend to resemble a normal distribution. The figure also shows the presence of extreme scores.

The Q-Q plot of the same data (figure 14) shows that the data tend to adhere to the normal distribution value despite the presence of a minor deviation. The figure also shows the presence of one extreme value.

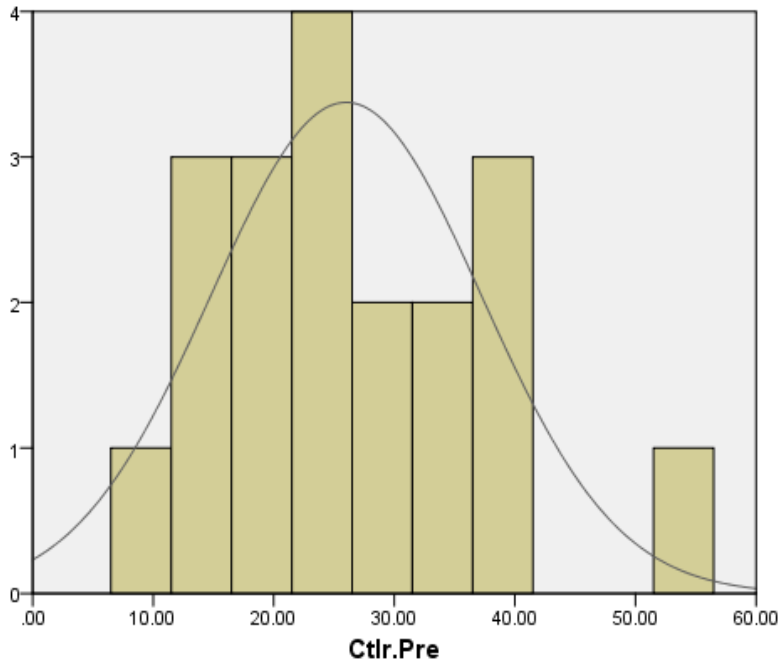

Figure 13: A Histogram of the Control Group's Pre-test Data

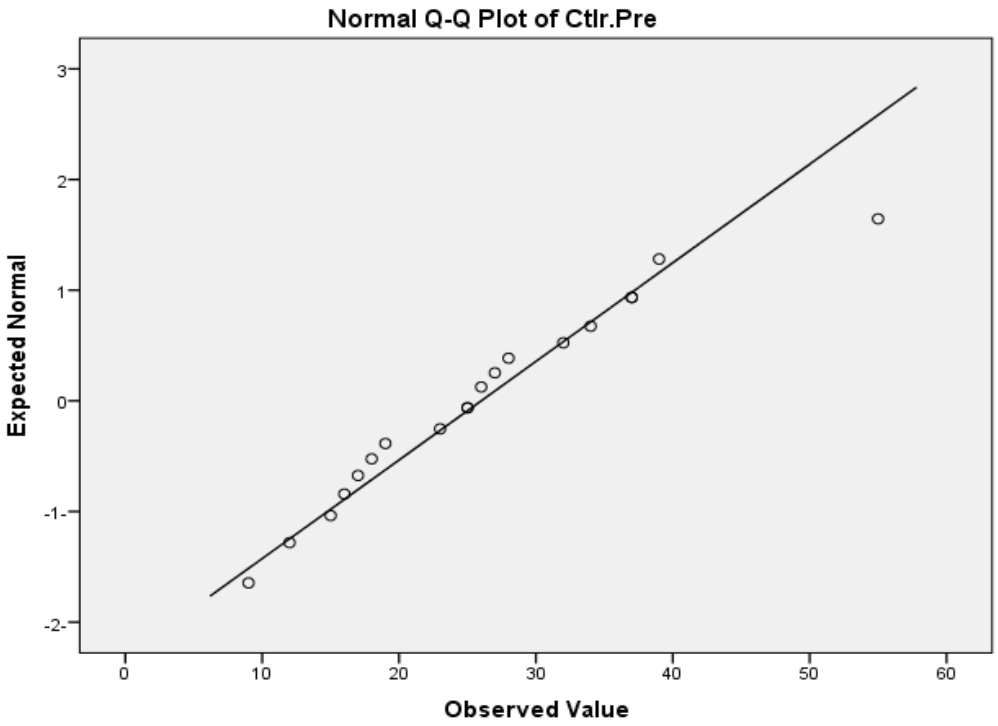

Figure 14: Q-Q Plot of the Control Group's Pre-test Data

To examine the normality of the data distribution numerically, we ran the Shapiro-Wilk normality test (Shapiro & Wilk, 1965); a formal statistical test which is "the best for sample sizes under 50" (Larson-Hall, 2010, p. 94). To test normality, we set our alternative hypothesis as:

- H_1: There is a statistically significant difference between our pre-test data and a normal distribution.

The null hypothesis which we tested was:

H_0: There would be no statistically significant difference between our pre-test data and a normal distribution.

Table 13 shows the results of the Shapiro-Wilk normality test.

Tests of Normality

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Exp. Pre	.138	19	.200	.956	19	.500
Ctlr. Pre	.114	19	.200	.953	19	.437

Table 13: Normality Test of the Pre-test Data of the Experimental and the Control Groups

As the table shows, the pre-test data of the EG and the CG were normally distributed for the purpose of running a t-test: the observed p-values in the Shapiro-Wilk test of the two groups were: p=.500>.05; and p=.437>.05, respectively. Since the p-values were higher than the level of significance (95% confidence interval in all the tests we conducted), the null hypothesis was confirmed, suggesting that the first assumption was met.

4.3.1.2. Equality of the Pre-test Data Variance

The second assumption that should be met has to do with the homogeneity of variance[1] among the data of the two groups; i.e., the group variances should be equal.

Homogeneity is examined visually through the inspection of boxplots. This is done by considering and comparing the lengths of the boxes which represent the data side by side; if they are equal, or roughly equal, the two sets of data are said to be homogeneous (Larson-Hall, 2010). Boxplots also allow for spotting any outliers in the data. Figure 15 shows boxplots generated for the pre-test data of the EG and the CG, side by side. As demonstrated in the figure, the lengths of the two boxes are roughly equal. Hence, we can assume that the homogeneity of variance assumption was met.

To confirm this observation, homogeneity was tested numerically using Levene's F test. As the results of this test are closely related to the interpretation of the t-test results, they are automatically generated side-by-side with the results of the t-test results-table when it is

[1] By definition, *Variance* is " a measure of the amount of variability around a mean. Formally, it is the average squared distance from the mean to any point" (Larson-Hall, 2010, p. 403).

run with SPSS. According to Larson-hall (2010), “for t-tests, there are estimators that can take unequal variances into account” (p. 251). As a rule of thumb, if Levene’s F test is not significant (i.e., the groups are not homogeneous), the p-value that appears in the second line in the t-test results table should be considered. For this reason, our Levene’s test results will be reported later with the results of the independent samples t-test.

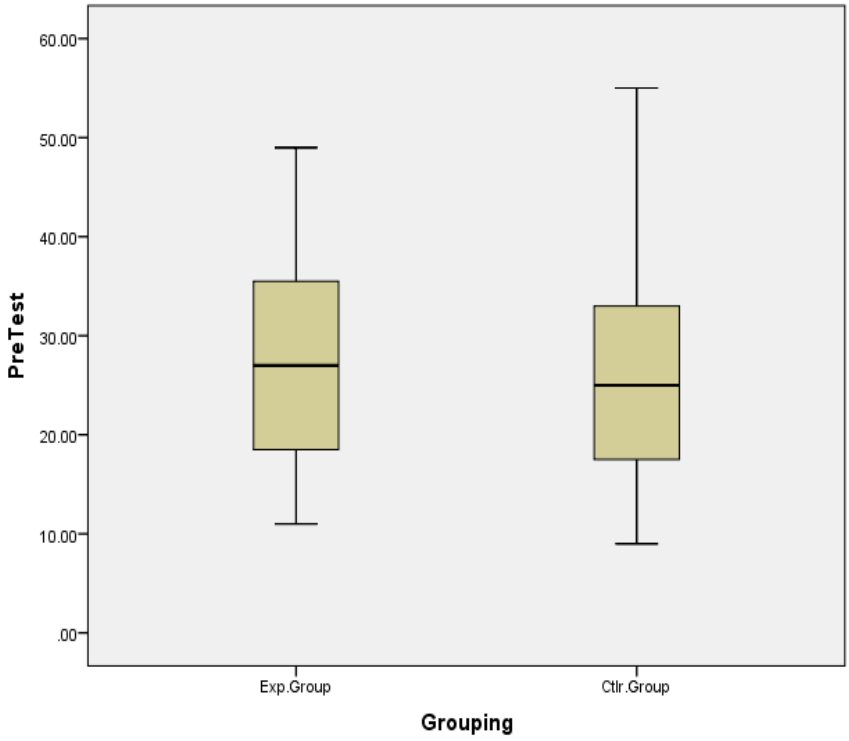

Figure 15: Boxplots for the Experimental and the Control Groups’ Pre-test Data

4.3.2. Independent Samples t-test between the Pre-test Means

After considering the required assumptions, we ran a t-test to compare the means of the EG and the CG in the pre-test. The alternative hypothesis was:

H_1: There would be a statistically significant difference between the mean scores obtained by the EG and the CG in the pre-test.

The null hypothesis was:

H_0: There would be no statistically significant difference between the scores obtained by the EG and the CG in the pre-test.

It should be noted that, in hypothesis testing, the null hypothesis H_0 is the one which is actually checked, rather than the alternative hypothesis H_1. The assumption is that no significant difference exists between the groups unless there is enough evidence that proves the opposite is true. The results of the independent samples t-test are summarised in table 14.

Independent Samples Test

		Levene's Test for Equality of Variances		**t-test for Equality of Means**						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
PreTest	Equal variances assumed	.049	.825	.365	36	.718	1.31579	3.60884	-6.00327-	8.63485
	Equal variances not assumed			.365	35.988	.718	1.31579	3.60884	-6.00336-	8.63493

Table 14: Independent Samples T-test between the Pre-test Means of the Control Group and the Experimental Group

As mentioned earlier, Levene's test of equality of variances should be first considered to interpret the results of the independent samples t-test. Levene's F results indicated that the two groups had equal variances; $F(36) = .049$, $p = .825$. Hence, based on the numerical test, our second assumption for the t-test was met. As for the results of the t-test, the statistical findings suggested that there was no statistically significant difference between the mean scores of the EG and the CG in the pre-test; $t\ (36) = .365$, $p = .718 > .05$.

4.4. Results of the Post-Test

After the treatment, the post-test was administered to the two groups in order to find out whether there was any significant improvement in the performance of the group which received the special instruction on the CS features compared to the group which did not. In this section, the results of the post-test will be reported together with the descriptive statistics of the data.

4.4.1. **Results of the Experimental Group**

The participants' answers to the dictation task and the cloze task in the post-test were analysed separately and, then, scored. The final scores of the EG are presented in table 15.

Students	Post-Test		Total Score
	Cloze Text	Dictation Test	
S1	29	8	**37**
S2	35	6	**41**
S3	58	11	**69**
S4	51	9	**60**
S5	39	7	**46**
S6	46	11	**57**
S7	39	8	**47**
S8	44	11	**55**
S9	48	11	**59**
S10	36	10	**46**
S11	51	6	**57**
S12	32	9	**41**
S13	36	7	**43**
S14	41	6	**47**
S15	39	12	**51**
S16	33	6	**39**
S17	41	7	**48**
S18	48	7	**55**
S19	28	10	**38**

Table 15: The Experimental Group's Post-test Raw Scores

As the table shows, similar observations can be made about the participants' post-test scores compared to those of the pre-test. Most of the members of the EG obtained relatively lower scores in the dictation task compared to the cloze task. The maximum score was 69 and the minimum score was 37.

Figure 16 shows the pre-test and the post-test scores obtained by the members of the treatment group side-by-side. A visual inspection of the scores demonstrates that all the participants obtained comparatively higher scores in the post-test. No participant obtained a score below 30. Three participants obtained scores between 30 and 40, eight between 40 and 50, seven between 50 and 60, and only one above 60.

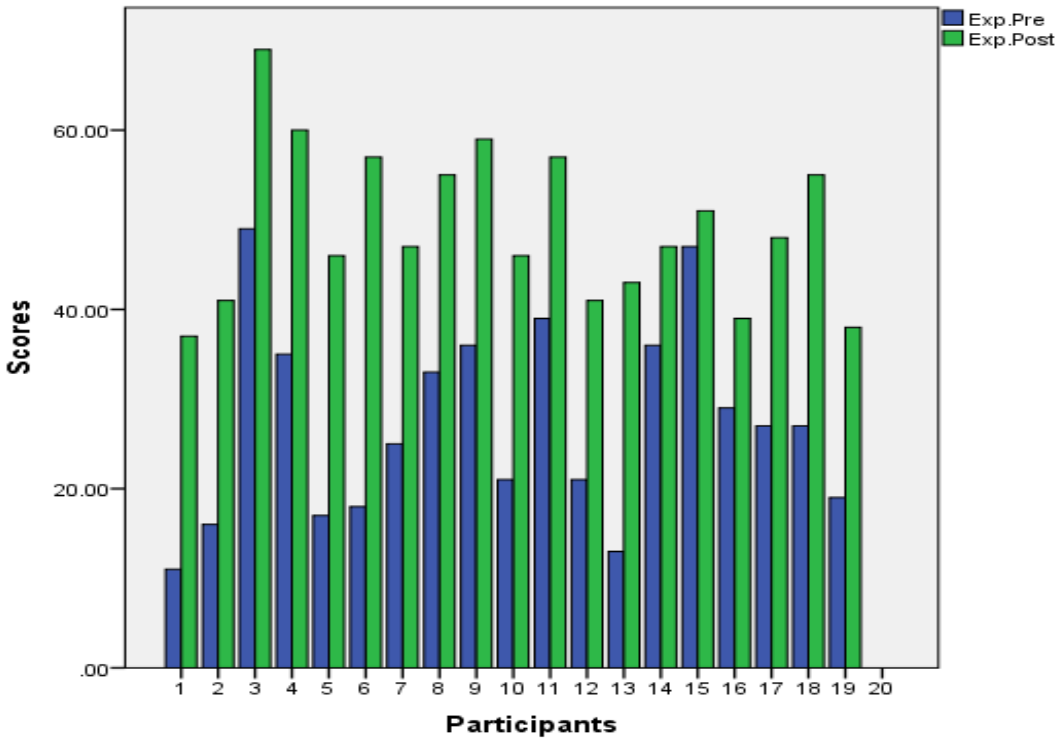

Figure 16: The Pre-test versus the Post-test Scores of the Experimental Group

Table 16 presents the descriptive statistics of the EG's post-test data. The group obtained a numerically higher mean score compared to the pre-test; *M*=49.2632 (*N*=19,

SD=8.75495). Despite this, the post-test mean was very low compared to the overall scale of the test.

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Exp.Post	19	37.00	69.00	49.2632	8.75495
Valid N (list wise)	19				

Table 16: Descriptive Statistics of the Experimental Group's Post-test Scores

4.4.2. Results of the Control Group

Table 17 presents the scores obtained by the CG in the post-test. We can observe that the scores obtained in the dictation task were very low in comparison to the overall scale of this task (total=21). A similar observation can be made about the cloze task scores. Overall, the post-test results of the CG did show a relative improvement compared to the pre-test. A visual inspection of the pre-test and the post-test data may demonstrate this.

Students	Post-Test		Total Score
	Cloze Text	Dictation Test	
S1	30	5	**35**
S2	39	9	**48**
S3	18	3	**21**
S4	17	3	**20**
S5	34	7	**41**
S6	18	2	**20**
S7	27	9	**36**
S8	23	3	**26**
S9	37	11	**48**
S10	26	4	**30**
S11	26	3	**29**
S12	16	3	**19**
S13	10	4	**14**
S14	26	2	**28**
S15	28	9	**37**
S16	28	4	**32**
S17	11	4	**15**
S18	16	1	**17**
S19	28	7	**35**

Table 17: The Control Group's Post-test Raw Scores

Figure 17 presents the pre-test and the post-test data of the CG. It can be noted from the figure that more than 50% of the participants obtained higher, or equal, scores in the post-test compared to the pre-test. The maximum and the minimum scores were 14 and 48. We note that the scores varied greatly among the participants: six participants obtained

scores between 10 and 20, five between 20 and 30, five between 30 and 40, and three between 40 and 50. No one obtained a score above the average.

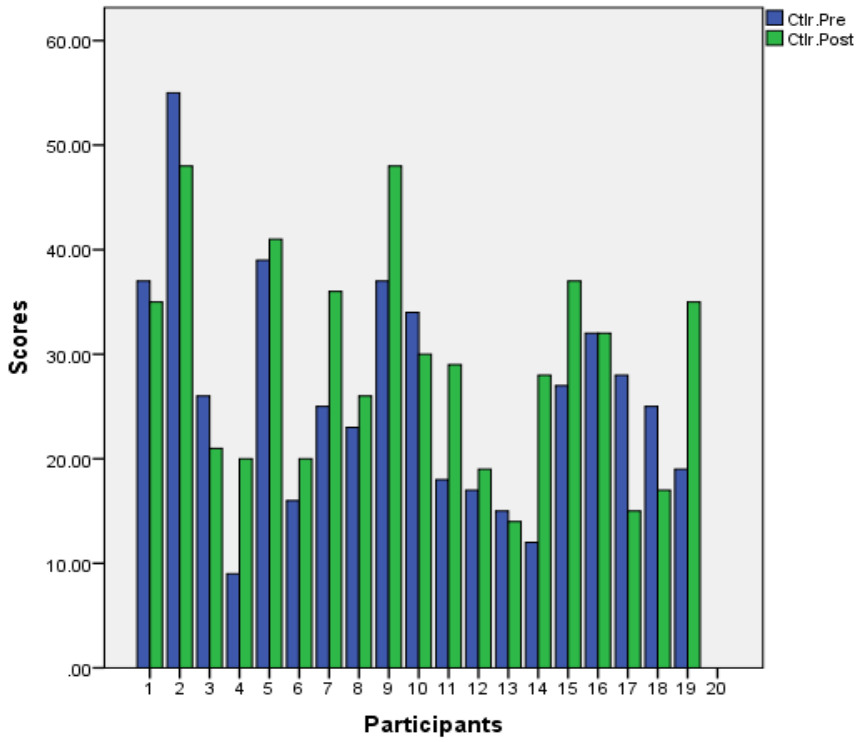

Figure 17: The Pre-test versus the Post-test Scores of the Control Group

Descriptive statistics (Table 18) show that, in the post-test, the CG obtained a numerically higher mean score than in the pre-test –*M*=29.0000 (*N*=19, *SD*=10.44031) – but it was still very low in comparison to the total test-scale.

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Ctlr.Post	19	14.00	48.00	29.0000	10.44031
Valid N (list wise)	19				

Table 18:Descriptive Statistics of the Control Group's Post-test Scores

4.5. Comparing the Post-test Scores

The descriptive statistics of the post-test mean scores of the two groups revealed that the EG obtained a numerically higher score in the CS features test compared to the CG. This is clearly noticeable from the bar graph of the two sets of data (figure 18), as most of the EG's scores were higher than the CG's ones. To find out whether this difference in the means had any statistical significance, we ran an independent samples t-test between the two sets of data. As mentioned earlier, this test has two basic assumptions (normality and homogeneity of variance) that should be met for the results to be reliable. In the same procedure followed with the results of the pre-test, we will first deal with these assumptions, and then, report the results of the t-test.

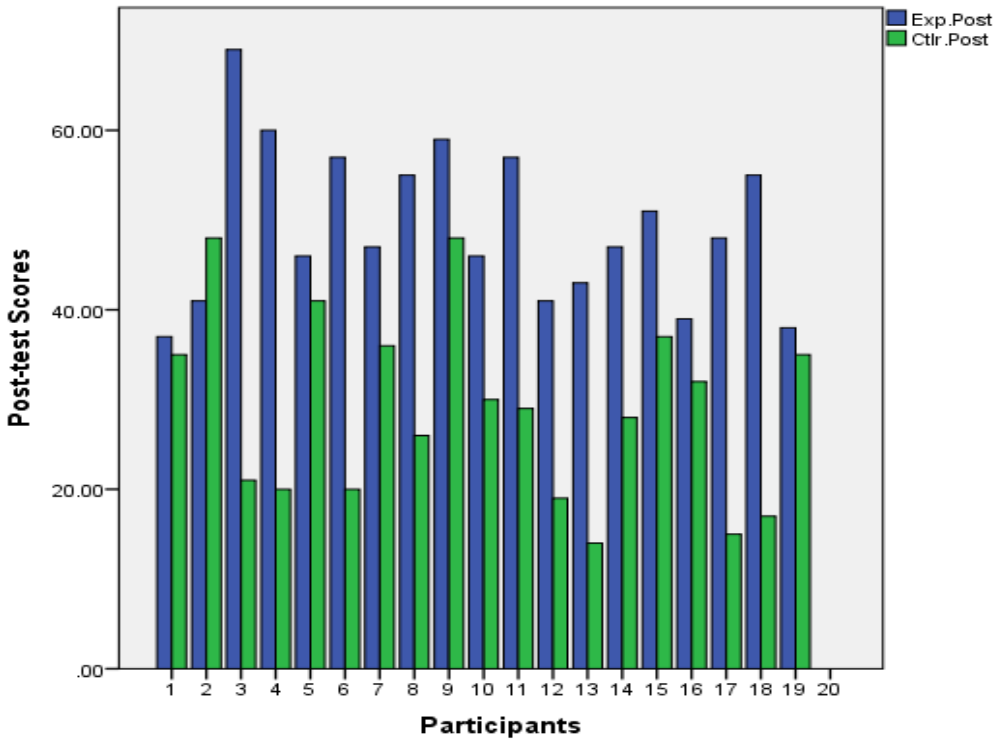

Figure 18: Comparing the Post-test Scores of the Experimental Group and the Control Group

4.5.1. Test Assumptions for the Post-test Data

As the comparison of the mean scores of the two groups was done by running an independent samples t-test, this section will deal with the assumptions of the test that the data should satisfy.

