Second language listening in an academic context: lexical, perceptual, and contextual cues to word recognition.

Doctor of Philosophy

Department of English Language & Applied Linguistics

University of Reading UK

Jane Ward

August 2018

Declaration

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

This thesis has been proofread by an approved third-party proofreading service in accordance with the guidelines of the University of Reading.

Candidate: Jane Ward

_____

Acknowledgements

I would like to thank the participants of this study, their tutors, and the Course Directors at the University of Reading's International Study and Language Institute for their roles in my data collection.

I am also grateful to my dear friends, advisors, and proof readers Dr. Lynda O'Brien, and Associate Professors Colin Campbell and Joan McCormack.

I am similarly appreciative of my supervisors, Dr. John Field and Dr. Jackie Laws for all they taught me over the years and to my examiners for their valuable contribution to this thesis.

I am indebted to my friends for their support, including Pauline, who I lost along the way.

Finally, my sincerest thanks to my parents.

Abstract

Second language (L2) listening processes are generally thought to be under researched, especially in academic listening contexts, despite the importance of listening to the academic performance of international students in English-speaking countries. Additionally, research in the area is often not informed by theory from other fields. This research addresses both issues and investigates the word recognition processes of L2 listeners using methods informed by psycholinguistic theory. It explores the relationship between the words listeners' report having heard and the evidence on which they rely, in the form of cues, within the speech signal. In addition, by ascertaining how much perceptual information is available to L2 listeners, the degree of their ability to build meaning and construct an accurate interpretation of the overall discourse can be implied, processes crucial in an English for Academic Purposes context.

A paused transcription tasks is used, which is a method designed to induce the processes used in real-life listening. In this study, 171 university pre-sessional student participants listened to a recording taken from a pre-sessional listening course book which replicated a section of an academic lecture. Pauses were inserted irregularly into the recording at the end of target phrases. Whenever a pause occurred, participants reported immediately in writing the last four or five words heard. The outcome was a set of responses in the form of a series of words. Participants' correct responses were analysed to establish the lexical features they may have relied on in their accurate word recognition processes, and their incorrect responses were analysed to establish which cues they made use of in their erroneous word recognition processes. This enabled insight into participants' cognitive processes employed during the listening process.

The findings were extensive and addressed several gaps in the knowledge of L2 listening. The major findings were i) word frequency was a key factor in accurate word recognition; ii) perceptual cues were more heavily relied on than contextual cues when listening broke down; iii) incorrect responses were informed by more than one single cue; iv) co-textual cues should be viewed separately from contextual cues; v) co-text is highly influential to L2 listeners; and vi) participants' lack of confidence in their incorrect responses indicate they are behaving strategically.

The findings support the notion that traditional teaching methods are not appropriate for L2 mid-level listeners, especially those preparing for academic study. Consequently, a call is made for a re-evaluation of listening training both in an academic context and more generally and suggestions for doing so are put forward.

List of Appendices

List of Tables

List of Figures

List of Abbreviations and Glossary

BNC	British National Corpus
CEFR	Common European Framework of Reference
EAP	English for Academic Purposes
L1	First Language
L2	Second (Additional) Language
MSS	Metrical Segmentation Strategy
NES	Native English Speaker(s)
NNES	Non-Native English Speakers(s)
SLA	Second Language Acquisition
TEL	Technology enhanced learning

Conceptual cues	Cues provided by contextual and co-textual information.
Contextual cues	Information provided by various types of listener knowledge, i.e. knowledge of the topic, the world, and the listener's interpretation of the text so far. Co-textual cues are often included within the notion of 'context' in second language listening literature.
Co-textual cues	Cues provided by the words immediately surrounding the utterance. Co-textual cues are often included within the notion of 'context' in second language listening literature.
Formulaic sequence/chunk	Highly frequent strings of speech.
High level listening processes	See 'conceptual cues'. Cues provided by contextual and co-textual information.
Low level listening processes	See 'perceptual cues'. Auditory cues provided by the speech signal
Perceptual cues	See 'low level cues'. Auditory cues provided by the speech signal
Syntactic cues	A type of co-textual cue provided by the listener's knowledge of the grammar of the language.
Semantic cues	A type of co-textual cue provided by the listener's understanding of the immediately surrounding words.

1. Introduction

1.1. Overview

The focus of this research is the behaviour of second language (L2) listeners of English in their word recognition processes. It explores the lexical features of words which L2 listeners are able to employ successfully and identifies those features which influence their word recognition when listening in the L2. Additionally, it explores the relationship between listeners' breakdowns in understanding and the evidence on which they rely, in the form of perceptual and contextual cues, within the speech signal. By ascertaining how much perceptual information is available to L2 listeners, the degree of their ability to build meaning and construct an accurate interpretation of the overall discourse can be implied. Finally, the study investigates listeners' confidence in their word recognition decisions in order to consider the extent to which they are forced to listen strategically. The research takes place in an academic context, a setting which, in addition to the challenges of general L2 listening, imposes additional, specific demands on listeners. The nature of these additional demands is considered throughout this study. The findings, therefore, can be related not only to general English L2 listening contexts, but also to the context of academic English L2 listening.

General L2 listening processes seem to be under researched, and academic L2 listening even more so (e.g. Lynch, 2011). The latter is especially alarming given the importance of listening to the academic performance of international students in English-speaking countries. This lack of research is especially noticeable when contrasted with the substantial number of studies which investigate academic writing and reading. Furthermore, the lack of research informed by theory from other fields, such as child language acquisition, cognitive linguistics, and psycholinguistics, has been noted, and Ellis (2003:47) points out the benefits that such interdisciplinary studies would bring to the field of second language acquisition (SLA). The research reported here responds to Ellis' (ibid.) comments. It is an interdisciplinary study that takes advantage of the extensive, existing knowledge of L1 listening processes and applies insights from empirical psycholinguistic research in order to increase understanding of the listening processes of L2 learners in general, and in academic contexts specifically. It applies knowledge gained through the findings to the pedagogic aspect of L2 listening with the aim of re-examining how the skill is taught.

A great deal of psycholinguistic research (e.g. Cutler, 1995; Cutler & Butterfield, 1992; Grosjean, 1985; McQueen, 1998, 2004) has examined how L1 listeners are able to distinguish words in connected speech where phenomena such as elision, assimilation, resyllabification, and reduction take place. Processes, such as lexical segmentation, lexical recognition, and parsing, have been extensively researched, and operations to build meaning and integrate utterances into the wider discourse have been discussed. This theoretical knowledge of expert listening is not often applied to an L2 context, despite its great potential to promote understanding in that area and to inform pedagogical practices. It is also rarely applied to academic L2 listening contexts where understanding lectures is crucial to academic success, and training students to do so is a fundamental aim of pre-sessional, academic English provision.

1.2. Context and rationale

This study is set in the context of a UK university, and specifically relates to the provision of pre-sessional, academic English courses. The participants were drawn from a population of non-native English speaking (NNES) learners preparing for academic study in UK universities, and the research design centred on a task which involved lecture-style recorded material designed to prepare learners for that context.

Reputable pre-sessional courses should, of course, strive to use teaching materials and methodologies informed by research. However, Hyland (2006:5) suggests that:

> many EAP [English for Academic Purposes] courses still lack a theoretical or research rationale and textbooks too often continue to depend on the writer's experience and intuition rather than on systematic research.

Graham (2011:5) concurs, and posits that teaching listening is based on traditional course book design, rather than being informed by theoretical principles. The problem highlighted during the course of this research is that it appears that text books to teach listening *cannot* be informed by 'systematic research' and 'theoretical principles' because L2

general listening is distinctly under-researched, and even more so, in an English for Academic Purposes (EAP) context (e.g. Lynch, 2011). In addition, what little research there is seems to be largely uninformed by research from other fields (e.g. Ellis, 2003). Thus it is questionable whether course books designed to teach listening can adequately fulfil their goal.

The manner of teaching listening posited here is one reflected in both teacher and student attitudes to teaching and learning. In fact, the author's dissatisfaction in both teaching L2 listening and learning to listen in an L2 was the fundamental motivation for this study. As a teacher, I felt inadequate simply telling students that their comprehension answers were incorrect, playing the recording again to illustrate, and moving on to the next exercise. As a learner, my teacher would do the same to me, and when I still could not segment the stream of words after multiple hearings, I would feel dispirited. It seems that this view is not uncommon, and teachers of listening have reported that they are uncertain about how the skill should be taught (Graham, Santos and Vanderplank, 2008). The traditional, comprehension approach to teaching listening involves learners listening to a text and completing a related exercise, which is then corrected. Field (2008d:79) refers to this as "extended practice" rather than teaching. This suggests that listening "takes the form of an activity to be 'delivered' rather than a skill to develop in its own right" (Graham, 2011:5). Furthermore, general online advice for students seems be similarly lacking, even from reputable organisations. For instance, a web page by The British Council (Ahmed, 2015) discusses "five essential listening skills for English learners", all of which relate to top-down strategic processes, such as prediction or inferring meaning, and none relate to word recognition.

The views of students learning to listen appear to be similarly disheartening; listening is often reported by students as the most challenging of the macro skills (e.g. Graham, 2006; Vandergrift, 2007). In research specifically targeted to investigate the learner's perspective of listening, Graham (2006) reports that students feel very demotivated in listening classes, and they even give up as a result. A distinctive study by Bekleyen (2009) highlights a further worrying impact of the inadequate teaching of listening using typical methodologies, such as the comprehension approach. Bekleyen (ibid:671) investigates Turkish trainee English teachers who reported a great deal of anxiety in learning to teach listening in English. It appears that this anxiety stemmed largely from "the low priority placed on listening in their previous language training". Graham (2011:5) suggests that this study implies that "these

would-be teachers dealt with their own listening anxiety as students by simply practising more". It was suggested above that research should inform pedagogy, however, this dissatisfaction expressed by both teachers and students implies that this may not be happening in the context of L2 listening.

1.3. Objectives

The success of pre-sessional EAP students in their future academic studies is influenced by the English training they receive, certainly in terms of listening as their ability to understand lectures is crucial. In addition, listening comprehension directly impacts the development of other skills in L2 learning (Vandergrift, 2007). Brown (1986:286) expresses a need for "some sort of method of investigating the student's problems" in order to teach listening effectively, and it is the objective of this study to provide such a method. Accordingly, it investigates the behaviour of L2 listeners using methods informed by psycholinguistic theory. The findings are discussed in the context of both general and academic L2 listening. Ultimately, the aim is to contribute to the body of knowledge in SLA, and to enable a re-evaluation of the way listening is taught in both general English teaching contexts and on EAP courses. In an academic context, if NNES students have effective listening training on their EAP courses, their chances of success in their degree programmes should be enhanced due to their improved lecture comprehension.

1.4. Outline of the thesis

This initial chapter introduces the study by providing an overview of the research. The context of the study and its rationale are discussed, along with its objectives.

In the second chapter, the psycholinguistic processes of expert listening are discussed in order to provide the theoretical background for the study. These processes are then explored in the context of L2 listening, and possible reasons for listening breakdowns are put forward. Given that the ultimate aim of the study is to promote re-evaluation of L2 listening training, a brief background of research in this field is provided. This allows several gaps in

the literature to be identified and outlined, and subsequently leads to the presentation of the research questions.

Chapter 3 describes the methods used. This includes the data-gathering methods, sampling procedures, materials and their design, procedures for administering the task, ethical considerations, and collation of the data.

Chapter 4 addresses Research Question 1, i.e., the investigation into the influence of three lexical features on participants correct word recognition. The data analysis procedures are provided, and the results of the investigation are presented and discussed.

Chapter 5 begins with details of the categorisation of participants' incorrect word recognition in order to address Research Questions 2, 3, and 4. Then the results of the investigation are presented, some of which are illustrated through examples from the data, and the findings are discussed in the context of how they promote understanding of L2 word recognition.

Research Question 5 forms the final investigation, whereby the supplementary qualitative analysis explores listeners' confidence and strategy use, and is reported in Chapter 6. The data processing is presented, followed by the results. The implications of the findings in terms of the analysis of the incorrect responses, and on L2 listener behaviour in general, are also considered.

Chapter 7 reflects on how the results of this study could impact L2 listening training. Firstly, it presents the findings in the context of the limitations of traditional teaching methods, then it suggests alternative approaches, which include technology enhanced learning methods, to promote effective learning.

Finally, Chapter 8 reviews the findings in light of all the research questions, discusses limitations, and provides recommendations for future research in this field of study. It closes with a call for L2 pedagogy to be re-evaluated, for course materials to be revised, and for teachers to be trained to meet these changes.

2. Literature review

2.1. Introduction

There are two assumptions generally accepted in the field of psycholinguistics on which the current research is centred. The first is that there is a complex set of phases which represent the cognitive processes necessary for a listener to extract meaning from speech. For instance, listeners need to recognise and segment words in the speech stream, identify the meaning imposed by the grammatical structure of the utterance, and comprehend how a series of utterances builds into an extended discourse. The second is that listening is both tentative and highly interactive. Listeners make use of multiple cues in the listening process, which might be perceptual, lexical, syntactic (i.e., related to parsing), or broadly contextual (i.e., related to higher-level conceptual processes), and 'trade cues', i.e., they may ascribe greater significance to some cues than to others. Cues in the speech stream may be more or less influential depending on circumstances. For instance, in cases where there is a reduction in the amount of perceptual information available from the speech stream due to "ambiguity, inconsistency or lack of clarity on the signal" (e.g. Field, 2008d), listeners may hypothesise a lexical match from higher level cues, such as syntax, context, or knowledge of the text so far (Bond, 1999). In the field of psycholinguistics this direction of processing is referred to as top-down processing, meaning that higher level cues are relied upon to inform lower level processes. Alternatively, bottom-up processing refers to instances where listeners use lower level cues to inform higher level processes, such as prosody informing word recognition, or syntactic cues informing meaning building. This interactivity means that listeners continuously update and revise their hypotheses as the speech signal progresses (e.g. Lieberman, 1967; McQueen & Cutler, 2010).

Any exploration of what cues are available to L2 listeners at different levels, and which of these cues are most relied upon in the processes of word recognition, must first identify a cognitive model of listening which provides a complete account of these processes in expert listening. Given that the listening skills of L2 listeners are impacted by a lack of linguistic knowledge, and a lack of experience in undertaking these processes in an L2, the

model can then be used for reference when describing L2 listeners' performance, and to provide a benchmark from which L2 listening can be compared. The model should:

> i) represent the full range of phases of the listening process, and the cues within them, which account for listeners' abilities to transform acoustic cues into linguistic units, and then into meaning; and,
> ii) take into account the interactive relationship between the cues.

To this end, the next section reviews the relevant literature surrounding psycholinguistic models of expert listening processes, from which the study emerges, and presents the framework of analysis for the current research. Subsequent sections include a discussion of lexical cues to word recognition, and the perceptual and contextual cues used by listeners. The focus of the chapter then shifts to discuss L2 listeners. A review of the literature researching both general and academic L2 listener behaviour is provided. Given the academic context within which the study is set, the specific demands of L2 academic listening are discussed, along with problems of word recognition which impede the higher-level processing necessary for successful lecture comprehension. As there is a noticeable paucity of investigations into academic L2 listening, the contribution of the current study is demonstrated. Owing to the pedagogical motivation for the study and the aim to improve listening teaching for EAP students, the chapter then presents a range of classroom-based investigations into L2 listening teaching which are relevant to both general and academic listening environments. A number of gaps in the literature are identified throughout this chapter, and these lead to the final section which presents the research questions that guide the study.

2.2. The listening skill

Listeners extract meaning from speech within three broad phases, which can be referred to as:

> i) decoding: transforming input in the form of acoustic cues into linguistic forms;
> ii) parsing: forming chunks and clauses which are transformed into a proposition; and
> iii) meaning construction: transforming linguistic forms into abstract notions.
>
> (Field, 2013)

The first phase has been extensively researched in the field of psycholinguistics since the 1960s. Very early on in speech perception research, Liberman, Cooper, Shankweiler and Studdert-Kennedy (1967) put forward a model which illustrates how phonemes are perceived; Pollack and Pickett (1964) investigated speech decoding and found that as auditory information increases, so does the listener's ability to decode what has been heard previously. In contrast, full accounts of the listening process and the relationship between the cues, which also include how listeners build meaning, are less common. The few models of the listening process which address all three phases have a similar architecture; however, scholars differ in their identification of the cognitive processes within each phase.

To achieve the aim of the study a framework is called for which can provide a comprehensive model of expert listening which L2 listeners can aspire to attain. Within such a framework, the behaviour of the non-expert participants can be investigated. Field's (2013:84) model of listening, designed to be used for testing listening, is an "empirically attested model of … [listening] as employed by users under non-test conditions", i.e., it articulates the cognitive processes employed by experts in real-life listening. The model has been used by Harding, Alderson and Brunfaut (2015) as a framework for research into L2 listening diagnostic assessment. Crucially for the current study, Harding et al. (2015:328) point out that the model is an appropriate tool to aid identification of "the point of a breakdown in the listening comprehension process as a whole".

Field's (2013) model is summative as it is based on two earlier models. The earliest of the two is Anderson's (1995) model, which groups listening processes into three phases:

i) Perceptual processing: word recognition;
ii) Parsing: words form syntactic structures which then generate a mental representation;
iii) Utilisation: mental representation is augmented by the receiver interpreting the words (e.g., inferring)

(Anderson, 1995:55-60, 379-404)

The phases are broad and, as such, the cognitive processes, or cues, within the phases are not fully specified. In addition, the utilisation phase seems to conflate several processes. For example, it includes both building meaning from a series of propositions and building an

entire discourse by adding new information to previously assembled meanings of the discourse so far.

The second of the two models is that of Cutler and Clifton (1999), who expand on Anderson's (1995) model, and describe four stages of the listening process and cues within each:

i) Decode: separating speech from other peripheral noise, transforming an input consisting of patterns of signal variation over time into an abstract representation;
ii) Segment (overlaps with stage 3): comprising a) segmental and suprasegmental analysis of the input; b) the use of segmentation cues indicating to where word boundaries fall;
iii) Recognise: comprising a) activation of lexical candidates that form potential matches to the input, and competition between them b) utterance level processing;
iv) Integrate: adding the derived message to the discourse model.

(Cutler & Clifton, 1999:151)

Cutler and Clifton's (1999) model focuses on lower level processes, including parsing, but does not present higher levels of processing in full detail, nor does it thoroughly articulate the 'integrate' phase.

With regards to interactivity, these two models appear, essentially, to assume that the listening process takes place in one direction, namely from the bottom-up, and do not specifically account for possible top-down influences. For instance, Anderson's (1995:379) phases are broadly noted as "by necessity partially ordered in time; however, they also partly overlap". Cutler and Clifton (1999) state that the listening processes cannot be dichotomised as broadly interactive or broadly autonomous due to the failure of researchers in the field to clarify the precise nature of its interactivity. That is to say, there may be interactivity between some stages of the process, but not others. Discussing whether higher level processes can influence prelexical decisions, the authors conclude "it might … seem unnecessary … to build top-down information flow into … the listener's word recognition system" (ibid:53).

These limitations, in the context of the current study, are addressed in Field's (2013) model of listening. This model is more comprehensive as, unlike Anderson (1995) and in line with Cutler and Clifton (1999), it divides the perceptual phase into i) phonological and ii)

lexical processes, and this reflects the difference between the phonological and lexical knowledge sources used in these processes. It also refers to the use of prosodic cues in both input decoding and lexical segmentation processes. Unlike both other models, it acknowledges that parsing, what Cutler and Clifton (1999) refer to as 'utterance level processing', includes two distinct elements, namely working out word meaning in relation to a wider context, and imposing syntactic patterns. Finally, the model is in line with the discourse analyst view (e.g. Gernsbacher, 2013) which sees two major phases in processing meaning, i.e., it distinguishes between adding interpretive meaning to an utterance that has been successfully parsed, and then incorporating that unit of meaning into a wider discourse representation. These two phases are conflated in the models of Anderson (1995) and Cutler and Clifton (1999). The five phases of Field's (2013) framework are outlined below, diagrams of each, as well as the cues which support them, can be found in Appendix A. Each of the phases is discussed fully later in the chapter (see Sections 2.3.1 to 2.3.5).

i) *Input decoding*: (effectively, Cutler and Clifton's (1999) 'decode'). The listener transforms acoustic information into groups of syllables, some marked for stress and others not (in languages where stress is a feature). Listeners identify which syllables carry stress and use this information in the 'lexical search' phase to enable lexical retrieval.

ii) *Lexical search*: (equivalent to Cutler and Clifton's (1999) 'segment' and the lexical part of 'recognise'). The listener maps groups of syllables, including prosodic information, to stored forms of spoken words in the lexicon. The listener is assisted in this by segmentation cues which may vary from one language to another. As there are no regular pauses between words in connected speech, the listener must decide the position of word boundaries. In English, prosodic cues based on syllable stress help the listener to achieve segmentation as most content words in English running speech are either monosyllabic or have initial syllable stress (Cutler, 1990, 2012). Thus, listeners segment based on the knowledge that stressed syllables are likely to indicate new content words, whereas weak syllables are unlikely to be word initial. This stage of the process also informs the input decoding stage retroactively because tentative word matching influences the perception of phonemes. A number of possible matches are activated and compete with each other as the utterance continues, until one word is identified as the best

match. As well as phonological information, a lexical search is supported by the listener's awareness of the frequency of the word in speech as more frequent words will be more highly activated. A further process is utilised in this phase, namely 'spreading activation', which relates to the notion that words are stored in the mind by intricate lexical networks (Aitchison, 2003), and therefore words which are associated with each other will be more highly activated.

iii) *Parsing*: A group of completely or partially identified words is carried forward in the mind of the listener until such time as the speaker briefly pauses and/or appears to complete a grammatical unit (phrase or clause) which may be marked by intonational features as well. A syntactic pattern is then recognised in that unit. (The syntactic part of Cutler and Clifton's (1999) 'recognise').

iv) *Meaning construction*: The listener employs world knowledge and inference to add to the raw meaning of the message.

v) *Discourse construction*: The listener makes decisions on the relevance of the new information and how congruent it is with what has gone before and, if appropriate, integrates it into a representation of the larger listening event. (Cutler and Clifton's (1999) 'integrate').

(adapted from Field, 2013:97-101)

The model fully integrates lexical, perceptual, and contextual cues, and identifies the interactive nature of the process. Clearly, these factors are essential to the investigation of the relative strength of cues that L2 listeners might rely on in word recognition processes.