4.5.1.1. Normality of the Pre-test Data Distribution

As far as the normality of the data distribution is concerned, we inspected the data visually by analysing histograms and Q-Q plots. Figure 19 is a histogram generated for the post-test data of the EG. As the figure shows, the data roughly approximated a normal distribution.

The Q-Q plot of the same data (figure 20) shows that there was a noticeable deviation of the data from the normal value. It does not show the presence of outliers, but there is an extreme value.

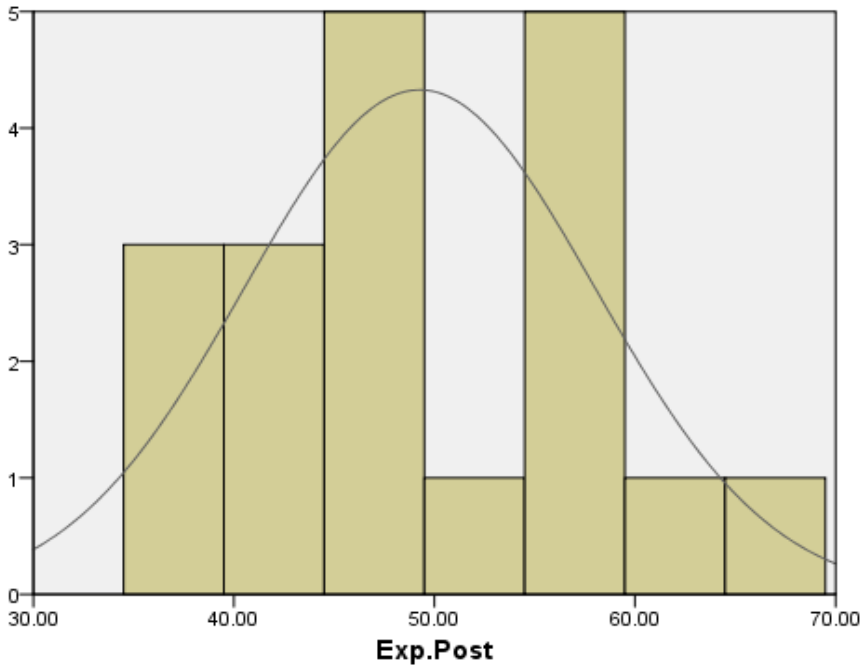

Figure 19: A Histogram of the Experimental Group's Post-test Data

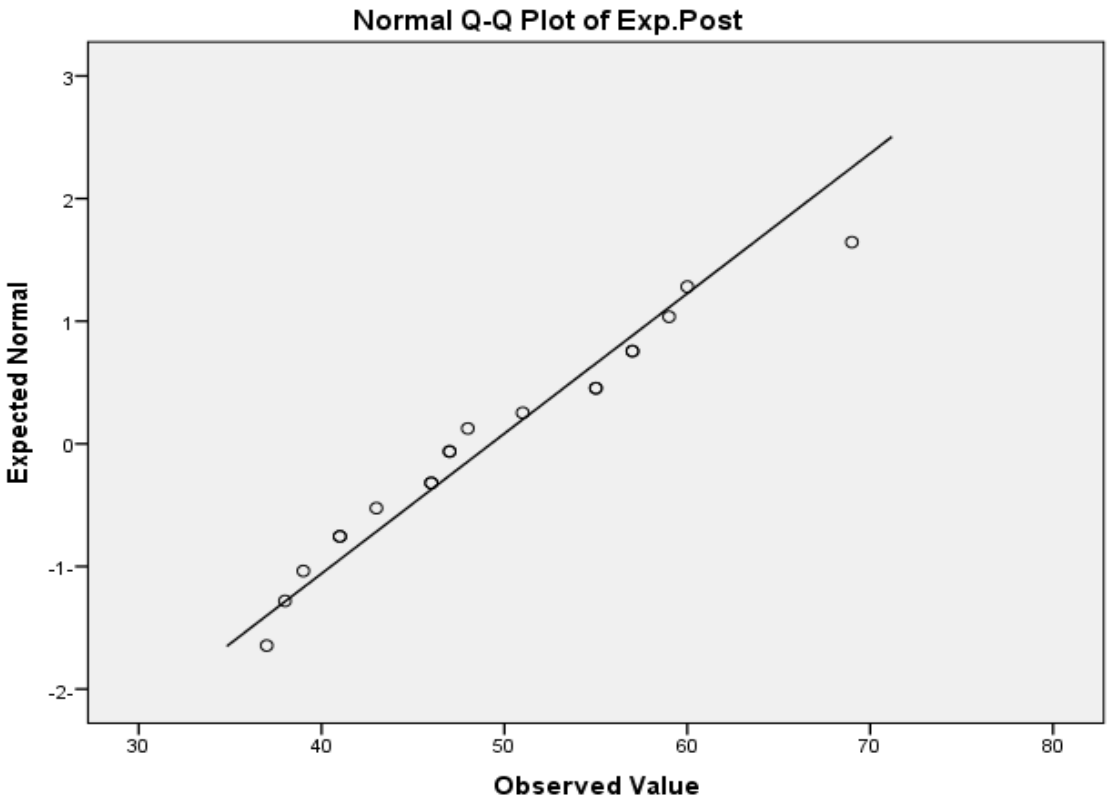

Figure 20: Q-Q Plot of the Experimental Group's Post-test Data

The histogram in figure 21 stands for the post-test data of the CG. It shows that the data were slightly positively skewed. They only approximated, but did not totally resemble, a normal distribution curve. The figure also suggests the presence of extreme scores. The data partially deviate from the normal value (figure 22), but no outliers can be noticed.

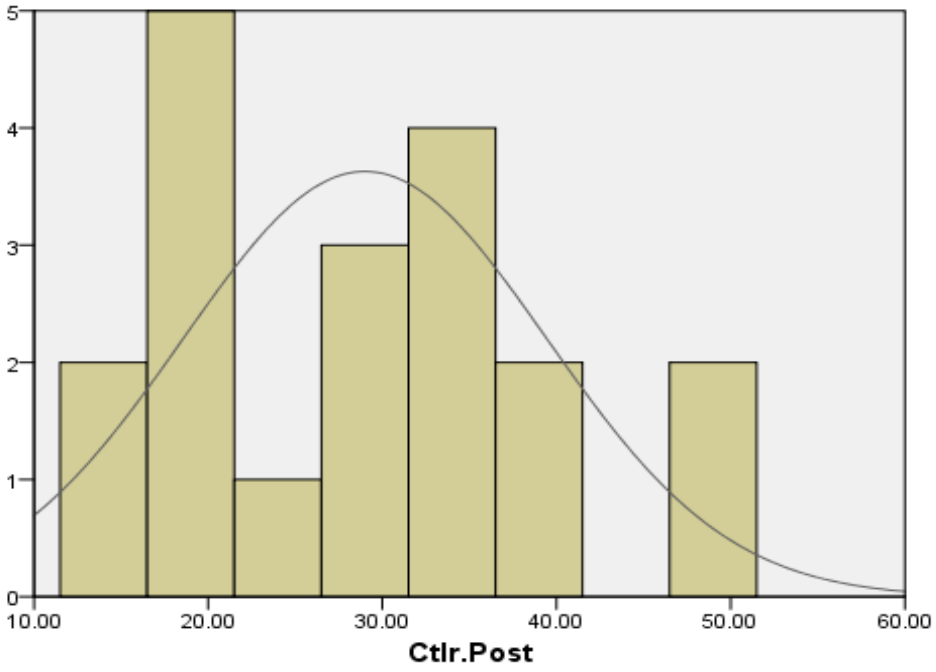

Figure 21: A Histogram of the Control Group's Post-test Data

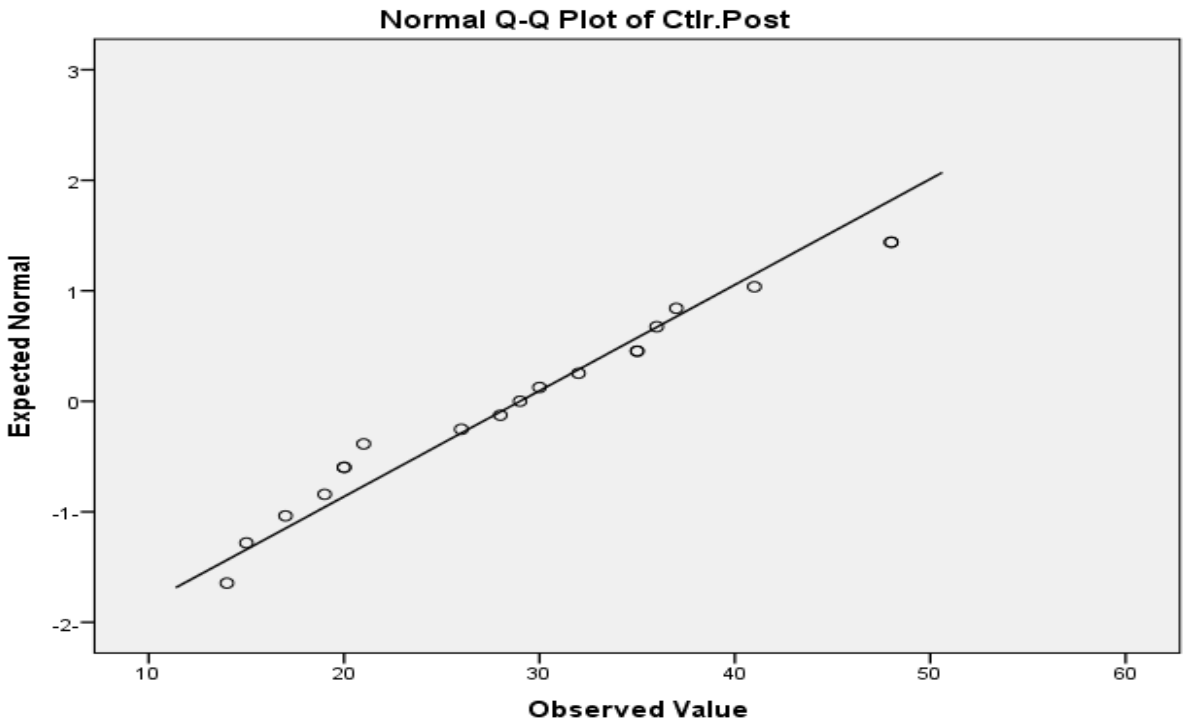

Figure 22: Q-Q Plot of the Control Group's Post-test Data

To analyse the data numerically, we ran the Shapiro-Wilk (1965) normality test based on the two sets of data. Table 19 summarises the results of this test for the two groups.

Tests of Normality

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Exp.Post	.136	19	.200	.952	19	.422
Ctlr.Pre	.114	19	.200	.953	19	.437

Table 19: Normality Test of the Post-test Data of the Experimental and the Control Groups

As the table shows, the results of the Shapiro-Wilk normality test indicate that the two sets of data were normally distributed. For the EG, the observed p-value was: p=0.422>.05. The observed p-value for the CG was: p=.437>.05. Hence, the results suggest that there was no statistically significant difference between each of the two groups' post-test data and a normal distribution. Consequently, we could assume that our post-test data were normally distributed for the purpose of conducting a t-test (Larson-Hall, 2010).

4.5.1.2. Equality of the Post-test Data Variance

The last step in the visual inspection of the data involved the analysis of boxplots in order to spot any outliers in the data, and to compare the group variances. The two boxplots in Figure 23 stand for the post-test data of the two groups. Despite the skewness in both of the boxplots, the figure shows no outliers which might lead to the distortion of group means. The medians[1] in both of the boxes were not centred in the middle of the boxes. There is no noticeable difference between the lengths of the boxes which appear to be roughly equal. This is an indicator of an equal variance between the two sets of data (Larson-Hall, 2010).

[1] The median is the line inside the box which represents "the point at which 50% of the scores are above and 50% of the scores are below" (Larson-Hall, 2010, p. 245).

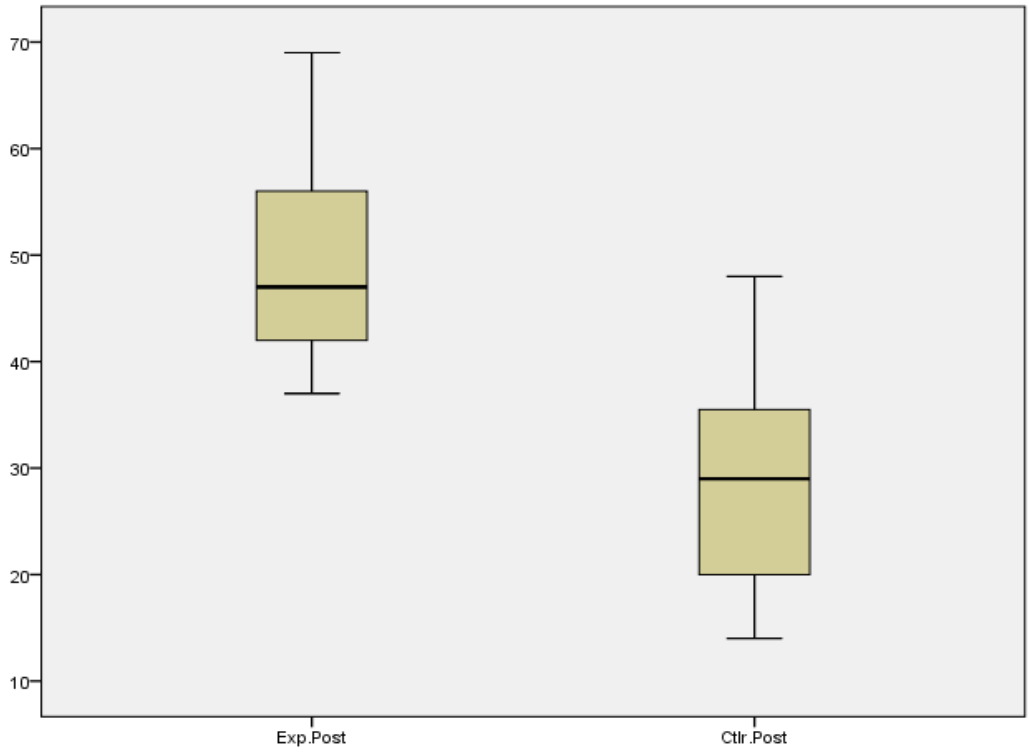

Figure 23: Boxplots for the Experimental and the Control Groups' Post-test Data

4.5.2. **Independent Samples t-test between the Post-test Means**

After considering the required assumptions, an independent samples t-test was run between the post-test data of the two groups to find out whether or not the numerically higher mean score obtained by the EG had any statistical significance. So, our main hypothesis was:

H_1: There would be a statistically significant difference between the data obtained by the EG and the CG in the post-test.

The null hypothesis was:

H_0: There would be no statistically significant difference between the data obtained by the EG and the CG in the post-test.

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
PostTest	Equal variances assumed	.578	.452	6.482	36	.000	20.26316	3.12586	13.92362	26.60270
	Equal variances not assumed			6.482	34.939	.000	20.26316	3.12586	13.91693	26.60939

Table 20: Independent Samples t-test between the Post-test Data of the Experimental Group and the Control Group

The results of the t-test are summarised in table 20. To interpret the results, Levene's homogeneity of variance test should be considered first. The observed p-value for Levene's F test indicates that, as far as the post-test data are concerned, the two groups had equal variances; F (36) =578, p = .452. Hence, we can assume that the second assumption of the test was numerically verified and satisfied. The observed p-value of the independent samples t-test suggests that there was a statistically significant difference between the two groups; t(36) =6.482, p = .000< .05, with a large effect size, d = 2.11 according to Cohen's guidelines (Cohen J. , 1992). Hence, the null hypothesis H_0 was rejected in favour of the alternative hypothesis H_1.

4.6. Comparing the Pre-test and the Post-test Results: Paired Samples t-test

The descriptive statistics revealed that both of the EG and the CG obtained higher mean scores in the post-test compared to the pre-test. Although the results of the independent samples t-test revealed a significant difference, we needed to consider whether the improvement in the mean scores also had any significance. To this end, we ran a paired-

samples t-test between the pre-test and the post-test data of each group to examine if the numerically higher mean scores of the post-test had any statistical significance. A paired samples t-test is used to compare the two sets of data obtained by the same group in two different points of time, especially in a study which involves a given treatment condition (Dörnyei, 2007; Larson-Hall, 2010). Table 21 summarises the results of the test. The hypotheses of the test were as follows:

-The Alternative Hypothesis H_0: There would be no statistically significant difference between the pre-test and the post-test mean scores of both the EG and the CG.

-The Null Hypothesis H_1: There would be a statistically significant difference between the pre-test and the post-test mean scores of both the EG and the CG.

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	**Exp. Post – Exp. Pre**	21.94737	7.83492	1.79745	18.17106	25.72368	12.210	18	.000
Pair 2	**Ctlr. Post – Ctlr. Pre**	3.00000	8.45248	1.93913	-1.07397-	7.07397	1.547	18	.139

Table 21: Paired Samples t-test between the Pre-test and the Post-test Means of the Experimental Group and the Control Group

As the table shows, the homogeneity of variance assumption is not verified in this test because it is "assumed to be true" (Larson-Hall, 2010, p. 251). The firs and the second lines represent the EG's and the CG's test results, respectively. Concerning the EG, the results show that there was a statistically significant difference between the pre-test and the post-test mean scores; t (18) =12.210, p=.000<.05. Thus, the null hypothesis H_0 was rejected in favour of the alternative hypothesis H_1. For the CG, however, the test results revealed no significant difference between the pre-test and the post-test mean scores; t(18)=1.547, p=.139>.05. Hence, the null hypothesis was accepted.

4.7. Discussion of the Experiment Results

The results of the pre-test revealed a serious problem in decoding speech by the participants in this study. The very low scores obtained by the two groups in the pre-test indicate that CS features form a barrier for them in decoding naturally spoken English. Hence, in an answer to the first question of the study, it could be said that the participants' results obtained in the pre-test do show that CS features constitute one of the obstacles that the learners have in listening to naturally spoken English. These results go in line with the findings of previous studies on EFL learners' ability to decode CS features (see chapter two).

The independent samples t-test of the pre-test results revealed no significant difference between the two groups. This implies that, at the outset of the study, the two groups did not differ in their ability to decode CS features before the treatment. Both of them showed a poor decoding ability as far as the CS features are concerned. This is significant in the sense that the results obtained after the treatment would not be biased. The chance of having any difference between the groups after the instruction due to unwanted variables was limited, especially a possible inherent inequality between the two groups in decoding CS.

For the CG, the post-test results improved numerically compared to those of the pre-test. However, as indicated by the results of the paired samples t-test, the fact that this numerical difference had no statistical significance suggests that no real improvement in their ability to decode CS can be traced. The lessons the CG received, which dealt with listening and TD processing without a focus on CS features, did not have an impact on their ability to decode the features of CS in the post-test. On the other hand, the numerically higher mean score obtained by the EG in the post-test did have a statistical significance as indicated by the results of the paired samples t-test of this group's scores. A statistically significant difference suggests that the higher mean score cannot be attributed to chance variables,

especially if we compare this difference to the results of the CG. This implies that the lessons the EG received –which focused on diagnosing and addressing decoding problems caused by CS – did have an impact on the participants' performance on the post-test; i.e. on their ability to decode CS features. The results obtained support the hypothesis that the diagnostic approach would be significant in improving learners' ability to decode with CS in naturally spoken English.

The independent samples t-test of the two groups' post-test results has also put the EG ahead of the CG. The EG's performance in the post-test was significantly higher than that of the CG. This gives additional support to the previous results, and confirms that, unlike the CG, the instruction the EG received did have an impact on their ability to decode CS features.

Conclusion

This chapter presented the results of the experiment, including the results of the tests on CS features decoding which the two groups received. The EG and the CG obtained very low mean scores in the pre-test (M=27.3158 and M=26.0000 respectively) compared to the overall test scale (total=101). This implies a very weak ability to decode CS features. It also suggests that the learners do have a serious problem in coping with the features of natural speech. While the CG made no significant improvement in the post-test, the EG did; the post-test mean score of the latter was statistically significantly higher than that of the pre-test, which suggests an effect of the treatment they received on their performance. Concerning the post-test results of the two groups, the EG outperformed the CG in the post-test significantly, and this cannot be attributed to chance factors. Accordingly, the results support the hypothesis that a LC session which balances between TD and BU skills in a way

that allows for diagnosing and addressing learners' decoding problems is significant in improving learners' ability to decode CS features.

Chapter Five

Qualitative Data Analysis and Discussion of Students' Problems and their Perceptions of the Instruction

Chapter Five
Qualitative Data Analysis and Discussion of Students' Problems and their Perceptions of the Instruction

Introduction

This chapter will deal with the qualitative analysis of the findings of both the participants' answers in the pre-test and the pose-treatment interview. As for the pre-test results, the common mishearings resulting from the misperception of the CS features will be categorised to spot any systematic patterns. These patterns will be analysed to track the effects of the different CS featured on the decoding of the utterances in the test. In addition, the patterns of mishearings will be utilised to trace any apparent application of the TD and the BU modes in decoding the words in the utterances, and the extent to which the participants were successful in their choice and use of the related strategies. Concerning the post-treatment interview, the answers provided by the interviewees will be analysed to find out how they perceived the lessons they received during the instruction.

5.1. Effects of Connected Speech Features: Results and Discussion

To depict the students' problems in decoding the CS features, the answers they provided in the pre-test were subject to a qualitative analysis of the misperceptions. The participants' answers were analysed to find out:

1- the effects of the different CS features on the decoding of the utterances in the test, with a focus on those highlighted in the literature review, and;
2- the common patterns of misperception which could be attributed to either of the two processing modes, TD and BU. We searched for any indications of a

reliance/overreliance on a particular processing mode to decode the CS features or to overcome the decoding failure.

5.1.1. **Effects of Weak Forms**

In chapter II, some theoretical considerations were raised concerning the characteristics of WFs that hinder their decoding in speech and which lead to problems in listening. The participants' answers provided in the pre-test were analysed to trace the major effects of WFs on their ability to decode the words and utterances in the test, with a focus on their ability to:

- recognise the different reduced versions of the function words.
- segment speech in the presence of WFs.
- differentiate between neutralised WFs of different function words.
- recover from the effects of other CS features that WFs may undergo .

As for the recognition, most of the WFs were proved to be difficult to decode from the part of the participants. This is despite the variability in the ability to decode them which was found not only across the different categories (prepositions, auxiliaries…etc.) but which also depended on the phonological environment and the manner of reduction. The number of incorrect recognitions by all of the participants was used as an indication of the extent to which a given reduced form of a function word represented a barrier to recognition. This was compared to the contexts where the same function word was reduced differently, or occupied a different position in the utterance, which could have made it easier/more difficult to decode. Instances where other aspects, such as assimilation, were involved in the modification of the word shape were also considered. Table 22 presents the numbers of

incorrect recognitions[1] and/or non-recognitions across the different word categories of the function words.

Categories	Function Word	Reduced Form	Position in the Utterance	Other Features Involved	Incorrect recognitions Recognitions	
					N=°	%
Pronouns	you	/jʊ/	initial	/	19	50
		/tʃə/	middle	coalescent assimilation	16	42
		/jə/	middle	/	11	29
		/jʊ/	middle	linking / ʲ/	37	97
		/jʊ/	middle	//	24	63
		/juː/	initial	/	1	3
	them	/əm/	middle	assimilation	19	50
	us	/əs/	middle	/ ʷ/ linking	34	89
		/əs/	middle	/ ʲ/ linking	24	63
	her	/ə/	final	elision	25	66
		/ə/	middle	elision	37	97
		/ər/	final	elision	34	89
	he	/ɪ/	middle	elision	37	97
		/ɪ/	middle	elision	37	97
		/hɪ/	middle	(h not elided)	15	39
		/hɪ/	initial	/	1	3
		/ɪ/	middle	elision	38	100
	him	/ɪm/	middle	elision	16	42
		/ɪm/	middle	elision	21	55
	his	/hɪz/	middle	elision	36	95
		/ɪz/	middle	elision	38	100
Verbs	were	/wər/	middle	/	8	21
	was	/wəz/	middle	/	37	97
		/wəs/	middle	assimilation	14	37

[1] The participants' answers can be categorised into "*incorrect recognitions*" when there was an incorrect attempt to fill in the blank, and "no-recognition" when the blank was not filled at all. As both of them demonstrated an inability to decode the input, they, hence, were rated as wrong; the term "*incorrect recognition*" is used to account for both instances.

Categories	Function Word	Reduced Form	Position in the Utterance	Other Features Involved	Incorrect recognitions	Recognitions
					N=°	%
		/wəz/		/	32	84
		/wəs/	middle	assimilation	2	5
	been	/bɪn/	middle	/	32	84
		/bɪn/	middle	/	37	97
	are	/ə/	initial	/	30	79
		/ə/	middle		31	82
		/ə/	middle		31	82
		/ɑː/	initial	(strong)	1	3
	should	/ʃəd/	middle	/	0	0
	would	/əd/	middle	/	37	97
	shall	/ʃˢl/	middle	/	31	82
	is	/ɪs/	middle	assimilation	12	32
		/s/	middle	contraction	13	34
		/z/	middle	contraction	30	79
		/s/	middle	assimilation	26	68
		/s/	middle	assimilation contraction	21	55
		/ɪz/	middle	/	15	39
	must	/məs/	middle	elision	31	82
		/məs/	middle	elision	28	74
	can	/kən/	middle	/	19	50
	have	/əf/	Middle	elision assimilation	38	100
		/ə/	middle	elision	36	95
		/ə/	middle	elision	37	97
		/əv/	middle	elision	28	74
	has	/əs/	middle	elision assimilation	37	97
		/əz/	middle	elision	38	100
		/əs/	middle	elision assimilation	36	95
		/hæz/	middle	/	1	3
	had	/əd/	middle	elision	34	89
		/həd/	middle	elision	25	66
Articles	**a**	/ə/	middle	/ʳ/ linking	37	97
		/ə/			32	84

Categories	Function Word	Reduced Form	Position in the Utterance	Other Features Involved	Incorrect recognitions	Recognitions
					N=°	%
		/ə/			36	95
		/ə/	middle	/	30	79
		/ə/	middle	/	26	68
	the	/ðə̣/	middle	/	10	26
		/n̪ə/	middle	assimilation	5	13
		/ðə/	middle	/	19	50
		/ðə/	middle	/	2	5
		/ðɪː/	middle	/	6	16
		/ðə/	initial	/	23	61
Prepositions	from	/frəm/	middle	/	32	84
	of	/v/	middle	/	34	89
		/əv/	middle	/	35	92
		/əv/	middle	/	30	79
	to	/tə/	middle	/	20	53
		/tʊ/	middle	/ʷ/linking	38	100
		/tʊ/	middle	Assimilation (simplification) Linking /ʷ/	1	3
		/tə/	middle	/	16	42
		/tə/	middle	assimilation (simplified)	29	76
	at	/ət/	middle	/	20	53
		/ət̪/	middle	assimilation	23	61
	for	/fə/	middle	/	14	37
		/fə/	middle	/	25	66
		/fər/	middle	/r/ linking	31	82
Conjunctions	and	/n/	middle	elision	12	32
		/ən/	middle	elision	20	53
		/ən/	middle	elision	32	84
		/ən/	middle	elision assimilation	32	84
		/ən/	middle	elision	22	58

Categories	Function Word	Reduced Form	Position in the Utterance	Other Features Involved	Incorrect recognitions	Recognitions
					N=°	%
	but	/bəˀ/	middle	/	21	55
	than	/ðəm/	middle	assimilation	32	84
Determiner	**some**	/səm/	middle	/	19	50
		/səm/	middle	/	29	76

Table 22: Wrong Recognitions of Weak Forms across the Different Contexts and Word Categories

We ought not to forget that the pre-test quantitative results suggested a weak ability from the part of the participants to decode WFs. As table 22 demonstrates, this weakness is palpable in the numbers of incorrect recognitions of most of the WFs in the test. It should be mentioned that the table presents not only the numbers of incorrect recognitions of the WFs which were considered in rating the participants' answers, but also those instances where the function words were pronounced in their SF. This allows for comparing some of the answers to check whether the SFs did, or did not, hinder decoding compared to the WFs.

A close observation of the table, which counts 87 cases of function words (including both strong and WFs), shows that there are 65 cases where at least half of the participants failed to decode the WF of a given word. In 42 of these instances, incorrect recognitions exceeded the level of 75%. These include all categories of function words. Among them, there are 5 instances where *all* of the students failed to decode certain WFs:

- WFs where the initial /h/ sound is elided: *he* /ɪ/, *his* /ɪz/, *have* /əf/ *(assimilated)* and *has* /əz/.
- The WF of *to* /tə/.

In addition, there are 7 WFs for which incorrect recognitions reached 97% (=all participants except one) for each of them. These include the WFs of 'you' /yʊ/, 'her' /ə/, 'he' /ɪ/, 'was' /wəz/, 'would' /əd/, 'have' /ə/ and 'a' /ə/.

On the other hand, certain function words (both strong and weak) resulted in relatively low rates of incorrect recognitions. The table counts only 22 cases where the rate of incorrect recognitions of functors (both strong and weak) represented less than 50%. Of these, 11 instances of functors did not exceed the level of 25% of incorrect answers: ‘you’ /yuː/, ‘he’ /hɪ/, ‘were’ /wər/, ‘was’ /wəs/ (assimilated), ‘should’ /ʃʊd/, ‘would’ /əd/, ‘has’ /hæz/, ‘the’ /n̪ə/ (assimilated), ‘the’ /ðə/, ‘the’ /ðɪ/ and ‘to’ /tʊ/. Only one function word (should /ʃəd/) was correctly recognised by all of the participants.

There was some disparity in the numbers of incorrect recognitions of the words being reduced in different ways. This was noticed mainly in cases where incorrect recognitions of strong versus WFs were compared. Generally, the function words which underwent fewer modifications resulted in a lower rate of incorrect recognitions, and vice versa. For instance, some pronouns and auxiliary verbs reduced with the elision of the initial /h/ and the reduction of the central vowel had a higher rate of incorrect recognitions than those where the weakening was limited to vowel reduction:

- he /ɪ/ : 39-100% incorrect recognitions – he /hɪ/ only 3% incorrect recognitions
- has /əz/ or /əs/: 95-100% incorrect recognitions – has /hæz/ only 3% incorrect recognitions
- had /əd/ 89% incorrect recognitions – had /həd/ 66% incorrect recognitions
- his /ɪz/ 100% incorrect recognitions – his /hɪz/ 95M incorrect recognitions

The pronoun *they*, which does not have a WF, also had a high rate of incorrect recognitions reaching up to 100%. Despite having no WF, this pronoun undergoes a special type of reduction through the elision of the final element of the closing diphthong; /ðeɪ/→/ðe/.

Similarly, other cases demonstrated that the number of incorrect recognitions was higher in functors pronounced in their WFs compared to the cases where the SFs were used.

It should be recalled that, as the main focus in testing was on the WFs, the pre-test did not include instances of strong pronunciations for all function words. Despite this, in all the cases where a SF happened to have occurred in the test, the rates of incorrect recognitions were very low compared to those cases where the WF was used:

- you /jʊ/ or /jə/ 29-97% incorrect recognitions – you /juː/ only 3% incorrect recognitions
- are /ə/ 79-82% incorrect recognitions – are /ɑː/ only 3% incorrect recognitions
- the /ðə/ 50% incorrect recognitions – the /ðɪː/ only 16% incorrect recognitions
- has /əz/ 95-100% incorrect recognitions – has /hæz/ only 3% incorrect recognitions

In this regard, the findings, so far, are compatible with the results reported by Henrichsen (1984) and Ito (2001; 2006) which suggests that the presence of CS features affects the decoding of speech by non-native listeners.

Two of the function words present in the test were not really problematic across most of their WFs occurrences. In three of the four cases where the WF of the article 'the' was used, the rate of incorrect recognitions did not exceed 26%. Similar observations were found for the auxiliary 'is' where the highest rate of incorrect answers did not go beyond 39% in half of the WF occurrences. On the other hand, the rates of incorrect recognitions for the rest of the test items all showed a great disparity from one context to another, and form one reduction pattern to another. As displayed in table 22, the different rates of the following items were a case in point: 'you', 'was', 'and', 'a', 'to' and 'for'.

Whether or not the additional feature of assimilation was involved in the modification of the function word does not appear to have affected the numbers of recognitions across the different instances. In some cases, the numbers of incorrect recognitions for assimilated WFs did not differ much from those where there was no assimilation (e.g. 100% incorrect

recognition in ‘has’ /əz/ versus 95 incorrect recognition in *has* /əs/). In others, however, even in those instances where assimilation was involved, incorrect recognition rates were lower than in the cases where no assimilation occurred (e.g. 97% incorrect recognition in ‘was’ /wəz/ versus 37 incorrect recognition in ‘was’ /wəs/). This being said, the qualitative analysis of the participants’ mishearings has revealed a noticeable effect of assimilation on decoding the function words out of their surroundings. Table 23 shows examples of such effects with common answers provided by the test takers. The effects of the other features will be highlighted in the discussion of the other types of mishearings which resulted in the problems of LS and WFs confusion.

Function Words	Examples	Transcription	Assimilation type	Common Mishearings
and	but nice and fresh	/bəˀnaɪsəm̩freʃ/	place assimilation	but icing fresh
was	…in July was stolen	/dʒʊlaɪwəstəʊlən/	devoicing	July stolen
that you	that you can talk	/ðətʃəkəntɔːk/	coalescence	chicken talk
is	there is some soup	/ðərɪsəmsuːp/	devoicing	this soup this some
is	it is starting	/ɪtstɑːtɪŋ/	devoicing	it starting
them	send them anything	/sen̪əmenɪθɪŋ/	manner assimilation	send him anything
as	…know her, as simple as…	/nəʊə.əsɪmpləz/	devoicing	know her simple as…
have	might have found	/maɪtəfaʊnd/	devoicing	might find
had	had yours	/hədʒjɔːz/	coalescence	hedg yours

Table 23: Effects of Assimilation on Weak Forms' Decoding

In cases where WFs were not recognised, the test takers often did not leave the blanks unfilled, but rather transcribed what they heard or what they inferred from the context. Such answers represented examples of mishearings that were analysed to see what major effects a given non-recognised WF had on decoding the rest of the utterance, namely the

surrounding content words. The analysis revealed two main patterns: 1) WFs were recognised as parts of other longer English words, and sometimes non-words, or; 2) lexical gaps were deleted both before and after the WF and the neighbouring content words were merged together. In both cases, the mishearings show evidence of LS problems; deleting word borders or inserting others irrespective of what the input suggests.

Function Word	Example Utterances	Common Mishearings	Description
has	-I think Mark has been … -The number has been …	I think Marks The numbers *have* …	*Lexical gap deleted before the function word. The function word was recognised as an inflected form.*
was	You knew their house was burgled	..knew their houses burgled	
him	I asked him for some money	I Asking for some money	
and	He wants to come and see use at home	He wants to coming see…	
is	…the boat is useless	…the boats useless	
are	They are for Jane –	Therefore Jane.	*Lexical gap deleted before and after the function word. The function word was recognised as a part of a longer word.*
a	Send Frank a card	Sen Franka – Senfran Cacards – Senfraka	
as	… a couple of annoying people, as simple as that	..people simples / simpliza / simplice	*Lexical gap deleted and the function word was recognised as a part of a new non-word.*
have	You have already had yours	Evoloy handeons – you volong	
can	…you can talk to the person you are with…	Chicken talk to …	*Lexical gap deleted and the function word was recognised as a part of a new word.*
there	There is a comma after that	This a comma /these a comma	
have	You have already had yours	You eventually / even	
has	-I think Mark has been paid quite a lot of redundancy money	I think Mark Spin / spend / explained	
them	…and I did not send them anything	… and I did not understand anything.	
as	… a couple of annoying people, as simple as that	…people simplest that	
to	…you really get to know her	… you really get an hour	
is	There is a comma after that	This a camera - does it camera	
us	The experiment sounds absolutely terrible to us today	..terrible yesterday	
of	…people out of their money	… people after their	
must	There are some new books I must read	…books science read / silence read	
they	They are second hand	Their /there second hand	
her	… you taking him to see her	…taking him this year	

	…you really get to know her	…get an hour – to our	
are	They are for Jane. There are some new books…	There /their for Jane There some new books…	
his	Read his book and write some notes.	Read spoke and write …	

Table 24: Lexical Segmentation Problems Caused by Weak Forms Misperception

Table 24 shows some prevalent patterns of LS problems resulting from the misrecognition of function words. A common pattern is when a lexical gap was deleted before the function word, and the latter was recognised as a part of either the content or the function word that precedes it. The result of the modification was that the function word was considered as:

- a part of an existing content word, rendering it into an inflected form (e.g. The number has been →the numbers been/ *are being*),
- a part of a totally new content word (e.g. that you can talk → chicken talk); or,
- as a part of a non-word (as simple as that → simpliza).

Another pattern was when a lexical gap was deleted both before and after the WF and, thus, the latter was recognised as a part of a longer word:

- they are for →therefore.

Unlike what the cohort theory suggests in terms of the activation of new candidates as a result of the perception of the first sounds of a word (Marslen-Wilson & Tyler, 1980), some mishearings in the test occurred despite that there was no sound matching on word onsets. Such mishearings did not start with a sequence of sounds included in the original item or the stream of sounds, while the end part of the word did. In other words, it was the last part of the word which was perceived, and the test takers activated, or perhaps inferred, the beginning of a mistaken word based on it. If we take the example of the phrase "to us today", which was perceived as 'yesterday', we find that the beginning of the transcribed

word has no matching with the stream of sounds in this utterance. Thus, the first sequence of sounds was the cause of activating the word ‘yesterday’. The latter was rather mistakenly inferred based on the last sequence of the sounds after deleting borders around the function word ‘us’. This can also be interpreted as a cognitive lexical effect (Rost, 1990) since the listeners opted for modifying the result of the auditory perception (inserting segments and syllables) based on expectations rather than what the input suggested.

Whatever the pattern was, as LS problems resulted in a modification to the wording of the utterance, the message transmitted through the utterance was consequently affected. However, as the following examples show, the extent of this effect in the test was relative:

- **A:** “He wants to come and see us at home” transcribed as “He wants to coming see us at home”)
- **B:** “and you really get to know her” transcribed as “and you really get an hour”.

There is a difference between what the original utterance suggested and what the testees transcribed in both of the above examples. However, the extent to which the basic utterance meaning was affected varied significantly. Although the conjunction ‘and’ and the verb ‘come’ in example **A** were merged to create an inflected form, the basic meaning was only slightly affected, and the listener could still get the intended meaning. On the contrary, the misperception of the function word ‘to’ and the shifting of lexical gaps in **B** led to a major change to the basic meaning. Accordingly, it could be said that the effects of misperceiving function words could potentially have devastating effects on the utterance’s basic meaning.

One possible interpretation of the second example (**B**) is that, as a consequence of lexical effects, the listeners modified the results of the auditory analysis to meet their own expectations. The phrase “and get an hour” which only partially matched the auditory input may logically have reflected a cognitive effect. However, any analysis should not lose sight

of the effects of the phonological modifications. In this example, a complex modification pattern led to a major effect on both meaning and form: the simplified elision of /t/ in "get to" made it unlikely for the listeners to recover from its effect. The vowel /ə/ in /tə/ was linked to the first sound of 'know' and another lexical gap was mistakenly inserted after /n/, resulting in "get an hour" (instead of "get to know her"). We can also notice that the diphthong /əʊ/ was ignored in favour of the new word 'hour' when another lexical gap was deleted before 'her'. The point is that, although the mishearing in this case may be explained as a result of the cognitive lexical effects, the phonological modifications that the words underwent should not be ignored. We will shortly return to this issue when we discuss the mishearings resulting from the feature of assimilation.

The other problem discussed in chapter two in relation to the function words' decoding is that of neutralisation, i.e., when two or more function words share similar WFs. There were some cases where most of the test takers confused between neutralised function words:

- the function words 'his' and 'is' which share the WF /ɪz/
  - e.g. He stood his gun → He stood is gun
- the function words 'had' and 'would' which share the WF /əd/.
  - e.g. ɪf they had searched → ɪf they would *search* ˙.
- the function words 'her' and 'a' which share the same WF /ə/.
  - e.g. I wanted her to stay → I wanted a to stay

Apart from these, the analysis of the participants' answers did not reveal other examples of function words being confused with other ones which share the same WF. Confusion, however, was much more evident in WFs which share at least one single sound. The phonological context also contributed to this. Table 25 gives examples of this type.

Function Words	Confused with…	Utterances
and	on can	He wants to come **and** see us at home. Read his book **and** write some notes.
there	their that they this	**There** are some new books I must read I don't care if **there** is no imported caviar. **There** were a lot… **There** is some soup in the fridge
was	is will	…he **was** taking part in this... Algeria **was** split …
would	could	…it **would** get better
shall	'should'	I **shall** take as much as I want
as	is– (assimilated)	….**as** simple as that
might (have)	may to	I **might** have left my keys…
him	them (in) his	I asked **him** for some money …taking **him** to see her
has	is – was	The number **has** been engaged

Table 25: Mishearings of Partly Similar Weak Forms

As Table 25 shows, some mishearings may partially resemble the original version. So, it could be said that wrong recognitions may be triggered based on the partial decoding of the stream of sounds in the utterance. However, the analysis of the mishearings revealed another pattern where confusion in transcribing the function words was not sound induced, and where the linguistic context also constrained the recognition process.

The participants showed a tendency to fill in the missing words with ones which demonstrate a reliance on the surrounding words. Such mishearings show that the main basis for the inferences was the syntactic role that a given structure word had in a particular position within the utterance regardless of the sound input. This is because, on the one hand, such guesses did not match the exact words or the sequence of sounds uttered in the recording; on the other hand, however, they were correct and logical in light of the linguistic context of the utterance. In other words, instead of leaving the blank of a non-recognised word unfilled, the test takers opted for the best fit word by taking into consideration available syntactic information. In the following examples, syntactic effects manifested themselves in mishearings where a misperceived function word was replaced with another word (*a*

syntagma) that, in most of the cases, shared the same word category despite being just about completely different in pronunciation:

- You will come /was *coming* over for dinner soon (original "*must* come").
- If he had searched more carefully …(original "*you* had ").
- I wanted it/you/him to stay. (original 'her').
- You should talk to the person… (original "*can* talk").
- It is a reproduction of the digital file (original "*a* digital file").
- Read this book and write… (original '*his*').
- Apart from the couple of annoying people (original "*a* couple of").

The first of the examples shows that syntactic information was deployed by most of the listeners to compensate for the poor decoding ability, especially that most of them failed to recognise the WF /məs/. The linguistic context suggested that the word preceding the verb 'come' should be an auxiliary or a modal verb. Given this, many possible candidates could be activated at the same time, but the choice of the modal 'will' out of the wide range of possibilities was far from being arbitrary. One possible explanation is that, in addition to the syntactic effects, schematic effects also contributed to the limiting of the number of candidates to only one. The presence of the adverb 'soon' which was not blanked out in the answer sheet could have hinted the possibility of having a model verb that denotes the future. It is worth noting that, when 'must' was replaced with 'was', the following content word was also misheard as 'coming' (you must come over→ you was *coming* over), which gives further evidence to the role of syntactic effects and/or lexical effects after a failure in decoding function words. This issue will be further discussed when we analyse the effects of the TD and the BU processing modes.

5.1.2. **Effects of Assimilation**

The analysis of the mishearings of the WFs in which assimilation was involved revealed that the latter had a prominent effect on the test takers' decoding process. In this section, the results of the analysis of the other instances of assimilation, particularly those present in the dictation test, will be presented and discussed. It is important to recall that, with reference to the results of the pre-test presented earlier, the participants' scores obtained in the dictation test were very low. That is to say, the participants failed to decode most of CS features in the pre-test, including assimilation.