The comprehensive and interactive nature of this model is consistent with Stanovich's (1980) interactive compensatory model of reading. He, along with other scholars (e.g., Adam Just & Carpenter, 2004; Perfetti & Roth, 1980; Rumelhart, 2004) argues against strictly bottom-up and top-down models because, as with listening, multiple sources of information are available to readers, and are processed concurrently. The compensatory nature of Stanovich's (1980) model of reading is particularly pertinent to this study as Stanovitch considers the impact of reading proficiency on the behaviour of readers by assuming that "a deficit in any particular process results in a greater reliance on other knowledge sources, regardless of their level in the processing hierarchy" (ibid.:63). Thus, a reader who is less

proficient at word recognition may rely more on contextual information than a reader who is more proficient in this skill. The current study investigates L2 listeners who are less proficient at word recognition, and thus, may also rely on this compensatory strategy. For example, a listener faced with difficulties in word recognition may draw on higher level information in the form of syntactic or contextual cues. Therefore, if listeners lack confidence in the accuracy of their decoding decisions, they may be more likely to act strategically and rely on higher level cues, as suggested by Field (2008d).

2.3. Factors affecting word recognition in listening

This section discusses word recognition in connected speech and the implications of the framework presented in Section 2.2 above, in relation to the different types of cue that might affect recognition. It first considers perceptual cues, and then discusses the lexical factors which influence word recognition. It also considers the effects of syntactic and contextual cues on the listener's ability to identify words regressively in connected speech.

2.3.1. Input decoding

Input decoding involves the mapping of the sound of connected speech, i.e., the acoustic signal, onto words. However, it is likely that the main unit of analysis is not the phoneme, but that listeners map directly from the acoustic-phonetic signal, i.e., the speech stream, to words or chunks. Pisoni and Luce (1987:23) report that "phonemes are rarely, if ever, realized in the speech waveform as a linearly-ordered sequence of discrete acoustic events". That is to say, individual phonemes are not pieced together in a linear fashion until a word is recognised, rather listeners recognise larger sections of speech. This is because, even in single words, phoneme forms vary due to co-articulation, which reflects the movement of the articulators from the proceeding phoneme and on to the next (Roach, 2009). This was established in the early days of speech perception research, when spectrograms of the acoustic signal illustrated the varying articulation of individual phonemes, and showed that phonemes merge with each other within syllables (Delattre, Liberman, & Cooper, 1955; Liberman, 1957). In addition, the manner of pronunciation of phonemes in the speech stream varies considerably due to several factors, such as speech rate (e.g. Miller, 1981), and speakers' physical characteristics (e.g. Johnson, 1996; Liberman et al., 1967).

Possibly the most apparent reason that there are no consistent acoustic correlates of phonemes in isolation, from which a listener might build words, is the phonological complexity of connected speech (Roach 2009:145) (see Appendix B for a glossary of common aspects of connected speech). Thus, the assimilatory processes in connected speech and the lack of consistent boundaries make perception of single words challenging for listeners. For example, a sample of natural speech from Cauldwell (1996) showed that reduction in the pronunciation of the word *where* can lead to the vowel, which is pronounced as a diphthong /weə/ in its citation form, being pronounced as a monophthong /we/ in fast connected speech. Also, Shockey's (2003) analysis of natural speech showed the word '*screenplay*' pronounced as /skri:mpleɪ/, with the /n/ assimilating to /m/ before the labial consonant /p/.

Evidence of the role of whole words in decoding is provided by phoneme restoration research which illustrates that perceptual cue information from the acoustic signal is overruled by word level expectation. This phenomenon was first noted in early research by Warren (1970) who manipulated the word '*legi**s**latures'* in the sentence '*the state governors met with their respective legislatures convening in the capital city*', replacing the phoneme /s/ (in bold) with a cough. Nineteen of his 20 native speaker participants, on reporting whether any sounds were absent, did not notice the absence of the phoneme. In addition, in the example given in the previous paragraph ('*screenplay*' pronounced as /skri:mpleɪ/), Shockey's (2003) experiment showed that although her L2 participants initially reported (in writing) hearing the word '*scream'*, as the utterance continued they reanalysed the input and regressively recognised 'screenplay'. This illustrates that top-down processing in the form of word knowledge (i.e., the meaning of the word, its frequency, and spreading activation), as well as syntactic and semantic cues, serves to compensate for phoneme variability. As such, word, syntactic, and contextual knowledge are discussed in the following sections.

2.3.2. Lexical search

The 'lexical search' phase of the analytical framework involves mapping groups of syllables and prosodic information to spoken forms of words stored in the lexicon (Aitchison, 2003). A number of possible matches are activated by the incoming speech, referred to as a cohort. The candidate words, which are those considered compatible with the incoming

speech, are subject to various degrees of activation, competing as the speech progresses until one word, the most highly activated, emerges as the best match. This process is known as *the competition effect* (Cutler, 2012; McQueen, 2007). The likelihood of a word being present in the utterance is determined by a number of cues, stored as part of the word's lexical entry. At this point, it should be noted that lexical segmentation is discussed both as an element of competition, and as a separate process. In the study reported here, lexical segmentation is treated separately, and discussed in the next section.

Word frequency

The major cue in many accounts of the competition for lexical matching is lexical information in the form of word frequency cues. Knowledge of the frequency of a word is stored in the mind as part of its lexical entry and it has long been established that the most frequent words are those most easily processed (e.g. Marslen-Wilson, 1990; Morton, 1979), resulting in quicker, and more successful, recognition (Kirsner, 1994; Luce, 1986a). This is illustrated in a gating experiment by Grosjean (1980:270) who found that "low-frequency words were isolated [recognised] in 274 milliseconds on the average as compared with 213 milliseconds for high-frequency words". For instance, the sentence '*Stephen worked on a tr–'* initially activated in his listeners more frequent modes of locomotion, i.e. '*train, truck, tram'.* It was only when more phonetic information was forthcoming that these words were deactivated and the less frequent word '*trawler'* was matched when the final /ɘ/ was heard. However, Tyler (1984) found that effects of frequency are not clear cut. The results of her gating experiments showed that, even though more high than medium or low frequency words were activated, words across the whole frequency range were stimulated. Additionally, she found no evidence of frequency effects 200 milliseconds into the word. The results from these two experiments illustrate that there is no simple model of word recognition, and support the notion discussed in Section 2.2 that multiple cues, which are subject to varying degrees of activation, contribute to listening processes. A final note with regards to frequency is its relationship with contextual cues. Cutler (2012:68) cautions that the effect of frequency is "dependent on the uncertainty of the response set". For instance, the topic of a discussion may mean that a low frequency word is more likely to appear. She presents the example of the word '*ford'* which is low frequency, unless the discussion involves cars.

Word stress

A further cue that may contribute to activation includes the extent to which the word matches prosodic evidence in the input. A word's stress pattern is stored in the mind as part of its lexical entry, so it is a particularly important cue to word recognition (Grosjean, 1985). During input decoding, listeners perceive that syllables are marked for stress, and use this cue, alongside others, to recognise words. For instance, in the word /ˈfəʊtəˌgrɑːf/, listeners perceive initial syllable stress and words with stress in this position are activated. The influence of stress in word recognition processes is highlighted by evidence suggesting that the vowels of stressed syllables are rarely mistaken in slips of the ear (Bond, 1979; Bond, 1999; Grosjean & Gee, 1987b) as their prominence is heard "against a background of … weak syllables" (Roach, 2009:95).

The prominence of stressed syllables in tone units also assists lexical recognition in longer utterances, where the articulation demands of connected speech impact how words are pronounced. In every tone unit, one syllable is more prominent (Roach, 2009:163) and prominence is a feature of stressed syllables, largely due to loudness, duration, pitch, and quality (ibid:95). The stressed syllable in a tone unit "assists the listener by foregrounding the most critical words in an utterance" as it provides more reliable input in relation to its phonemes than others due to its prominence (Field, 2013:87). For example, in the spoken utterances 'I'll **think** about it', and '**thanks** very much', the speaker placed focal stress (in bold) on the most important words (Lynch, 2004). A consequence of focal stress foregrounding key words is that syllables in a tone unit which do not bear focal stress are likely to be articulated less clearly, and the longer the tone unit, the more likely that reductions in form will take place, including syllables bearing focal stress (Laver, 1994). For example, the common formulaic chunk *do you **know** what I **mean*** contains two tone units, with a monosyllabic word bearing focal stress in each (in bold). In an extreme form of reduction, this chunk could be articulated simply by pronouncing variations of the two words bearing focal stress, namely 'know' and 'mean', such as /nʔmiːn/. This clearly illustrates two factors which greatly impact the listening process. Firstly, the syllables of least importance in a tone unit may become reduced to such an extent they could be unrecognisable, or may even disappear, and secondly, the syllables bearing focal stress, i.e. those of greatest importance, may also be reduced to such an extent they become difficult to identify (Field 2008).

Word length

Evidence in the input in the form of the number of syllables of a word provides a further cue that may contribute to activation processes. According to Grosjean (1980), the effect of word length has not been fully articulated in word recognition models, although Marslen-Wilson and Welsh (1978) suggest that shorter words are recognised more quickly than longer words if no context is provided. However, Cutler and Norris (1979) suggest that there may be a delay in recognising short words in sentential context as listeners may expect further acoustic information. That is to say, on hearing a monosyllabic word, listeners may not make a confident match until further syllables are heard. For example, on hearing the syllable '*part'*, a monosyllabic word and the initial syllable of many polysyllabic words (e.g., particle, partner, partisan), a confident match could not be made until further acoustic information is heard. On the other hand, the offset of longer words may be identified as they occur, and identified with confidence. Luce (1986b) shares this view, suggesting that short words may not be identified in connected speech until the subsequent word has been identified.

Networks of association

A further type of lexical knowledge used in word recognition makes use of the way in which words are stored in the mind in connected networks of association (Aitchison, 2003). According to Fodor (1983:80), "the mental lexicon is a sort of connected graph, with lexical items at the nodes with paths from each item to the other". Each 'path' is a type of semantic link, and Aitchison (2003:86) states that the four most common, according to word association tests, are:

- coordination, e.g., salt → pepper, black → white
- collocation, e.g., salt → water, bright → red
- superordination, e.g., insect → butterfly, colour → red
- synonymy, e.g., starved → hungry

Networks of association have a top-down role in lexical search as words within a similar semantic field are activated and recognised more quickly than those that are not (Aitchison, 2003); this concept is known as spreading activation. For example, in a

conversation about driving, words such as '*car'* and '*road'* will be activated and recognised more quickly than words outside the lexical network. Also, slips of the ear "almost always preserve the word class of the target" (Aitchison, 2003:103), thus verbs are replaced with verbs, and so on. Clearly, networks of association are not identical for all language users (Meara, 2009), e.g., the better their knowledge of the topic under discussion, the more effective spreading activation will be.

Delayed word recognition

Although the lexical search phase of the listening process involves recognising words in the speech stream, this matching may not always take place on a word by word basis, as represented by views of the competition effect.

In the early days of L1 listening research, an influential study by Pollack and Pickett (1964) found that the intelligibility of the first word in a spoken sentence was largely unrecognised until "at least up to three [further] words" had been heard. In addition, the researchers excised single words from the speech stream and found that listeners were largely unable to decode them, likely due to the assimilatory effects of connected speech (as discussed in Section 2.3.1). In contrast, single words read from a spoken list were decoded far more accurately. As the number of concurrent words excised from the speech stream increased, so did the decoding success rate. This is likely because cues at different levels became available, such as prosody, in the form of syllable stress and the focal stress of a tone unit, and syntax.

In further significant research using the gating paradigm (Grosjean, 1980), whereby "one truncates all but a small amount of the beginning of an utterance, then re-introduces the deleted material in small increments (gates) until the entire utterance is heard" (Shockey, 2003:95), Grosjean (1985) found that a great deal of word recognition takes place retroactively, that is to say, words are identifiable only after subsequent words have been heard. Using the same paradigm, a study by Shockey (2003:103) using natural speech is particularly revealing. Native English speaker (NES) participants heard the phrase '*and they arrived on the Friday night*' in 50 millisecond gates. She concludes from her results that it was only between gates 9 and 10, half a second after the utterance began, that "sensible interpretation" took place, and no "viable hypotheses" as to the words of the utterance took place until then. Even at gate 10, only 40% of participants had accurately recognised the

words. These studies provide evidence of top-down processing whereby cues, not available in single words, retroactively aid word recognition. Thus, it seems that word recognition may operate at levels higher than single words. An important consideration is that retroactive word recognition may result in a reduction of competition effects in lexical recognition. That is, if a word is not perceived accurately until later in the utterance, competition effects will not be available at the time the word is uttered.

Co-occurrence effects

Also stored in the mental lexicon are formulaic sequences, or 'chunks', of frequently co-occurring words, e.g., 'in front of' and 'I don't know' (Field, 2008d). These are stored as a single entry, and speakers find it more efficient and effective to use already formed prefabricated strings of words, rather than create new ones (Wray, 2002). Thus, expert listeners' knowledge of highly frequent strings of speech, referred to by Miller (1956) as 'formulaic lexical chunks', and throughout this thesis as 'chunks', is an important factor in lexical search processes. For speakers, chunks are efficient to use, they require less processing as they are stored fully assembled as part of the lexicon, and hence they "constitute single choices, even though they might appear to be analysable into segments" (Sinclair, 1991:110). For listeners, chunks are easy to process; a string of speech in the form of a chunk is accessed automatically from long term memory as one item, rather than as a series of individual items. Thus, working memory resources, where acoustic information is stored temporarily while a lexical match is sought from long term memory (Baddeley, 1992), are free for other processing tasks.

Spoken chunks often form tone units, in which less important words are less prominent and are affected by the reductions in form mentioned above. Speakers automatically assume that chunks are part of listeners' lexicons and, as such, are spoken quickly, imprecisely, and with pauses, usually at tone unit boundaries (Wray, 2002). Despite this, chunks aid lexical recognition as they may be easier to recognise than single words due to evidence being available to the listener at several levels, i.e., phoneme, syllable, word, and co-text. In addition, the listener's familiarity with the chunk may override reduction, such as omitted phonemes, as illustrated by the gating experiments, discussed above, which provide evidence of delayed word recognition.

According to Conrad and Biber (2005:57), there are several characteristics of chunks which appear to be most recognised by researchers in the field, namely: "fixedness; idiomaticity; frequency; length of sequence; completeness in syntax, semantics, or pragmatics; and intuitive recognition by native speakers of a language community". In terms of classifications of chunks, numerous suggestions have been put forward relating to their function in discourse, and Wray (2002:262) created a "macro model … to accommodate the mechanisms of those [previous] processes within a single framework". Her 'heteromorphic distributed lexicon' combines previous models and identifies five lexicons by the functions to which they relate. The lexicons and examples of the formulaic word strings given for each are as follows:

i) Grammatical: *in order to, on account of, out of*
ii) Referential: *give* NP *to* NP, *highly likely, half past (number)*
iii) Interactional (routine): *great to see you, look out!, I'll give you* NP *for it*
iv) Memorized: Hamlet's soliloquy, times table, songs
v) Reflexive: *Bloody hell, goodness gracious, what the*!

Wray (2002:263)

In an academic discourse the context of the study reported here, chunks occur almost as frequently as in conversational settings. Conrad and Biber (2005) report that three and four word chunks make up approximately 20% of academic discourse, as opposed to 28% in conversation. Of this 20%, around 60% are noun and/or prepositional phrases, and most of these are referential phrases used to specify attributes or aspects of information. Examples include '*in the case of*', '*on the basis of*', and '*at the same time*' (ibid.:66). It is evident, therefore, that referential chunks play an important role in the delivery of precise information in academic lectures.

An element related to the lexical search phase of the analytical framework is collocation, a type of chunk explored in the current research. According to Barnbrook, Mason & Krishnamurthy (2013:3) there is "considerable variation" in the use of the term 'collocation' which is evident in the literature. Sinclair (1987) differentiates between formulaic chunks and collocation by referring to their stability, he states that formulaic chunks are fixed to some degree, whereas collocations are more fluid, allowing more variation, e.g., 'hard luck/hard work/hard facts'. It is evident in this example that 'hard'

cannot be interpreted correctly without the subsequent word. Halliday (1966, cited in Barnbrook, Mason & Krishnamurthy, 2013:39) discusses the lexical-semantic nature of collocations and refers to the relationship of the synonyms 'strong' and 'powerful'. 'Strong tea' is acceptable in terms of syntax and lexicality, however, 'powerful tea', although syntactically correct, would be marked by an expert user of the language as 'unlexical'.

Finally, Aitchison (2003:91) refers to collocation as words likely to be found together, ranging on a cline from words likely, but optionally, found together (e.g., fresh-faced youths), to idioms. This notion of a continuum is generally accepted throughout the literature (e.g., Barnbrook et al., 2013; Wray, 2002), hence formulaic chunks, including collocation, are a crucial element of language processing. However, unlike formulaic chunks, which are processed as single entries, collocations usually retain their separate senses, with one sense influencing the other. Exceptions to this include collocations in which a revision of the sense of one of the pair is called for, for instance, 'heavy smoker' (Field, 2008d). In this example, the implication is not that the smoker is heavy, rather that the smoker smokes a lot.

2.3.3. Lexical segmentation

A crucial aspect of word recognition is the listener's ability to segment words in the speech stream (Cutler & Norris, 1988). Many accounts of listening view lexical segmentation as interacting with competition effects, discussed above (e.g., Norris, McQueen and Cutler Norris, McQueen, & Cutler, 1995), in that the likelihood of a word being present is partly determined by the presence of lexically stressed syllables that are potentially word initial. Due to the complexity of segmentation processes, they are discussed separately from lexical search in the current section.

As a result of aspects of connected speech, word boundaries in continuous speech are not reliably marked, thus listeners must determine the boundaries. As the utterance continues and further cues become available, listeners may need to regress and reanalyse the speech stream, including the position of word boundaries. Expert listeners use many cues to word boundary position, and speech recognition and segmentation models which discuss the cognitive processes are widespread. For example, in Norris' (1994) connectionist model, 'Shortlist', word recognition takes place at two autonomous levels, phonemic and lexical, and

does not include higher level cues, such as context. In this model, in a given segment of input, the first stage of recognition results in a shortlist of lexical candidates being identified following a lexical search, and the second stage uses phonemic information to narrow the competition until a 'winner' is identified and subsequent segmentation can occur.

Cutler (1991) identifies what she terms a 'metrical segmentation strategy' (MSS), which is founded on the notion that 90% of lexical (content) words in English running speech are either monosyllabic, or have initial syllable stress (primary or secondary). Therefore, it is suggested that expert listeners segment words in the speech stream based on the knowledge that stressed syllables likely indicate new content words, whereas weak syllables are unlikely to be word initial (Cutler & Butterfield, 1992; Cutler & Norris, 1988). This is evidenced repeatedly in several experiments by Cutler and Butterfield (1992). In one experiment, they presented phrases of alternating 'strong/weak' (S/W) syllables to NES at a low volume, and participants reported, in writing, what they heard. Many examples of boundary misperception were noticed due to boundaries being inserted before stressed, instead of unstressed, syllables. For example, in the phrase '***sons** ex**pect** en**list**ment*' (stressed syllables in bold), the syllables alternate S/W/S/W/S/W, with two content words bearing unstressed initial syllables. One listener responded '*suns expectant listen*', showing that although she segmented the first word boundary successfully, she expected the stressed syllable '*list*' to be word initial.

A similar view of segmentation is put forward by Grosjean and Gee (1987b:144), with an additional feature. They identify two parallel processes, the stressed syllable initiates a lexical search and the weak syllables surrounding the stressed syllable are identified via "a pattern-recognition-like analysis", aided by listeners' knowledge of phonotactic rules and common chunks of language. Thus, it seems that "the heightened acoustic clarity of stressed syllables … facilitates … lexical and/or segmental processing" (Cutler, Dahan and van Donselaar1997a:172).

Cutler (1990:118) points out that it is not in fact syllables that guide segmentation, but the "strong and weak vowels" within them; "strong syllables contain full vowels ... weak syllables contain reduced vowels" (Cutler & Norris, 1988:114). This reflects evidence that listeners use duration, as well as stress, as a cue to lexical segmentation (Smith, Cutler, Butterfield, & Nimmo-Smith 1989), because vowel phonemes in stressed syllables show an increased duration. The terms strong/weak and stressed/unstressed seem to be used

interchangeably in the literature of Cutler and her colleagues and, for the sake of clarity, this research views syllables as being lexically stressed or unstressed, and vowels as being full or reduced depending on their duration (Roach, 2009).

Reflecting this view of the role of stress in segmentation, the MSS was subsequently incorporated into the Shortlist model (discussed above) after McQueen, Norris and Cutler (1994a) conducted experimental simulations to investigate lexical competition and prosody in segmentation. They found that, in cases where there were a high number of activated lexical candidates, lexical stress greatly aided segmentation. The authors revised Shortlist, and put forward a new model "combining competition between alternative lexical candidates and sensitivity to prosodic structure" (Norris, McQueen & Cutler 1995:1209). Subsequently, experiments by van Donselaar, Koster and Cutler (2005) offer support for the notion that segmentation derives from both lexical and prosodic information.

The findings of Cutler and her colleagues, discussed above, were based on word spotting tasks, whereby short texts are manipulated to include prompt words or parts of words, and these involved individual, unconnected words, rather than natural speech. Thus, the results may not be generalisable to natural listening scenarios. However, research by Mattys, White and Melhorn (2005), based on real life listening, provides evidence that listeners perform segmentation by weighing multiple "lexical (knowledge-derived) and sublexical (signal-derived) cues" (ibid.:477), depending on the listening conditions. The model of lexical segmentation by Mattys et al. (2005) brings together not only lexical and prosodic cues, but also refers to the phonotactic constraints of the language, which moderates competition in segmentation. According to Mattys et al. (2005:483) "the speech system favours segmentation solutions that are lexically plausible", thus these constraints guide listeners to where in the speech stream words may begin or end. This view is supported by experiments by McQueen (1998), who reports that because some groups of phonemes do not occur within the same syllable, for instance, the phonemes /lv/ and /mr/, Dutch listeners assume they must signal syllable and, as a consequence, possible word boundaries. He noted that his 52 Dutch listeners found it easier to identify words at the beginnings of non-word sequences when the words constituted a phonotactic boundary (e.g., *pil*, pill, in /pɪlvrem/) than when they were misaligned (e.g., *pil* in /pɪlmrem/) (ibid.:26). He also found that phonotactic constraints were more influential in segmentation processes than prosodic cues, although his experiments were laboratory-based word spotting tasks, rather than based on real

life listening, and so may not be representative of the latter. Accordingly, Norris, McQueen, Cutler and Butterfield (1997) further revised the Shortlist model (Norris, 1994), to incorporate a cue to segmentation based on prosody, alongside constraints on what may constitute a potential word of a language in the form of the Possible Word Constraint (PWC). Among activated candidates, the PWC rejects those parses in which implausible groups of sounds remain between the end of a candidate word and the likely location of a word boundary, for instance, /lv/ and /mr/ in Dutch.