In addition to the mishearings, which represented unsuccessful attempts to decode a given utterance where the assimilation feature was present, incorrect answers also included those instances where the test takers left blanks on their answer sheets. Concerning the mishearings, the analysis revealed that both LS and WR problems did occur as a result of the different types of assimilation. This was confirmed after analysing the effects of the assimilated sounds on decoding the content words and on inserting or deleting lexical gaps. In what follows, we will report the common mishearings for each test item, and discuss the causes in the light of the different types of the obligatory phonological processes: Coalescent assimilation, place assimilation, voice assimilation and manner assimilation. The analysis revealed many instances of mishearings including all the sub-types of segment or syllable modifications categorised by Rost (1990) (insertion, deletion and transversal) in his discussion of how cognitive effects may constrain speech perception. It should be made clear that, in addition to the missegmentation and WR problems, the main focus of the analysis will be on those mishearing instances which show a certain effect of the phonological modifications on the participants' decoding of the test items.

The following are examples of common mishearings resulting from coalescent assimilation:

1. You have already *had yours*.
    - Examples of common mishearings: hand use, hand juice, what Jose
2. She *had university* students.
    - -Examples of common mishearings: *changed/ hedj / change / shed/John's …*university students.

As the examples show, the result of assimilation is a new sound which shares the characteristics of both neighbouring sounds at word borders. In both of the above examples, the assimilation of the conditioning sounds /d/ and /j/ resulted in the sound /dʒ/. For example 1, the participants' answers demonstrate that most of them were unable to recognise the phonological modification occurring at word borders. The assimilating sound /dʒ/ was recognised as a part the second word while the first one was not, or was partly, affected (had /*hand* Jose, Juice). The mishearing in this case represents the phonological error of insertion (Rost, 1990) as the result the auditory analysis included the insertion of new segments or syllables. The second example shows an instance of syllable insertion but in the opposite direction. The second word was correctly decoded (university), but the assimilated sound affected the recognition of the first one; the participants transcribed a new word or a non-word (changed, hedj), the beginning of which did not match the sounds in the original utterance. In terms of the cognitive effects, some mishearings (e.g., she had university students→ Charles university student***s***/ should university students/ She is a university student) show evidence of a lexical effect which is, according to Rost (1990), a modification or a reinterpretation of the auditory input based the listener's expectations, rather than what the auditory input suggests.

The following are examples of coalescent assimilation where /t/ and /j/ meet at word borders resulting in /tʃ/:

1. You taught yourself French.

- Examples of common mishearings: You touch / touched yourself /search/told just

2. I cannot beat you at this game.
    - Examples of common mishearings: I can/cannot bechaire/ be chaired/bicha/ bitcher/ be sharing *at this game.*
3. So that you can talk to the person you are with…
    - Examples of common mishearings: so chicken *talk to the person*…

In all of the examples, coalescent assimilation resulted in the activation of either a new word or a non-word. That is, the problem of WR resulted from the failure to recognise the modification that the word borders underwent. Assimilation in the first example affected the recognition of the first word (taught). Even the vowel /ɔː/ was replaced with /ʌ/ in favour of the new word *touch* and, thus, the mishearing can be interpreted as the result of a lexical effect. This is because it is very unlikely that the participants confused /ɔː/ with /ʌ/ which have totally different parameters. So, this is another example where a problem in decoding the last part of the word affects the recognition/decoding of the first part. In example 2, assimilation caused a LS problem as the assimilating sound /tʃ/ was recognised as a part of a longer non-word (e.g. bechaire), or a lexical gap was mistakenly inserted (e.g. be chaired). This is also evident in the last example where a new word was transcribed after deleting lexical gaps, and the assimilating sound /tʃ/ activated a new item which is not part of the utterance (that you can →chicken). The listeners perceived /tʃ/ as an onset of a new word after failing to recover from the assimilatory effect.

It could be said that some lexical effects come *due to* the failure to decode a given CS feature, and they cannot be interpreted without analysing the results of the phonological modifications. The above examples of mishearings are cases for lexical effects on WR. The listeners expected and transcribed words which do not match, wholly or partially, the stream

of sounds in the input. But the fact that the mishearings in most of the cases have traces which are the results of segments or syllables being modified shows that the phonological modifications and the lexical effects are, most of the time, as inseparable as two sides of a coin. Based on the analysis, it could be said that phonological modifications provided reliable data for interpreting the mishearings as they do have a role to play. As it is often difficult to decide whether a given mishearing constitutes a lexical effect, a syntactic effect, a schematic effect, or an interaction between all of them, the analysis of phonological modifications may provide important clues.

Similar observations can be made about the effects of place assimilation. Table 26 presents examples of mishearings for this type.

Assimilated Sound	Utterance	Common Mishearings	Assimilation Effects
/t/ before a bilabial sound	They shot bears.	they shop bears – cheap bears – shape wars – chop beart – they're charpners – chap lars – *shop ears*	New words activated
	They cheat people out of their money.	they cheep / ship / chop / ship / choop sheep….. *people*	New word activated – non-word activated
/t/ before a velar sound	… and art galleries	(not recognised as /k/) all galleries other galleries	No visible effect, but the word was not recognised at all
	I taught classes this morning.	took classes, took courses – talk classes	New word activated (affecting the decoding of the following item)
/n/ before a labiodental	…not spicy but nice and fresh.	But Some fresh – is some fresh	Lexical segmentation
/n/ before a bilabial	The sun burned my neck.	the sumbem / resembled /sunbon somber / december sumber … *my neck*	New words activated Non-words activated Lexical segmentation
	…pay for that rather than buying a digital file.	…them by	New word activated

Table 26: The Effects of Place Assimilation on the Participants' Decoding

Table 26 shows instances of regressive place assimilation form the pre-test. In all of the cases, assimilation resulted in problems of WR, while LS was evident in only two examples. Concerning the assimilation of /t/ before bilabials, the problem of WR was evident across the two instances, before /b/ and /p/. The examples of mishearings show that the participants failed to recognise the first word which includes the assimilating sound (e.g. shop bears, cheep people), or both neighbouring words (e.g. shop ears). In the second case, the assimilated sound led to the activation of a new candidate instead of the one included in the utterance. The same observation applies to the sound /t/ when it occurs before a velar; /t/ in *taught* was perceived as /k/ and most of the participants misrecognised both neighbouring

words as an effect of assimilation (e.g. took courses). Here again, the test takers confused the vowel /ɔː/ with /ʊ/: taught →took.

The analysis of the mishearings confirmed that misrecognition and missegmentation errors also occur as a consequence of the assimilation of /n/ before bilabials. Unaware of the assimilation process, the test takers failed to recognise the phonological modification, and transcribed exactly what they heard. Lexical gaps were deleted after the assimilating sound /m/ and the participants' mishearings were either: 1) new words which show a reformulation of what is in input (December my neck?), or 2) non-words resulting from a direct mapping of sounds to words (eg. The sumbem/sumber my neck). For the new words, the appearance of the assimilated sound /m/ in the result of the auditory processing suggests that a slight reformulation to the stream of sounds took place as a result of a lexical effect. Writing a non-word, on the other hand, may imply lack of confidence about one's decoding skills. It is also very likely that the listeners' uncertainty about their linguistic knowledge led them to think that what they heard was not within their vocabulary knowledge and, consequently, they wrote the sequence of sounds on the belief that it was an unknown word.

Both voice assimilation and, to a lower degree, manner assimilation resulted in recognition and segmentation problems in the participants' decoding of the utterances in the test. When a given sound was devoiced, the listeners failed to segment the words, especially if the modification involved a pair of a voiced and a voiceless consonants being harmonised together. As discussed earlier, the effect of voice assimilation was clear on the decoding of the function words like *has, is,* and *was* which were either misrecognised or confused due to the devoicing of the final /z/ sound. The analysis of the participants' transcriptions of the dictation test revealed similar results:

- The boat is useless without the oars.

  Example of common mishearings: The boats useless

- It has seen suppression, revolution…
  - Examples of common mishearings: It seen / It is seem/ It seems

Other mishearings, however, revealed that devoicing at word borders made it difficult to recognise the neighbouring content words while the function word was correctly decoded:

- Algeria was split into departments
  - Examples of common mishearings: was planting – was please – was plating – was pletting – was pattern

These mishearings show that the participants were able to decode /wəs/ as a modified version of /wəz/. However, this is not an indication of a successful detection of the assimilatory effect at the word border because, simply, the assimilating sound in the neighbouring word was not decoded. One possible reason was that there is a complex modification which starts by assimilation and is then followed by elision through simplification where the /s/ sound was lengthened to indicate the presence of two, rather than one, sounds. Most of the listeners, however, were unable to notice this.

The three items which involved a phonological modification through manner assimilation showed a relative impact on decoding.

- …taking part in this TV show.
  - Example of common mishearings: taking part in TV
- There is some soup in the fridge.
  - (no mishearings)
- …I did not send them anything.
  - Examples of common mishearings: send him / send anything/understand anything

All of the above items involved the modification of the manner though which the initial sound /ð/ is pronounced so as to accommodate with the previous sound, and to insure a smooth transition across the word borders. All of the test-takers failed to recognise the word ‘this’ in the first example. It could be said that the modification of the initial sound and the inability to recognise the assimilatory effect made it impossible for the listeners to activate the word in their lexicon. In contrast to this, all of the test takes correctly decoded the word ‘the’ in the second example despite that it underwent exactly the same assimilatory process. This may be attributed to the comparatively high frequency of articles which renders them more predictable and, thus, easier to decode under different modification conditions. The last item in this category showed a complex modification process of elision followed by manner assimilation. While some answers showed evidence of a lexical effect where the input was modified to match the listener’s expectations rather than the input (e.g. I did not understand anything), some answers demonstrated a clear effect of manner assimilation. The replacement of the original pronoun ‘them’ with ‘him’ (e.g. I did not send him anything) may have taken place due to two possible reasons: 1) as the initial sound /ð/ was misperceived, the similarity between the words created confusion (/em/ Vs. /ɪm/, and this was supported by the fact that, 2) the two words shared the same word category and, in the absence of sufficient context, they could both be predictable and linguistically appropriate within the utterance.

5.1.3. **Effects of Elision**

The analysis of the cloze test answers in the pre-test highlighted some effects of consonant elision on the decoding of structure words, mainly the /t/ and /d/ sounds in some functors, and the initial /h/ in some pronouns and auxiliary verbs. Due to some reasons, this process of elision in function words should be disentangled from that through which consonants are deleted in content words. At first, it is inappropriate to attribute a

misrecognition of a function word like ‘must’ (/məs/) to the process of elision only because the modification involved not only consonant elision, but also vowel reduction. Tracking the effects of elision in this case is a challenging task. On the contrary, such a combination of modification processes is not found in content words and, thus, the effect of elision at this level can be isolated. Second, the rules which govern consonant elision in content words are generalizable to any word, including function words. However, those rules which govern consonant elision in function words are less systematic, and they do not necessarily apply to content words. Accordingly, to trace the real effects of elision, it is necessary to consider their occurrence in content words. Table 27 displays the participants’ common mishearings of words involving elision, and the associated problems.

Elision Type	Test Item	Common Mishearings	Description
/d/ elision	Send Frank a card	Senfrank Sun fake San franca	-Segmentation failure -Word segmentation: New word resulting from consonantal elision
	The sun burned my neck	Some by my Some born my Sumber my December my Sumber my Sumbern my	-LS failure -WR problem: New word resulting from a complex modification which involves both elision and assimilation
	Send them anything	//	//
/t/ elision	I cannot beat you	I can I come – came Can't	-WR problem: New word resulting from a complex modification which involves both elision and assimilation
	If they had searched more…	If they search more	WR problem: inflected form not recognised
	Get to know her	Get an hour	WR problem: preposition not decoded as a result of elision through simplification
	… docked the deep blue sea	Don't the deep Judge the deep Touch – George – Inch	-WR problem: New words in accordance with the original Rhythmic pattern
/ɪ/ second part of a closing diphthong	They are for Jane	Therefore Jane The are for Jane	-LS problem resulting from elision. -WR problem

Table 27: Common Mishearings of Words Involving the Feature of Elision

In addition to the elision of the *vowel* /ɪ/, the table includes the two most common types of elision in spoken English; the elision of /t/ and /d/ occurring at the end of words, in the centre of three consonant clusters. Similar to the results of the mishearings concerning the features discussed earlier, a close observation of the mishearings in the table reveals that LS and WR problems did occur as a result of elision.

Concerning the elision of /d/ in clusters, most of the participants failed to recognise the modified version in the first example in the table, and their answers included either new words, or the original words without the final sound (e.g. send Frank→ Sen Frank). The second example in the table represents a complex modification process involving both

elision and assimilation. The participants not only failed to detect the elided sound but also could not recover from the assimilatory effect in the preceding gap (sun burned my→ some born my). The results of the auditory analysis represented by the mishearings showed evidence of LS as well as WR problems. Lexical gaps were either deleted or erroneously inserted, and most of the participants tended to write exactly what they heard without being able to spot the phonological modifications occurring at word borders. With such failure, lexical effects extended to the rest of the utterance, showing different expectations of the listeners which did not match the auditory input (e.g. Send Frank a card →Sun fake account/ San Francisco).

The mishearings resulting from the lexical effects, which differed completely from the input, may suggest that the problem did not lie at the level of the perceptual processing phase. This is because it is possible that the decoding of sound segments was first done correctly but eventually ignored in favour of the listeners' expectations. This is supported by the fact that many mishearings showed a total deviation from the actual words in the original utterances. However, wrong decoding of segments could also be the result of lack of confidence in one's decoding skills. Minor modifications to the input (e.g. send →sun/san) raised serious questions about the learners' decoding ability at the segmental level which was done during the perceptual processing phase. It is very likely that the erroneous processing of segments occurred due to weak ability in recognising the segments and differentiating between them by the learners. In the analysis of the participants' mishearings, it was found that, apart from failure to decode the consonants which were subject to CS modifications, most of the perceptual errors at the segmental level included the replacement of vowel sounds with others.

As Table 27 shows, problems of segmentation and recognition also occurred when phonological modifications involved /t/ elision. The most problematic item was that which

included a combination of CS processes (assimilation and elision), and which led to both segmentation and recognition errors (e.g. I cannot beat you →I come beat you). Even with recovery from the assimilatory effect, distortion of the utterance meaning took place due to elision alone (I cannot beat you → I can beat you). For the rest of the items, the participants misperceived the elided words and mistakenly inserted, deleted or shifted lexical gaps. The special type of elision where the pronunciation of the final /t/ was prolonged at the word border (get to know) resulted in a major shift not only in the words and the gaps but also in the meaning of the phrase (most of the transcriptions included *get an hour*). As an inflected form, the final /t/ in verbs was also misrecognised, leading to a modification in the tense of the verb (If you searched →If you search).

Not only consonant elision, but also vowel elision was found to cause decoding errors. In the last item in the table, the elision of the second element of a closing diphthong caused either a segmentation error (they are for → therefore) or a WR problem (they are for → the are for). It should be mentioned that the errors of segmentation logically implied WR errors. Mistakenly shifting, inserting or deleting gaps will consequently modify the shape of the surrounding words. However, errors of WR do not necessarily lead to wrong segmentations. Most of the time, the test takers failed to recognise certain words, but their answers were compatible with the rhythmic template of the utterance regardless of the extent to which the transcribed words differed from the original ones.

5.1.4. **Effects of Liaison**

Recognising linking sounds as phonemic was found as one of the major obstacles in decoding the utterances by the participants. The analysis of the answers revealed that the modifications made at word borders to make smooth transitions for the speakers created serious obstacles for the listeners in this study in the processes of depicting word borders and activating the exact lexical items.

Linking Type	Test Items	Common Mishearings
Linking /w/	Are you into golf	are you wearing to golf/ win the golf are you want are you win to
	Do I owe you anything?	do why do are do aware
	Do I owe you anything?	you were anything want anything
	…starting to ache	to wake
	…terrible to us today	//
Linking /j/	…without the oars	the yours the others the yalls
	…come and see us	…yesterday
	…over the ears	over the years the ears (4)
	…coming up to see her	this year
Intrusive /r/	There is a comma after that.	camera after that comer after that
	Law and order	lore and order/ nor and order
C+V linking	..split into departments	splitting plating planty splitting today

Table 28: Common Mishearings of Words Involving the Feature of Linking

Table 28 highlights the most common mishearings among the participants' answers to the test items which included the different types of linking. The analysis of the mishearings provides a logical explanation as to why the participants' scores in the dictation test were very low as far as the CS aspects are concerned. The different linking processes exerted prominent effects on decoding the utterances in the test. By inserting a consonant *phoneme* that did not exist in the original utterance, most of the test-takers fell in the trap of thinking the different linking sounds could be:

a- the beginning of a new word for which they have to search in their lexicon. In this case, most of the participants transcribed new words with respect to the surrounding sounds, or even wrote non-words.

b- the centre of a word for which they had to define the borders within the sounds in the input by (mistakenly) inserting, deleting or shifting lexical gaps.

The two common types of linking discussed in chapter two were found to cause decoding errors. Most of the test-takers were confused by the linking sound /ʷ/ which was commonly perceived as phonemic. By failing to notice that the pronunciation of the linking sound is not as elongated as when the sound is phonemic, or due to lack of unawareness of such modification process, the testees opted for transcribing the linking sound as a phoneme which constitutes an element in an individual word. The mishearings also suggested a possible thematic effect on decoding, especially in the first example "*Are you into golf?"*. As most of the test takers transcribed the section which includes linking as "you win to", it makes it possible that the word 'golf' helped in activating the words related to sports in general. This also raises questions about the linearity of processing by the participants. Did the listeners in the study decode the word 'win' first and then confirmed their decoding after reaching the last word in the utterance (golf)? It is possible that the word 'win' was activated only after the decoding of the word 'golf'? It is difficult, in this example, to confirm in what order the processing happened exactly, whatever the processing pattern is, the aspect of linking did contribute to the two main errors in decoding (WR and LS).

Unlike the first and the second test items presented in the table, the third item of /ʷ/ linking includes only one common mishearing made by all the participants. All of them wrote the linking /ʷ/ in "It is starting to ache" as phonemic, resulting in 'wake'. One would say that the two options are possible given that no context was available to decide between the two. However, it should be reminded that, from a productive point of view, the pronunciations of the linking /ʷ, ʲ/ and the phonemic /w, j/ sounds are acoustically different (Hunt, 2009); the former are pronounced more slightly and they are less elongated (Cruttenden, 2014). In addition, at least for the participants in this study, the fact that none

of them (38 participants) wrote the version in which /ʷ/ is recognised as *linking* raises the issue of awareness concerning this CS aspect among them. No one was able to notice that the two phonemes /ʊ/ in 'to' and /eɪ/ in 'ache' may be in hiatus and that the /ʷ/ was inserted only to fill the gap and link them together. This suggests that the word candidates which involve the modification processes (including linking, assimilation and elision) are seldom activated due to a lack of awareness of such modifications and a poor decoding ability. From another perspective, if we consider the fact that all the transcriptions included the same word reduces the possibility that the participants were able to activate another candidate in addition to the word 'wake'. That is to say, their capacity to consider more than one option or all of the possible options and to hold them in the STM for an evaluation of fitness in the context is very limited.

As far as the /j/ linking is concerned, errors of decoding did occur due to the wrong recognition of this linking sound as being phonemic. Different words which include the initial /j/ sound were activated and transcribed by the participants as a result of the process of linking. While some of these wrong words were syntactically and/or semantically acceptable within the utterances where they were used (e.g. …coming to see her → …coming this year), others were not (e.g. …without the oars →…without the yours). A closer look at the mishearings reveals that major changes in the meanings of the utterances occurred as a result of the failing to decode the linking aspect.

The other type of linking involves inserting the intrusive /r/ between word borders where vowel sounds meet. While this type of modification influenced the way the words were decoded and recognised by the participants, it did not have an effect on the segmentation process of the utterances. Concerning WR, the intrusive /r/ was perceived as a single phoneme by almost all of the test takers. In the two test items, the participants linked the intrusive /r/ to the first word rather than the second. Other observable effects included

the modification of the auditory input in a way that favoured the new word shapes imposed by the new sounds. For instance, after linking the /r/ to the first word, the majority of the participants transcribed “comma after” as “camera after”. The replacement of the vowel sound /ɒ/ with /æ/ in this case suggests, in addition to an obvious effect of CS, a lexical effect on word access as the listeners abandoned the auditory input, and the end product of processing reflected the listeners’ expectations.

Defining where words start and where they end was found problematic in the case of consonant to vowel linking. In the test item where this type of linking occurred, lexical gaps were deleted and the words were merged together (e.g. split into departments → splitting). In many examples of mishearings, the syllables at the word borders were restructured, and lexical gaps were shifted or deleted (e.g. splitting to departments/ splitting today part/splitting to day pat among). As discussed earlier, errors of segmentation logically lead to WR errors, and this is obvious in the mishearings in this type of linking. These errors altered both the utterance shape and meaning.

5.2. Processing Modes and Strategy Application

The participants’ mishearings were analysed to find out any common patterns that have traces of the application of the TD or the BU processing modes. The main focus of such analysis was to explain how systematic patterns occurred and to find out whether a link could be established between the listening errors and the listeners’ choice and use of the processing modes. Resorting to one specific mode under certain constraints and in a systematic way may reflect a strategic choice by the listeners, especially when decoding problems arise or when the other mode seems more appropriate. However, as it will be shown, the extent to which such strategic choices could be successful cannot be well-defined.