The elements discussed above are incorporated in a framework positing a hierarchy of cues in lexical segmentation by Mattys et al. (2005), which may go some way to inform the analysis in the study presented here. In an experiment attempting to grade the influence of multiple cues on the segmentation of naturalistic sentences, namely stress, sentential context, and coarticulation, their forty-eight NES participants were assigned to groups with intact, mild noise, moderate noise, and severe noise conditions. Each group was presented with 180 sentences manipulated to explore the use of the various cues, e.g., sentences were semantically appropriate or inappropriate, and had stressed or unstressed initial syllables. The sentences were played through headphones and a sequence of letters from the sentence appeared concurrently on a screen, at which point the participants decided whether the sequence of letters was a whole word. The results enabled the authors to grade the cues in order of their influence on segmentation, and they state that "the contribution of each cue … was dependent on its position in the hierarchy" (ibid.:491). The framework is comprised of, in order of importance:

i) sentential context (pragmatics, syntax, semantics)
ii) lexical knowledge
iii) phonotactics, acoustic-phonetics
iv) word stress

(Mattys et al., 2005:488)

At the top of the hierarchy, in optimal listening conditions, both lexical-semantic and lexical information were favoured during segmentation over sub-lexical cues and word stress. Word stress was used in a compensatory manner when contextual cues were not available, and it is placed at the bottom of the hierarchy in favourable listening conditions. Conversely,

in severe noise conditions, word stress was relied upon most heavily by listeners, and to a much greater degree than lexically driven cues. This was also found to be the case in the researchers' other experiments based on single words and reported in the same paper. In the mild and moderate noise conditions, there was no significant effect of cues on segmentation. It was also found that word stress either supports or refutes word activation, a view supported by the PWC (Norris et al., 1997). However, the results should be treated with some caution. Mattys et al. (2005) point out that this hierarchy may be a somewhat basic account of segmentation as the experiments were largely based on speech containing conflicting cues, and thus did not account for the effects of cues working together as in real life listening. In addition, the experiments used highly manipulated material, rather than natural speech.

This section has demonstrated that a wide range of cues contribute to the operations of lexical segmentation, and thus of word recognition. The importance of syntactic cues in lexical segmentation has been mentioned; the role of syntax in overall listening comprehension is discussed in more detail in the next section.

2.3.4. The role of syntax

Analyses of the syntactic structure of speech can be used regressively in a top-down manner to aid word recognition and segmentation, as discussed in Sections 2.3.2 and 2.3.3 (e.g., Grosjean, 1980, 1996, Pollack and Pickett, 1964, Mattys et al., 2005). In addition, syntactic parsing involves listeners monitoring incoming speech until they recognise a completed syntactic unit, at which point they impose upon it a grammatical structure (Field, 2008c). Subsequently, the syntactic unit becomes an abstract proposition and "only an approximation of the original unit [can] be reconstructed" (Jarvella, 1971:413). According to Cutler and Clifton (1999:141), syntactic analysis is "guided by a language user's knowledge of the structure of his or her language", and so is informed by factors such as knowledge of word order, word class, and inflections. However, the frequency by which inflections are affected by aspects of connected speech often renders them unreliable syntactic cues as they are regularly of weak quality, contracted, or omitted. For instance, auxiliary verbs, such as *have/has*, are often weak or elided (Field, 2008d), and the /t/ and /d/ of regular past tense forms of verbs are more often elided between consonants than they are pronounced (Brown, 1977:68).

Parsing is assisted by prosodic cues which indicate where tone units end. As such, it is common that tone units and syntactic units "have similar domains, since they will be mutually reinforcing" (Brown, 1977:89). One such prosodic cue is the pauses in natural speech which speakers use to plan their utterances. As utterances are generally planned in complete clauses, these pauses can provide listeners with a signal that the utterance between pauses is a complete clause to be parsed (Klatt & Cooper, 1975). For example:

Clause 1: *I didn't want to go*
Pause
Clause 2: *but I knew I should*.

A further prosodic cue to aid parsing is speech rate, which usually slows at the end of an utterance (Klatt, 1976). However, in major phrase boundaries, pitch is more salient than speech rate to listeners (Price, Ostendorf, Shattuck-Hufnagel and Fong1991) as focal stress often features late in an intonation group. For example, " 'at 10 o'clock on MONday' " and " 'in front of the SUpermarket' " (Field, 2013:186). Support for the notion that prosodic information helps listeners recognise that groups of words are connected syntactically is provided by Bond (1999:130), who reports that slips of the ear regularly "maintain a global rhythmic similarity", for example 'doggie' misheard as 'donkey' or 'when their condition' misheard as 'when air condition' (Bond & Shockey, 2014:109-110).

The prosodic features of syntactic units can be used to resolve global ambiguities, i.e., at clause or sentence level (Cutler, Dahan, & Van Donselaar, 1997b). In the ambiguous phrase *Paula phoned her friend from Alabama* listeners are more likely to attach the prepositional phrase *from Alabama* to *Paula phoned* if there is a pause between *friend* and *from*. If there is no pause, the prepositional phrase is more likely to be attached to the noun phrase *her friend* (Schafer, 1995). However, it is not always possible to disambiguate multiple interpretations if the prosodic structures of the competing interpretations of the utterance are similar. Cutler et al. (1997b:162) also point out that "listeners' ability to resolve such ambiguity seems to vary depending on whether speakers were aware of the possibility of multiple interpretations when they produced the utterance". In the context of the research reported here, which involves L2 listeners, this notion could be applicable as non-expert

listeners may not be "aware of the possibility of multiple interpretations when they *heard* the utterance" (ibid.), and thus they may not have the linguistic ability to make use of these cues.

Listeners do not wait until syntactic units have ended before decoding speech (Cutler, 2012); their hypotheses about the incoming speech stream may need to be re-evaluated if incoming syntactic cues prove the initial word level analysis to be ungrammatical or implausible, known as 'garden path' effects (Davis, Marslen-Wilson & Gaskell, 2002). In addition to prosody, input from higher level cues greatly influences parsing, for example functional language may not be accurately interpreted without surrounding conceptual cues. For instance, Field (2008d:192) refers to the word *will,* or its contraction *'ll,* which could be interpreted as "a prediction [e.g., it'll rain later], a threat [e.g., I'll scream if you don't go away], an order [e.g., you'll get on with your work now] or an offer [e.g., I'll take your bag] depending on the context".

It has been determined that once a syntactic unit has been parsed it becomes an abstract proposition. Consequently, the proposition must be incorporated into the context intended by the speaker, and into the discourse as a whole; it is these operations which are discussed in the following section.

2.3.5. Higher level processes

Higher level processes have received little attention in psycholinguistic models of listening processes, as discussed in Section 2.2. In lexical search processes, "the listener's knowledge of the world and of the immediate situation" play a crucial role in that they are used strategically as a cue to word recognition and segmentation in a top-down manner (Grosjean & Gee, 1987b:145-146). In this setting, they are often referred to generically as 'context' (e.g., Mattys et al., 2005). The impact of context on word recognition was illustrated in Grosjean's (1980) experiment using a gating task. He found that as context becomes more constraining the isolation time of words, i.e., when words are confidently matched, decreases substantially. Thus, it seems that higher level processes are important contributors to "the speed and efficiency of on-going word-recognition" (Marslen-Wilson & Welsh, 1978:61).

Although the study reported here focusses on word recognition and cues used to aid this operation, higher level processes are discussed here for the sake of completeness and in order to present the listening process in its entirety. Furthermore, it has been reported that the challenge of word recognition in connected speech by L2 listeners means that they are less able to attend to higher level cues (e.g., Segalowitz, Trofimovich, Gatbonton, & Sokolovskaya, 2008), and so it is informative to know what elements of the listening processes are impaired by word recognition difficulties.

The model of the listening process adopted for the current study (Field, 2013) addresses higher level processes in detail and divides them into two phases: 'meaning construction' and 'discourse construction' (see Appendix A).

Meaning construction

The proposition the listener is left with after parsing must be aligned within the context intended by the speaker. This phase in the analytical framework is referred to as meaning construction. At this stage listeners build meaning into the proposition based on their recall of the text so far, to do this they use several types of information, detailed below. However, it should be pointed out that only the first two types of information are used during word recognition processes, providing compensatory cues in a top-down manner.

- *Contextual*. The listener relates the proposition to the context in which it occurs by making use of: a) world knowledge, knowledge of the speaker, and knowledge of the situation; and, b) recall of what has been said so far.
- *Semantic*. The listener draws upon world knowledge of entities and ideas that have been mentioned by the speaker.
- *Inferential*. The listener supplies details that the speaker has not felt it necessary to include.
- *Pragmatic*. The listener interprets the speaker's illocutionary intentions, using knowledge of the pragmatic forms of the language. Interpretation may go beyond the forms of language used and take account of context, speaker knowledge, etc.

(Field, 2013:102)

This type of meaning-led processing involves processing a representation of the overall structure of the discourse. It corresponds with Johnson-Laird's (1983) 'mental model', and van Dijk and Kintsch's (1983) 'situation model', which both refer to "all the knowledge that is left implicit in the text or otherwise presupposed" (ibid.:338). According to these models, in order to understand discourse it is necessary to "imagine a situation in which certain individuals have the properties or relations indicated in the text" and "understand the relations between the local facts and the global facts to which the text refers" (ibid.:337). If knowledge of the situation is not immediately ready to be employed because the situation has never been previously experienced, knowledge can be constructed by analogy from "partly relevant existing models" if the situation is somewhat similar to previous experience (Carbonell, 1982, cited in van Dijk and Kintsch (1983:337). For instance, the discourse model relating to a visit to a solicitor may be similar to that of a visit to a doctor as both are formal and both are situations in which advice is sought and given. This illustrates the subjectivity of the knowledge used to construct meaning, as well as the extent of the cognitive processing required to do so successfully.

With regards to the study reported here, the concern is with the kinds of contextual and semantic cues provided by the wider discourse that retroactively aid word recognition and segmentation.

Discourse construction

Once a meaning representation of an utterance has been established, listeners relate it to the *whole* listening event. For example, students listening to a lecture need to trace the line of an argument by relating sub points to the main point, and the main point to the whole lecture. Field (2013) identifies four main processes:

i) *Selection*. The listener needs to decide upon the relevance of a new piece of information to the discourse as a whole. The listener also judges the information in relation to what are perceived to be the goals of the speaker – and in relation to the listener's own goals, which may be rather different.
ii) *Integration*. The listener needs to add the new item of meaning to the developing discourse representation.

iii) *Self-monitoring*. Part of integration entails comparing a new piece of information with what has gone before, to ensure that it is consistent. If it is not, then the listener must decide whether to reserve judgement about the accuracy of the new item or whether to question his/her understanding and recall of what was said before.

iv) *Structure building*. As more and more information is acquired, the listener has to take account of the relative importance of each item. On this basis, he/she constructs a hierarchical pattern of what has been said, consisting of a set of major points with subordinate points attached to them.

(Field, 2013:102-3)

The integration stage, above, is similar to Anderson's (1996) notion of 'inference', although his discussion of this refers only to reading. Field (2013) differentiates between bridging inferences, where connections are made with preceding text, and elaborative inferences, where external information, based on, for example, world knowledge or recall of the line of argument, is added to a text.

2.4. The second language listener

2.4.1. Introduction

Comprehending natural speech is perhaps the most challenging task language learners face. Speech must be processed by the listener in real time, at a speed controlled by the speaker, and listeners are required to segment the speech into words and turn these words into propositions. These highly complex procedures, discussed above in Sections 2.2 and 2.3, form the core of this research, which investigates the word recognition processes of L2 listeners in an academic context. The following sections examine what is known about L2 listening processes, both in general and academic contexts. They provide a synthesis of research which has influenced and informed the current study, and the discussion includes strategies used to compensate for listeners' failures in word recognition or understanding. Recently, there has been an interest in listening instruction to promote word recognition (e.g. Linebaugh & Roche, 2015; Siegel & Siegel, 2013), and, considering the pedagogical

motivation for the current study, classroom-based research in this area is also reviewed here. Firstly, it is necessary to discuss the need for automatic word recognition in order for L2 listening to be successful.

2.4.2. Constraints on second language listening: automaticity

Automaticity can be defined as a process which "occurs without intention, without giving rise to any conscious awareness, and without producing interference to any other ongoing mental activity" (Posner & Snyder, 1975:56); automaticity is vital in operations at lower levels of the listening process. During input decoding operations, expert listeners automatically map acoustic-phonetic sounds to words; in lexical search operations, the form of words uttered are automatically associated with their possible senses; and, in parsing operations, the grammatical forms of words and phrases are automatically associated with the possible concepts they signify (Cutler, 2012).

The reason automaticity is essential for successful listening relates to the nature of working memory, defined as "a brain system that provides temporary storage and manipulation of the information necessary for such complex cognitive tasks as language comprehension, learning, and reasoning" (Baddeley, 1992:556). Working memory capacity is limited (e.g., Baddeley, 1999; Klatzky, 1984), and this is a critical factor to avoid overloading the memory with extraneous information (Logie, 1999). In listening, this limited capacity is not a problem for experts, whose automatic decoding abilities leave them with sufficient cognitive capacity to undertake further mental activities, for example, higher-level processing, such as meaning building. In addition, expert listeners also have sufficient cognitive capacity to revise their provisional word recognition as the speech stream continues (discussed in Section 2.1), a process key to successful listening. However, for less proficient L2 listeners these lower level processes are not automatic as their abilities are limited by a lack of linguistic knowledge, and by a lack of experience in undertaking these processes in an L2. This impairs successful listening to a great degree. Field (2013:137) discusses this issue in the context of testing listening, asserting that:

> …one might expect a test taker at Levels A2 and B1 [approximately pre-intermediate to intermediate] to need to focus a great deal of attention at the more local levels of processing (input decoding, lexical search and

> parsing) and to have little spare attentional capacity to give to the wider areas of meaning construction and discourse construction. As test takers progress …, one can expect increasing automaticity in local-level perceptual processing and hence an increasing capacity for handling complex meaning-related processes such as inference, interpreting speakers' intentions or building a wider discourse structure.
>
> Field (2013:137)

A further effect of the lack of automatic word recognition is that the increased demand on working memory may limit the number of words L2 listeners are able to retain in working memory until they are parsed (Cook, 1979; Meara, 1980). This was illustrated in a study by Field (2011) using the paused transcription method. This method requires listeners to transcribe small sections of a recording into which pauses have been irregularly inserted. At each pause, the last four to five words heard are transcribed. The results indicated that the final word in a phrase was likely to be produced more accurately than others, and it is suggested that "L2 listeners at this level may … retain [in working memory] less than L1 speakers" (ibid. 2011:104). However, no studies have been identified which use the paused transcription method in relation to L1 listening, such studies may have offered evidence to support this suggestion. Increased demand on working memory may also cause earlier decay of the words retained (Call, 1985). Thus, although L2 listeners may successfully decode words in an utterance, they may not be able to retain them in working memory long enough to parse them and impose a meaning representation.

In sum, the effect of the lack of automatic word recognition seems to be that L2 listeners use limited working memory resources in a manner *dissimilar to experts*, they appear to need to focus their limited cognitive resources on decoding, rather than on building meaning. This profoundly impedes their ability to comprehend natural speech. In fact, Segalowitz, Trofimovich, Gatbonton, and Sokolovskaya (2008) view automatic word recognition as being so important to language learning that it is a reliable indication of listening proficiency.

2.4.3. Cues to resolve misunderstanding

It was discussed in Section 2.2 that current psycholinguistic theories of listening generally support the notion that it is highly interactive and that listeners rely on a wide range of cues to understand speech. Despite this, a somewhat divergent view of L2 listening seems to be prevalent in the field. Commentators in L2 listening (e.g., Rost, 2011; Vandergrift, 1997) commonly refer to listening in terms of 'bottom-up' and 'top-down' processing. Vandergrift (2002:1) states that they "are two distinct processes", and Rost (2011:36) notes that listeners process speech in a "sequential fashion, word by word". This gives the impression that speech is decoded step by step, from small units to large, and implies a hierarchical view of the stages of the listening process. However, although psycholinguistic research has shown that there are indeed two levels of processing, i.e., perceptual (or low level) and conceptual, often referred to as contextual (or high level), the terms bottom-up and top-down in fact refer to possible *directions* of processing, rather than suggesting a linear, fixed set of processes. For example, syntax influencing lexical segmentation is processing in a top-down direction, and prosody influencing lexical segmentation is processing in a bottom-up direction (Field, 2004a).

In the recent past, a considerable amount of research, with varying results, has investigated L2 listeners' use of bottom-up (i.e., perceptual) and top-down (i.e., contextual) processing, and which sources of information are more trusted by listeners in situations where cues at one level conflict with cues at another level (e.g., Field, 1997, 2004a; Tsui & Fullilove, 1998). Hansen and Jensen (1994) analysed the listening placement test responses of 233 mixed nationality L2 listeners. The text was a semi-authentic, introductory academic lecture requiring responses to short answer comprehension questions. They found that less skilled listeners relied more heavily on bottom-up processing to compensate for perception difficulties, and that the higher the proficiency level of the listener, the more higher level processing took place. The results support Field's (2013:137) comments, noted in Section 2.4.10, that less skilled listeners lack the cognitive capacity to build meaning due to their lack of automatic word recognition. Interestingly, this notion is reported by L2 listeners themselves in a study by Airey and Linder (2006). They questioned 22 Swedish university students, who were receiving physics instruction in both Swedish and English, about their listening difficulties when attending lectures in English. The researchers video-recorded

lectures in both languages and used these as stimulated recall prompts during student interviews. Despite the students initially reporting that lectures delivered in English were not problematic, several students responded in ways which implied the contrary. For instance, one student reported focusing more on word level understanding which meant, for this student, "maybe the general message of the physics or maths gets lost a bit more" (ibid.:4). This could be interpreted as listeners having difficulties in meaning building and discourse construction, which will greatly impede their understanding of the lecture. Furthermore, students reported similar issues regarding note-taking, i.e., that they spent more time focussing on the words than following the overall meaning.

Contradicting this view, Tsui and Fullilove (1998), in their large scale study of Hong Kong listeners' behaviour in tests, found that less skilled listeners relied more on top-down processing to compensate for perception difficulties, using contextual cues to guess answers to comprehension questions. However, this study was based on data reported under test conditions with written responses to written questions, and so the results are not indicative of natural listening processes. However, in line with Hansen and Jensen (1994), above, they reported that higher level listeners were more successful at answering global questions than were those of a lower level. Similarly, evidence from Field (1997) suggests that less skilled L2 listeners rely on top-down processing, in the form of schemata regarding the topic of a listening text, to supplement weaknesses in bottom-up processing. Field (2008c:28) refers to "the inverse correlation between an adequate interpretation of the signal and the amount of compensatory top-down processing that is brought to bear".

Research has also illustrated that L2 listeners' behaviour when compensating for perceptual difficulties may not be generalisable. Field (2008c) used a paused transcription task to investigate L2 listeners' decoding strategies when dealing with unfamiliar strings of phonemes in speech. The data showed several differing types of behaviour, two of which were most noticeable. When dealing with an item not in their lexis, one group of participants ignored bottom-up evidence in the signal in an attempt to fit the unknown word into the discourse representation they had built so far, what Field (ibid:26) refers to as "suppressing phonological evidence in the interests of a consistent account". Conversely, the second group did quite the opposite. They ignored top-down evidence and preferred a word with an approximate phonological match, despite the word not fitting in with the context of the text. Field (2008c:27) concludes that L2 listeners' behaviour alters depending on the specific

situation, “reflecting factors such as the perceptual saliency of the target item or the transparency of the co-textual information (i.e., the text immediately surrounding the utterance)”.

Students’ own reports of their listening difficulties also shed light on their use of cues in their listening processes. Goh (2000) investigated forty tertiary-level students in China. All students self-reported on general listening difficulties in learner diaries and in interviews; twenty-three students additionally reported on the use of immediate retrospective verbalisation methods, i.e., they listened to a short text and commented on their perceived difficulties. The reported problems were categorised within Anderson’s (1995) three stage framework of listening processes, namely perception, parsing, and utilisation (discussed in Section 2.2). She found that half of the reported problems involved the use of perceptual cues, a further three were problems with parsing, and two involved utilisation. However, the limitations of self-reported data become evident here, it could be suggested that listeners may believe, and therefore not report, that they have an accurate global understanding of the text, but unless the meaning of the text is discussed with them, they cannot identify areas of misunderstanding. In addition, it is possible that the listeners may have focussed mainly on perception issues as it seems unlikely that they can have established a global understanding of a text having experienced these perceptual difficulties.

In a similar study, a large proportion of Hasan’s (2000:146) eighty-one participants in an EAP context in Syria reported difficulties using perceptual cues; they stated in questionnaires: “I find it difficult to understand the meaning of words which are not pronounced clearly”; and, “I find it difficult to understand well when speakers speak too fast”. The researcher does not report his research design or a framework within which he designed his questionnaire, however, it seems that the questions do not address all levels of the listening process. For instance, he does not include reference to the use of stress, likely due to the difficulties of self-reporting problems with these cues. However, he does include exercises which focus on the use of stress in his recommendations for remedial classroom activities. Again, the limitations of self-reported data impact the usefulness of these results. For example, it seems surprising that the majority of participants report that they only seldom, or sometimes, find that “difficult grammatical structures interfere with my listening comprehension” (ibid.:142). It could be suggested that listeners can only judge the accuracy

of their comprehension if they are provided with feedback, which, in this research report, it seems they were not.

Flowerdew and Miller (1992) explored 30 students' perceptions of their ability to understand undergraduate lectures delivered in English in a Hong Kong university. The proficiency of the students is reported by means of the Hong Kong Certificate of Education grade, but from the authors' description this appears to be similar to that of the participants in the current study, i.e., pre-sessional level. Using a variety of self-report methods, such as diaries, interviews, and questionnaires, the students reported missing points in the lectures, getting lost, and being unable to take notes as not enough was understood. One reason put forward for this by the researchers, unsurprisingly, was that the speed of the lecture delivery meant that perceptual cues to aid word recognition were not available to listeners. In a study similar in aim, but on a much larger scale, Graham (2006) investigated UK high school learners' of French, she questioned 595 of them about their perception of their listening comprehension difficulties. The results were also related largely to speed of delivery. A summary of key research into learners' own perceptions of their listening difficulties is presented in Table 2.1 below.

Given the problems of automatic word recognition discussed in Section 2.4.2, the results of these studies are largely predictable, and they support the aim of the study reported here, i.e., to investigate in depth the word recognition abilities of L2 listeners. With this aim in mind, the following section focusses on research which investigates how listening processes differ in an L2 from those of expert listeners, it does so by referring to the phases of the listening process discussed in Section 2.2.

Author	Context	Participants	Aim and method	Results
Flowerdew and Miller (1992)	BA TESL course Hong Kong	30 students Mixed levels	Students' perceptions of their difficulties with lecture comprehension. Questionnaires, diary studies, classroom observation, and in-depth interviews.	The main problems reported were with the speed of delivery, the new terminology and concepts introduced, and with concentration.
Hasan (2000)	EAP Syria	81 students Intermediate level	Learners' perceptions of difficulties with listening comprehension. Questionnaire relating to specific listening difficulties.	Listeners largely reported decoding difficulties due to speakers' speed and lack of articulation.
Goh (2000)	Tertiary education China	40 ESL students Level not specified	Investigation into the comprehension problems of L2 listeners. Self-report: diaries, interviews, retrospective verbal reports.	Less skilled listeners reported more difficulties using lower level cues. Listeners mainly reported decoding problems. Few reported problems with parsing or 'utilisation'.
Graham (2006)	UK high school	595 learners of French Mixed level	Learners' perceptions of difficulties with listening comprehension. Questionnaires (N=595) and interviews (N=28).	The main difficulties identified were coping with the speed of the speech, word recognition and segmentation, and building meaning from words identified.

Table 2.1: A chronological summary of key research into learners' own perceptions of their listening difficulties.

2.4.4. L2 listening processes: input decoding

It was established in Section 2.3.1 that the phoneme is unlikely to be the main unit of analysis due to its variation in the speech stream. Thus, for expert listeners top-down information in the form of words and chunks influences the interpretation of phoneme level information (Field, 2004a). L2 listeners may be uncertain of some phoneme values in connected speech due to this variation, having been largely exposed to the citation forms of phonemes. This notion is supported by episodic views of speech recognition, grounded in exemplar-based models of linguistic categorisation (e.g., Goldinger, 1996; Johnson, 1996), which suggest that mental representations of language consist of stored long-term memory traces of specific tokens of a category. For example, if the category is the word *told* , a listener will store memory traces of tokens both with the final /d/ pronounced and with it deleted (Bybee, 2000), and so the listener will be able to recognise the word in speech pronounced in either manner. Every token encountered is stored as a separate exemplar of that category, and so more tokens will be stored of frequent categories than of infrequent categories. The frequency of similar exemplars of a category is also stored in long term memory, for instance, continuing with the example of the word *told*, it is more frequently pronounced without the final /d/ than with (ibid.), and is stored as such. This means that "listeners can recognize frequent variant forms more effectively than infrequent forms" (Connine & Pinnow, 2006:241-2). In addition to variations due to connected speech features being stored, Johnson (1996:1) reports that "speech is highly variable both within and between talkers". One reason for this variability is gender, and he presents spectrograms of a man and a woman saying the word *cat,* the vowel of which is realised with a great degree of variance between speakers. Consequently, it is suggested that exemplars are stored for individual speakers.

Given that "exposure to a variant [phonological] form underlies the development of lexical representations" (Connine & Pinnow, 2006:243), it seems likely that L2 listeners' limited exposure to the variations of phoneme pronunciation will result in fewer stored memory traces. This means that they will be less efficient at decoding alternate variations of sounds in the speech stream, and so matching perceptual cues in the input to words will be less accurate, slower, and will require more effort.

The difficulty in recognising phonemes may lead to insecurity as to whether a word has or has not been correctly identified. L2 listeners may ignore phonetic evidence in the speech stream in their eagerness to find a lexical match, and rely more on cues at higher levels of processing (Field, 1997, 2008c). This issue is illustrated in L2 listeners' self-reported listening difficulties (Goh, 2000) (see Table 2.1), whereby half of their problems involved perceptual processing, and weaker listeners were far more likely to report difficulties at this level of processing.

The position of phonemes within a word may also influence word recognition. NES rely on the probability that, in connected speech, it is largely word onsets which are most reliable as assimilation effects are more often regressive than progressive (Cruttenden, 2014:308). Thus, the initial phoneme in the initial syllable is less likely to be affected by assimilation. Interestingly, Field (2004a) found that his 48 low level L2 listeners were similarly inclined, placing more confidence in their perception of word onsets than in that of vowels and word offsets. Further evidence revealing a lack of reliance on word offsets is put forward by Tauroza (1993), who tested his 42 L2 listeners' ability to recognise two groups of words, the first with the final consonants fully articulated, and the second with them deleted. He found that there was no significant difference in the participants' rate of recognition between the two groups.

A further pertinent aspect of perceptual research in SLA is cross language speech perception. A great deal of research has found that the way L2 listeners perceive perceptual information is influenced by the phonology of their L1 (e.g., Flege, 1995; Strange, 1995). Drawing on Segui, Frauenfelder, and Halle (2001), Sebastian-Galles (2005:547) categorises these influences into three types: i) deafness (a difference in phonemes cannot be heard); ii) mirage (phonemes are created that are not actually present); and iii) mutation (listeners change a sound that does not exist in their L1 to one that does). Consequently, L2 listeners could be interpreting a subset of phonemic cues differently than their L1 counterparts. Nonetheless, this research into phoneme recognition does not seem to have involved experiments using longer sections of speech, and instead uses only individual words. Thus, the research does not replicate authentic listening processes where top-down influences, such as syntax or co-text, are available to listeners and can be used in a compensatory manner.

Given the unreliability of phoneme values in connected speech and the notion that top-down processing informs retroactive recognition of phonemes, the study reported here does not investigate phoneme level decoding.

2.4.5. L2 listening processes: lexical search

Incoming speech activates a number of possible matches in a listener's lexicon and these matches compete as the speech progresses until one word is identified as the best match. In expert listeners, this competition is influenced by several features of lexical knowledge, such as word frequency, networks of association, lexical stress and knowledge of frequently co-occurring words (see Section 2.3.2). It has been noted that lexical search processes in L2 listening have not been widely researched, and Rost (2011:168) assumes that they are similar processes as in expert listening. However, the basis for competition, an integral part of word recognition, is entirely different in an L2 for several reasons. For instance, the listener has a smaller L2 vocabulary, limited exposure to phonological variants and a reduced awareness of what is or is not frequent. In addition, Wesche and Paribakht (2009:6) make the point that in L2 lexical processing of written text the L1 is a "persistent presence"; it impacts even highly proficient L2 readers. Bearing this in mind, it seems reasonable to suggest that L2 listening processes are similarly compromised by this issue.

L2 listeners face great challenges in word recognition due to the standard citation forms of phonemes being affected by aspects of connected speech, discussed in Section 2.4.4. Whereas for L1 listeners, "competition [in lexical search processes] is … a highly efficient mechanism for rapid processing of multiply ambiguous signals such as speech", L2 listeners are "likely to suffer from extensive unnecessary lexical activation and hence added competition" due to their difficulties in identifying phonemes (Broersma & Cutler, 2008:29-30). That is to say, misperception of phonemes may lead to an inaccurate cohort of lexical candidates being activated, or, due to listeners' confusion, more words being activated (Broersma & Cutler, 2008). Furthermore, L2 listeners are likely to either narrow the activated set and so identify a candidate more slowly than expert listeners, or will make an incorrect match which only partially fits the evidence in the signal (Field, 2008e). A further complication is that learners' limited lexicons may result in phantom words being activated during lexical search operations (Broersma & Cutler, 2008:29). The researchers found that L2

listeners accepted non-words, such as *groof* and *flide*, as possible words, an issue which severely impedes the competition for lexical recognition.

Further difficulties in lexical search operations are caused by the limitations of L2 listeners' aural lexicons, as it was found that L2 listeners' knowledge of words in their phonological form is far lower than in orthographic form (Milton, 2009), likely due to the variation of spoken language compared to the standard form of the written word (see Section 2.3.2). Thus, a word stored in the lexicon in written form may not be recognised in connected speech.

This view is supported by evidence from a series of spoken word recognition experiments by Pemberton (2004), who asked undergraduate second semester EAP students to transcribe recordings using paused transcription tasks. He found that, despite listeners knowing the orthographic form of the words in the recording, one in four of the 1000 most frequent English words were not recognised in connected speech after repeated listening. In a subsequent experiment, participants were able to pause the recording and listen repeatedly, and Pemberton (ibid.) used keystroke software to identify the effect of this repeated listening on word recognition scores. He found that one in three words were recognised when participants listened the first time, and this number increased to two in three words after participants listened the second time. Even with the speech rate of the recording slowed by 20 percent in a further experiment, half of words were not recognised when participants listened the first time. These results illustrate that word frequency cues may be employed with limited success by L2 listeners due to the difficulties they face recognising words in the speech stream. Pemberton (2004:106) predictably concludes that this is due to the effects of aspects of connected speech.

An additional cue to word recognition by expert listeners is knowledge about the frequency of sequences of words co-occurring (see Section 2.3.2). These sequences, or lexical chunks, "result from memorizing the sequence of frequent collocations", and the more often language users are exposed to exemplars of the chunk, the more likely they will recognise it as a unit (Ellis, 2003:10). Lexical chunks are stored in long term memory as a single entry, and so are more efficient to retrieve than strings of single words (Wray, 2002). In an L2 context, there has been extensive research into L2 learners' production of lexical

chunks using spoken and written corpora (e.g., De Cock, 1998; Nesselhauf, 2003; Siyanova & Schmitt, 2008). Results show that L2 learners use lexical chunks at least as often as expert speakers, although, predictably, with less accuracy. With regards to recognition of lexical chunks, research investigating recognition in orthographic form found that the 11 mixed nationality EAP participants more easily recognised those lexical chunks which were more frequent (Ellis, Simpson-Vlach, and Maynard, 2008). Aural recognition of lexical chunks appears to be under-researched, however a study by Koster (1987) found that words forming part of a collocation were recognised more easily by L2 listeners. However, the notion that high frequency words and chunks are articulated in a reduced form more often than lower frequency words, as determined by (Shockey, 2003), could negatively impact their recognisability.

It was established, in Section 2.3.2, that a key feature of word recognition is lexical stress, and listeners store prosodic information about words in their mental lexicon (Aitchison, 2003). This point may have great bearing on the listening processes of L2 listeners who may not recognise a word's stress pattern and store it accurately. No studies have been identified which explore this issue in relation to L2 word recognition in natural speech, and so this study aims to shed light on the issue.

2.4.6. L2 listening processes: lexical segmentation

Section 2.3.3 discussed the need for listeners to segment continuous speech into words and chunks. Expert listeners appear to use many cues to word boundary position, namely sentential context (pragmatics, syntax, semantics), phonotactics and word stress in English (Mattys et al., 2005:488), and a great deal of research by Cutler and her colleagues and others has investigated L2 listeners' segmentation processes (e.g., Content, Kearns & Frauenfelder, 2001; Cutler, Mehler, Norris, & Segui, 1987). As the syllable plays an important role in segmentation, L2 listeners' lexical segmentation operations are determined by their ability to match initial syllables. The MSS was put forward by Cutler (1990) as a strategy for lexical segmentation of all languages (see Section 2.3.3), however, contrary to what she intended, the term seemed only to be applied to English (Cutler, 2012), as English metrical forms are stress based. Thus, NES listeners segment words in the speech stream based on the fact that stressed syllables likely indicate new content words, whereas weak

syllables are unlikely to be word initial. On the other hand, listeners of syllable-based languages, such as French, are likely to segment continuous speech at syllable onsets which "provide privileged entry or reference points for segmentation and lexical access", referred to as the *onset hypothesis* (Content et al., 2001:197).

It has been suggested that these differences in language specific segmentation strategies impede segmentation in an L2, and Cutler et al. (1997a:148) report that "listeners will apply their native language-specific procedures to foreign language input, even in cases where the procedures may not operate efficiently at all". Evidence is put forward for this view by Cutler et al. (1986), who investigated the role of the syllable in the online lexical segmentation processes of French, a language in which the syllable aids segmentation, and English, where segmentation is aided by, among other cues, the presence of a strong syllable. In this laboratory-based research, 24 English and 20 French speakers were presented with recorded lists of unrelated words and non-words in both languages, they were asked to monitor within each list for a specified word-initial sequence of sounds. Results indicated that "native speakers of French appear to use syllabification to segment whether they are listening to familiar easy-to-syllabify French words or unfamiliar and hard-to-syllabify English words" (ibid.:395). Similar findings were put forward by Golato (2002:441). His research, also laboratory-based, used a different research design to that of Cutler et al. (1986); instead of participants either responding or not responding to a stimulus, i.e., processing online, Golato's (2002) participants decided whether a single word they had heard fulfilled one of two criteria, and thus it was post-perceptual processing that was under investigation. Despite this difference, his conclusion was in line with Cutler et al. (1986), in that he found, for his 21 French-dominant French/English bilinguals, "the development of a segmentation routine for English appeared to be unattainable". Thus, it seems that French learners of English attempt to segment English words using the syllabic cues appropriate in their native language, rather than identifying strong syllables as likely indicators of the presence of a new word, a reliable cue in English.

The view presented above suggests that once a segmentation strategy is acquired in infanthood, it remains the method of parsing speech, and that acquisition of foreign language segmentation strategies may not take place implicitly. However, contrasting findings are put forward in an extensive study by Field (2001). According to the MSS (Cutler, 1990), in English, strong syllables are taken to be boundary markers and weak syllables are attached to

preceding strong ones. Consequently, much segmentation research by Cutler and her colleagues (e.g. Cutler, 1990; Cutler & Butterfield, 1992; Cutler & Clifton, 1999; Cutler & Norris, 1988) has focussed on the role of strong syllables. In contrast, Field's (2001) investigation focussed on native English and French listeners' segmentation decisions regarding the insertion of boundaries between strong and weak syllables. Identifying a word boundary in this position, where the weak syllable is detached from the strong syllable, implies the weak syllable has been perceived to be a weakly-stressed functor, a prefix, a derivational or inflectional suffix, the weak syllable of a content word, or a hesitation marker (Cutler & Carter, 1987). The focus of Field's (2001) research was how listeners decide whether the weak syllable is a functor or a prefix. A further contrast to previous segmentation research, mentioned above, is that Field's (ibid.) study explored segmentation of chunks of speech, where there are several possibilities for boundary insertion (as in natural speech), as opposed to segmentation of single words. Among the extensive findings was evidence that both native English and French listeners used similar segmentation strategies in that they attached a weak syllable to the right of a strong syllable. Field (ibid.) concludes that sensitivity to the prosody of a stressed-based L2 can be acquired as a developmental process. An additional similarity was found in that both native English and French listeners were inclined to associate weak syllables with functors, possibly, Field suggests, a transfer from the similar expectation in the French listeners' L1.

Differences in the research designs of Field (ibid.) and Cutler et al. (1986) may account in some way for these contradictory findings, as the former study used larger units of speech than the latter. However, research by Sanders, Neville and Woldorff (2002) also found that their 55 Japanese and Spanish listeners who started to learn English after the age of 12, applied native speaker-like segmentation rules, segmenting at the presence of strong syllables. They posit, therefore, that "non-native speakers are able to learn new segmentation cues" (ibid.:10). This is one of the issues that the study reported here investigates.

Turning to the role of phonotactics in segmentation, Section 2.3.3 presented evidence that L1 listeners use the phonotactics of their L1 to aid segmentation. Research has shown that L2 listeners' ability to segment is influenced by the phonotactics of their L1. Weber and Cutler (2006) investigated the use of phonotactic constraints by 48 advanced German learners of English. They compared reaction times in a word spotting task, whereby participants are

asked to identify a word embedded in a group of sounds. For instance, in the example below, listeners had to spot the word embedded amongst four types of boundary classifications:

Common boundary	English boundary	German boundary	No boundary	Embedded word
/dʒimlæns/	/ðiʃlæns/	/blɔɪslæns/	/θiplæns/	lance
/fumlɔft/	/prarʃlɔft/	/forslɔft/	/zarplɔft/	loft

(Weber & Cutler, 2006:606)

The researchers found that listeners responded a great deal faster, and more accurately, to words where boundaries were acceptable in German, but not in English.

Al-jasser (2008) used a similar method in his study of 40 Arabic EFL undergraduates who had been learning English for six years. His results also showed that reaction times were quicker in instances with word boundaries permissible in Arabic, but not English. Interestingly, evidence suggests that learners can learn to apply phonotactic rules from the L2 to varying degrees. Al-jasser's (ibid.) participants made some use of English phonotactic constraints after approximately six years of learning English. However, the limitations of the research tasks for both studies should be borne in mind. In natural connected speech, word recognition draws upon cues from co-text and context as well as from perceptual sources, and word forms are often reduced. Therefore, tasks involving only phoneme-based segmentation could be considered as lacking in ecological validity, a limitation addressed within the study reported here.

The tentative nature of listening means that as the speech stream progresses and further cues become available, listeners may need to regress and reanalyse the speech stream (Shockey, 2003), what Broersma and Cutler (2008:30) refer to as "adjusting the competitor population". The manner in which L1 and L2 listeners revise their incorrect segmentation hypotheses was investigated by Field (2008e). He used a type of gating task (Grosjean, 1980) where participants transcribed input of gradually increasing length and found that, unlike his 112 L1 listeners, his 152 L2 listeners were disinclined to change their incorrect segmentation decisions as the signal progressed despite evidence that the decisions were inaccurate. Field

(2008e:48) suggests that L2 listeners may not have realised that alternate segmentation hypotheses were available as they may have failed to recognise the lexis, despite the strong likelihood of it being within their lexical range. He points out that "acquisition of a vocabulary item is not the same as the ability to recognise it when it occurs in running speech" (ibid.:48). In addition, it is possible that L2 listeners lack the confidence to revise their original interpretations based on further input in the speech stream. Although top-down evidence in the signal is available to L2 listeners, it seems likely that the major obstacle to revising segmentation decisions is the cognitive demands this would enforce on listeners who would be obliged to "carry forward a set of partially activated competitors once a lexical or segmentation decision has been made" and reactivate the competition processes (Field, 2008e:49).

2.4.7. L2 listening processes: parsing

Imposing a grammatical structure on incoming speech, as discussed in Section 2.3.4, requires a mastery of the syntax of the language which L2 listeners are unlikely to have fully achieved. Little research has been identified which investigates L2 listeners' use of syntactic cues to aid word recognition, a lack highlighted by Brown (2006). However, one such study by Kim (1995) investigated the effect of pauses at syntactic boundaries on L2 listening comprehension. Two listening texts were used, one with regular pauses at phrase boundaries, and one with longer and more frequent than average pauses, although still at phrase boundaries. The latter text, thus, had more prominent tonic syllables. As expected, the results showed that L2 listeners' performance in comprehension tests was higher in the second text, and the author suggests that listeners may have been able to shift their focus from lexical level decoding to make use of top-down influences on word recognition in the form of syntax. However, it could be suggested that instances of aspects of connected speech would have also been reduced due to additional pauses, which no doubt would have aided word recognition.

The extent of L2 listeners' reliability on syntactic cues over other levels of input in the speech signal was highlighted by Shockey (2003) who suggests that L2 listeners process speech similarly to L1 listeners, but they need more input from the speech stream to process accurately. She reports that L2 listeners "depend heavily on syntactico-semantic information to arrive at an understanding rather than using phonological context" to resolve perceptual

difficulties (ibid.:122), what Brown (1977) refers to as a processing lag. However, results of a study by Harley (2000) are in contrast. She investigated the reliance on syntax compared to sentence stress of 56 Cantonese and 33 Polish listeners of English, as well as 30 NES. Manipulating 10 sentences so that the syntactic and phonological cues to the structure of the sentences were at odds, she found that all three sets of participants were less sensitive to syntax than to prosody. However, her participants were young learners, and given that prosodic structure is "one of the earliest features to which language users are sensitive" (Jusczyk, Cutler, & Redanz, 1993:684), it may not be accurate to extrapolate these findings to adult learners.

Considering the limited number of studies surrounding this issue, it is evident that further investigations into the extent of the influence of syntactic cues on L2 word recognition would be a useful contribution to the literature, and the study reported here aims to do so.

2.4.8. L2 listening processes: the role of context and co-text

Research into L2 listeners' use of higher level cues has largely investigated whether higher level cues, in general, are relied upon in a compensatory manner to a greater or lesser degree than lower level cues (discussed in Section 2.4.3). Nonetheless, a small number of studies have been identified which explore the impact of specific high level cues on decoding at lower levels.

Early research by Koster (1987) investigated the use of surrounding co-text to aid the lexical recognition of three groups of listeners: intermediate non-native listeners; advanced non-native listeners; and native listeners. He found that the lexical recognition of the intermediate non-native listeners improved most when co-textual cues were available, i.e., when listeners heard the surrounding text. Similarly, evidence from Field (2008c), discussed in Section 2.4.3, showed that some of his 27 Arabic speaking L2 listeners favoured co-textual and contextual cues in their attempts to decode a new lexical item over phonological cues. However, conversely, Field (2004a, 2008c) also found that some L2 listeners ignored co-text when confronted with a new lexical item and attempted matches that were at odds with the co-text syntactically and semantically, but were an approximate phonological match to

known words. In response to evidence of this differing behaviour, Field (2008c:27) suggests that rather than there being a general *type* of L2 listener behaviour, listeners have "what could be termed a personal listening style". However, considering the small data set (N=27), he calls for further investigation into this area.

In a study of how topic knowledge affects the listening comprehension of L2 learners, Long (1990) presented her 188 mixed nationality learners of Spanish with two listening texts, one relating to a topic with which they were likely to be familiar, and the second relating to a topic likely to be less well known. As expected, comprehension of the former text was much more successful than the latter. However, methodological issues may have had some impact on the results as the comprehension task asked listeners to write summaries of the text after listening, and so memory effects may have been at play. Nevertheless, it seems clear that L2 listeners use higher level cues to inform decoding in the listening process.

These higher-level operations are extremely cognitively demanding, and it seems that lower proficiency listeners may be so heavily focused on decoding and word recognition that they do not have the cognitive capacity to take account of anything other than very broad contextual cues. In contrast, expert listeners' ability to decode automatically leaves capacity for higher level processing, discussed in Section 2.3.5 (Brunfaut & Révész, 2015; Field, 2008d). In line with this view, Brown (1977:151) suggests that L2 listeners need to "learn to control the phonological code of the target language…with sufficient ease to provide a constrained input for the 'top-down' inference-driven interpretation to be constructed". It seems, therefore, that lower proficiency listeners may have little capacity to process at higher levels, and so bring higher level information to bear on word recognition, as their attention is likely to be focused on lower level processing. This possible focus on lower level processing is investigated in the current study.