Two main patterns of errors raised in the previous section show a tendency from the participants to approach the input in a BU fashion. The first of these is the *direct matching of sounds to words* whereby lexical selection is done more or less arbitrarily based on the mental representations which result from the linear decoding of sound sequences. The BU matching process implies that the listeners scan these sequences to find any possible words irrespective of the words and gaps present in the original utterance. Consequently, the utterance's original gaps will be deleted or shifted resulting in a distortion of the wording, despite that the sounds are preserved or slightly modified. The following categories of mishearing errors caused by the different CS modifications demonstrate the matching process:

- New words were formed by merging two or more words:
  - they are for → therefore (elision)
  - Send frank a card → senfrank a card / senfran cacard (elision and assimilation)
  - Split into → splitting to (consonant to vowel linking)
  - So that you can talk → so chicken talk (assimilation and reduction)
- Lexical gaps inserted around short words imbedded within or across longer ones:
  - the sun burnt my → *De*cember my (assimilation and elision)
  - but nice and fresh → but some fresh (assimilation)
  - get to know her → get an hour (elision and reduction)
  - I cannot beat you → I can beacher /bitcher (assimilation and reduction)
  - You taught yourself → you touch a self (assimilation)
- Recognising linking, assimilatory and intrusive sounds as phonemes:
  - Do I owe you → Do why owe

- The sun burnt my neck → the sumbem /December my neck (assimilation and elision)
- I cannot beat you → I come beach (elision and assimilation)
- There is a comma after → there is a camera/comer after (linking)
- Law and order → lore an order (linking and elision)
- Rather than buying →rather them by (assimilation)
- They shot bears →they shop pairs/ they shop ears (assimilation)
- You taught yourself → you touch yourself (assimilation)
- They cheat people → they cheap /cheep people (assimilation)

These errors show how the phonological modifications may affect the way in which words are recognised in speech. The modified shapes of words created confusion for the participants in the process of lexical access. Based on such modifications, new words were activated and accessed instead of the ones present in the input, resulting in a relative modification to the wording and the basic meaning of the different utterances.

Minor changes made by the listeners to the sounds which precede or follow the phonologically modified ones have some implications as to the participants' decoding process and their linguistic knowledge. As far as the knowledge of the sound system is concerned, the sound replacements suggest a weak ability to recognise segments and to differentiate between them. This is evident mainly in the cases when the replacement involves some sounds which, like minimal pairs, share similar characteristics.

On the other hand, when the decoding errors include the replacement of one sound with a totally different one, confusion cannot be considered the cause of the error. Given this, one would question the learners' confidence in their listening skills and the reasons why they tend to modify the result of their auditory processing. Many mishearings show that some portions of the input were modified only to satisfy an agreement with other parts of

the text, most of which were erroneously decoded CS features. In other words, errors which took place because of failure to decode a given CS feature led to *decoding* the surrounding sound which were not input-based. This suggests lack of confidence from the participants to decode segments which, combined with lack of awareness concerning the CS features, created a serious barrier in the processes of WR and LS.

The second type of errors has to do with the LS strategy discussed in chapter two, which holds that accented syllables in the stream of speech are more likely to be the beginnings of words since most of the words in English receive a primary stress on the first syllable (Cutler & Carter, 1987). This implies inserting lexical gaps before accented syllables and deleting them before unaccented ones. Since segmentation in this case is done based on acoustic cues, the strategy principally follows a BU mode. Because our main concern in the analysis of the mishearings was directed to the CS features, the main focus was on the function words due to their weak and unaccented forms.

The analysis revealed that a number of errors of segmentation occurred due to the omission of lexical gaps before unaccented syllables. WFs were merged with content words resulting in three main patterns. The first pattern includes instances of missegmentation where lexical gaps were deleted before structure words which made up inflected forms of different content words:

- Function words ending in /z/ or /s/ (assimilated) were recognised as inflections (plural form or the third person ‘s’ in the present simple):
  - The boat is useless → The boats useless
  - I think mark has been … → I think marks being/ been
  - The number has been…→ The numbers *have* been
  - Their house was burgled → Their houses burgled

- Gaps were deleted before function words which include the nasals /n/ and /m/, resulting in an '-ing' form word (gerund or present participle):
  - …come and see us → coming see us
  - …but nice and fresh → but icing fresh
  - I asked him for → I *am* asking for / I asking for …

The second pattern includes those instances where functors were recognised as parts of preceding content words or other function words. The result of this category of missegmentation was the insertion of new words which were not present in the input:

- Clusters of two or more function words were merged together creating new words:
  - They are for Jane → therefore Jane
  - Everyone knows me as Hag → Everyone know Miss Hag
  - They are second hand →Their second hand
  - …the person you are with → …the person your with
- Functors merged with the preceding content words resulting in new words:
  - There is a → This a / these a
  - Read his book and write → Read *spok*en write
  - There is some → this some
  - …to know her → an hour

The last pattern resulted in the insertion of non-words after deleting gaps before function words:

- As simple as that → simplice that / simpliza
- Send Frank a Card →Sen Franka card /Francacard
- You have already → Evoloy
- Terrible to us →terribles

All of the above patterns involve the LS error that occurs when a lexical gap is deleted before unaccented function words. The question that arises is that, since the errors show a clear tendency to merge unaccented syllables with the preceding content/function words, can we assume for sure that the participants applied the LS strategy mentioned above? It seems from the mishearing patterns that an assumption can apply to certain function words and under certain conditions. For instance, the conjunction 'and' was recognised as an *–ing* form only when it came after a verb. Similarly, the functors which end in /z/ or /s/ were merged only when the preceding words were nouns (plural form) or verbs (third person singular forms). On the other hand, in other contexts, these words were not subject to such lexical missegmentation. Consequently, as the linguistic context seems to have played a role in the segmentation process, one should be very cautious in assuming a LS strategy taking place, even if most of the instances of function words missegmentation seem to support this claim. This being said, however, the analysis does reveal a serious problem of dissecting WFs of function words out of their surroundings and recognising their borders in running speech. The above mishearings also reflect a poor application of the BU mode in segmenting speech, characterised by the linking of sound sequences across word borders and the shifting of lexical gaps.

The discussion of the role of context in shaping the way words are recognised and segmented in speech takes a special dimension in describing how the listeners approached the auditory input. It is true that the auditory input should be the primary basis for the processes of recognition and segmentation to operate effectively, but this does not rule out the possibility of deploying non-auditory information to create a mental representation which makes the basis for further processing. From this perspective, the listening process implies the coordination of both the TD and the BU modes to decode speech. For the participants in this study, the results of the analysis revealed that the handling of these two modes

effectively was far from being successful. The analysis of the mishearings that resulted from the inability to decode CS features revealed not only a defective BU ability to decode speech, but also a poor application and coordination of some strategies in a TD mode, especially when these strategies stand on inaccurate (BU) decoding of speech.

As the participants failed to decode words by means of deciphering the speech signal (BU mode), they resorted to what is known as the compensation strategies. This is at least what some patterns suggest. The application of a TD mode in the process of WR manifested itself in the apparent reliance on thematic information to *infer* words (e.g. the lexical effects discussed earlier) or in the use of linguistic knowledge to *anticipate* structures, grammatical forms and even words. This is evident in the typical mishearings that reflect a reliance on linguistic cues, rather than input, in the processes of WR and LS. The listeners showed a tendency to deploy such cues to guess and to make expectations about the speakers' words, and the sentence structures which will be used. In what follows, typical examples of TD mode use (or misuse) and compensation strategies' application will be presented and discussed.

Based on the decoding of earlier parts of the utterances, be it correct or not, anticipatory processes were often triggered to predict the language forms or the words that would be used. As Oxford (1990) put it, inferential processes rely on two main types of knowledge: Linguistic knowledge and general background knowledge. In this regard, one pattern of mishearings shows the effects of CS features misperception on the decoding of the remaining of a given utterance. As soon as the beginning of an utterance took a new form as a result of misperceiving (function) words, the way the rest of the utterance was decoded was greatly affected because anticipatory processes would follow the *new* form rather than the original one. Examples of the participants' mishearings confirm that they not only tried to anticipate the words and structures (using linguistic knowledge) based on the modified

shapes of utterances but also failed to monitor and to readjust the results of their processing whenever contradictory evidence was revealed by input. In other words, a wrong decoding of words at the beginning of utterances created constraints according to which the incoming words and phrases should comply. For the participants, these constraints would guarantee that whatever is coming should satisfy previously decoded speech grammatically and/or syntactically.

Example 1

Original:	**Common Mishearings:**
The number has been engaged for over an hour	**(1a)** The numbers have been engaged for over an hour **(1b)** The number is being engaged for over an hour. **(1c)** The number was being engaged for over an hour. **(1d)** The numbers were being engaged for over an hour

The above examples of mishearing show that the testees followed a BU mode in decoding by mapping the sounds they received to the words they have. However, we can also notice hard evidence for a combined use of both BU and TD modes. In (1a), the function word 'has' was misperceived by mapping the sounds to words, resulting in 'numbers'. Linguistic knowledge was also stimulated to anticipate the structure of the utterance. When the following word 'been' was perceived, the learners inserted a new word before it to account for the correctness of the tense used, and the agreement between the subject and the predicate in number: "*The numbers have been engaged for over an hour"*.

In (1b), (1c) and (1d), the auxiliary 'has' was misrecognised as a separate item, either as 'was', 'is' or 'were'. In this case, the WF of the following function word 'been' was modified and perceived as 'being'. It is worth noting that no participant wrote "the numbers have being" or "the number is/was been". This suggests that the misperceived items, through predicting and inferential processes, do impose certain structures according to which subsequent items are expected to agree. What is important to note is that the learners

appeared to be very reluctant to revise their decoding and segmentation of the stream of speech against the incoming input, while first judgements about the earlier parts of the input were retained and they were not open to readjustment.

Example 2

Original:	**Common Mishearings:**
If he stopped smoking, it would get better.	**(2a)** If **you** stop smoking, **you** get better. **(2b)** If you **stop** smoking, you **will** get better. **-**

In (2a), most of the test takers assumed that the second clause starts with the pronoun 'you' as a result of failing to decode a WF in the first clause and replacing 'he' with 'you'. This is despite that the pronoun is not present in both clauses of the original utterance. Bearing in mind that the 'you' and 'it' have completely different pronunciation forms, the mishearing suggests that the pronoun in the second clause was inferred based on the text so far. The listeners compensated for their inability to decode the word by inferring the *missing* word based on information sources other than the sound input. That is to say, they followed the TD mode to guess the word. In (2b), not only the pronouns were misperceived, but also the modal verb 'would'. As the elision of the final /t/ (searched) in the first clause was difficult to decode, the listeners failed to notice the verb was in the past form. Based on this, the listeners were probably obliged to refer to the rule of the first conditional which would be a logical explanation for the replacement of the modal 'would' with 'will'. Alternatively, the replacement might have been the result of mere confusion between the 'will' and 'would' which share the first sound /w/.

Example 3

Original:	**Common Mishearing:**
Because he was taking part in this TV show	**(3a)** Because it takes part in this TV show

Most of the participants misperceived the auxiliary ‘was’ in the example above (3.a). As the auxiliary was completely deleted, the form of the following verb was systematically modified. By applying linguistic knowledge, the listeners modified the result of their auditory perception of the verb ‘take’ by replacing the suffix ‘-ing’ with ‘-s’ to insure a grammatical agreement with the singular subject ‘it’. This is another example of an application of linguistic knowledge in a TD mode where a weak decoding ability negatively affects the decoding of words in the utterance, namely content words.

Example 4

Original:	**Common Mishearings:**
Read his book and write some notes	**(4a)** Read *some* books and write some notes. **(4b)** Read, speak and write some notes. **(4c)** Read this book and write some notes.

As the mishearing in (4a) shows, the misperception of the WF ‘his’ had a direct effect on the decoding of ‘book’. When the word was replaced by the quantifier ‘some’, the listeners mistakenly transcribed the following noun in the plural form. In (4b), as the listeners failed to decode the phrase “his book”, they filled the gap with the word ‘speak’. In this case, they resorted to the strategy of inferring; they relied on the surrounding words ‘read’ and ‘write’[1], which refer to language skills, to infer the word of the same lexical field: ‘speak’.

Example 5

Original:	**Common Mishearing:**
If they had searched more carefully, they might have found the Jewels.	**(5a)** If they search more carefully, they might find the jewels.

In example 5, except for two participants, all of the test-takers could not decode the auxiliary ‘had’. The fact that they transcribed the following verb as ‘search’ instead of ‘searched’ may be a direct result of the failure to decode the auxiliary ‘had’ which would

[1] The words ‘read’ and ‘write’ were not blanked out in the answer sheet.

suggest the past participle form for the verb that follows. On the other hand, the two participants who decoded 'had' correctly wrote 'searched' rather than 'search'. Unlike the majority of the participants, the latter probably applied their linguistic knowledge to note the elision in the word 'search' which is, in this particular situation, difficult to note without reference to the co-text. Similarly, most of the testees failed to perceive the auxiliary 'have' in the second clause. So, relying on the first clause which *imposed* the first conditional as the correct structure for the utterance, the participants not only failed to anticipate and perceive the auxiliary 'have' but also modified the verb 'found' to become a stem form ("might find" instead of "might have found"). Hence, a series of complex modifications took place after the misperception of CS features at the beginning of the utterance, and a failure to apply the strategy of monitoring to verify the results of earlier decoding of speech. Instead of monitoring their understanding, the participants tended to stick to earlier judgements about their processing assuming that it was correct and that what followed should have been congruent with it.

Example 6

Original:	**Common Mishearing:**
You should have called the police immediately.	**(6a)** You should call the police immediately.

As a result of elision, all of the participants transcribed the verb in (6a) as 'call' instead of 'called'; they could not note the elided 'd'. Recovering from the elision process could have been possible through the application of linguistic knowledge in a TD mode by considering the previous words "should have". However, this was not possible for most of the testees due to the inability to decode the auxiliary 'have' in the first place. In this respect, the stem form becomes more appropriate.

In contrast to previous examples in which TD strategies were overused, this example reveals that inferences were inhibited by the inability to decode CS features. If they had been

correctly decoded, the latter would have had the potential to activate appropriate linguistic knowledge to anticipate relevant sentence structures. This example indicates that without sufficient and reliable decoding skills, TD processing may not function appropriately.

What is special about this particular example of mishearing is the devastating effect it has on the meaning of the original utterance. The illocutionary force of the misheard utterance (6a) would induce the listener to take the phone and call the police, ignoring the fact that it is probably too late to do so.

Example 7:

Original:	**Common Mishearings:**
I wanted her to stay	**(7a)** I wanted you to stay. **(7b)** I wanted it to say. **(7c)** I wanted him to stay.

For example 10, the mishearings include words which would fit into the utterance both structurally and semantically although they have no acoustic matching with the input. Instead of leaving the answer sheets blank, the participants compensated for their weak decoding skill by inferring the word they could not decode based on its position in the utterance. As such, it is not the sound signal that was the basis of recognition, but rather the linguistic knowledge (TD mode).

Example 8

Original:	**Common Mishearing:**
There are some new books I must read.	**(8a)** There is a new books I must read.

Example 13 gives further evidence for the participants' reluctance in revising the result of their decoding. They appear to be disinclined to modify their wrong decisions about decoded words when sufficient contradictory evidence is revealed. The words "are some"

were misheard and *replaced* with "is a". Despite that the word 'books', which was not blanked out, clearly indicates the plural form, the participants retained their wrong answers. This reluctance in revising the results of their own processing of speech indicates that the participants did not use the monitoring strategy to assess their processing of speech. The listeners did not revise their processing probably because they did not possess this strategy. That is, in listening to the target language, they approached the input in a linear way, and ignored those strategies that they had already used in the first language. Alternatively, it is probable that they did not know when and how to use these strategies. The last explanation stands on the fact that some answers suggest a backwards impact where the last portions of the text were deployed to give clues that were used in transcribing the first portions. The following set of examples (**9, 10, 11, and 12)** provides some evidence about the participants' application of the monitoring strategy.

Example 9:

Original:	**Common Mishearings:**
You must come over for dinner soon.	**(9a)** He was coming over for dinner soon. **(9b)** You will come over for dinner soon.

Example 10:

Original:	**Common Mishearings:**
Aren't there some letters for her to open?	**(10a)** Aren't Sent letters to open **(10b)** Are Send letters for to open **(10c)** Wrote letters for to open

Example 11:

Original:	**Common Mishearing:**
I asked him for some money.	**(11a)** I am asking for some money.

Example 12:

Original:	**Common Mishearing:**
The number has been engaged for over an hour	**(12a)** The numbers have been engaged for over an hour.

The mishearings in examples **9, 10, 11** and **12** show a special type of inference which takes place in a regressive direction; i.e. from right to left. Like the previous examples, lexical items were systematically inferred and inserted within the utterances. However, what appears to be the source of inference were not words which precede the inferred item, but rather the ones that follow.

Example 9b clearly indicates a process of monitoring taking place. The replacement of the modal 'must' with 'will' –which differ in pronunciation – can by no means be attributed to sound confusion. This being said, the modal 'will' is in a paradigmatic relation with 'must' and, like other modal verbs, it can occupy the same position in the utterance as 'must'. If this is true, why is it that all mishearings include 'will' but not the other modal verbs (e.g. would, may, might, should)? One logical explanation is that the listeners in this example relied on the word 'soon' which comes at the end of the utterance to infer and to *select* the modal 'will' out of the range of possible items. We cannot know for sure whether they waited until the last part of the utterance to do so, but the fact that most of them wrote 'will' reduces the likeliness of an arbitrary choice being made.

If we take 10a, it is clear that the mishearings 'write', 'sent' and 'send' were inferred based on the word 'letter' which comes after them. As they failed to decode the function word in the blank, the listeners had to find evidence in the unfolding text to fill in the missing word. That is, after recognising the word 'letter', they went back in a right to left direction to reconsider their processing of earlier parts of the utterance (in case it had already been done) or to simply fill in the blank. In both cases, going back to earlier parts of the text and reconsidering them based on new evidence is evident of a monitoring strategy taking place.

Similar observations can be made from examples 11 and 12. In 11a, the auxiliary 'am' was inserted without any clear justification as far as the sound input is concerned. To understand the cause of this mishearing, it is important to consider the words that follow

which provide clues to this. The participants could not decode the WF of 'him' and merged it with the verb 'asked' resulting in 'asking'. The fact that the resulting phrase "I asking" is not grammatically correct may have led the listeners to readjust their decoding of the earlier part by inserting the auxiliary 'am'. Similarly, the auxiliary 'have' in 12a was also inserted to comply with the grammaticality of the utterance.

The following two examples show that mishearings may take place by extending thematic aspects from one part of the utterance to another. Inferences about difficult parts of the utterance were made based on previous parts. In 13a, the listeners extended the meaning of *liking* expressed in the first part of the utterance to the second part after failing to decode the assimilated word 'art'. Likewise, in (14a), the negative form in "is not spicy" was extended to apply for the second part 'fresh' → "not fresh". This is the result of misperceiving the conjunction 'but' (which expresses contrast) with the conjunction 'and' which makes logical the extension of the negative aspect to the following part of the utterance. Thus, it could be said that the misperception of logical connectors may affect the decoding of subsequent text and allow for wrong inferences to persist, especially in the absence of an adequate monitoring strategy.

Example 13:

Original:	**Common Mishearing:**
I like going to museums and art galleries.	**(a)** I like going to museums and I love galleries.

Example 14:

Original:	**Common Mishearing:**
I like plain food, that's not spicy, but nice and fresh	**(a)** I like plain food, that's not spicy and not fresh.

5.3. Qualitative Analysis and Discussion of the Interview Results

The main purpose behind the post-treatment interview was to uncover the participants' reactions to the lessons they received, and to allow them to reflect on their experience in learning about CS features through the diagnostic approach. We tried to elicit their opinions, in addition to their own evaluation of the lessons and of their oral language skill(s) before and after the instruction. As it was mentioned before, the semi-structured interview was chosen due to its flexibility in allowing for an open scope in dealing with the participants' answers. A number of questions formed the starting point of the inquiry but, by the time the interactions took place, additional issues imposed themselves too, and were deemed worth questioning.

Four main questions were planned before the interviews:

- Before the oral expression lessons, did you have an idea about CS aspects?
- How do you evaluate your listening skill before the instruction you received?
- In your view, what made listening to English difficult/easy for you?
- How do you evaluate your listening skill after the instruction you received?

Surprisingly, at least as an initial answer, all of the students said that they had not known about CS aspects before the OE lessons. This is despite that CS features are usually presented in the first year in the phonetics module. When they were reminded of that, most of them said that the lessons they received were theoretical, with reference to the way the features had been presented in comparison to the lessons they received in the OE module. This may explain the reason why all of them said that they had not had an idea about CS before; they did not have enough practice as 7 of them complained. Some said the lessons they had received focused on defining the features and that they did not go any deeper. For them, the focus was on pronunciation and speaking, without paying attention to the listening

skill. One explained that, since the module was phonetics, the focus must have been on speaking rather than on listening.

Despite that this is a students' perspective in describing the way they had dealt with CS in phonetics, which may not have been completely accurate, it gives a general idea about their own understanding, how they perceived the phonetics lessons and the extent to which they benefited from them. In addition, it reveals that they did not realise the importance of the features from a listening perspective. On the other hand, three students said that they got to know about CS features through personal work by reading pronunciation books, or listening to news channels. Only one student acknowledged that the teacher of OE in the first year "spoke about CS" frequently, and gave them "a general idea about CS". As for the features they knew about, assimilation was cited 12 times, elision 6 times, linking 5 times and WFs only twice. Of these, assimilation was rated as the most important feature to know about because, for them, it facilitates both listening and speaking.

As for the second question, most of the students (79%) acknowledged that they had problems understanding spoken English before the lessons. One student said that she did not have problems in listening and explained this by experience in listening and frequent exposure to naturally spoken English through TV. As for the problems that made listening difficult according to them, fast speech rate was blamed as the major obstacle by most of the students. In their explanations, some of them tried to make a link between speech rate and CS. The following are some of the explanations they provided:

- "I thought that native speakers speak fast, but now I recognise that the problem is [a] matter of linking words."
- "I don't know what they are speaking [about] unless they talk slowly so that each word is alone."
- "Native speakers speak quickly and they use assimilation and WFs."

- "[listening to English was] difficult because of assimilation […] I faced a difficulty."

Interestingly, some of the students' descriptions of their listening problems exactly matched those revealed in the qualitative analysis of the common mishearings. Locating word borders and segmenting utterances have been identified as difficult tasks to perform, especially under the pressure of fast speech rates. One student described that "when native speakers combine between two words, I think that it is a new word." Combining words in this case may involve any of the CS features that modify word borders. The linking between words may create confusion and force the learners to make erroneous decisions in the word segmentation process. One student explained: "I used to search for words, but I usually found that the words I was looking for did not exist. So, I divide them." By searching for words, the listener makes a linear matching of sounds to words by taking a BU mode; the latter was highlighted earlier in the qualitative analysis as the cause of many mishearings.