2.4.9. Summary of L2 listening research

Sections 2.4.3 to 2.4.8 have reviewed the extensive literature which investigates L2 word recognition processes. Limitations to this research, as well as under-researched areas, have illustrated the gaps in the current body of knowledge, and it seems that accounts of L2 listening are often incomplete. It is evident that a considerable range of work, from both

Cutler and her colleagues (although largely in L1 listening) and Field, has contributed greatly to current knowledge in the field, and these studies, as well as that of Pemberton, have been highly influential to the study reported here. The current study aims to build on this research, as well as other research discussed above, and this is highlighted in the summary of key research into L2 listening processes, provided in Table 2.2 below. The far-right column shows the differences in each study compared to the current study, and so illustrates how the current study develops knowledge in the field. For example, it can be seen that the work of Broersma and Cutler (2008) and Cutler (1986) were both small-scale and did not involve long sections of speech, being artificially designed. Thus, the range of cues available in normal listening situations to support word recognition in a top-down manner, such as the sentence stress and syntactic cues found in co-text, were absent. The results, therefore, do not provide insight into L2 listening behaviour as a whole process. Hansen and Jensen (1994) used academic lecture material in their study, however, they investigated only L2 listeners' ability to answer detailed and global questions, and the global questions were posed after a second hearing of the text. As such, there was little insight into L2 listening processes as a whole and in natural context, i.e., a single hearing. Pemberton's (2004) research into word recognition was influential to the study reported here as it used a similar methodology to identify errors in word recognition of large sections of natural speech. However, his study investigated only the extent to which words were misidentified, and it was not his aim to identify the causes of listening breakdowns and the cues relied upon. Finally, Field's research into L2 listener behaviour provides a wide range of investigations and extensive insight into the topic. Early studies (e.g., 2001) involved the kind of short stimulus that Cutler and her colleagues employed (e.g., Broersma & Cutler, 2008). More recent studies were more naturalistic, however, they were small-scale (e.g., Field, 2008a).

Author	Context	Participants	Aim and method	Results	Differences in comparison to the current study
Cutler (1986)	General English UK and French universities	24 English speakers 20 French speakers	Lexical segmentation in French and English. Monitor lists of unrelated words for a specified word-initial sequence of sounds.	Native speakers of French appeared to use syllabification to segment whether they were listening to familiar easy-to-syllabify French words, or unfamiliar and hard-to-syllabify English words.	Laboratory-based experiments with unrelated words. Small scale
Long (1990)	Academic US university	188 mixed nationality students of Spanish	Investigation into how background knowledge affects the listening comprehension of L2 learners. Summaries after listening and true/false statements.	Listeners with less background knowledge about the topic of the text relied on lower level processes, whereas listeners with background knowledge showed signs of having processed more globally.	General English Texts were scripted, thus did not replicate natural speech. Focused on schema activation. Memory effects may impact results.
Hansen and Jensen (1994)	Academic US university	233 L2 EAP students	Investigation into listeners' comprehension of lectures. Segments of a semi- authentic lecture. Short answer comprehension questions after first play, and global questions after second play.	More skilled listeners scored better on the global questions than the less skilled listeners. Less skilled listeners relied more on top-down processing as compensation for perception difficulties.	The investigation did not include listeners' ability to parse the speech. The global questions were asked after a second play of the text, and so does not represent real life lecture listening.

Tsui and Fullilove (1998)	Tertiary education Hong Kong	The responses of approx. 20,000 secondary level candidates in public examinations to 177 test items.	L2 listeners' use of schema.	Students with the highest overall grades were most skilled at top-down processing. Less skilled listeners rely more on top-down processing as compensation for perception difficulties.	Texts were scripted, thus did not replicate natural speech. MCQ was given orally, before the text was heard. Focused on schema activation.
Field (2001)	British state secondary schools, and British and French private English language teaching institutions	Various experiments involving groups of 97 to 189 NES and 29 to 91 French speakers.	Investigation into the segmentation strategies of French learners of English. A gating task and a transcription task.	The French listeners were as likely to segment English speech at the presence of strong syllables as native listeners, despite this strategy not being applicable in their native language.	
Pemberton (2004)	Academic A Hong Kong university	27 L2 undergraduate students and university staff Mixed levels	Word recognition Paused transcription	Listeners recognised the orthographic form of the words in the recording, but only one in three of the 1000 most frequent English words was recognised when they listened the first time. One in four was not recognised after repeated listening. With the speech rate of the recording slowed by 20 %, half of the words were not recognised when participants listened the first time.	General English Investigated word recognition. Small scale.
Field (2004a)	General English UK private language school	31 students lower intermediate, 17 high elementary Mixed L1	Whether L2 listeners rely more on top-down or bottom-up information when transcribing an	Results suggest that both bottom up and top down dependence are evident depending on various features of the speech, and listener preference.	General English Transcriptions of separate, unconnected sentences. Small scale

			unknown word. Paused transcription		
Broersma and Cutler (2008)	General English Dutch and UK universities	Experiment 1: 24 Dutch-speaking students and 24 English-speaking students. Experiment 2: 36 Dutch speaking students and 36 English speaking students.	2 studies investigating the activation and competition processes of candidate words in spoken-word recognition of L1 and L2 listeners.	Experiment 1: lexical decision study. L2 listeners accepted spoken non-words (e.g., groof and flide) as real English words, but L1 listeners did not. Experiment 2: priming experiment. The same spoken non-words made recognition of real words (e.g., groove, flight) easier for L2 listeners, but not for L1 listeners. This suggests that, for the L2 listeners, the real words had been activated by the spoken non-word input.	Laboratory-based experiments with artificially designed cues. Small scale
Field (2008a)	General English UK private language school	48 students Intermediate or upper intermediate mixed L1	Whether function or content words are processed more accurately by L2 listeners. Paused transcription	All levels and nationalities were more able to recognise content words than functors.	General English Small scale
Field (2008c)	General English UK summer school	27 Arabic primary school teachers of English Level: assumed proficient	Investigation into listeners' patterns of decoding errors when attempting to transcribe unknown words. Paused transcription	Results suggest listeners' behaviours vary depending on several aspects, e.g., perceptual saliency, saliency of co-textual information. Also, it seems that individual listeners favour particular decoding techniques.	General English Small scale
Field (2008e)	General English	112 native speaker students 39 students French L1,	Investigations into how first and second language	L1 listeners changed their segmentation decisions quickly on the basis of incoming evidence. L2 listeners were much less likely to do so.	General English Short segments of speech.

		113 students mixed L1 Intermediate or upper intermediate	listeners adjust their segmentation assumptions as new perceptual evidence comes in. Transcribe short sequences of 2 to 3 words.		Small scale
Field (2011)	General English UK private language school	47 students B2 level	Investigation into how much of the input is successfully processed. Paused transcription	In handling new items of lexis, some listeners transcribed an approximate phonological match in the form of a known word, even though the match was syntactically and semantically inconsistent. The final word in a phrase was likely to be produced more accurately than others.	General English Small scale

Table 2.2: A chronological summary of studies focused on L2 listening behaviour.

2.4.10. Second language academic listening

Having established the importance of automatic word recognition in listening, and the considerable effect of failings in this area, it is crucial to the current research to consider the specific demands faced by L2 listeners in academic contexts where English is the language of instruction. This will promote understanding of the position within which the participants in the current study are situated. Academic discourse takes place in many scenarios, such as tutorials, seminars, and discussions, with lectures as the main form of instruction. This study considers academic listening to refer largely to lecture listening. This section firstly analyses the features of lecture input, and then considers the cognitive processes of lecture listening and the challenges faced by L2 listeners.

Academic listening is considerably different to more general listening situations, and its nature gives rise to its complexity. The practicalities of lecture listening present complications not commonly found in general listening scenarios, and can be summarised as follows:

- in most cases, listeners have only one chance to listen to a lecture
- listeners must focus for long stretches of time
- lecturers assume listeners understand, thus listeners are accountable for their own understanding and, consequently, their own learning; misunderstandings may go unnoticed by both parties
- listeners have a largely passive and less interactive role, and so have little opportunity to negotiate meaning for clarification

(Flowerdew, 1994; Rost, 2011)

These points indicate the need for listeners to take away an accurate representation of the message lecturers wish to convey in a single, lengthy listening situation.

Academic listening is also distinct in terms of the cognitive processes involved. The psycholinguistic model of listening adopted for the current research (outlined in Section 2.2 and presented in Appendix A) is a model of general listening, thus, as the context of the study

reported here is academic, it is crucial to discuss how the processes might differ between the two settings. The five phases of the model of the listening process, input decoding, lexical search, parsing, meaning construction, and discourse construction (Field, 2013), are relevant to all listening situations, however, lecture listening involves a great degree of processing at higher levels, i.e., those of meaning and discourse construction. As Benson (1994:192) reports, academic listening is not simply learning a lists of facts, but is learning "complex systems of ideas". There are a number of components involved in higher level listening processing, such as augmenting the parsed input by employing knowledge of the world, the speaker and the speaker's intentions, the context, and the text so far; these stages were discussed in full in Section 2.3.5. It is crucial to note, however, that there are several features of higher level processes of particular significance in lecture listening. Field (in press) argues that academic listeners must be capable of:

- identifying the current main point;
- judging whether a new piece of information is central, secondary or irrelevant;
- distinguishing macro-propositions from micro-;
- linking points of information (especially where the links have to be inferred);
- integrating new information into a developing discourse representation;
- monitoring the developing discourse representation for consistency;
- building an overall discourse structure which represents the lecturer's line of argument.

(Field, in press)

The set of cognitive processes employed in structure building of this kind are discussed by Gernsbacher (2013), who refers to the need for comprehenders (as she refers to listeners) to build a coherent mental representation of incoming information, and either to map subsequent coherent information onto existing representations, or to build new representations if subsequent information is not coherent. It is evident that this process is cognitively complex, and so it is unsurprising, given the limited capacity of working memory, that Gernsbacher (ibid.) argues that less skilled comprehenders' abilities to remember recent input is detrimentally affected during this process. The result is that they build new representations more often than more skilled comprehenders as they are less able to recognise the coherence of new information with old, and thus less able to integrate new information

into existing mental representations. In lectures, for example, less skilled listeners are less likely to recognise a coherent hierarchy of points and sub points in an argument and, as such, are more likely to view each point as separate units (Field, 2011).

It is evident that the higher-level processing called for in lecture listening is extremely cognitively demanding, and automatic decoding, discussed in Section 2.4.2, is necessary in order for listeners to retain sufficient cognitive capacity to do so. However, less skilled L2 listeners, whose abilities are limited in terms of language knowledge and listening expertise, focus heavily on word recognition and parsing (Field, 2013). These skills are less likely to be automatic, and so require more conscious processing. Thus, less skilled L2 listeners struggle with information processing at the levels of meaning and discourse representation (Brown, 1977; Field, 2011), and L2 *academic* listeners struggle to construct an accurate representation of lecturers' discourse.

An additional characteristic of lecture listening, which should be considered, is the need for listeners to divide their attention between listening, looking at visual aids, reading handouts, and taking notes (Flowerdew & Miller, 2005). Clearly, in scenarios such as this, while expert listeners may have the cognitive capacity to achieve this with a high degree of success, the impact of limited automatic decoding, and the inherent strain on cognitive capacity, means that L2 academic listeners are likely to be less able to do so.

It is evident that the demands of lecture listening are vastly different from most general listening situations. For L2 academic listeners, whose automatic word recognition skills may not be honed to the degree of expert listeners, these challenges are likely to hinder their understanding of lectures, and consequently their success in their academic study.

2.4.11. Confidence levels and strategy use

It has been illustrated that L2 listening differs greatly from L1 listening due to the impact of various elements, such as a lack of automatic word recognition, gaps in understanding due to poor perceptual processing, and lack of language knowledge. The decoding difficulties that ensue, and the resulting reliance of L2 listeners on perceptual cues, means that they are forced to listen strategically. Strategy use in L2 listening is an extensively researched topic (e.g., Goh, 2002; Oxford, 1990; Vandergrift, 1997). However, there is a

noticeable lack of agreement on definitions, categories, and terms, and L2 listening commentators have been known to conflate the terms ‘strategy’, ‘skill’ and ‘subskill’, or even use them interchangeably. For example, Hasan (2000:149) reports “EFL learners were in some respects poorly equipped with effective strategies, skills, and activities to help them to improve their listening comprehension”, however, he fails to differentiate between the two terms ‘strategies’ and ‘skills’. Further, Ridgway (2000) uses all three terms interchangeably, although he acknowledges that “it is very difficult to define strategies or subskills, or to differentiate them in any empirically significant way” (ibid.:182).

The current study uses the term ‘strategy’ and follows (Cohen, 1998:4) definition as “action taken to enhance the learning or use of a second or foreign language, through the storage, retention, recall and application of information about that language”. A frequently cited categorisation of strategies designed for use in the field of SLA by O'Malley, Chamot and Küpper, (1990:44) divides them into three types:

i) metacognitive: “higher order executive skills that may entail planning for, monitoring, or evaluating the success of a learning activity;
ii) cognitive: strategies which “operate directly on incoming information”; and,
iii) socio-affective: strategies involving transacting or mediating with others.

Only cognitive strategies are relevant to the current research, with its psycholinguistic theoretical background, and these strategies are discussed as they are defined within the field of cognitive psychology, i.e., to refer to mental operations undertaken in order to carry out cognitive activities (Gellatly, 1986). Some commentators have also been known to conflate the notions of *processes* and *strategies*. Following Field (2008b:2), the current research views listening processes as “the cognitive operations which underlie all listening”, and listening strategies as “compensatory techniques that are used to fill gaps in word recognition or in understanding”. As such, listeners’ use of strategies in listening is not a process, but the listening process itself is strategic (Field, 1998).

Strategy use in listening augments incomplete input by drawing on other levels of information in the speech signal. For example, if perceptual information in the speech stream is unreliable due to noise interference, listeners may hypothesise a lexical match from other cues, such as co-text or context (Bond 1999, Field 2006). In L2 listening, listeners who lack

confidence in their decoding decisions, that is to say, they are aware that their decoding decisions may be inaccurate, are likely to engage in similar strategy use for this purpose. Therefore, L2 listeners' use of strategies is influenced by the awareness that their decoding decisions may not be accurate and by the confidence they have in these decoding decisions.

Although much importance is accorded to the role of awareness in SLA processes (e.g., Schmidt, 1992; 1990), its vital role in strategy use seems to be rarely discussed in the field of L2 listening, and little research has been identified which investigates this issue. In the context of testing listening,Yule (1988:84) refers to measuring the degree of listeners' confidence in listening tests as a means of identifying learners who "select answers based on effective self-monitoring and those whose answers are based on poor self-monitoring". Effective self-monitoring implies strategic behaviour, i.e., that a listener is aware, in that instant, that their interpretation is approximate, and therefore needs to be checked against incoming speech.

By way of an illustration of strategy use in L2 listening, the strategic behaviour of two L2 listeners of contrasting listening proficiency is reported in research by Graham, Santos and Vanderplank (2008). Both listeners reported strategic behaviour in that they relied on cues from co-text and context when they lacked confidence in their ability to recognise words. Unsurprisingly, the lower-level listener reported a lack of confidence in her decisions far more often than the higher-level listener, on numerous occasions she expressed that she was not sure and guessed word matches. In addition, the researchers found that both listeners were forming hypotheses about their understanding of the text. However, the lower level listener was largely unable to revise her hypotheses on the realisation that she had misunderstood sections of the text, whereas the higher-level listener was aware of the tentativeness of his comprehension and was able to revise his hypotheses, a crucial aspect of successful listening (see Section 2.1). The researchers concluded that he seemed much more able to decode at speed, i.e., automatically, which is likely to have allowed him the cognitive capacity for such hypothesis revisions. Support for this conclusion can be found in a study by Tsui and Fullilove (1998), who found that only the highest level listeners were able to correctly answer detailed questions where initial schemata were subsequently contradicted as the text continued, thus illustrating their ability to revise their initial hypotheses. However, less skilled listeners were less able to revise their hypotheses, likely due to the time constraints imposed by the need to process spoken input in real time (Koster, 1987), and to

their lack of cognitive capacity. Furthermore, Field (1997) also found that L2 listeners were disinclined to revise their schemata despite conflicting bottom-up evidence, preferring to adjust their account of perceptual cues to suit their predetermined schemata, rather than the opposite. The case studies by Graham et al. (2008) are indicative of the impact of L2 listeners' confidence in their decoding decisions on their use of strategies, and further research into this topic would be informative to understand L2 listening processes.

Further research into strategy use by Field (2008c:29) suggests strategic behaviour in both high and low level processing is affected by aspects of the input, i.e., "perceptual saliency, speech rate and frequency of pausing, problems of lexical segmentation and degree of transparency of context and co-text". Thus, it seems that effective strategy use is directly related to the nature of the problems in comprehension, and so "a strategy cannot be considered independently of the circumstances which give rise to it" (Field, 2008b:6). An additional point worthy of note is that both Field (2008b) and Graham et al. (2008) suggest that strategy use varies depending on listeners' individual style, one feature of which is their willingness to make strategic decisions. Field (2008b:5), based on his analysis of data gathered from L2 listeners when listening breaks down, states that "in most groups of individuals, there will be risk takers, ready to form hypotheses on the basis of partial evidence, and risk avoiders, who rely upon having decoded most or all of the input".

With regards to what influences individual variations in strategy use, several researchers have offered suggestions. Locastro (1994:410) suggests it may be linked to the individual's learning environment, namely "the social system of values embedded in an educational context". Meanwhile, Macaro (2001) suggests age may be a factor, and gender is put forward as a possible influence by Oxford (1990). This notion suggests that research which attempts to make generalisations about strategy use may not be possible.

Research into language learners' strategy use commonly uses self-report methods of data gathering. For example, Graham et al. (2008) data were gathered via verbal reports, Goh (2000) used questionnaires, learner interviews and diaries, and Hasan (2000) used questionnaires. Also, Vandergrift (1997) used a think aloud procedure for data gathering, pausing the recording and asking listeners to report on what strategies they used to comprehend the text. However, data gathered from learner reporting methods in the context of investigating listening processes should be treated with some caution as cognitive

processes are generally highly automatic, and therefore unconscious (e.g., Cutler & Clifton, 1999; McQueen, 2007), which makes them very difficult to report. In addition, the concern whether "the act of verbalisation disrupts the process of listening" has also been noted, as well as "whether subjects can in fact remember the strategies they have used" (Graham et al., 2008:55). Both Hasan (2000) and Goh (2000) discuss strategy training to improve L2 listeners' self-reported perceptions of their own difficulties. However, it could be suggested that the success of strategy training is dependent on an accurate interpretation of these listeners' difficulties, and there is a danger with self-reporting in this context in that participants may misinterpret reasons for their listening breakdowns, or may not have the meta-language to describe them accurately. Flowerdew (1994:97) suggests that "self-report will provide clues to the subject's mental representations, and not a clear 'reading' of them". Hasan (2000:137) concurs and notes the helpful caveat that "learners' perceptions of their listening problems may or may not correspond to what actually happens, as different factors which the listener may not be aware of may interact and influence learners' perceptions". It seems that data gathered during the process of listening, without interrupting that process to self-report, as in the current study, may prove to be more ecologically valid than self-reported problems, and so make a valuable contribution to the literature. Despite the limitations of self-reporting, it is interesting to note that results of learners' perceptions of their difficulties identified through self-reporting methods are largely consistent, that is to say, the majority of participants report low level processing problems.

A study by Field (2011) is particularly pertinent to the current research as his participants' English ability was similar to those of the current study. His 47 level B2 participants were studying at a private U.K. language school with the aim of moving to an academic context. His investigation highlighted the limited amount of input successfully processed by the participants. Results showed that listeners were able to decode far less than expected, and Field (ibid.:104) concluded that listeners at this level must be "quite heavily reliant upon compensatory strategies to supply sections of the text that they have been unable to match to words". It could be anticipated that the findings of the current study may be in line with Field's (ibid.)

Given Graham's (2011) assertion that L2 listeners' self-efficacy may be enhanced by listening strategy instruction, it seems clear that research exploring L2 listeners' confidence in their decoding decisions and subsequent strategy use would be of great interest to L2

listening researchers and teachers, and the current study addresses this issue. Earlier researchers who have used any type of transcription task (e.g., Field, 2008a; Pemberton, 2004) do not appear to have considered that listeners' reports of what they have heard may be the product of two distinct processes, namely mishearings or strategic operations, which aim to compensate for cases where the listener is unsure of what has been heard, or is aware that s/he has failed to understand it. Furthermore, no research has been identified which addresses the issue of listeners' confidence in their abilities to recognise words in long sections of connected speech. Therefore, a small-scale, supplementary enquiry in the study reported here will address this question.

The discussion so far has reviewed what is known about L2 listener behaviour, and in the next section, the focus alters. The current research context involves listeners whose second language is about to become their means of learning, in that their programmes of study will be undertaken in English. They have a short time frame for their EAP studies to prepare them for this. The following section reviews literature which seems to illustrate a possible shift in traditional listening teaching practices (see Section 1.2), and this shift seems to reflect the importance of word recognition highlighted throughout this chapter.

2.4.12. L2 listening instruction research

Although this study is not primarily a pedagogical one, its fundamental motivation was entirely pedagogical. Therefore, knowledge of L2 listening gained through the study, both in general and academic contexts, should inform listening teaching practice by challenging the received views, expressed even today in teacher training and some materials (see Chapter 7), that 'context saves the day' (Field, 2008d:127). An understanding of where problems of word recognition occur, and what cues in the speech signal are most relied upon by L2 listeners in word recognition processes, would enable English teaching practitioners and materials writers to target their techniques in the way that Field (ibid.) advises, rather than continuing to rely on the comprehension approach. These implications will be discussed in Chapter 7, however, in order to contextualise the discussion, it is necessary first to provide a brief review of the current status of L2 listening teacher training.

L2 listening lessons generally revolve around comprehension questions and have tended to emphasise the use of high level cues, such as context and prior knowledge, to solve decoding problems. As such, there seems to be little concern in the field about addressing, in the classroom, what Shockey (2003:123) refers to as "the importance of variability in phonological input". Lynch (2006:91) reports that "teaching materials for L2 listening have overemphasized schema-based strategies, and students are encouraged to engage in strategy use of this nature in order to aid the decoding process". An example of this is research by Hasan (2000), which identified participants' listening problems as largely related to decoding. However, despite his proposal that teachers should "determine the causes which make comprehension break down and design remedial tasks for each problematic area" (ibid.:149), he suggests that tasks largely aimed at the improvement of higher level processing will equip listeners with strategies to compensate for weak decoding skills.

Regular exposure to the L2 is beneficial in that listeners become accustomed to its sounds and lexis and the link between the two, as well as the variations in pronunciation of words in natural connected speech. Despite this, if learners are exposed to large portions of input they find incomprehensible, their progress may be inhibited by a decrease in motivation and over-use of compensatory strategies (Field, 2008d). The results of such comprehension-based exercises only *test* listening comprehension, and Brown (1977) argues in favour of *teaching* listening, calling for "some sort of method of investigating the student's problems" in order to teach listening effectively (Brown, 1986:286).

In accordance with these views, Field (1998:111) proposes several different approaches to teaching listening, including a diagnostic approach where "wrong answers can be seen to be of more significance than correct ones". By analysing wrong answers, teachers can diagnose where listening breaks down, and address the problem in the classroom. This approach provides insights into the processes of listening, rather than attending only to the product, as in the comprehension approach (Field, 2008d:81). The diagnostic approach requires classroom tasks to aid instructors with error diagnosis and frameworks within which to classify errors. Subsequently, remedial tasks are required to address the problems identified (ibid.). The rationale behind the remedial tasks is that learners need to be exposed to features of the language with which they have difficulties. For example, if listeners' problems are due to inaccurately using syntactic cues, such as not hearing inflexional morphemes, micro-dictation tasks can be designed to highlight these structures.

This approach is advocated by Wilson (2003:335) whose teaching methodology, referred to as 'discovery listening', allows students "to discover and then prioritize their own listening difficulties". He achieves this by using three-stage tasks, summarised as follows:

1) listening: take notes, self-assess, listen again while taking notes
2) reconstructing: reconstruct text in groups
3) discovering: compare text with the original, classify causes of mistakes

Students are provided with categories of potential errors which refer to cues from lower to higher levels:

a) I couldn't hear which sound it was
b) I couldn't separate the sounds into words
c) I heard the words but couldn't remember their meaning quickly enough
d) This word was new to me
e) I heard and understood the words but not the meaning of that part of the sentence
f) Other problems

(ibid.:340)

Clearly, the students are self-reporting their problems, the limitations of which were discussed in Section 2.4.3. However, Wilson (ibid.) allows students to compare their written responses to the original text and, as such, there is tangible evidence of breakdowns, rather than simply the learners' view of the cause of the breakdowns. This should provide more accurate data than students reporting their perceptions of their problems. Unfortunately, despite what seems to be a highly effective task, the author does not suggest remedial exercises, but posits that students 'noticing' the causes of breakdowns improves their listening skills. This may be the case to a limited degree, however, regular remedial exercises are likely to promote automatic word recognition, and thus more successful listening.

This notion of remedial tasks aimed at improving decoding (as opposed to overall comprehension) seems to be gaining momentum of late, although, as Vandergrift and Goh (2012) note, the impact of such tasks has not been fully established. Several classroom-based

research studies into this topic have emerged and it seems the focus of teaching L2 listening may now be shifting towards perceptual training. Much of this classroom-based research examines participants' improvements in decoding and shows largely positive results. However, research investigating overall listening comprehension is less common. A summary of several studies of L2 listening instruction research can be found below in Table 2.3.

Author/s	Research aim	Participants	Method	Data collection	Findings
Yeldham (2009)	Effect of bottom-up and top-down training approaches.	12 Taiwanese EFL students, various levels	1 group received top-down 'strategies-based' training, 1 group received bottom-up training.	A variety of pre and post tests	False-beginner levels need to focus more on top-down strategies. Lower-intermediate levels need a fully interactive approach to cover the diverse needs of learners with different top-down and bottom-up listening styles. Intermediate levels show less need for explicit skills and strategies instruction than the lower two levels.
Kuo (2010)	Effect of partial dictation.	31 Taiwanese EFL university students	Experimental group	Pre and post tests	Listening comprehension improved (students were also receiving standard listening lessons).
Marzban and Abdollahi (2013)	Effect of partial dictation.	60 intermediate Iranian EFL university students	Experimental and control groups	Pre and post tests	Experimental group performed better than the control group.
Siegel and Siegel (2013)	Effects of bottom up listening activities.	33 Japanese university students	Experimental group received instruction for one semester.	Pre and post dictation tests	Experimental group performed better than the control group.
Ahangari, Rahbar and Maleki (2015)	Effect of training in production to aid listening	42 Iranian students	Experimental and control groups	Pre and post tests	Experimental group performed better than the control group.
Ahmadian and Matour (2014)	Effect of connected speech instruction.	Iranian university students	Experimental and control groups	Pre and post tests	Experimental group performed better than the control group.

Alameen (2014)	Effect of linking instruction.	45 university students	Experimental and control groups	Pre, post, and delayed post tests (dictation)	Experimental group performed better than the control group.
Baghrahni, Shariati and Tajadini (2014)	Effect of assimilation and elision instruction.	42 Iranian junior high students	Experimental and control groups	Pre and post tests	Experimental group performed better than the control group.
Khaghaninezhad and Jafarzadeh (2014)	Effect of reduced forms instruction.	50 Iranian EFL students	Experimental and control groups	Pre and post tests	Experimental group performed better than the control group.
Linebaugh and Roche (2015)	Effect of training in production of problematic L2 sounds enhances perception.	46 Omani university students	Experimental and control groups	Pre and post tests	Experimental group performed better than the control group.

Table 2.3.: A chronological summary of key research into L2 listening instruction research.

It is noticeable that few of these studies appear to have adopted a coherent psycholinguistic model of listening processes to inform the features of listening training investigated, or suggested one based on their findings. Consequently, it seems that the features were not selected with reference to a model of expert listening to which L2 listeners could aspire. Yeldham (2009) is an exception. His study examined the effect of what he refers to as bottom-up and top-down approaches on the success of listening training. His aim was to provide insight into the most appropriate blend of the two types of processing in the classroom, and to create a pedagogical framework aimed at learners of varying listening proficiency. The research was a small-scale case study involving six participants receiving each of the two types of instruction. The findings relating to the intermediate level participants suggested that there was less need for strategy instruction than at lower levels, a view in line with that of Field (2008d). Instead, their gains seem largely due to increased speed of processing, likely due to improving automaticity of decoding. The findings of Yeldham (2009) imply that training in lower level processing may be most appropriate for mid-level learners. The study reported here adopts an empirically attested psycholinguistic model to explore L2 listening processes. Thus, the results can be used to inform future research investigating the teaching of word recognition.

2.5. Summary of key issues which provide a rationale for the current study

This chapter has identified a view of listening which suggests that, in word recognition processes, listeners pay attention at multiple levels within an utterance, and that varying importance is accorded to the cues within those levels. Expert listeners recognise words automatically and use high level cues to build meaning and disambiguate where more than one interpretation of an utterance is possible. In contrast, due to their incomplete mastery of the language, the cues relied upon by L2 listeners are approximate, and they make strategic decisions as to which cues are most reliable. In addition, they may lack the cognitive capacity to process speech at higher levels, which, especially in an academic context, diminishes their ability to fully comprehend lectures. Therefore, it seems that the cues which are heeded by L2 listeners may be different from those heeded by L1 listeners. The emerging issue is that L2 listening processes diverge from those of experts, a consideration which the

literature seems not to have fully addressed, and one which is clearly problematic in academic listening contexts when many students are NNES.

The current body of knowledge reviewed in the first half of this chapter illustrates that the processes of listening in an L1 have been explored extensively in the field of psycholinguistics. However, a recurrent issue throughout the second half has been that not enough is known about L2 listeners' behaviour. Conclusions have largely been drawn based on:

i) analyses not based on a full and complete model of the listening process;
ii) studies of single processes (i.e., word recognition or segmentation);
iii) studies using self-report data;
iv) studies using small sections of speech or individual words;
v) small-scale studies; and,
vi) studies that do not consider whether listeners errors in word recognition are the result of mishearings or of strategic behaviour.

Furthermore, there has been very limited discussion of the impact of findings in relation to the specific demands of listening in an academic context and in academic listening pedagogy, a point noted by Lynch (2011:79), who refers to the "low profile of listening research", and the "even lower" profile of academic listening research. In light of this, and of the limitations of previous research into L2 listening processes noted above, the study:

i) is founded on a framework that fully depicts the listening process according to psycholinguistic theory;
ii) investigates all levels of the listening process;
iii) analyses and draws conclusions from tangible evidence of what participants perceive;
iv) uses long sections of speech of an academic genre;
v) is a large-scale inquiry; and,
vi) is triangulated with a qualitative task to add further insight into listeners' strategic behaviour.

Although the study is not primarily pedagogic, its motivation was. Its pedagogic aim supports Field's (2008d) notion of the diagnostic approach to teaching listening, outlined in Section 7.3, whereby classification of errors is the first stage, before the design and administration of remedial tasks.

A summary of the aspects of L2 listening about which little is known include:

- the lexical characteristics of words that influence lexical recognition in L2 listening;
- the cognitive processes that influence L2 listeners' word recognition;
- the extent of L2 listeners' strategy use; and,
- whether L2 listeners make generalisable strategic decisions, in terms of cues relied upon, in order to compensate for their word recognition difficulties.

This summary provides the rationale for the current study which investigates these aspects, and which is based on the notion that L2 listeners use a range of cues to compensate for breakdowns in understanding, yet little is known of the relative importance of these cues, i.e., which listeners rely upon most and which they tend to ignore. Thus, the central issue under investigation is the relative contributions of lexical, perceptual, and contextual factors in L2 word recognition processes. Understanding this issue will provide insight into what degree L2 listeners in an academic context are able to build meaning and construct an accurate interpretation of lecture discourse, processes crucial to academic study. Furthermore, the results of the current research should provide a taxonomy of generalisable listening difficulties related to listeners of approximately similar levels, around which tasks can be designed for classroom use. As such, the current research ultimately aims to inform teaching practices in both EAP and other contexts. While this section has summarised the key issues at play, the following section provides a broader summary of the chapter as a whole and leads to the formulation of the research questions.

2.6. Chapter summary

This chapter has reviewed psycholinguistic research into expert listening processes and examined how these processes may differ for L2 listeners. In order to fully appreciate the complexity of academic listening, its differences were compared to general listening, and the likely impact of these complexities on L2 listeners' lecture comprehension was suggested. L2 listeners' awareness of their word recognition difficulties was considered from the perspective of the strategies used to compensate for these difficulties. It was demonstrated widely, during the course of the second half of this chapter, that research into L2 listening, especially in an academic context, is lacking. As a consequence, this study proposes that L2 listeners' pedagogical needs are not being met, and the L2 listening instruction research reviewed above supports this proposal. Hence, the aim of the current research, as illustrated by the research questions, is to increase knowledge of L2 listening processes, apply this knowledge to an academic context, and ultimately reform L2 listening instruction.

2.7. Research questions

This study addresses the key issues summarised in Section 2.5 and about which there are gaps in the current literature, it does so by identifying five research questions.

Research Question 1

What word level information contributes to the word recognition of L2 listeners in a British EAP context?

This research question investigates word level information associated with:

i) word frequency;

ii) word length; and,

iii) syllable stress.

Word frequency cues are stored in the mind as part of a word's lexical entry and aid the word recognition processes of expert listeners (discussed in Section 2.4.5). Consequently, one of the lexical factors explored as part of the first research question is whether L2 listeners are more able to recognise higher than lower frequency words.

The second lexical factor explored is whether the effect of word length, i.e., number of syllables, impacts L2 listeners' word recognition. It is plausible that longer words may be more easily recognised, given that more syllables mean more perceptual evidence in the signal. Therefore, a further area investigated is whether L2 listeners are more able to recognise multisyllabic words than monosyllabic words.

The third lexical factor explored is the effect of initial syllable stress on lexical recognition. Whilst it is acknowledged that lexical stress is a phonetic feature, it is treated here also as a lexical phenomenon because of its role in lexical segmentation, in accordance with the MSS (Cutler & Norris, 1988) (discussed at length in Section 2.3.3) and, consequently, word recognition. In relation to L2 listeners' ability to apply stress-based segmentation strategies when listening in English, there are conflicting views (as observed in Section 2.4.6). Thus, the third area investigated is whether L2 listeners are more likely to recognise words bearing initial syllable stress, rather than a weak initial syllable.

No L2 listening research has been identified investigating these three features as possible factors assisting word recognition.

Research Question 2

Do L2 listeners in a British EAP context rely more on perceptual cues, or the types of cue widely referred to as 'contextual', when compensating for words that have not been accurately recognised?

It was established in Section 2.4.3 that L2 listening processes are often referred to by the very simplistic and somewhat misleading terms 'bottom-up', meaning lower level processing, and 'top-down', meaning higher level processing (e.g. Rost, 2011; Vandergrift, 1997). This constraining view seems to have influenced L2 listening research in the recent past, and has resulted in perceptual and contextual cues being pitted against each other by investigations into which are more likely to be used in listening breakdowns (e.g. Field, 2004a; Hansen & Jensen, 1994; Tsui & Fullilove, 1998). With the exception of insight from small-scale studies by Field (e.g.2008a, 2008c; 2011), very little is known about what perceptual, lexical, co-textual, and contextual cues are available to L2 listeners and the extent to which each type is used.

This view has also influenced L2 listening teaching methods and materials, and may have led to the widespread assumption that, with limited linguistic knowledge and uncertain phoneme values, L2 listeners rely heavily on what is loosely referred to as 'context' in word recognition. Previous research into this area has either not been large-scale, or has not elicited naturalistic listening processes. Nor has it considered that both types of processing may, in fact, influence word recognition simultaneously. This study addresses all three issues, comparing the relative parts played by perceptual and contextual information in participants' erroneous word recognition.

Research Question 3

To what extent do L2 listeners in a British EAP context rely on word frequency cues when compensating for words that have not been accurately recognised?

This research question examines participants' perceptual errors, identified in Research Question 2, in a further lexical investigation of word frequency cues. Whilst Research Question 1 investigates the influence of frequency on accurately recognised words, Research Question 3 examines words which have been inaccurately recognised and replaced by an alternative, perceptually similar word, for instance, a participant responds to the word '*optimum'* with '*option'*. The aim is to ascertain whether the alternative word is likely to be more frequent than the target word, thus establishing the impact of word frequency cues on participants' erroneous word recognition.

Research Question 4

Which of the identified perceptual cues at syllable and word level are most heavily relied upon by L2 listeners in a British EAP context when compensating for words that have not been accurately recognised?

Investigating low level processing as a single category, i.e., 'perceptual cues', limits the insight available as expert listeners make use of several cues within lower level processing, for instance, prosodic cues. The ability to use all lower level cues contributes to successful word recognition, therefore, this research question looks more closely at the type of perceptual information that is available to L2 listeners. By examining participants'

perceptual errors, identified in Research Question 2, the most influential type of perceptual cue can be established, thus providing a deeper insight into L2 word recognition processes.

Research Question 5

To what extent are L2 listeners in a British EAP context aware of their errors of word recognition?

The final research question is a supplementary issue addressed as part of Research Questions 2 and 3 and is a minor investigation of the strategy use of L2 listeners. Whilst Research Questions 2 and 3 examine evidence of strategic behaviour in terms of the use of cues to compensate for word recognition difficulties, this research question addresses whether or not L2 listeners are aware of their inaccuracies of word recognition and views any such awareness as evidence of their strategy use. By establishing listeners' confidence in their responses, it is hoped to establish the extent to which they simply mishear a word without realising it, and to identify incidents where the awareness of their decoding difficulties leads them to operate strategically, but then they draw upon incorrect cues in the input.

The methodology and the tools required to answer the research questions are explored in the next chapter.

3. Research design

3.1. Introduction

This chapter has two purposes. The first is to discuss the methodology used, and the second is to present the rationale behind the decisions made in the design of the study. To this end, it provides details of the methods employed to address the research questions presented in the previous chapter. It briefly discusses the research paradigm which determines the methods considered appropriate for gathering suitable data, and then provides detail about those methods and the tools chosen. The participants and the context of the study are then presented, and this is followed by a discussion of the selection and design of the task materials. Data gathering procedures are described, followed by consideration of related ethical issues.

3.2. Research paradigm

The nature of the study reported here demands a pragmatic paradigm as its epistemological nature is viewed best from a practical perspective, where what is known provides solutions to problems, rather than a description of reality. In spite of this apparent interpretivist inductive way of looking at the world, i.e., one which allows the researcher to view the world through the experiences and perspectives of others (e.g., the research participants), pragmatism also allows a positivist deductive view of the world (Creswell & Clark, 2007).Thus, a pragmatic paradigm can combine both quantitative and qualitative research methods, depending on the research questions being investigated (Johnson & Onwuegbuzie, 2004).

Exploring the internalised nature of listening requires a researcher to first establish what the listener understood and then to view this in accordance with specific research aims. In the study reported here, participants' responses to a listening task are viewed as signal of the cognitive processes they employed in order to make sense of the speech. The main task investigates this relationship between the input, in the form of the recordings the participants'

hear, and the output, in the form of participants' responses, and in order to obtain more reliable and generalisable results, the task is quantitative (Cohen, Manion, & Morrison, 2007). This research method suits the study's positivist epistemological stance, as it is data-driven, and interprets this relationship with no *a priori* expectations (ibid.).

Alongside the quantitative task, an ancillary qualitative investigation took place, and triangulation of data in this manner is a major motivation for a mixed method research design (Johnson & Onwuegbuzie, 2004). A major section of the quantitative analysis assumes that listeners behave strategically in terms of their errors in decoding, rather than simply reporting mishearings, and so the qualitative investigation sheds additional light on the practical implications of the quantitative findings. This use of triangulation in order to answer the research questions more reliably illustrates a key feature of mixed methods research in that clearly defined research questions were central to the identification of a suitable research design, rather than a hypothesis to be tested (Teddlie & Tashakkori, 2010).

This combination of quantitative and qualitative data originating from mixed research methods is reported not only to enhance the various features of the area under investigation, but also to enable insights into new aspects of the area (Teddlie & Tashakkori, 2010), a notion particularly pertinent to the current, under-researched topic. Research in a closely related field, i.e., listening testing, has similarly employed quantitative and qualitative methods to investigate the cognitive processes of listening (Field, 2012). The participants in Field's (ibid.) study, numbering 29, were taken from a similar population as the current study, i.e., pre-sessional EAP students, and were of a similar proficiency. Verbal reports were used after a quantitative listening task because these were viewed to be the most suitable method by which to enhance the quantitative data in the establishment of "the processes which a listener employs in decoding input" (ibid.:397), an aim similar to the current study. Field reported no limitations in relation to the research methods.

With regards to the quantitative data analysis, an inductive method was used to create categories for coding the data in order to determine patterns which could be quantified. This evolution of categories is reminiscent of Grounded Theory, defined by Strauss and Corbin (1994:273) as "a general methodology for developing theory that is grounded in data systematically gathered and analysed". In this case, while not developing theory, the inductive nature of Grounded Theory methods allows the data, rather than the researcher, to

drive the analysis (Charmaz, 2014). During the analysis, emergent categories were developed by classifying patterns of participants' responses within the data until a set of categories was established to explain their behaviour, supported by the psycholinguistic framework of analysis.

3.3. Research method

The focus of the study reported here is the behaviours involved in the listening process of L2 listeners in an EAP context. The aim is to explore four different possible sources of information on which L2 listeners can draw by investigating two aspects of the listening process, and to ascertain on which they most rely. Firstly, the research investigates the relationship between three lexical features of words and listeners' ability to recognise words. The features examined are the frequency (according to a corpus of spoken language), stress, and the length of words, and it is the impact these features have on successful word recognition that is explored. Secondly, the research examines the relationship between listeners' breakdowns in understanding and the evidence on which they rely within the speech signal, in the form of perceptual and contextual cues. It does so by analysing incorrectly recognised words to ascertain which cues influenced their selection. Investigating these aspects of L2 listening should provide a deeper insight into the challenges faced by L2 academic listeners.

Rather than hypothesising, the study aims to answer research questions, drawn from a review of the literature, which centre on exploring the cognitive processes of L2 listeners. This is achieved by examining participants' correct and incorrect responses to a task in order to gain insight into the lexical, perceptual, and contextual cues they use. As listening is a set of internal, subconscious processes, accessing these processes is extremely challenging and calls for a research method which not only prompts naturalistic behaviour, but also provides evidence of this behaviour. In addition, given that the dominant paradigm was quantitative, sufficient evidence was required to enable generalisations. An observational method was used, namely the paused transcription method, and the following section discusses its suitability.

3.3.1. The paused transcription method

Paused transcription requires participants to listen to a recording of a section of a text into which pauses have been inserted at the end of target phrases. Selection of these target phrases clearly depends on the nature of the investigation. Whenever a pause occurs, participants report immediately, verbally or in writing, the last four or five words heard. The target phrases to be reported on should consist of a small number of words, and this is due to the nature of working memory, discussed in Section 2.4.2. Expert listeners monitor incoming speech, i.e., retain words in working memory, until they recognise a complete syntactic unit, at which point the words are parsed (Field, 2008c) and become an abstract proposition (Jarvella, 1971:413). Retaining the number of words in a phrase verbatim is, therefore, generally within the capacity of expert listeners who can automatically decode the speech. However, word recognition is likely to demand more attentional effort in an L2 listener, thus it is not realistic to expect recall to be as extensive as for the L1 listener (Cook, 1979; Meara, 1980). It is therefore prudent to limit the transcription to only the last four to five words, in the form of a complete phrase or clause, such as '*presentation is very important*'. Support for this view can be found in the use of paused transcription tasks by Pemberton (2004), who required his participants to type long sections of speech. He reports this as a limitation of his method as his participants were not able to retain all the words in working memory long enough to report them, and he amended his recording for subsequent experiments so that shorter phrases were reported on.

The paused transcription method evolved from small-scale L1 transcription tasks involving single words, which provided useful data with varying research aims. For example, listeners reported words excised from natural speech (Pollack & Pickett 1964) and identified words embedded in non-words (McQueen, Norris, & Cutler, 1994b). However, these studies used short sections of material recorded specifically for the research and, as such, natural listening processes were not being engaged. L2 researchers have also used dictation tasks to investigate word recognition, for example, Koster (1987) used this in his study into the effect of semantic constraints on word recognition (see Section 2.4.8). Nevertheless, there are limitations to using small-scale dictation when investigating the entire cognitive processes of listening, as cues at all levels are not available to listeners. Also, in conventional dictation tasks, where listeners transcribe every word, they are likely to listen systematically to each

word, and so focus on perceptual information only at word level, thus excluding cues at levels higher than single words.

In contrast to much of the laboratory-based research and conventional dictation tasks that have been used in investigations into listening, the paused transcription method more closely replicates the processes used in real-life listening, and it is these processes that the study reported here aims to investigate. As such, listeners should be listening in a naturalistic manner, that is to say following the wider meaning of speech, as in real life listening, rather than focusing only on word level understanding. Such ecologically valid data can be obtained using the paused transcription method for two reasons. Firstly, using longer sections of speech means that cues at all levels, such as co-text and context, are available to listeners, a factor particularly pertinent when investigating processes during longer utterances, such as academic discourse, the context of the current study. In addition, the paused transcription method does not require listeners to transcribe every word, and the irregularity of the pauses at which they report the last phrase heard also discourages them from focusing at word level as they cannot predict when a pause will occur. Results obtained by Field (2008c) using a paused transcription task illustrate that using longer texts provides listeners with cues at levels higher than word level, his participants' responses included words which matched syntactically and semantically, but not phonologically. For example, in transcribing the phrase '*I found out that the **thud** was the cat*' (in the text, a cat had just been hit by a car), with the target word in bold, responses included '*the sound was the cat*', and '*where was the cat'.* This illustrates listeners' use of semantic and syntactic co-text.