Moreover, any lack of linguistic knowledge may add more obstacles and render the LS task even more difficult. In this respect, one student linked the problem of LS to vocabulary knowledge: "I did not know errr how to separate words. Sometimes I thought it is one single word that I did not know. So, I said it is a new word. " This explains well the big number of non-words written by the test-takers in trying to transcribed what they heard during the listening tests. The mishearing in the last example would result in deleting lexical gaps that do exist as a result of: 1) inability to decode the CS feature involved; and 2) lack of confidence created by limited vocabulary knowledge.

The above descriptions show that the learners are aware of the problems they face during listening, and of the errors they may commit in segmenting speech. They may not know where the problem lies, but they are aware that they are unable to recognise words and define their borders. When decisions have to be done about word borders, the learners tend

to take the risk in the process of segmentation even if they are not completely sure of the result of their speech processing. This gives further evidence that the mishearings often take place in a systematic way and that they may not be completely arbitrary.

All of the learners, including the one who reported having experience in listening to English, expressed their satisfaction about the lessons they received in the practitiatory, and noted a relative improvement in their listening skill after the treatment. It was possible to notice how motivated and engaged the students were during the lessons, especially in the post-listening phase. Similar observations were reported by Rosa (2002) and Underwood & Wallace (2012) concerning the learners' motivation in learning about WFs. Working in the practitiatory which was supported with technological devices should have played a role in engaging the students, facilitating the teacher's task and in promoting learning. While several students said that there was a big difference in their ability to listen to naturally spoken English, others spoke about a little improvement. Despite this, they all had positive views about the LC lessons they received. Several students noticed a relative change in the lesson structure compared to the period before the treatment, especially the post-listening phase. Before the treatment, as they commented, that phase was dedicated to listening again and again to the listening script without having a task at hand. Two students said that they gained more confidence in listening to native speakers thanks to the lessons they had received. When they were asked to specify how their listening improved, they explained with reference to the different CS features and by citing examples. Speaking about WFs, for instance, some students commented:

- "I can catch WFs."
- "I can distinguish between 'her' and 'are', 'as' and 'is' [in CS]"
- "when they say /kæn/ [in CS]…before the lessons, I could not know whether they *can* or *cannot*. Now, I understand when they say /kæn/ they mean the *cannot*."
- Speaking about final /d/ elision, one student said: "I can distinguish between words…know if a verb is [in] the past or the present."

Other students spoke about more accuracy in segmenting speech, and in recognising words within short utterances.

In general, the students considered the lessons to be '*practical'*, '*beneficial'*, '*helpful'*, and '*important'* for them to improve their oral skills. It was very clear from their answers that they were not aware of the importance of CS in listening before the instruction. Once they knew about them and realised how crucial they are, they developed a desire to work on them. One student commented: "Before the lessons, I didn't know […] about CS and so on… and I will focus on them from now… […] they are very helpful." Many students reported that they are planning to work on the features again and to have more practice to improve their oral skills. One student explained: "I use these rules of CS in listening, and err I was shocked before …. I just [used to] listen as anyone who never studied English, but after using CS, I feel that I am capable to understand". Speaking about the importance of teaching CS, one student said that if she became a teacher, she would teach her learners about CS speech. She added: "I will not let them suffer from what I suffered from."

One student spoke about the TD strategies, namely the use of schemata and background knowledge to anticipate and to overcome problems while listening. She referred to the lesson phases to explain how she benefited: "After the steps you have taught us …about background knowledge and so on, I gather them all and when I hear a sentence, I

don’t rely only on listening. I rely on what I have learnt before…if I hear a sound, I try to know not necessarily the exact word, but something like it.” While this way of dealing with listening seems more inclined towards relying on TD strategies, it implies a combination of both TD and BU modes. The learner starts from the sound, and makes the best use of the different sources of knowledge stored in the long-term memory to create meaning. This is exactly what learners should be helped and encouraged to do.

Despite that the main focus of the lessons was to address the learners’ listening skill, almost half of the students thought that the lessons helped them in developing their speaking skill. More specifically, the benefits they mentioned include:

- Adopting an English-like rhythm and speaking *real* English
- Speaking English with ease and improving their oral skills
- Having more fluency in delivering messages

5.4. Summary and Discussion

The results of the qualitative analysis of the pre-test results confirm two of the research hypotheses. First, CS features have the potential of causing decoding problems for the students. The different errors of perception –including insertion, deletion and transferring– do take place following the failure to decode a given CS feature. Second, both LS and WR problems take place as a consequence of failing to decode the CS features. The phonological modifications that take place across word borders had devastating effects on both the form and the meaning of utterances. They were found to hinder the process of lexical access for the participants by increasing the number of activated candidates. In addition, the participants’ answers have seldom shown an ability to recover from the effects of the different CS features, and the more CS features were involved in an utterance, the more difficult it was for them to decode them.

In general, the findings suggest lack of awareness from the part of the participants concerning the modification processes that take place in spoken English, and a very weak ability to decode them. In addition, whenever CS features are involved, any lack of linguistic knowledge adds more problems to the segmentation and recognition processes. Poor vocabulary knowledge and lack of confidence make the matters worse as the listeners may put the blame on their limited vocabulary knowledge whenever they fail to decode words that undergo CS modifications. Even if they know the words, the fact that the latter are *distorted* by means of phonological modifications makes them *strange* and, consequently, difficult for the listeners to decode. In addition, it was found that a weak ability to perceive sounds at the segmental level adds more barriers to the decoding process.

Concerning the second hypothesis, it was found that the listeners do resort to TD strategies to overcome their weak BU decoding skills. The analyses revealed that weakness in decoding the CS features may provoke the listeners to resort to information sources other than the input as a strategic choice. In many cases, instead of ignoring the portions of the text they found difficult to decode, the listeners showed a tendency to deploy linguistic and background knowledge in a TD fashion to compensate for their poor decoding ability. However, the extent to which these strategic choices could be successful was very limited. This was evident in the mishearings which showed an apparent unsuccessful reliance on thematic information (from the co-text or by activating background knowledge) or linguistic knowledge to make (wrong) inferences and predictions about the difficult words or about the sentence structures. Despite this, the data gathered cannot exactly depict the extent to which the TD mode was resorted to; while the use of this mode could be spotted clearly from the wrong answers (some typical mishearings), it is difficult to decide whether the correct answers were the result of pure decoding of sounds or they came as a result of any compensation strategy.

This being said, the use of the TD mode did not prove to be a reliable alternative for the basic decoding skills. Both qualitative and quantitative data support this claim. On the one hand, the overall performance of the participants in decoding the CS features in the pre-test was relatively very weak despite that the qualitative data revealed a reliance on inferential and anticipatory processes. On the other hand, inability to decode the CS features was found to inhibit the appropriate use of TD compensatory strategies. Relying on poorly decoded lexical items to make expectations and inferences implies the building of meaning on a rather shaky and unstable basis. Sufficient evidence was found for this in the mishearings which reflect the anticipation of irrelevant/wrong words and structures on the basis of defective decoding of earlier parts of utterances. That is, the effect of failing to decode a given CS aspect was not only limited to words which undergo the modification process. It was negatively extended to influence the *decoding* and the recognition (or the inferencing) of the words in the unfolding text.

To make things even worse, poor monitoring strategies from the listeners hindered the possibility of revising the wrong decoding of earlier parts of the text when sufficient contradictory proof was raised. The listeners appeared to lack this flexibility which allows them to reconsider the decisions they had made in listening to a given utterance. The analysis revealed some evidence in support of the use of the monitoring strategy by the listeners in some cases, while it was not used in others. This suggests that this strategy was not used appropriately and the participants probably did not know when or how to use it successfully.

The results provided clear evidence against the view which holds that CS features are of a lower importance and that TD compensatory strategies can solve the problem. Even those structure words which are usually claimed to bring subtle information to the utterance were found to affect the decoding of the surrounding content words if they are not accurately decoded. Moreover, it was shown that with the absence of reliable decoding skills, not only

decoding errors would take place, but also the TD strategies may not operate effectively. This is, at least, what the typical mishearings in this study suggest.

Concerning the last hypothesis, the interview results indicate that the members of the EG had positive reactions to the lessons. They reported an improvement in their oral skills (both listening and speaking) and they were motivated to continue working on CS features to improve their skills.

Conclusion

This chapter was devoted to the analysis of the pre-test results and the post-treatment interview. The findings were congruent with the quantitative results reported in the previous chapter. The CS features were found to be a major obstacle for the students in decoding speech by inhibiting the processes of LS and WR. All of the CS features analysed caused almost similar problems for the students.

Concerning the listening modes, it was found that the participants suffered from a poor listening ability characterised by a deficient bottom up skill on the one hand, and an inappropriate application of the TD mode, on the other. The mishearings showed an overreliance on TD strategies whenever the basic decoding skills suffer. Moreover, inaccurate decoding of words and segments caused by the inability to decode the CS features led to wrong anticipation of structures and inferences of words, i.e. such poor decoding provided an unreliable basis for TD strategies to operate and, eventually, distorted the intended meanings (of the utterances).

The analysis of the mishearings also came in line with the findings of the interviews as the interviewees reported a number of problems that were actually manifested in the mishearings. They also reported improvement in their oral skills after the treatment and appreciation of the lessons they received.

General Conclusion

Understanding spoken language is considered as one of the most difficult and challenging language skills for learners to master. Despite that much time and effort is devoted to improving learners' listening skills in the oral skill module, a great majority of learners would still consider a nightmare any encounter with naturally spoken English outside the classroom, to the extent that they may lose confidence in their own abilities or in the way they go about learning to listen. This study focused on certain problems at the lower level bottom-up processes of listening; it aimed at addressing learners' ability to decode connected speech by concentrating on, and unveiling, the problems related to decoding some connected speech features, and examining how the ability to decode the features could be improved so as to allow for more automaticity and accuracy in deciphering the speech input.

The review of the literature related to listening comprehension research, which was the subject of the first two chapters of this study, highlighted two main trends in addressing the learners' listening skills; one focuses on the higher level top-down processes and the other on the lower level bottom-up processes. Despite that there is much consensus in the literature that successful processing and comprehension of speech is the result of coordinating processes at both levels, classroom practice and research is said to be biased towards over emphasising the higher level top-down processes. This is despite the fact that there have been many attempts along with a growing body of research in support of the effectiveness of developing learners' bottom-up processing to improve their listening skills, and to free their attention from form on the hope that sufficient attention capacity would be reserved for more challenging tasks. These attempts, in their turn, have also over concentrated on BU processes by addressing the lower level skills in isolation.

Following the trend that top-down and bottom-up processes should be regarded as complementary, the main concern of this study was to examine the effectiveness of a teaching framework –which allows for integrating the teaching of connected speech features into the listening comprehension lessons– in improving second year university students' of English decoding skills without losing sight of the key top-down processes. The aim was one of addressing the learners' decoding problems in an insightful manner by diagnosing and, then, addressing them following the diagnostic approach (Field, 2008a). In addition, we attempted to explore the types of problems resulting from the inability to decode some major connected speech aspects and the extent to which weakness in decoding the features could have any consequences on their processing of utterances. The hypotheses this study attempted to verify were: 1) Connected speech aspects would pose problems for the students in decoding naturally spoken English and in making lexical segmentation; 2) In addition to the poor bottom-up skills in decoding connected speech features, the top-down processing mode would have negative effects on the participants' recognition and segmentation of connected speech; 3) if students receive connected speech instruction that focuses on awareness-raising and provides practice following the principles of the diagnostic approach, their ability to decode connected speech phenomena would improve.

The method and the research tools used to verify the hypotheses were discussed in chapter three. A pre-test and a post-test were administered to an experimental group and a control group to test their ability to decode some connected speech features, and to analyse the types of problems they encounter in listening to connected speech. The experimental group received listening comprehension lessons that focused on connected speech features following the principles of the diagnostic approach. Each lesson included a pre-listening and a listening phase that focused on the higher level top-down processing, and an extended post-listening phase to address the connected speech features diagnosed as problematic during

the listening phase. The control group received similar lessons without a specific focus on connected speech. Following the instruction, a post-test was administered to the two groups to examine the effectiveness of the instruction on the performance of the experimental group. In addition, a post-instruction interview was conducted with the members of the latter to find out about their reactions to the instruction their received.

The results of the pre-test confirmed the hypothesis that connected speech features would pose problems for the learners in decoding speech, namely in the processes of word recognition and lexical segmentation. The modified sounds activate new words other than the ones present in the input and, eventually, affect both the meaning of the utterance and the segmentation of the neighbouring words. Factors such as lack of vocabulary and lack of knowledge at the segmental level render these processes more challenging.

The effects of the modifications become worse when the new words, through spreading activation and guessing strategies, form the basis for inferring the incoming words and structures. Among the problems revealed by the analysis of the participants' mishearings, it was found that in the absence of a reliable decoding skill, the learners may resort rather unsuccessfully to top-down processing to overcome their weak decoding skills. Inaccurately decoded connected speech at earlier portions of utterances potentially constrains the way in which the unfolding text would be processed, leading to lexical and syntactic cognitive effects. Based on poorly decoded connected speech, wrong predictions about the wording and the structure of the utterance could be triggered, and the listeners showed reluctance to revise them even when contradictory evidence was revealed by the text to come. These results are congruent with those found by Field (2008b). It could be said that, at least for the participants in this study, top-down compensatory strategies cause negative effects on speech processing when problems of decoding arise. In addition, some

mishearings suggest weakness in applying and coordinating top-down strategies, namely predicting and monitoring. This being said, however, the extent and the scale of the top-down mode effects on the participants' processing cannot be depicted by means of the tools adopted in this study. This is why, we believe, a think aloud protocol would have been much more informative, and it would have given more insights into the learners' listening problems and the strategies they use in relation to connected speech processing.

Concerning the experiment, the statistical analyses were in favour of the effectiveness of the instruction in improving the participants' decoding of connected speech. The participants showed positive reactions to the lessons they received and reported 1) an improvement in their overall listening skill, 2) higher confidence in listening and, 3) much enthusiasm to continue working on connected speech features to improve both listening and speaking skills.

In addition to the think aloud protocol, mentioned above, a large scale study which encompasses a bigger number of participants would provide more data and deeper insights into the learners' problems in processing speech. Moreover, in addition to the connected speech features test, the study could have included a listening comprehension test. This would have allowed for measuring, with statistical data, the effects of improved decoding skills on the participants' overall listening comprehension ability.

We believe that, without the active and informed attempts to diagnose and address the learners' actual problems in listening at the different levels of processing, classroom practice would continue to be an act of testing rather than teaching. The belief that teaching listening should focus on developing learners' higher level processes while the lower level ones will automatically take care of themselves is, in a way or another, promoting this idea of testing. It is also ignoring the primacy of the sound input in a manner that places "the cart

before the horse" (Norris, 1995, p. 47). On the other hand, raising learners' awareness to the aspects of spoken language and providing the necessary training on those decoding areas that cause their frustration while listening is strongly recommended. Just like in L1 listening, decoding connected speech features for language learners should reach a certain level of automaticity so as to free their attention from the formal aspects of the language, and to allow more room for concentrating on the higher levels of meaning. The learners are not only required to know about the different forms and to pronounce them correctly, but they also need to develop their decoding skills to be able to accurately and automatically decipher them while listening. To this end, the onus, we believe, rests on the listening comprehension teacher.

The diagnostic approach in this study focused on the connected speech-related listening comprehension problems only. However research should be by no means limited to this aspect of decoding. Further research could be extended to account for the different problems at the linguistic level such as vocabulary syntactic or grammatical knowledge. Both researchers and practitioners should allow a degree of openness to anticipate and diagnose all sorts of language-related barriers which have the potential of putting the learners' comprehension of spoken language at stake.

Bibliography

Aguilar, M. (2008). *Metadiscourse in Academic Speech: A Relevance-theoretic Approach.* Bern: Peter Lang.

Alanen, R. (1995). Input enhancement and rule presentation in second language acquisition. In R. Schmidt (Ed.), *Attention and Awareness in Foreign Language Learning* (pp. 259-302).

Al-jasser, F. (2008, December). The effect of teaching Enlish phonotactics on the lexical segmentation of English as a foreign language. *System, 36*, 94-106.

Altarriba, J., & Graves, D. F. (2013). Semantic processing. In P. Robinson (Ed.), *The Routledge Encyclopdeia of Second Language Acquisition* (pp. 578-581). Routledge.

Altmann, G. (1991). Modularity and interaction in sentence processing. In J. L. Garfield (Ed.), *Modularity in Knowledge Representation and Natural-language Understanding* (pp. 249-257). The MIT Press.

Anderson, A. H., Yule, G., & Brown, G. (1984). Hearer-effects on speaker performances: The influence of the hearer on speakers' effectiveness in oral communication tasks. *First Language, 15*(3), 23-40.

Anderson, A., & Lynch, T. (1988). *Listening.* Oxford: Oxford University Press.

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behaviour, 22*(3), 261-295.

Anderson, J. R. (1985). *Cognitive Psychology and its Implications* (2nd ed.). New York: Freeman.

Anderson, J. R. (2010). *Cognitive Psychology and its Implications* (7th ed.). Worth Publishers.

Anderson, R. C., Pichert, J. W., & Shirey, L. L. (1979, April). Effects of the reader's schema at different points in time. *Center for the Study of Reading Technical Report*(119).

Anderson, R. C., Reynolds, R. E., Schallert, D. L., & Goetz, E. T. (1977). Frameworks for comprehending discourse. *American Educational Research Journal, 14*(4), 367-381.

Asher, J. (1969). The total physical response approach to second language learning. *The Modern Language Journal, 53*(1), 3-17.

Asher, J. (1981). The total physical response: Theory and practice. *Annals of the New York Academy of Sciences, 379*(1), 324-331.

Avery, P., Ehrlich, S., & Jull, D. (1992). Connected speech. In P. Avery, & S. Ehrlich, *Teaching American English Pronunciation* (pp. 73-90). Oxford University Press.

Bansal, R. K., & Harrison, J. B. (Eds.). (1994). *Spoken English: A Manual of Speech and Phonetics.* Orient Longman.

Batie, B. D., & Bradley, D. C. (1995). Resolving word boundaries in spoken French: Native and non-native strategies. *Applied Psycholinguistics, 16,* 59-84.

Bloomer, A., Griffiths, P., & Merrison, A. J. (2005). *Introducing Language in Use: A Course Book.* Psychology Press.

Bowler, B., & Cunningham, S. (2003). *Upper-intermediate New Headway Pronunciation Course.* Oxford University Press.

Brindley, G. (1998). Assessing listening abilities. *Annual Review of Applied Linguistics*, 171-191.

Broughton, G., Brumfit, C., Flavell, R., Hill, P., & Pincas, A. (1980). *Teaching English as a Foreign Language* (2nd ed.). Routledge.

Brown, G. (1990). *Listening to Spoken English.* London: Longman.

Brown, G., & Yule, G. (1983). *Teaching the Spoken Language.* Cambridge Unibersity Press.

Brown, J. D., & Kondo-Brown, K. (2006). Introducing connected speech. In K. K.-B. James Dean Brown (Ed.), *Perspectives on Teaching Connected Speech to Second Language Speakers* (pp. 1-15). National Foreign Language Resource Center, University of Hawai'i at Mānoa.

Brown, J. D., & Kondo-Brown, K. (2006). Testing reduced forms. In J. D. Brown, & K. Kondo-Brown (Eds.), *Perspectives on Teaching Connected Speech to Second Language Speakers.*

Brown, J. I. (1987). Listening—Ubiquitous yet obscure. *International Listening Association Journal, 1*(1), 3-14.

Buck, G. (2001). *Assessing Listening.* Cambridge: Cambride University Press.

Bygate, M. (1987). *Speaking.* Oxford: Oxford University Press.

Byrnes, H. (1984). The role of listening comprehension: A theoretical base. *Foreign Language Annals, 17*(4), 317-329.

Call, M. E. (1985). Auditory short-term memory, listening comprehension, and the input hypothesis. *TESOL Quarterly, 19*(4), 765-781.

Carrell, P. L. (1983). Some issues in studying the role of schemata, or background knowledge, in second language comprehension. *Reading in a Foreign Language, 1*(2), 81-92.

Carrell, P. L. (1984a). The effects of rhetorical organisation on ESL readers. *TESOL Quarterly, 18*(3), 441-469.

Carrell, P. L. (1984b). Evidence of a formal schema in second language comprehension. *Language Learning, 34*(2), 87-112.

Carrell, P. L., & Eisterhold, J. C. (1983, December). Schema theory and ESL reading pedagogy. *TESOL quarterly, 17*(4), 553-573.

Carrier, K. (1999). The social environment of second language listening: Does status play a role in comprehension? *The Modern Language Journal*, 65-79.

Carrol, J. B. (1972). Defining language comprehensioon: some speculations. In R. O. Freedle, & J. B. Carrol (Eds.), *Language Comprehension and the Acquisition of Knowledge.* New York: Wiley.

Celce-Murcia, M. (1995). Discourse analysis and the teaching of listening. In G. Cook, & B. Seidlhofer (Eds.), *Principle and Practice in Applied Linguistics* (pp. 363-377). Oxford: Oxford University Press.

Celce-Murcia, M. (2001). Language teaching approaches: An overview. In M. Celce-Murcia (Ed.), *Teaching English as a Second or Foreign Language* (pp. 3-11).

Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (2010). *Teaching Pronunciation: A Course Book and Reference Guide* (éd. second). Cambridge: Cambridge University Press.

Celce-Murcia, M., M.Brinton, D., & Goodwin, J. M. (1996). *Teaching Pronunciation: A Reference for Teachers of English to Speakers of other Languages* (First ed.). Cambridge: Cambridge University Press.

Chaudron, C., & Richards, J. C. (1986). The effect of discourse markers on the comprehension of lectures. *Applied Linguistics, 7*(2), 113-127.

Cheng, H.-f. (2004). A comparison of multiple-choice and open-ended response formats for the assessment of listening proficiency in English. *Foreign Language Annals, 37*(4), 544-555.

Chenjun, D., & Li, L. (2012, April). The effectiveness of explicit instruction of certain decoding skills in improving chinese EFL listeners' general comprehension performance. *Chinese Journal of Applied Linguistis (Quarterly), 35*(2), 243-255.