As well as the irregularity of the pauses being crucial to the soundness of the research design, the length of the pauses is also significant. In his research, Field (2008a) planned the length of the pauses to allow time for only 4-5 words to be transcribed, in order to encourage his participants to follow the task instructions. In addition, this approach meant they did not have time to revise their responses, so their initial decoding decisions were captured, enabling a more accurate insight into their listening processes.

The phrases targeted in paused transcription tasks are selected by the researcher to provide specific linguistic difficulties determined by the research aims. For example, in research using a paused transcription task to compare listeners' ability to decode content and function words, Field (2008a) inserted pauses at phrases which contained a number of

examples of these two categories of words. He was then able to analyse the responses to investigate participants' listening processes when decoding each category. Furthermore, in research using this method to investigate participants' recognition of frequent words, Pemberton (2004) used a text principally made up of frequent words and, as such, inserted pauses at the end of random phrases with the knowledge that all phrases would include frequent words. The study reported here calls for data which illustrates how listeners respond to decoding difficulties caused by unknown words and perceptually unsalient phrases. Therefore, pauses were inserted in phrases containing low frequency words, and in phrases which were perceptually unsalient, for example where aspects of connected speech reduced the saliency of words, and where segmentation difficulties would likely occur (discussed in detail in Section 3.5.1. below).

In a written paused transcription task, such as the one used in the current study to gather quantitative data , the output is a set of written responses to each targeted phrase, and therefore the researcher has a permanent record of participants' responses for analysis. The responses are a series of words which the participant transcribes, any of which may either be correct, blank, or incorrect. The responses can then be coded and quantitatively analysed in terms of a cognitive analytical framework. In the study reported here, the participants' correct responses were analysed to investigate what lexical information contributes to the word recognition, and their incorrect responses, referred to in this study as miscues, were analysed to establish which cues participants made use of in their erroneous decoding. Details of the collation of data are reported in Section 3.10.

A further aspect where the paused transcription method may reflect real-life listening is by the researcher playing the recording only once, thereby imposing natural, real-time constraints on listeners. Of the studies mentioned above, Field (2008a, 2008b) imposed this constraint. In contrast, Pemberton (2004) allowed listeners to replay recordings several times, thus his data did not reflect real-life listening where recursion does not naturally take place. However, he did use key-stroke software to enable him to access his participants initial responses. In the study reported here, the procedure follows that of Field, so the recordings are played only once.

The research mentioned above illustrates several features of the paused transcription method which reflect its suitability for gathering data in response to naturalistic listening

processes. However, a potential flaw in this method is that listeners are required to report in writing what they have heard (or in Pemberton's (2004) case, by typing). Consequently, evidence of what they have heard is compromised by their competence in this skill, for example by their spelling abilities, or their writing speed. In this study, following the paused transcription research discussed above, (Field, 2008a, 2008c; Pemberton, 2004), a degree of tolerance was exercised on spelling accuracy if it seemed that the participant had simply made a mistake, for example writing '*qestion'* in response to recognising the word '*question'*. This decision also follows Buck (1988:31), whose view of dictation tests is that "spelling mistakes should be ignored in cases when it is obvious that the mistake is indeed a simple spelling mistake".

No further limitations per se are reported by the two researchers, i.e., Field (2008a, 2008b) and Pemberton, (2004), who have used the paused transcription method, however one caveat in relation to the post-perceptual nature of the paused transcription method has been reported in that…

> it demonstrates, not how listeners process the percept at the moment of hearing, but what interpretations they finally derive. The goal of the enquiry is to establish the quality of the bottom-up information available to L2 listeners when constructing a meaning representation of a passage. Evidence of on-line processing is thus not strictly relevant; what is important is the extent and accuracy of the linguistic information that listeners extract from the input.
>
> (Field, 2008c:17)

Given that the current research investigates listeners' online processes immediately after they take place, this limitation is not of relevance.

The discussion of the paused transcription method demonstrates its suitability as the research method chosen to meet the aims of the current study. It seems that this method could be employed more widely in research which investigates the cognitive processes of listening, specifically L2 listening, most of which has attempted to gain insight into L2 listening processes using participants' self-reported data. A crucial benefit of the paused transcription method lies in the transcription itself, i.e., listeners report on their listening in the form of a

transcription, this provides the researcher with hard evidence of their interpretation of a text. Participant errors can then be analysed by a researcher trained in the field, who can use the data to identify why listening has broken down. Conversely, in the context of self-reporting, although listeners can generally report reliably on metacognitive processes which involve control over how they listen or make decisions, they are less able to report reliably on many cognitive processes, as these are highly automatic and therefore not necessarily accessible to report upon.

This view is illustrated in research involving self-report methods, where there is a danger that participants may misconstrue or inaccurately describe reasons for their listening breakdowns, or they may not have the meta-language required to describe them accurately, a concern echoed by Pemberton (2004). This could lead the researcher, unwittingly, to misinterpret them. For example, in a study by Goh (2000:61), a listener reported "Listened to TV news. Some words sounded familiar but I can't remember their meanings. So I must develop my reaction speed". The listener felt s/he recognised words but could not automatically assign a meaning to them. The researcher suggested that "word-referent relationships might not be automatised… the students 'knew' the words but were slow when activating this knowledge" (ibid.). However, it is possible that the listener simply *thought* a word sounded familiar, when in fact s/he mistook the word for another, or failed to segment it correctly. Similarly, Hasan's (2000:142) listeners reported "unfamiliar words interfere with my listening comprehension". Although these words may indeed have been items of vocabulary not in their lexicon, they may also be known words which the listener simply did not recognise in connected speech, perhaps due to segmentation problems or coarticulation effects. In sum, paused transcription data provide accessible evidence of listeners' interpretations of a text, and so the need for participants to diagnose the reasons for any errors is unnecessary.

Having demonstrated the suitability of the written paused transcription method to provide quantitative data to address Research Questions 1 to 4, the next section details the use of this method to provide qualitative verbal report data.

3.3.2. Oral paused transcription and verbal reports

The research design incorporated a second, and subsidiary, task, qualitative in nature, designed to address Research Question 5. The task aimed to establish whether participants were certain of their decoding decisions, or whether they were aware they were making guesses. This enabled insight into three issues: i) participants self-awareness as listeners; ii) whether their incorrect responses were simple mishearings, or if they were strategic attempts to match an unrecognised sequence to a known word; and iii) whether their oral responses were similar to those obtained in writing in the main task reported in Chapter 5. Participants who had undertaken the earlier written paused transcription task were asked to volunteer for this additional task, and 21 did so.

For this second paused transcription task, qualitative data in the form of verbal reports were gathered through the use of a stimulated recall procedure. Stimulated recall entails assisting recall of a recent event or behaviour with the assistance of prompts, and it is widely believed (e.g., Bloom, 1954, cited in Gass & Mackey 2000; Ericsson & Simon, 1987) that this method enables respondents to report more accurately on metacognitive and cognitive processes, provided they are relatively recent. Stimulated recall is a common research tool used in SLA (Gass & Mackey, 2000). For example, Leeman (1999) used stimulated recall to investigate learners' use of recasts to promote language development. Also, Mackey, Gass, and McDonough (2000) successfully elicited participants' initial perceptions about a previous interaction using this method. All of these uses of stimulated recall involved establishing "when and if particular cognitive processes are being employed" (Gass & Mackey, 2000:21), thus it is appropriate to provide insight into participants' self-awareness of their decoding decisions, as in the current study.

It was necessary for the stimulated recall task to take place immediately after a paused transcription task, as immediate reporting has been found to strengthen the prompt, promote accuracy, and limit memory effects (Bloom, 1954, cited in Gass & Mackey 2000). The verbal reports (see Appendix C for transcriptions) were given in response to a paused transcription task, as discussed in Section 3.3, but with participants reporting orally to a recording, rather than in writing. The task, including participants' responses, was recorded and at the end it was played back to the participants as a prompt to enable them access to the decoding

processes they had used while listening. This procedure is similar to that of Mackey et al. (2000), who recorded their participants' initial task and used the recording as a prompt immediately after. Details of this procedure are provided in Section 3.6.2.

A possible limitation of this method is that "people often cannot report accurately on the effects of particular stimuli" (Nisbett & Wilson, 1977:233). However in the current research this possibility of inaccurate reporting was mitigated as participants were questioned only as to the certainty of their responses, rather than anything more complex, thus accurate reporting seemed more likely.

3.4. Participants

The written paused transcription task was large-scale and undertaken by 171 participants. of approximately B1 to B2 level, according to the proficiency scales of the Common European Framework of Reference for Languages (CEFR). They are referred to throughout this thesis as 'mid-level', and descriptors of learners at this level can be found in Appendix D. The participants were part of a group of students preparing to enter postgraduate courses at the University of Reading, or other UK universities, by taking part in three summer pre-sessional courses. The courses varied in length from five to 12 weeks, and took place at the International Study and Language Centre (ISLC), University of Reading, in 2011.

It was necessary to control for listening proficiency as the intention of the study was to focus on the behaviour of mid-range learners. To ascertain what constituted a mid-level learner, it was necessary to examine the University's listening placement test results from 2008 to 2010 to establish suitable selection criteria for participants (this placement test is used to allocate students into broadly homogenous groups for study). The mean result of the test was 14.32 and the standard deviation was 3.41. Therefore, it was decided that participants who scored under 11 out of a possible 23 in the test, i.e., approximately one standard deviation below the mean, would be excluded from the study data. Students with a listening placement test score of 18 or over out of a possible 23, i.e., approximately one standard deviation above the mean, were also excluded as they could not be considered mid-level listeners. Participant response sheets, presented in Appendix E, included a space for their

name so that their placement test scores could be ascertained and eligible participants identified. The listening placement test scores and nationalities of all participants can be found in the spreadsheet of coded data in Appendix F.

It was important that participants should be controlled their for length of residence in an English speaking country as listening proficiency can develop very rapidly after a short period in the target language environment. In addition, previous extended exposure to the target language may have resulted in higher levels of progress in the participants' language learning, which may render as outdated the early placement test results obtained at the beginning of courses. Therefore the participant response sheets asked whether participants had resided in an English speaking country in the past. Responses from participants who had done so for more than two weeks were excluded on the grounds that they would have had extensive exposure to the target language. As such, all participants had resided in an English-speaking country for two weeks or less.

While proficiency level was controlled for, it was not possible in practical terms to control for L1 since testing took place in class groups. In the event, the majority of participants proved to be speakers of Mandarin Chinese (N = 132). Other first languages included Thai (20), Vietnamese (6), two each of Japanese, Greek, Korean, Turkish, and Turkmen (total of 10), and one each of Urdu, Mandarin Taiwanese, and Arabic (total of 3).

The research design required participants to be divided into two groups, in order to respond to two different recordings, and this was necessary for two reasons. Firstly, one of the research aims was to strengthen the data by exploring whether the same behaviour was observed with two different recordings, and hence different content and speakers were required. Secondly, the verbal report required a further paused transcription task to take place. If the participants of this task were responding to the same recording twice, they would have become familiar enough with the recording to render their responses on the second hearing unrepresentative of real-life listening. There was a difference between the participant numbers for the first task because it took place in participants' listening classes whilst they were studying on the pre-sessional course, and class sizes were not equally divided. Also, it was not possible to predict which students would meet the criteria required to participate in the study, so it may have been that more students who did not meet the criteria happened to be in a particular group. The participants were divided into two groups as follows:

Paused transcription task

i) Group A.

77 participants using recording 1.

ii) Group B.

94 participants using recording 2.

Oral paused transcription and verbal report

The participants were two groups made up of subsets of the main groups A and B.

i) A subset (n=11) of Group A using recording 2

ii) A subset (n=11) of Group B using recording 1.

This is illustrated below:

Paused transcription	Group A (n=77) recording 1	Group B (n=94) recording 2
Oral paused transcription and verbal report	Subset of group A (n=11) recording 2	Subset of group B (n=11) recording 1

Table 3.1: Division of participants into groups for the two tasks.

3.5. Materials

3.5.1. Recordings

Pilot study recordings

Pilot studies were undertaken to ascertain the materials best suited to the main study. Participants were pre-sessional University of Reading students who were studying in the

same context as those participating in the main study, and were of approximately the same level of proficiency. Each pilot study was undertaken with 11 participants. In order for the task to be as naturalistic as possible, using a section from a fully authentic lecture was most appropriate. A recording from the British Academic Spoken English (BASE) corpus was identified, one which had been recorded at the University of Reading. The subject of the lecture was typography and the topic concerned differences between reading a screen and reading print. This lecture was selected for use in an attempt to control for participants' prior knowledge of the subject; there had been only one typography pre-sessional student at the university in each of the previous two years. Additionally, the topic was likely to be generally familiar to participants and the selected section did not include any technical terms. Therefore, a degree of external knowledge was likely which would allow for top-down processing, important in listening processes (e.g., Cutler, 2012), to take place as it would in a natural listening situation.

It was evident from the data gathered that an authentic lecture was too complex to provide informative data. Although the language used by the lecturer was likely to have been familiar to the participants', the speed of delivery of an authentic lecture seemed to be too challenging for them. They were not able to decode with sufficient competence to report a semblance of the target phrase, instead they either failed to respond at all or wrote words which they had heard randomly throughout the excerpt rather than immediately before the pause. In addition, several participants failed to respond to the first pause, expressing surprise and/or confusion when the pause occurred. However, by the second pause most participants had responded as expected. As a consequence, using a fully authentic lecture was deemed too complex for students of this proficiency.

A different class of the same cohort of students took part in the second pilot study. In order to address the two problems mentioned above, two new recordings were designed. They were semi-authentic lectures and included an 'example pause', which was played to the participants before the actual recording began as part of the teachers' instructions. Data gathered in response to this task was meaningful, and the recordings were used in the main study. The nature of the recordings is discussed in the next section.

Main study recordings

The two recordings that were chosen for the study were extracts from lectures within the EAP course book *English for Academic Study: Listening* (Campbell & Smith, 2009). All lectures within the course book had been identified as being appropriate in terms of content and delivery for EAP pre-sessional students, such as the participants in this study. The lectures involved different speakers and different topics; transcriptions of both texts and the positions of the pauses can be seen in Appendix G. Several factors were considered during the selection of the recordings and these are now discussed.

1. **Authenticity**: Authentic lectures would have provided the most ecologically valid data since they are real-life listening events, but unfortunately the first pilot led to the conclusion that using authentic lectures was not possible (see Section 3.5.1). The next best option was semi-authentic lectures. The lectures in the course book were considered semi-authentic because:

   i) they were based on "authentic lectures from a range of academic fields [and] … have been recorded in a genuine academic environment" (Garnet Education, n.d.)
   ii) they were based on authentic degree programme lectures that had been delivered at the University of Reading and which had been audio-recorded in their entirety in a normal lecture setting.
   iii) recording extracts were transcribed without changes so as to retain features such as false starts, hesitations, and reformulations. The extracts were re-recorded verbatim by professional actors because of the variable quality of some of the recordings. If the actors made any mistakes while recording, these were left in and the transcripts altered to reflect these minor changes (Campbell, 2009, personal communication).
   iv) the actors were speaking in a natural and unforced way, but at a speed and volume appropriate for the English level of the audience.

2. **Lecture topic:** Recording 1 was on the topic of health, and Recording 2 discussed using questionnaires for research; both topics were considered to be general enough to be accessible to all participants.

3. **Identifying a suitable segment of lecture:** Each entire recording was analysed to identify a segment of a similar length in each which fulfilled the following criteria:

   i) The segments were coherent after being isolated from the rest of the text.
   ii) They had a syntactic or phrasal structure for their entirety in order to enable participants to build a discourse representation as they would in a real life listening event (see Section 3.3.1).
   iii) They presented a similar number of examples of the two types of decoding difficulties which would form the 'targets', namely perceptually complex or lexically complex, as discussed below.

4. **Length of extract:** Five to six minutes were deemed appropriate for the following reasons:

   i) Longer recordings were impractical as participants were students whose classes were interrupted for this study.
   ii) Participants' responses may have been affected by fatigue had they been asked to listen and complete the task for a longer period.
   iii) The nature of the listening process. In order to enrich the meaning of a speaker's words, listeners build a 'discourse representation' by using their recall of the text so far and their knowledge of the topic under discussion (Field, 2013). Clearly, a whole lecture requires listeners to carry forward a much more complex discourse representation and, considering the participants' lack of mastery of the language, this may impose unreasonable cognitive demands and affect the validity of the data. On the other hand, recordings of five to six minutes, which remained coherent after isolation from the rest of the text, were sufficient to provide participants with the co-textual and contextual cues which would be available in a natural lecture environment, and thus required them to build a less challenging discourse representation.

Choice of targets

The recordings discussed in the previous section included specific targets which were likely to present two types of decoding difficulties. The targets consisted of content words in clauses which were considered to be either:

i) perceptually complex, i.e., not salient due to word boundaries within the clause being unclear, for example 'not **attacking'**, where resyllabification takes place and the initial syllable is unstressed; or,
ii) lexically complex, i.e., words likely to be unknown to the participants, for example 'how to **cure** it'. It should be noted that, given the pervasiveness of connected speech features in natural speech, lexically complex words may also be affected by features such as linking, and so also be considered perceptually complex.

In this thesis, 'knowing' a word refers to the ability to recognise the word aurally while attributing the appropriate meaning to it (adapted from Kelly, 1991).

The targets were short sections of no more than five words, so as not to exceed participants' working memory capacity, estimated by Klatzky (1984) and Miller (1956) to be seven (+/-2) units of information. This decision is informed by the three previous studies which have used the paused transcription method (Field, 2008a, 2008b and Pemberton, 2004) (see Section 3.3.1). One of a series of experiments in the earliest of these required listeners to report all words before the pause, which the researcher deemed to be too much of a cognitive challenge and reliant on memory (Pemberton, 2004), and the second two studies targeted four to five words, which resulted in useful data (Field, 2008a, 2008c)

The length of each recording was five to six minutes and ten phrases were targeted in each recording. The second pilot study illustrated that this number provided sufficient co-textual and contextual cues before each pause in order to replicate naturalistic listening, whereby listeners make use of multiple cues in the listening process (see Section 2.2, and the psycholinguistic model of listening processes in Appendix A). Recording 1 included 21 content words and Recording 2 included 20 content words. The phrases which included the targets are detailed below in Tables 3.2 and 3.3, and the targets are in bold. The table also includes a broad phonemic transcription of each phrase, to illustrate how features of connected speech affected the production of the phrases by the speaker.

Recording 1 Phrases with target words	**Broad phonetic transcription**	**Reason phrase targeted**
certainly **presentation** is very **important**	[ˈsɜːtn̩li ˌpɹɛzn̩ˈtʰeɪʃn̩ ɹz ˈveɹi ʲɪmˈpʰɔːtn̩tˀ]	linking, intrusion, potential segmentation difficulties
the **flow** of the **questionnaire**	[θə ˈfləʊ ʷə θə ˌkwestʃəˈneə]	linking, intrusion, potential segmentation difficulties
by the **market research process**	[baɪ θə ˈmɑːkɪt ˈɹiːsɜːtʃ ˈpɹəʊses]	fast speech
pretty personal questions really	[ˈpɹɪti ˈpʰɜːsənəɬ ˈkwestʃənz ˈɹɪəli]	fast speech
little bit of **trouble** about it	[ə ˈlɪt̬ɬ ˈbɪt̬ə ˈtɹʌbɬ̩əˈbaʊt̬ɹtˀ]	fast speech, linking, reduction, weak initial syllable, potential segmentation difficulties
need to **pre–test** the thing	[ˈniːdˀ tʰə ˈpriːˈteʂ θə θɪŋ]	fast speech, reduction
get **optimum response**	[ˈget̬ˈɒptɪməm ɹɪˈspɒns]	infrequent word, linking, weak initial syllable, potential segmentation difficulties
get the **implicit approval**	[ˈget̬ θi ʲɪmplɪsɪt̬əˈpɹuːvəɬ]	infrequent word, intrusion, linking, weak initial syllable, potential segmentation difficulties
want a **statistically valid sample**	[ˈwɒnt̬ə ˌstəˈtʰɪstɪkli ˈvælɪd ˈsɑːmpɬ]	infrequent words, weak initial syllable, linking, potential segmentation difficulties
the **style** of **particular** questions	[θə ˈstaɪl̬əv pəˈtʰɪkjʊlə ˈkwestʃənz]	infrequent word, weak initial syllable, linking, potential segmentation difficulties

Table 3.2: Target words Recording 1.

Recording 2 phrases with target words	**Broad phonetic transcription**	**Reason phrase targeted**
at least **controlling** many **illnesses**	[əˀt ˈliːs kənˈtrəʊlɪŋ ˈmeni ʲɪɬnəsɪz]	connected speech, potential segmentation difficulties
getting **sick** in the **first place**	[ˈgeʈɪŋ ˈsɪk ɪn̪ θə ˈfɜːs ˈpleɪs]	linking, reduction, potential segmentation difficulties
and that is **true**	[ən̪ ˈθæʈ̚ ɪz ˈtruː]	fast speech, linking, potential segmentation difficulties
we are not **attacking**	[wɪə ˈnɒt̚ əˈtʰækɪŋ]	linking, weak initial syllable, potential segmentation difficulties
the **same period** of **time**	[θə ˈseɪm ˈpʰɪəriəd̚ əv ˈtʰaɪm]	linking, potential segmentation difficulties
the **number** of **prescriptions**	[θə ˈnʌmbəɹ̚ əv pɹɪˈskɹɪpʃənz]	weak initial syllable, intrusion, potential segmentation difficulties, infrequent word
the **results** are **disastrous**	[θə ɹɪˈzʌlts̚ ə dɪˈzɑːstrəs]	reduction, linking, potential segmentation difficulties, infrequent word
in **terms** of **life expectancy**	[ɪn ˈtʰɜːmz̚ əv laɪfˌɪkˈspektənsi]	linking, weak initial syllable, potential segmentation difficulties, infrequent word
Norway, for instance, has **witnessed** a **drop**	[ˈnɔːweɪ fəɹ̚ˈɪnstəns həz ˈwɪtnəst̚ə ˈ dɹɒp]	infrequent word, linking, potential segmentation difficulties
how to **cure** it	[ˈhaʊ̚ tə ˈkjʊəɹ̚ɪʔt]	infrequent word, linking, potential segmentation difficulties

Table 3.3: Target words Recording 2.