Cintron-Valentine, M. C., & Ellis, N. C. (2007). Selience in second language acquisition: Physical form, learner attention, and instructional focus. Dans A. Blumenthal-Dramé, A. Hanulikova, & B. Kortmann (Éds.), *Perceptual Linguistic Salience: Modeling Causes and Consequences* (pp. 89-109). Frontiers.

Clark, H. H., & Clark, E. V. (1977). *Psychology and Language: An Introduction to Psycholinguistics.* Harcourt Brace Jovanovich.

Cohen, A. D., & Macaro, E. (2007). *Language Learning Strategies.* Oxford: Oxford University Press.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, 155-159.

Collins, B., & Mees, I. M. (2013). *Practical Phonetics and Phonology* (3rd ed.). Routledge.

Cook, V. (2008). *Second Language Learning and Language Teaching* (éd. 4th). Hodder Education.

Cruttenden, A. (2001). *Gimson's Pronunciation of English* (Sixth ed.). Edward Arnold.

Cruttenden, A. (2014). *Gimson's Pronunciation of English* (Eighth ed.). London: Routledge.

Cunningham, S., & Bowler, B. (2003). *Intermediate New Headway Pronunciation Course.* Oxford: Oxford University Press.

Cutler, A. (1997). The syllable's sole in the segmentation of stress languages. *Language and Cognitive Processes, 12*(5/6), 839-845.

Cutler, A. (2000). Listening to a second language through the ears of a first. *Interpreting*, 1-22.

Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence form juncture misperception. *Journal of Memory and Language, 31*, 218-236.

Cutler, A., & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language, 2*, 133-142.

Cutler, A., & McQueen, J. M. (1995). The recognition of lexical units in speech. In B. d. Gelder, & J. Morais (Eds.), *Speech and Reading: A Comparative Approach* (pp. 33-47). Erlbaum (UK) Taylor&Francis.

Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology, 14*(1), 113-121.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language, 25*, 385-400.

Cutrone, P. (2005). A case study examining backchannels in conversations between Japanese-British dyads. *Multilingua, 24*, 237-274.

Dalton, C., & Seidlhofer, B. (1994). *Pronunciation.* Oxford University Press.

Dirven, R., & Vespoor, M. (2004). *Cognitive explorations of language and linguistics.* Amsterdam: John Benjamins Publishing Company.

Doughty, C., & Williams, J. (1998). Issues and terminology. In C. Doughty, & J. Williams, *Focus on Form in Classroom Second Language Acquisition.* Cambridge: Cambridge University Press.

Dretzke, B. (1998). *Modern British and American English pronunciation: A basic textbook.* UTB.

Dretzke, B. (1998). *Modern British and American English Pronunciation: A Tasic Textbook.* UTB.

Dunkel, P. (1991). Listening in the native and second/foreign language: Toward an integration of research and practice. *TESOL Quarterly*, *25*(3), 431-457.

Dupuy, B. C. (1999). Narrow listening: An alternative way to develop and enhance listening comprehension in students of French as a foreign language. *System, 27*, 351-361.

Ellis, R. (2003). *Task-based Language Learning and Teaching.* Oxford: Oxford University Press.

Farrell, T. S., & Mallard, C. (2006). The use of reception strategies by learners of French as a foreign language. *The Modern Language Journal 90, 3*, 338-352.

Field, J. (1999, October). 'Bottom-up' and 'top-down'. *ELT Journal, 53*(4), 338-339.

Field, J. (2002). The changing face of listening. In J. C. Richards, & W. A. Renandya (Eds.), *Methodology in Language Teaching: An Anthology of Current Practice* (pp. 242-247). Cambridge University Press.

Field, J. (2003, October). Promoting perception: Lexical segmentation in L2 listening. *ELT Journal, 57/4*, 325-334.

Field, J. (2008a). *Listening in the Language Classroom.* Cambridge: Cambridge University Press.

Field, J. (2008b, October). Revising segmentation hypothesis in first and second language listening. *System*, 35-51.

Flowerdew, J., & Miller, L. (2005). *Second Language Listtening: Theory and Practice.* Cambridge University Press.

Foster, P. (2013). Units for analysing L2 speaking. In P. Robinson (Ed.), *The Routledge Encyclopedia of Second Language Acquisition* (pp. 665-667). Routledge.

Gilbert, J. B. (2005). *Clear Speech.* Cambridge: Cambridge University Press.

Glenn, E. C. (1989). A content analysis of fifty definitions of listening. *Journal of the International Listening Association, 3*(1), 21-31.

Goh, C. C. (2000, April). A cognitive perspective on language learners' listening comprehension problems. *System, 28*, 55-75.

Goss, B. (1982). Listening as information processing. *Communication Quarterly, 30*(4), 304-307.

Graham, S., & Macaro, E. (2008). Strategy instruction in listening for lower-intermediate learners of French. *Language Learning, 58*(4), 747-783.

Grittiths, C. (2008). Teaching/learning method and good language learners. In C. Griffiths (Ed.), *Lessons from Good Language Learners* (pp. 255-265).

Guan, Y. (2014). *The effects of explicit listening strategy instruction on the listening comprehension of English as second language (ESL) community college students.* Unpublished Ph.D thesis, University of San Francisco.

Harmer, J. (2004). *Just Listening and Speaking: For Class or Self-study. Intermediate.* London: Marshall Cavendish.

Harmer, J. (2007). *The Practice of English Language Teaching* (4th ed.). Pearson.

Harmer, J., & Elsworth, S. (1989). *The Listening File: Authentic Interviews with Language Activities.* Longman.

Harmer, J., & Lethaby, C. (2005). *Just Listening and Speaking: For Class or Self Study. Upper-intermediate.* Marshall Cavendish Education.

Hasan, A. S. (2000). Learners' perceptions of listening comprehension problems. *Language, Culture and Curriculum, 13*(2), 137-153.

Hauser, M. F., & Hughes, M. A. (1988). Defining the cognitive process of listening: A dream or reality? *International Listening Association Journal, 2*(1), 75-88.

Helgesen, M., & Brown, S. (2007). *Practical English Language Teaching: Listening.* McGraw-Hill ESL/ELT.

Henner-Stanchina, C. (1986-1987). Autonomy as metacognitive awareness: Suggestions for training self-monitoring of listening comprehension. *Mélanges Pédagogiques, 17*, 69-84.

Henrichsen, L. E. (1984). Sandhi-variation: A filter of input for learners of ESL. *Language Learning, 34*(3), 103-126.

Henrichsen, L. E. (1984). Sandhi-Variation: A filter of input for learners of ESL. *Language Learning, 34*(3), 103-126.

Hewings, M. (2004). *Pronunciation Practice Activities.* Cambridge: Cambridge University Press.

Hill, C., & Beebe, L. M. (1980). Contraction and blending: The use of orthographic clues in teaching pronunciation. *TESOL Quarterly, XIV*(3), 299-323.

Hulstijn, J. H. (2001). Intentional and incidental second-language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (Ed.), *Cognition and Second Language Instruction* (pp. 258-286). Cambridge: Cambridge University Press.

Hunt, E. H. (2009). *Acoustic Characterisation of the Glides /j/ and /w/ in American English.* Massachusetts Institute of Technology: Unpublished Ph.D dissertation.

Imhof, M. (2010). What is going in the mind of a listener? The cognitive psychology of listening. In A. D. Wolvin (Ed.), *Listening and Human Communication in the 21st Century* (pp. 97-126). Wiley Blackwell.

Ito, Y. (2001). Effect of reduced forms on input-intake process. *Second Language Studies, 20*(1), 99-124.

Ito, Y. (2006). The comprehension of English reduced forms by second language searners and its effects on input-intake process. In J. D. Brown, & K. Kondo-Brown (Eds.), *Perspectives on Teaching Connected Speech to Second Language Speakers.*

Janusik, L. A. (2010). Listennig pedagogy: Where do we go from here? In A. D. Wolvin (Ed.), *Listening and Human Communication in the 21st Century* (pp. 208-224). Wiley-Blackwell.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*(1), 122-149.

Kisno. (2012). *Phonetics and Phonology: Theory and Practice.* Jakarta: Halaman Moeka.

Krashen, S. D. (1996). A case for narrow listening. *System, 24*(1), 97-100.

Krashen, S. D., & Terrell, T. D. (1983). *The Natural Approach: Language Acquisition in the Classroom.*

Krug, M. (1998). String frequency. *Journal of English linguistics*, 286-320.

Labed, N. (2001, July-Septemer). Laboratory or practitiatory? *El-Mutargim*(2), 63-67.

Ladefoged, P., & Johnson, K. (2010). *A Course in Phonetics.* Wadsworth: Cengage Learning.

Ladefoged, P., & Johnson, K. (2011). *A Course in Phonetics.* Wadsworth: Cengage Learning.

Larsen-Freeman, D. (2000). *Techniques and Principles in Language Teaching* (2nd ed.). Oxford University Press.

Lass, R. (1984). *Phonology: An Introduction to Basic Concepts.* Cambridge: Cambridge University Press.

Levelt, W. J. (1993). language use in normal speakers and its disorders. *Linguistic Disorders and Pathologies*, 1-15.

Lindblom, B. (1990). Explaining phonetic variation: A scetch of the H&H theory. In W. J. Hardcastle, & A. Marchal (Eds.), *Speech Production and Speech Modelling* (pp. 403-439). Kluwer Academic Publishers.

Long, D. R. (1989). Second language listening comprehension: A schema-theoretic perspective. *The Modern Language Journal*, 32-40.

Long, D. R. (1990). What you don't know can't help you: An exploratory study of background knowledge and second language listening comprehension. *Studies in Second Language Acquisition, 12*(1), 65-80.

Lund, R. J. (1990). A taxonomy for teaching second language listening. *Foreign Language Annals, 23*(2), 105-115.

Lynch, T. (1995). The development of interactive listening strategies in second language academic settings. In J. R. D J. Mendelsohn, *A guide for the Teaching of Academic Listening* (pp. 166-185).

Lynch, T. (1998). Theoretical perspectives on listening. *Annual Review of Applied Linguistics, 18,* 3-19.

Lynch, T. (2006). Academic listening: Marrying top and bottom. In E. Uso-Juan, & A. Martinez-Flor (Eds.), *Current Trends in the Development and Teaching of the Four Language Skills* (pp. 91-110). Berlin.

Lynch, T. (2009). *Teaching Second Language Listening.* Oxford: Oxford University Press.

Lynch, T., & Mendelsohn, D. (2010). Listening. In N. Schmitt (Ed.), *An Introduction to Applied Linguistics* (2nd ed., pp. 180-196). Hodder Education.

Markham, P., & Latham, M. (1987). The influence of religioun-specific background knowledge on the listening comprehension of adult second-language students. *Language Learning, 37*(2), 157-170.

Marslen-Wilson, W. D. (1978, July 18). Sentence perception as an interactive parallel process. *Science, 189*(4198), 226-228.

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitiv Psychology, 10*, 29-63.

Marslen-Wilson, W., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition, 8*, 1-71.

Martinez-Flor, A., & Uso-Juan, E. (2006). Towards acquiring communicative competence through listening. In E. Uso-Juan, & A. Martinez-Flor (Eds.), *Current Trends in the Development and Teaching of the Four Language Skills* (pp. 29-46).

Mashahiro, Y., & Luan, N. L. (2012). The cognitive processes of Japanese foreign language learners in understanding videos. In S. Abdul-Mannan, & H. Abdul-Rahim (Eds.), *Linguistics, Literature and Culture: Millennium Realities and Innovative Practices in Asia* (pp. 394-423). Cambridge Scholars Publishing.

Matsuzawa, T. (2006). Comprehension of English reduced forms by Japanese business people and the efectiveness of instruction. In K. K.-B. James Dean Brown (Ed.), *Perspectives on Teaching Connected Speech to Second Language Speakers* (pp. 59-66). University of Hawaii, National Foreign Language Resource Centre.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18*, 1-86.

McQueen, J. M. (1998). Segmentation of continuous speech usin phonotactics. *Journal of Memory and Language, 46*, 21-46.

Mendelsohn, D. J. (1998). Teaching listening. *Annual review of Applied Linguistics, 18*, 81-101.

Meyer, B. J., & Freedle, R. O. (1984). Effects of discourse type on recall. *American Educational Research Journal, 21*(1), 121-143.

Meyer, R. (1984). "Listen my children and you shall hear...". *Foreign language Annals, 17*(4), 343-344.

Miller, G. A. (1994). The magical number seven, plus or minus two: Some limits on our capacity for processing informaiton. *Psychological Review, 101*(2), 343-352.

Morley, J. (2001). Aural comprehension instruction: Principles and practices. In M. Celce-Murcia (Ed.), *Teaching English as a Second or Foreign Language* (3rd ed., pp. 69-85).

Morton, J. (1969). Interaction of information in word recognition . *Psychological Review, 76*(2), 165-178.

Mueller, G. A. (1980). Visual contextual cues and listening comprehension: An experiment. *The Modern Language Journal, 64*(3), 335-340.

Murphy, J. M. (1991). Oral communication in TESOL: Integrating speaking, listening and pronunciation. *TESOL Quarterly, 25*(1), 51-75.

Nagle, S. J., & Sanders, S. L. (1983). Comprehension theory and second language pedagogy. *TESOL Quarterly*, 9-26.

Nation, I. S., & Newton, J. (2009). *Teaching ESL/EFL Listening and Speaking.* New York: Routledge.

Nation, I., & Macalister, J. (2010). *Language Cuurriculum Design.* New York: Routledge.

Nichols, R. G. (1947). Listening: Questions and problems. *Quarterly Journal of Speech, 33*(1), 83-86.

Nix, J.-M. I. (2016). Measuring Latent Listening Strategies: Development and Validation of the EFL Listening Strategy Inventory. *System*, 79-97.

Norris, R. W. (1995). Teaching reduced forms: Putting the horse before the cart. *English Teaching Forum, 33*, 47-50.

Nunan, D. (2002). Listening in language learning. In J. C. Richards, & W. A. Renandya (Eds.), *Methodology in Language Teaching: An Anthology of Current Practice* (pp. 238-241). New York: Cambridge University Press.

Nunan, D. (2015). *Teaching English to Speakers of other Languages.* New York: Routledge.

Omalley, J. M., & Chamot, A. U. (1990). *Learning Strategies in Second Language Acquisition.* Cambridge University Press.

O'malley, J. M., Chamot, A. U., & Kupper, L. (1989). Listening comprehension strategies and second langage acquisition. *Applied Linguistics, 10*(4), 418-437.

O'Malley, J. M., Chamot, A. U., Stewner-Manzanares, G., & Kupper, L. (1985). Listening strategy applications with students of English as a second language. *TESOL Quarterly, 19*(3), 557-584.

Onnis, L. (2017, May). Caregiver communication to the child as moderator and mediator of genes for language. *Behavioural Brain Research*, 197-202.

Osada, N. (2004). Listening comprehension research: A brief review of the past thirty years. *Dialogue, 3*, 53-66.

Pan, Z. (2017). *Assessing Listening for Chinese English Learners: Developing a Communicative Listening Comprehension Test Suite for CET.* Routledge.

Peterson, P. W. (2001). Skills and strategies for profecient listening. In M. Celce-Murcia (Ed.), *Teaching English as a Second or Foreign Language* (3rd ed., pp. 87-100). Heinle and Heinle.

Pica, T. (1992). Communication with second language learners: what does it reveal about the social and linguistic processes of second language learning? In J. E. Alatis (Ed.), *Goergetown University Round Table on Languages and Linguistics* (pp. 435-464).

Poeppel, D., & Idsardi, W. (2011). Recognising words from speech: The perception-action-mamory loop. In G. Gaskell, & P. Zwitserlood (Eds.), *Lexical Representation: A Multidisciplinary Approach.* De Gruyter Mouton.

Pouplier, M. (2003). The dynamics of error. *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 2245-2248). Barcelona: Universitat Autonoma de Barcelona.

Prince, P. (2014). Listening cmprehension: Processing demands and assessment issues. In P. Leclercq, A. Edmonds, & H. Hilton (Eds.), *Measuring L2 Profeciency: Perspectives form SLA* (pp. 93-108). Multilingual Matters.

Purdy, M. (1997). What is listening? In M. Purdy, & D. Borisoff (Eds.), *Listening in Everyday Life: A Personal and Professional Approach* (2nd ed., pp. 1-20).

Rast, R. (2010). First exposure. In M. Putz, & L. Sicola (Eds.), *Cognitive Processing in Second Language Acquisition* (Vol. 13, pp. 99-115). John Benjamins Publishing.

Renandya, W. A. (2013). The role of input- and output- based practice in ELT. In A. Ahmed, M. Hanzala, F. Saleem, & G. Cane (Eds.), *ELT in a Changing World: Innovative Approaches to New Challenges* (pp. 41-52). Cambridge Scholars Publishing.

Renandya, W. A., & Farrell, T. S. (2010). 'Teacher, the tape is too fast:' Extensive listening in ELT. *ELT Journal, 65*(1), 52-59.

Richards, J. C. (1983, June). Listening comprehension: Approach, design, procedure. *TESOL Quarterly, 17*(2), 219-240.

Richards, J. C. (1990). *The Language Teaching Matrix.* Cambridge: Cambridge University Press.

Richards, J. C. (2015). *Key Issues in Language Teaching.* Cambridge: Cambridge University Press.

Ridge, A. (1993). A perspective of listening skills. In A. D. Wolvin, & C. G. Coakley (Eds.), *Perspectives on Listening* (pp. 1-14). David Fulton Publisher.

Roach, P. (2009). *English Phonetics and Phonology* (Fourth ed.). Cambridge: Cambridge University Press.

Robert, L. (2013). Discourse pocessing. In P. Robinson (Ed.), *The Routlege Encyclopedia of Second Language Acquisition* (pp. 190-194).

Rosa, M. (2002). Don'cha know? A survey of ESL teachers' perspectives on reduced forms instruction. *Second Language Studies*, 49-78.

Rost, M. (1990). *Listening in Language Learning.* Routledge.

Rost, M. (2001). Listening. In R. Carter, & D. Nunan (Eds.), *The Cambridge Guide to Teaching English to Speakers of other Languages* (pp. 7-13). Cambridge University Press.

Rost, M. (2006). Areas of research that influence L2 listening instruciton. In E. Uso-Juan, & A. Martinez-Flor (Eds.), *Current Trends in the Developing and Teaching of the Four Language Skills* (pp. 47-74).

Rost, M. (2011). *Teaching and Researching: Listening* (éd. 2nd). Pearson Education.

Rost, M., & Ross, S. (1991). Learner use of strategies in interaction: Typology and teachability. *Language Learning, 41*(2), 235-273.

Rubin, J. (1995). An overview to a guide for the teaching of second language listening. In D. J. Mendelsohn, & J. Rubin (Eds.), *A Guide for the Teaching of Second Language Listening.* San Diego: Dominie Press.

Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical Issues in Reading Comprehension* (pp. 33-58).

Rumelhart, D. E., & Ortony, A. (1977). The representation of knowledge in memory. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.). Lawrence Erlbaum Associates.

Rylander, J., Clark, P., & Derrah, R. (2013). A video-based method of assessing pragmatic awareness. In S. J. Ross, & G. Kasper (Eds.), *Assessing Second Language Pragmatics* (pp. 65-97). Palgrave Macmillan.

Sato, C. J. (1988). Origins of complex syntax. *Studies in Second Language Acquisition, 10,* 371-395.

Schmidt-Rinehart, B. C. (1994). The effects of topic familiarity on second language listening comprehension. *The Modern Language Journal, 78*(2), 179-189.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human informaiton processing: I. Detection, search, and attention. *American Psychological Association, 84*(1), 1-66.

Shannon, C. E., & Weaver, W. (1949). *The Mathematical Theory of Communication.* The University of Illionois Press. Urabana.

Skandrea, P., & Burrleigh, P. (2011). *A Manual of English Phonetics and Phonology* (Second ed.). Narr.

Skehan, P. (1998). *A Cognitive Approach to Language Learning.* Oxford: Oxford University Press.

Skipper, J. I., Wassenhove, V. v., Nusbaum, H. C., & Small, S. L. (2007). Hearing lips and seeing voices: How cortical areas supporting speech production mediate audiovisual speech perception. *Cerebral Cortex, 17*(10), 2388-2399.

Stæhr, L. S. (2009). Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Studies in Second Language Acquisition, 30*(4), 577-607.

Stanovich, K. E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly, 16*(1), 32-71.

Thiele, A., & Scheibner-Herzig, G. (1983). Listening Comprehension Training in Teaching English to Beginers. *System*, 277-286.

Thompson, I., & Rubin, J. (1996). Can strategy instruction improve listening comprehension? *Foreign Language Annals, 29*(3), 331-342.

Thompson, K., Leintz, P., Nervers, B., & Witkowski, S. (2010). The integration of listening model: An approach to teaching and learning listening. In A. D. Wolvin (Ed.), *Listening and Human Communication in the 21st Century* (pp. 267-286). Wiley-Blackwell.

Thorndyke, P. W. (1997). Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology, 9*, 77-110.

Tode, T. (2013). Chunking and prefabrication. In P. Robinson (Ed.), *The Routledge Encyclopdeia of Second Language Acquisition* (pp. 75-77). Routledge.

Tomitch, L. M. (1988). Schema activation and text comprehension.

Towell, R., & Hawkins, R. (1994). *Approaches to Second Language Acquisition.* Multilingual Matters LTD.

Tsui, A. B., & Fullilove, J. (1998c). Bottom-up or top-down processing as a discriminator of L2 listening performance. *Applied Linguistics, 19*(4).

Tyler, M. D. (2001, June). Resource concumption as a funtion of topic knowledge in nonnative and native comprehension. *Language Learning, 51*(2), 257-280.

Underwood, P., & Wallace, M. (2012). The effects of instruction in reduced forms on the performance of low-proficiency EFL university students. *The Asian EFL Journal, 14*(4), 1-24.

Ur, P. (1984). *Teaching Listening Comprehension.* Cambridge University Press.

Vandergrift, L. (1997a). The cinderella of communication strategies: Reception strategies in interactive listening. *The Modern Language Journal, 97*, 494-505.

Vandergrift, L. (1997b). The comprehension strategies of second language (French) listeners: A descriptive study. *Foreign Language Annals, 30*(3), 387-409.

Vandergrift, L. (1999). Facilitating second language listening comprehension: Acquiring successful strategies. *ELT Journal, 53*(3), 168-176.

Vandergrift, L. (2004). Listening to learn or learning to listen. *Annual Review of Applied Linguistics, 24*, 3-25.

Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language Teaching, 40*(3), 191-210.

Vandergrift, L. (2011). Second language listening: Pressage, process, and pedagogy. In E. Hinkel (Ed.), *Handbook of Research in Second Language Teaching and Learning* (pp. 455-471). New York: Routledge.

Vandergrift, L. (2013). Listening. In P. Robinson (Ed.), *The Routledge Encyclopedia of Second Language Acquisition* (pp. 400-404). Routledge.

Vandergrift, L., & Goh, C. C. (2012). *Teaching and Learning Second Language Listening.* New York: Routledge.

VanPatten, B., & Jegerski, J. (2010). Second language processing and parsing. In B. VanPatten, & J. Jegerski (Eds.), *Research in Second Language Processing and Parsing* (pp. 3-23). John Benjamins Publishing Company.

Verspoor, M., Lowie, W., & De-Bot, K. (2009). Input and second language development. In T. Piske, & M. Young-Scholten (Eds.), *Input Matters in SLA* (pp. 62-80).

Vitevitch, M. S., & Luce, P. A. (2016). Phonological neighborhood effects in spoken word perception and production. *Annual Review of Linguistics, 2*, 75-94.

Volenec, V. (2015). Coarticulation. In D. Jasmine (Ed.), *Phonetics: Funcamentals, Potential Applications and Role in Communicative Disorders* (pp. 47-86). New York: Nova.

Wang, Y. T. (2005). An exploration of the effects of reduced forms instruction on EFL college students' listening comprehension. *Unpublished master's thesis, National Tsing Hua University, Hsinchu, Taiwan*.

Weber, A. (2000). The role of phonotactics in the segmentation of native and non-native continuous speech. In *Proceedings of the Workshop on Spoken Access Processes* (pp. 143-146).

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language, 50*, 1-25.

Weber, A., & Cutler, A. (2006). First-language phonotactics in second-language listening. *The Journal of the Acoustical Society of America*, *119*(1), 597-607.

Weinstein, N. (2001). *Whaddaya Say? Guided Practice in Relaxed Speech* (2nd ed.). Longman.

White, G. (1998). *Listening.* Oxford: Oxford University Press.

Wilson, M. (2003, October). Discovery listening: Improving perceptual processing. *ELT Journal, 57*(4), 335-343.

Wingfield, A., Alexander, A. H., & Cavigelli, S. (1994). Does memory constrain utilisation of top-down information in spoken word recognition? Evidence form normal aging. *Language and Speech*, 221-235.

Witkin, B. R. (1990). Listening theory and research: The state of the art. *Journal of the International Listening Association, 4*, 7-32.

Wolvin, A. D., & Coakley, C. G. (Eds.). (1993a). *Perspectives on Listening.* David Fulton Publishers.

Zwaan, R. A., & Brown, C. M. (1996). The influence of language proficiency and comprehension skill on situation-model construction. *Discourse Processes, 3*(21), 189-327.

APPENDICES

APPENDIX A

Pre-test - Listening Cloze Task

The following is a transcript of 34 utterances to which you are going to listen. As you notice, some words are blanked out. You are required to fill in the blanks on the basis of what you listen to. You will hear the number of the relevant utterance before it is played.

1. Read ……write some notes
2. ……lot not just ……..two
3. ……..……part in …… TV show
4. The ……engaged for over an hour
5. We went ……
6. ……second hand
7. ……some money
8. I…… to stay
9. ……Jane
10. ……against the wall
11. ……talk to the person …….. you really get ……
12. I think ……..quite a lot of redundancy money
13. ……more, carefully, they
……the jewels.
14. You ……..police immediately
15. ……soup …… fridge
16. ……key in the office
17. ……smoking ……get better
18. People will pay for that ……a digital file
19. …….. did not ……anything
20. The experiment sounds ……..today
21. The computer which we bought …….
22. You knew their ……
23. ……do is make it up

24. Yes, everyone ……Hag apart

……annoying people…..……

25. It is a ……file

26. ……..letters …… to open

27. I don't ……imported caviar

28. Much learning ……..……..today

29. I like plain food ……spicy ……fresh

30. ……take …… want

31. He wants to ……..……home

32. …….. new books ……read

33. When are you ……

34. …….over for dinner soon

APPENDIX B

Pre-test – Cloze-task Listening Script

1. Read ***his book and*** write some notes

2. ***There were a*** lot not just ***one or two*** two

3. …***because he was taking*** part in ***this*** TV show .

4. The ***number has been*** engaged for over an hour

5. We went ***to her room***.

6. ***They are*** second hand

7. ***I asked him for*** some money

8. I ***wanted her*** to stay

9. ***They are for*** Jane

10. ***He stood his gun*** against the wall

11. ***So that you can*** talk to the person ***you are with and*** you really get ***to know her.***

12. I think ***Mark has been paid*** quite a lot of redundancy money.

13. ***If they had searched*** more, carefully, they ***might have found*** the jewels.

14. You ***should have called the*** police immediately

15. ***There is some*** soup ***in the*** fridge

16. ***I might have left my*** keys in the office

17. ***If he stopped*** smoking ***it would*** get better

18. People will pay for that ***rather than buying*** a digital file

19. ***And I*** did not ***send them*** anything

20. The experiment sounds ***absolutely terrible to us*** today

21. The computer which we bought ***in July was stolen***.

22. You knew their ***house was burgled last year.***

23. ***All he had to*** do is make it up

24. Yes, everyone ***knows me as*** Hag apart ***from a couple of*** annoying people ***as simple as that.***

25. It is ***a reproduction of a digital*** file

26. ***Aren't there some*** letters ***for her*** to open

27. I don't ***care if there is no*** imported caviar

28. Much teaching and learning ***that goes on*** today

29. I like plain food ***that is not*** spicy ***but nice and*** fresh

30. ***I shall*** take ***as much as*** want

31. He wants to ***come and see us at*** home

32. ***There are some*** new books ***I must*** read

33. When are you ***taking him to see her?***

34. ***You must come*** over for dinner soon.

APPENDIX C

Pre-test – Dictation-task Listening Script

1. Algeria was split into department.

2. Are you into golf?

3. Do I owe you anything?

4. Docked the deep blue sea

5. He has hair over the ears.

6. I cannot beat you at this game.

7. I hate going to museums and art galleries.

8. I taught classes this morning.

9. It has seen suppression and uprising revolution.

10. It's starting to ach.

11. Law and order

12. Send Frank a card.

13. She had university students.

14. The boat is useless without the oars.

15. The Last Car

16. The sun burnt my neck.

17. There is a comma after that.

18. They cheat people out of their money.

19. They shot bears.

20. You have already had yours.

21. You taught yourself French.

APPENDIX D

Post-test - Listening Cloze Task

The following is a transcript of 36 utterances to which you are going to listen. As you notice, some words are blanked out. You are required to fill in the blanks on the basis of what you listen to. You will hear the number of the utterance before it is played.

1. Beaten against the wall .....

2. .....not get .....few days

3. but there is far .....flippancy

4. .....it puts a great deal of .....life

5. …weeping with despair .....pounds

6. I don't know what's the matter.....something wrong

7. .....actually eaten one myself

8. I have to .....by Friday

9. .....you

10. .....get them .....possible

11. .....what the fastest .....

12. .....if you paid me more

13. .....you , I .....careful when talking to him

14. If .....by plane.....simpler

15. In the .....how to catch and cook hedgehogs

16. is when ..... am sitting …….same table with .....Tobacco industry

17. .....driving .....ages

18. It must be hard to believe .....

19. It was far .....one

20. .....a savoury dish

21. I've .....holiday

22. Laura pattered .....other window

23. .....international issues

24. .....they have to get to my .....

25. Thanks but that will .....I just want

26.

27. The .....have is newer

28. They cook .....

29. They did .....paper

30. This is not .....

31. Especially .....have so many children

32. .....to live off the land and eat

33. Well they used to at the .....

34. Well, ...last week, do you know .....

35. What do you .....important environmental issues?

36. ..... the British Hedgehog preservation society.

37. You .....and .....soon.

APPENDIX E
Post-test – Cloze-task Listening Script

1. Beaten against the wall ***till he was dead.***
2. ***but he might*** not get ***them for few*** days.
3. but there is far ***less of that*** flippancy.
4. ***Did you find that it*** puts a great deal of ***strain on your private*** life.
5. …weeping with despair ***here are your twenty*** pounds
6. I don't know what's the matter. ***I must have said*** something wrong.
7. ***I have never*** actually eaten one myself.
8. I have to ***get them to him*** by Friday.
9. ***I must e-mail*** you.
10. ***I want them to*** get them ***as soon as*** possible.
11. ***I will ask her*** what the fastest ***way would be.***
12. ***I would work faster*** if you paid me more.
13. ***If I were*** you, I ***would be very*** careful when talking to him.
14. If ***she had gone by*** plane***, it would have been*** simpler.
15. In the ***past you were trained*** how to catch and cook hedgehogs.
16. …is when ***I have found I*** am sitting ***at the*** same table with ***someone from the*** Tobacco industry
17. ***It has been*** driving ***behind us for*** ages.
18. It must be hard to believe ***what has happened.***
19. It was far ***better from the last*** one.
20. ***It was quite*** a savoury dish.
21. I've ***been there on a*** holiday.
22. Laura pattered ***beside him to the*** other window.
23. ***Some of those are*** international issues.
24. ***Tell her that*** they have to get to my ***uncle as soon as possible.***
25. Thanks but that will ***take you too long*** I just want…
26. The ***version that you*** have is newer.
27. They cook ***worse than I do.***
28. They did ***not give her any*** paper.
29. This is not ***the attitude that we have.***
30. Especially ***we are happy to*** have so many children

31. ***We were taught*** to live off the land and eat.

32. Well they used to at the ***start, it was interesting.***

33. Well, I ***sent him the package*** last week, do you know ***when he got it?***

34. What do you see ***as the most*** important environmental issues?

35. ***You are the founder of*** the British Hedgehog preservation society.

36. You ***must come*** and ***see us again*** soon.

APPENDIX F

Post-test – Dictation-task Listening Script

1. Suddenly, she was sitting straight up in bed.

2. A major in the army.

3. You came to a sticky end.

4. Have you ever tried Belgian beer?

5. It's quarter to eight already.

6. It is made of Fur and leather.

7. I cannot speak Spanish.

8. Well, my mother was in the house.

9. Have you taken them from that box?

10. She had put them on the lawn to air them out.

11. You told me that you had your homework done.

12. I got a card for my aunt.

13. Hold the Dog!

14. We had a bad year.

15. She did not go to France that year.

16. Get in touch with me as soon as you get there, won't you?

17. You should send them to <u>h</u>im by priority mail.

18. It only took <u>h</u>er a second to realize what had happened.

APPENDIX G
Samples of the Participants' Answers (Dictation)

The Last Car	The sun burned my neck	There is a comma after that	They cheat people out of their money	They shot bears
The last ……	ressumbled my neck	there is a Camera after that	They cheat people after their money	They ……
the last car	the sunbon my kneeq	this is camera after that	they sheep the paper without money	they shap ears
the last car	they somber my neck	it is come after that	they cheat people after their money	they shope airs
Tha last car	The sun burn my neck	It's camera after that	They cheat people out of their money	They shop birds
the last car	OO	OO after that	they sheep people after their money	shope bears
the last …….	there are some of them in my neck	there is camer after that	they cheat people after their money	they chap air
the last come	by my neck	it is a camera after that	they cheep people …..money	they chap airs
The last car	……..	There is a comra after that	They cheat people out of their money	They shot bears
the last	December my neck	It is a comer after that	that choop people after their money	the chope bears
the last colour	the sambam my neck	this camera of that	they cheep people of money	they shop ears
the length …..	december my neck	descover avter that	they cheet the people ….their money	……
the samber my nack …….	…….	the camera after that	they cheep people after their money	they chop bears
the last car	the sun pet my neck	there is a carner after that	they cheat ….after their money	they shop bears
OOO	…. My neck	does it OO aftter taht	they cheat people OO	OO
The last car	The sumba …	This is a camera after that	They cheat people after their money	They cheape pears
the last came	……..	there is a camera after that	they cheat people out of they money	……
The last pen	The sun bet my neck	There is a comer after that	The cheat people after their money	The shot bears
the last car	the sun by my neck	there is ucomer after that	they cheat people after their money	they shape wars
The last ……	ressumbled my neck	there is a Camera after that	They cheat people after their money	They ……

I hate going to Muesumes and art galleries, too	**I coach classes this morning**	**It has seen oppression, revolution, uprising**
I hate go to the mesum and other	I talk anything	It seem opression revolution
I hate going to meusms don't galeres	I OO	it seen supression revolution
I hate going to museums and art galories too	I took classes this morning	It's sems seperation, revoluting and upraising
I hate go to musen and art galaries	I tought classes this morning	is seems OO revolution OO
I hate go to meusime and all galories	I took courses this morning	It seems
I hate go the museme in all galeries	I toke	It seems supression and revolution
I hate going to musims and all galleries	I took classes this morning	it seems pression revolution up rising
I hate going to Musims and all galleries	I took ……..this morning	It seems improtion revolution aprasing
I was going to mazims	I took koses this morning	it seem supression revolution aprizing
I hate going to mesium in all gallories	I …..;this morning	it's seem a pression revolution aprizing
I hate going to mesum is not galorion	I talk …….;;;this morning	it seem abvaling revolu uprising
I hate going to mesums and art galeries	I took carses this morning	It seems pression and revolution apprising
I h going to musims and art galories	I took courses this morning	its seem spression revolution up rising
I hate going to mesim	I talk ……..;this morning	It is seems deprechen
I hate go to mieusim and galoriesm	I took cassis this morning	it seems supression revolution apraising
I hate going to museums and art galeries	I tought …this morning	It seems supresing and revolution
I hate ggoing to musimes and old galories	I took classes this mornin	it seems suppression revolution
I hate going to Muesumes and art galleries, too	I coach classes this morning	It sems

APPENDIX H

Samples of the Participants' Answers (Cloze)

I don't care if there is no imported caviar.	Much teaching and learning that goes on today	I like plain food that is not spicy, but nice and fresh
I don't ------ imported caviar	Much learning in ,,,,closer …today	Plain food this not spicy in nice ….fresh
I don't OO imported caviar	Much learning and teaching are lying today	Plain food this not …spicy ……fresh
I don't care if no …..imorted caviar	Much learning and teaching that goes on today	Plain food is not ….spicy but nice and fresh
I don't care it is not imported caviar	Much learning OO today	Plain food is not ..spicy but nice and fresh
I don't care if it is imported caviar	Much learning goes on ……today	Plain food this noo spicy OO fresh
I don't like imported caviar	Much learning ** today	Plain food is not spicy but it is fresh
I don't care if doesn't imported cavdiar	Much learning goes on …..today	Plain food is not spicy but nice and fresh
I don't care if that an imported caviar	Much learrning OO today	Plain food is not spicy but is not fresh
I don't care for this imported caviar	Much learning and teaching goes on today	Plain food is not spicy but nice and fresh
I don't OO it's not imported	Much learning and teaching spposed on today	Plain food is not spicy a nice fresh
I don't care of this no imported caviar	Much learning that goes on today	Plain food is not spicy but fresh
I don't care if this no imported caviar	Much learning goes on today	Plain food is not spicy but OO some fresh
I don't carre this no imported caviar	Much learning and teaching OO today	Plain food this no spicy for fresh
I don't care of this imported caviar	Much learning and teaching go on today	Plain food OO spicy OO fresh
I don't care if this is not imported caviar	Much lreaning blows on today	Plain food ther's no spicy and icing fresh
I don't care imported caviar	Much learning goes on today	Plain food is not spicy OO fresh
I don't care if it's imported caviar	Much learning goes on today	Plain food is not spicy nor fresh.
I don't care if that's not imported caviar.	Much learning go ….OO today.	Plain food Is not spicy but it's fresh
I don't care if this imported caviar	Much teaching and learning OO today	Plain food is not spicy and not fresh

You know their house was burgled last year	**All he had to do is make it up.**	**Yes, everyone knows me as Hag, apart from a couple of annoying people, as simple as that**
You knew their houses	OO do is make it up	Yes, everyone known miss Hag apart ……people annoying people sample that
you knew their houses	OO do is make it up	Yes, everyone has those ….Hag apart OO annoying people simple s
You knew their house was burgle	What we have to do is make it up	Yes, everyone knows its Hag apart of couple of annoying people simple is that
You knew their houses OO	All you have to do is make it up	Yes, everyone knows Hag apart a couple of the annoying people simple that
you knew theire houses	What we had to do is make it up	Yes, everyone knows Hag apart about annoying people souple that
youknew their houses ….	what have you do is make it up	Yes, everyone knows Hag apart OO annoying people OO
You knew their houses ws OO	All you have to do is make it up	Yes, everyone knows his Hag apart from annoying people simple is that
you knew their OO	All what you have to do is make it up	Yes, everyone OO Hag apart form the annoying people simply as that
You knew their houses	all 've have to do is make it up	yes, everyone OO Hag apart from a couple annoying people simple as that
you knew their houses O	what OO do is make it up	Yes, everyone knows me Hag apart OO annoying people sympl OO
You kenw their house is burggle	Ii do i smake it up	Yes, evryone knows his Hag apart OO annoying people simeple is taht
you knew their hous is	all we had to do is make it up	Yes, everyone …….Hag apart ……annoying people……
you knew their house was	OO do is make it up	Yes, everyone knows his Hag apart ** annoying people simples that
You knwe their house's pergal ..	OO do is make it up	Yes, everyone knows me Hag apart OO annoying people simpliza
You kenw their house were OO	What do we have to do is make it up	Yes, everyone know OO Hag apart OO anooying people simple is taht
You knew their OO	OO do is make it up	Yes, everyone OO Hag apart OO annoying People….
You knew their house is boreger	all what you have to do is make it up	Yes, everyone OO Hag apart from a cople annoying people simple is that
You knew their house was OO	we have to do is make it up	Yes, everyone XX hag apart couple ennoying people…….
You knew their house is OO	We OO do is make it up	Yes, everyone OO Hag apart OO annoying people simplest OO

الملخص

تركز هذه الدراسة على الصعوبات التي يواجهها طلبة اللغة الإنجليزية بجامعة المسيلة والمتعلقة بفهم المنطوق و القدرة على فك الشفرة الصوتية للكلام المتصل. حيث تهدف الى تشخيص المشاكل المتعلقة بخصائص الكلام المتصل ومعرفة مدى تأثيرها على عملية الفهم، و معالجتها من خلال اعتماد المنهج التشخيصي. حيث أن هذا المنهج يسمح بإدماج الخصائص في حصص الاستماع وتطوير أسلوبي الاستماع من أسفل المستويات الى إلى أعلاها في اللغة والعكس معا. واعتمدت الدراسة على المنهج التجريبي وتحليل البيانات كمية والنوعية للإجابة عن الأسئلة المطروحة و التحقق من الفرضيات الموضوعة. تم اجراء اختبار قبلي على عينتين من الطلبة احداهما تجريبية والأخرى ضابطة. ويتمحور الاختبار أساسا على القدرة على فك الشفرة الصوتية في وجود خصائص الكلام المتصل. بعد الاختبار تلقت العينة التجريبية دروسا في خصائص الكلام المتصل باعتماد المنهج المذكور، بينما تلقت العينة الضابطة دروسا مشابهة دون التركيز على مميزات الكلام المتصل. أبرزت نتائج تحليل أخطاء الاستماع أن خصائص الكلام المتصل تمثل عقبة في عملية فك الشفرة الصوتية خاصة في عمليتي ادراك الكلمات و تقسيمها لدى العينتين وذلك بنسب متفاوتة. كما تبين أن الطلبة يميلون الى الاعتماد على بعض استراتيجيات الاستماع من الأعلى الى الأسفل والتي تبين أن لها تأثيرا سلبيا الى حد ما على قدرة الاستماع لديهم في غياب المهارات الأساسية —من الأسفل الى الأعلى– لفك شفرة الكلام المتصل. وبينت النتائج الكمية تحسنا ذا دلالة إحصائية في قدرات الطلبة الذين تلقوا الدروس فقط دون الآخرين. كما تم اجراء مقابلة مع طلبة العينة التجريبية حيث ثمَّنوا الدروس وتحدثوا عن تحسن نوعي في قدرتهم على الاستماع عموما.

Résumé

Cette étude porte sur l'efficacité de l'enseignement de la parole structurée (connected speech) à améliorer le décodage de l'oral anglais et l'attitude des apprenants envers cet aspect de la langue anglaise. L'étude est orientée vers l'organisation de sessions d'écoute équilibrées avec focalisation sur les deux modes d'enseignement ascendant et descendant en même temps. Par conséquent, la recherche tente d'intégrer l'enseignement de la parole structurée dans les séances de compréhension orale suivant une approche diagnostique du signal (suggérée par Field, 2003 et 2008a) afin de révéler les traits de la parole structurée qui sont à l'origine de l'échec des apprenants en matière de compréhension orale. Un groupe expérimental assiste à des séances d'écoute suivies d'un cours sur les traits de la parole structurée sources des problèmes lors de l'écoute. Un groupe de contrôle assiste aux séances d'écoute sans aucun apport théorique sur les traits de la parole structurée. Les résultats du pré-test montrent que l'aspect parole structurée crée de sérieux obstacles au décodage et à la segmentation du discours naturel avec une interférence potentiellement négative due à un traitement selon une approche descendante. Le groupe expérimental s'est statistiquement amélioré de manière considérable après le traitement. Ceci implique que l'intégration de l'enseignement explicite de la parole structurée est efficace pour améliorer la capacité de décodage des apprenants. Une interview post-traitement a montré que les apprenants sont en faveur de ce genre d'enseignement et manifestent une amélioration conséquente de leur compétence en compréhension orale de l'anglais.