Insertion of pauses

To insert pauses after the target phrases and clauses required the lectures to be converted from CD to MP3 files and then imported to Audacity (free open source software for editing sound files). After editing, the files were exported back to MP3 format for ease of use in the classroom. Presented below are the criteria used for the insertion of the pauses; transcriptions of the recordings can be found in Appendix G.

1. **Space between pauses:** The pauses were inserted irregularly. Previous research (Field, 2008a) shows that irregular pauses mean that participants cannot anticipate when they needed to focus attention at word level to transcribe what they hear, thus they listen to the recording and attempt to follow its meaning, as in real-life listening. In addition, no pause was inserted for the first 30 seconds to allow participants sufficient time to become normalised to the speaker and the topic (Pisoni, 1997).

2. **Time between pauses:** Pauses were insterted at least ten seconds apart and inserted at natural pause breaks in the speakers' deliveries. This was in order to provide sufficient co-text and context before the pause to encourage higher-level processing to take place, rather than unnatural focusing at lexical level (Field, 2008a:419).

3. **Length of pause:** Pauses needed to be long enough for participants to write the four to five words of the target phrase, but not long enough for them to change their minds as this would not be entirely representative of natural listening, and thus would have affected the quality of the data. The second pilot study, using the materials later selected for the main study with gaps inserted, showed that seven seconds were appropriate.

4. **Replay of recording:** Recordings in Field's (2008a, 2008b) research were played once only in order to replicate naturalistic listening situations by providing the processing constraints found in real life. Conversely, Pemberton (2004) allowed multiple replays, but analysed the initial responses rather than any changes made on subsequent listening. In line with Field (2008a, 2008b), this research sought to acquire data in conditions that were as ecological as possible and as such the recordings were played once only.

3.5.2. Participant response sheets

The response sheets were found to be successful during both pilot studies and were subsequently used in the main study. Copies can be found in Appendix E. At the top of the sheet, there were sections designed to: i) ascertain whether students fulfilled the selection criteria for the study (see Section 3.4); ii) obtain consent to use the response for the study (see Section 3.8); and, iii) provide task instructions (see Section 3.6). Under this section, for the written task only, lines numbered 1 to 10 indicated where participants should write their response to each pause.

3.6. Procedure

3.6.1. Paused transcription

Given the large number of participants, it was necessary for the task to take place in their classes. To satisfy the requirement of the course directors that the task should have pedagogical value, class tutors delivered a 30 minute lesson based around the recordings and designed by the researcher. This process enabled all pre-sessional students to take part in the task, but only the data from those who indicated agreement on their response sheet and who fulfilled the sampling criteria were selected for the study.

The procedure for the task required precise planning in order to ensure validity as it would be undertaken not by the researcher, but by 32 pre-sessional teachers teaching 436 students. The instructions given to teachers carrying out the task were piloted on two pre-sessional teachers, who also took part in the main study, in order to ensure they were unambiguous, and they were found to be suitable (copies can be found in Appendix H). The researcher advised teachers that all the documentation they needed was stored in electronic folders in a shared drive on all PCs. They included an MP3 audio file of the recording with pauses inserted, written instructions for the teachers, and the student response sheet (paper copies of which were also provided).

The task was administered within the first week of the participants arrival for their pre-sessional courses to ensure that they had not spend long periods of time in an English-speaking country (see Section 3.4). In the classroom, teachers read to their classes the students' task instructions which were printed on the participant response sheets, showed them the response sheet, and pointed out the section for the participant to sign their informed consent. Teachers played an example pause so that students knew what to expect, the need for which was identified in the second pilot study. The example pause was unrelated to the recordings in terms of speaker and topic and formed a separate sound file. Then they distributed the response sheets, allowing students time to read the instructions themselves. Teachers played the recording and the researcher collected the responses at the end of the class.

A threat to research design validity is reactivity effects, where participants alter their normal behaviour due to their awareness that they are being observed (Cohen et al., 2007). These were considered minimal as students are used to various types of learning exercises taking place in the classroom, especially what is essentially a listening gap fill activity. Also, students were assured that this was not a test, but that it was the first stage of a listening lesson. It is therefore likely that students' behaviour would not be out of the ordinary for them during the task.

After collecting the response sheets it was necessary to identify which of the students fulfilled the criteria necessary to participate in the study and who had signed agreement to their answers being used in this research. A total of 171 responses fulfilled the criteria and these made up the sample.

3.6.2. Oral paused transcription and verbal report

This was a small-scale task to shed light on the cognitive processes used by participants when decoding the target phrases. During this task, a small group of participants from those who had undertaken the main task were asked to comment on their responses to a second paused transcription task in which they reported orally. This provided some insight into whether their incorrect responses were the result of strategic behaviour or were simply

mishearings. One week after the quantitative paused transcription task, volunteers were called for to take part in a further task, and those who fulfilled the criteria necessary to take part in the study, discussed in Section 3.4, were selected.

This task was conducted on a one-to-one basis with only the researcher and one participant present. Each participant was given time to read and sign the response sheet, which can be found in Appendix E. Next, the researcher explained that the task was similar to the one undertaken previously, but this time they would speak their responses at the pauses, rather than write them down, and their responses would be audio-recorded. Then, they would listen to the recording again and the researcher would ask them questions at each pause. The oral responses acted as prompts (Gass & Mackey, 2000) and were in the listener's own voice, therefore they provided a reliable cue for subsequent recall, discussed in Section 3.3.2. The participants were given the opportunity to seek clarification of these instructions and when they indicated their readiness, the task took place.

All stages of this task were recorded on a digital voice recorder. The recording was played, and the participants responded orally. Next, the recording was played a second time during the stimulated recall stage as a prompt to assist participants in reporting the cognitive and metacognitive processes used during the first hearing. After each pause, the researcher paused the recording itself and asked, "are you sure those words are correct", and "are there any words you are not sure of". At this point, participants reported verbally on the confidence of their decoding decisions. For example a participant reported "I only know the first word". Generally, participants responded in various forms of "sure", "sure of them all", or "not sure of the first/last word".

3.7. Ethics

Guidelines by the British Educational Research Association (BERA) state that "all educational research should be conducted within an ethic of respect for The Person, Knowledge, Democratic Values, The Quality of Educational Research, Academic Freedom" (2011:4), and these values are mirrored in The University of Reading's Research Ethics Guidelines (n.d.). These guidelines state that permission to conduct research should be obtained from the University of Reading Research Ethics Committee, subsequently an

application was submitted to the Department of Applied Linguistics Ethics Committee and approved accordingly in August 2010. Subsequently, the following month, permission to carry out this research during the ISLC pre-sessional courses was sought and granted by the Director of the ISLC during a face to face meeting. A copy of the application for ethical approval and confirmation of its acceptance is presented in Appendix I.

3.8. Informed consent

According to BERA, one of the researcher's responsibilities to participants is the notion of informed consent, which is defined as "the condition in which participants understand and agree to their participation without any duress, prior to the research getting underway" (2011:5). In this research, task 1 was deemed to be of pedagogical value and therefore all students took part. In order to seek permission to use the responses for the research, information sheets and consent forms were supplied to all students during the class and these can be found in Appendix J. These included details about the researcher and the research project, notification of students' right to withdraw from the study at any stage and procedures for how to do so, and the contact details of the researcher's supervisor. Students were asked to sign the section on the participant response sheet to provide permission for their responses to be used for the research. A response from only one student who did not agree was removed and destroyed.

During the verbal report exercise, the informed consent procedures were identical to the quantitative task, expect that the forms were given to each participant individually before the task.

3.9. Privacy of data

A data set of the words transcribed was recorded by the researcher using an Excel spreadsheet on a home PC. The data were accessible to only three people electronically: the researcher, the researcher's supervisor, and a judge (a University of Reading staff member) who acted as inter-rater for the reliability check (see Section 3.11). The hard-copy responses from participants are kept in a locked cabinet in the researcher's office at the University of Reading. Anonymity was not possible in the initial stages as the researcher needed access to

the participants' identity in order to ascertain their listening placement test result, and the length of time spent in an English-speaking country. However, before the data was transcribed, the task answer sheets and recordings were coded to ensure anonymity, and codes were used in all references to respondents.

3.10. Collation of data

The paused transcription tasks discussed in the first half of this chapter provided data in the form of a set of written and oral responses from each participant of target sections of a recording. The first stage in collating the data involved the researcher transcribing each participant's responses to both tasks into to an Excel spreadsheet. Each participant's response sheet was allocated a number, which was transcribed along with details of the participant's first language and listening placement test result. The latter two items were noted in order to provide a full picture of the participants.

Responses were coded by the researcher as follows:

i) An alternative word to the target was transcribed in full. In the oral responses, the researcher transcribed her interpretation of the word spoken in an orthographic form, for example a response to the target '*flow'* was transcribed '*float'*. In the case of a non-word, her interpretation in an orthographic form was transcribed, for example a response to the target '*implicit'* was transcribed '*pressit'*.
ii) A correct response was transcribed as '1'. Spelling mistakes were ignored in cases when it was clear that a simple spelling mistake had been made (see Section 3.3.1).
iii) A blank response was transcribed as '0'.
iv) Where a participant left a seemingly deliberate gap in their response to Task 1. i.e., it was evident that they had heard a word, but had made no attempt to transcribe it, it was coded with a '-' (dash). The following examples of responses illustrate these instances:

_____ a drug (the participants drew a line)

market put on it (the participant left a clear gap in their written text)

An example of a coded participant response to one phrase can be seen below in Table 3.4. Coding in this manner proved straightforward and no difficulties were encountered.

L1	Pl tst L	P	Phrase to report (targets in bold) and responses			
			flow	of	the	**questionnaire**
Chinese	11	1	words	and	-	1

Table 3.4: An example of a coded participant response to one phrase.
Key: L1 = first language; Pl tst L = listening placement test score; P = participant number

Data for each recording were treated separately because two separate groups of participants were responding to two different recordings on two different topics. Thus, each recording comprised completely different sets of phonological stimuli, and it was therefore necessary for the quantitative analysis addressing Research Questions 1, 2, 3, and 4 to be undertaken separately for each recording. If the recordings were combined, the variance patterns of each feature of the analysis would be disguised. This separation of results also enabled a comparison of results across recordings, which informed their strength.

After collating and coding the data, overall, figures for written responses to the two recordings were calculated and are shown in Table 3.5 below:

	Correct responses (% of total)	**Blank responses (% of total)**	**Miscues (% of total)**	**Total**
Recording 1 (P=77; t=21)	862 (53.37)	472 (29.23)	281 (17.40)	1615
Recording 2 (P=94; t=20)	798 (42.40)	814 (43.25)	270 (14.35)	1882

Table 3.5: Task 1 (written) Global figures for responses to targets by recording.
Key: P# = number of participants; t = targets

Mean responses per participant were calculated and are presented in Table 3.6 below.

	Mean per participant of miscues	Mean per participant of blank responses	Mean per participant of correct responses
Recording 1 (P=77; t=21)	3.65	6.13	11.19
Recording 2 (P=94; t=20)	2.87	8.66	8.49

Table 3.6: Task 1 (written) Mean responses per participant by recording.
Key: P# = number of participants; t = targets

Having collated the blank, correct, and incorrect responses, the data sets relating to the latter two categories could be analysed separately to address the related research questions. The fully coded data for the written task can be seen in Appendix F.

3.11. Inter-rater and intra-rater reliability checks

Reliability refers to how dependable research is in terms of the procedures used, the data collected, and the data analysis conducted (Dörnyei, 2007). In the study reported here, after collation of the data, it was necessary to categorise participants' erroneous responses to identify the cues which misled them. The reliability of this categorisation was ensured through processes of inter-and intra-rater reliability checks, and this section describes those processes.

A proportion of the researcher's categorisation decisions were checked by an independent judge, who was a colleague with several decades of experience in EAP and Applied Linguistics. He also co-authored an academic listening course book which has been used on pre-sessional courses for many years. The judge was provided with the spreadsheets of participants' miscue responses to each target (N=551), transcribed but not categorised. He then categorised approximately 10% of the miscue responses (N=57), which he selected at random. In addition, 16 responses to two targets were also referred to the judge when the researcher sought to clarify categorisation decisions which may have been less clear.

For the first of these decisions, four participants had responded with the words '*second place'* to the target phrase *'first place'*, which were initially categorised as the participant being misled by co-textual cues. However, the whole phrase was '*stopping people getting sick in the first place*', and it was evident from the co-text that the four participants may have also been misled by perceptual cues from the words '*sick in*'. If this was the case, the responses should be categorised as perceptually and co-textually led. Clarification from the judge was sought and he independently categorised these responses as being a syntactic match. After subsequent discussion, it was agreed by both the researcher and the judge that these responses could only be categorised as being led by co-textual cues as it was not possible to ascertain with certainty additional cognitive processes of the participants.

The second categorisation decision checked by the judge concerned twelve participants who responded with the words '*talking' and* '*taking'* to the target phrase '*we are not **attacking'*** (target word in bold). These participants appeared, at first glance, to have been misled by perceptual cues, however there is not sufficient empirical evidence to support adding weak final syllables or individual phonemes to the framework of analysis as being influential in word recognition. After discussion with the judge, it was agreed that it was not possible to code these responses as perceptual matches.

A further check of coding decisions was undertaken by the researcher in the form of an intra-rater check. A further set of spreadsheets for each recording was prepared with the participant miscue responses transcribed, but not coded. The researcher coded 10% of miscue responses (N=59) a second time and these coding decisions were compared to the decisions previously made. The results of the intra-rater reliability check were that the coding decisions recorded the second time were the same as those initially recorded by the researcher, thereby confirming reliability. The coded data is provided in Appendix K.

3.12. Participants' confidence in their decoding decisions

The data gathered during the verbal reports required the 21 participants to report on their degree of certainty about their decoding decisions after they had responded to each pause in an oral paused transcription task (discussed in Section 3.6.2). Initially, recordings of

participants' oral responses were transcribed on spreadsheets set up in the same manner as for the written task, i.e., with the targets noted in the first column and the responses transcribed below. However, for this task each participant was allocated a separate spreadsheet as this enabled qualitative data to be recorded and coded. Participants' responses were then coded using the same categories as for the written responses, presented in Section 3.10, i.e., '1' for correct, '0' for no response, and an alternative word to the target was transcribed in full. Next, the responses to the researcher's question, "are you sure those words are correct", were transcribed underneath the corresponding words. These responses could generally be interpreted as 'sure' or 'not sure', e.g., participants commonly replied 'sure (or not sure) of all' or 'yes/no'. In some cases, responses were 'not sure of the word after/before X (a specific word they knew). Responses were also coded as to whether participants had correctly or incorrectly responded to the targets. Responses which were not straightforward to interpret, and words to which participants did not respond, were not included in the analysis and were categorised under the row 'not specified'. Therefore, five response types were possible and these are detailed below in Table 3.7:

Response type	Example responses
i) unsure of the accuracy of their incorrect response	unsure, no, not sure of the word after/before X
ii) sure that their incorrect response was accurate	sure, yes, sure of the word after/before X
iii) unsure of the accuracy of their correct response	unsure, no, not sure of the word after/before X
iv) sure that their correct response was accurate	sure, yes, sure of the word after/before X
v) not specified	

Table 3.7: Oral task - coding categories used for participants' degree of certainty about their decoding decisions.

An example of the coded data for one target phrase is illustrated in Table 3.8 below, and the full set of data collated for this task is presented in Appendix L.

Target phrase	**style**	**of**	**particular**	**questions**
Oral response to task	1	1	1	1
P response	sure		no	yes
sure (when response incorrect)				
not sure (when response incorrect)				
sure (when response correct)	1			1
not sure (when response correct)			1	
not specified				

Table 3.8: Oral Task - An example of coded data in respect of the degree of certainty held about decoding decisions.

3.13. Summary

This chapter has provided a detailed account of the design of this study, as well as the rationale behind each decision. It has included descriptions and explanations about the decisions made in the choice of research method, the tasks employed and their design, and in ethical considerations. In the second half, the chapter has presented details of the complex issue of the collation of data, as well as checks on the reliability of the coded data. The additional coding of participants' erroneous responses, the statistical analyses, and the results and discussions of each set of data are presented in the following three chapters.

4. Data processing, results and discussion: lexical cues

4.1. Introduction

This chapter presents the method of data processing related to the lexical features of the target words which contributed to participants' correct responses. It provides the results of the investigation and subsequently discusses them in the context of the aims of the study, i.e., how the they advance knowledge in the field of L2 listening and contribute to the development of L2 listening pedagogy. In doing so, Research Question 1 is addressed:

What word level information contributes to the word recognition of L2 mid-level listeners in a British EAP context?

The results for the written responses to targets in the two recordings, presented in full in Section 3.10, showed that 53.37% of responses to targets in Recording 1 were correct, and 42.40% in recording 2 were correct. The results are presented separately by each recording throughout this chapter as the lexical targets in each recording were different, as were the number of respondents.

4.2. Data processing

Research Question 1 examined the manner in which the characteristics of words informed participants' successful lexical recognition. It was established in Section 2.3 that lexical recognition draws heavily on i) word frequency, ii) word initial stress, and iii) word length. To this end, these three characteristics of the target words were investigated to ascertain the strength of their influence on participant responses. A full discussion of the rationale for selecting these features can be found in Section 2.3.

Word frequency

The first lexical feature of the target words analysed was frequency. Word frequency data were extracted using BNCweb (Lehmann, Hoffman & Schneider2002), a web-based

program for searching and retrieving lexical, grammatical, and textual data from The British National Corpus (BNC).

Given that this research investigated listening, only frequency figures from the spoken data of the BNC were used, which comprises two sub-components, i.e., demographically-sampled and context-governed, and is made up of about 10 million words. This is due to the notion that written forms are stored separately from oral forms in the mental lexicon (Field, 2004b). In addition, only frequency figures relating to the specific sense represented by the target word in the recording were used. This is because recognition of the word may be determined by its co-text as well as its form. For example, the target word 'pretty' was used in the recording in the sense of 'fairly', rather than 'attractive', therefore only the frequency figure for its use as an adverb was used. Finally, the frequency of each target word was calculated for all relevant inflected forms in the case of nouns and verbs (e.g., for the target '*witnessed',* the frequency figures included those for '*witness/es*' (n/pl), '*witness'* (v), '*witnessed'* etc.) and derivational forms in the case of adjectives and adverbs (e.g., for the target '*disastrous'*, the frequency figures included those for '*disastrously'*). The target words used in this analysis are listed in Table 4.1 below, along with their frequency figures.

Recording 1	
Target	**Frequency**
pretest	0
optimum	9
statistically	17
implicit	20
questionnaire	110
valid	128
approval	175
sample	182
flow	213
presentation	291
style	358
response	503
research	642
personal	669
process	854
pretty	1113
market	1443
trouble	1443
particular	2473
important	2886
questions	4876

Recording 2	
Target	**Frequency**
expectancy	20
witnessed	34
disastrous	39
illnesses	51
attacking	62
controlling	75
cure	75
prescriptions	93
results	457
sick	656
drop	713
period	1149
terms	1448
true	1831
place	2988
life	3010
number	5488
same	6842
first	10026
time	19057

Table 4.1: Target words and frequency figures in ascending order according to the spoken section of the British National Corpus by recording.

Data were extracted which related to the total number of correct responses to each target word to enable investigation into whether word frequency explained the variance in this factor, or not.

Word initial primary stress

The second lexical feature of the target words analysed was the impact of word initial stress, primary and secondary, on participants' recognition of the target words. According to

the MSS (Cutler, 1990), L1 listeners of English segment words in the speech stream based on the likelihood that either primary or secondary stressed syllables indicate new content words, whereas weak syllables are unlikely to be word initial. Therefore, the saliency of stressed syllables aids lexical segmentation (Cutler et al., 1997a:172) (see Section 2.4.6 for a full discussion). In order to investigate whether participants may recognise the saliency of stressed syllables and use them as indicators of new words in the same way as L1 listeners, data were analysed to calculate the difference between the number of correct responses to words with and without initial stressed syllables. Monosyllabic words were included in this analysis as they include a stressed vowel and, therefore, indicate the presence of a new word. Data were extracted relating to the total number of correct responses to each of the target words in Table 4.2 below.

Recording 1	
Targets with word initial stress (primary or secondary)	**Targets with non-word initial stress**
optimum	important
market	response
research	implicit
process	approval
pretty	statistically
personal	particular
questions	
trouble	
valid	
sample	
presentation	
questionnaire	
pre-test	
style	
flow	

Recording 2	
Targets with word initial stress (primary or secondary)	**Targets with non-word initial stress**
witnessed	results
illnesses	prescriptions
period	expectancy
number	attacking
cure	controlling
sick	disastrous
drop	
terms	
true	
place	
life	
same	
first	
time	

Table 4.2: Targets used for testing the difference between correct responses and position of stressed syllable.

Word length

Finally, data regarding the effect of word length, i.e., the number of syllables, on word recognition were collated as it is possible that longer words may have been more easily recognised, given that more syllables provide more perceptual evidence in the signal. However, it was argued in Section 2.3.2 that listeners process whole words rather than individual syllables and, as such, the number of syllables in a word may not impact the process of word recognition.

The target words were divided into two groups:

i) those with one or two syllables, and

ii) those with three or more syllables.

Data relating to the total number of correct responses to each of the words were collated to calculate the correlation between the number of correct responses to a word and the number of syllables in the word. The target words in each of the two groups are shown in Table 4.3 below.

Recording 1	
Group 1 (1 or 2 syllable targets)	**Group 2 (3 or more syllable targets)**
flow	optimum
style	important
response	questionnaire
market	implicit
research	approval
process	personal
pretty	presentation
questions	statistically
trouble	particular
pretest	
valid	
sample	

Recording 2	
Group 1 (1 or 2 syllable targets)	**Group 2 (3 or more syllable targets)**
sick	controlling
first	illnesses
place	attacking
true	period
same	prescriptions
time	disastrous
number	expectancy
results	
terms	
life	
witnessed	
drop	
cure	

Table 4.3: Grouping of targets used to test for a correlation between correct responses and number of syllables.

4.3. Results

This line of investigation explored whether the variables word frequency, initial syllable stress, and word length could predict participants' number of correct responses to the targets. Therefore, a multiple regression analysis was conducted. In the tables presented, the convention is used for significance levels where * $p<.05$, ** $p<.01$, *** $p<.001$, **** $p<.0001$.

Recording 1

The results, shown in Table 4.4 below, indicate that there is a moderate positive correlation between percent correct responses and frequency, where $r = 0.433$, n = 21, $p = 0.025$. Furthermore, there was a robust negative correlation between initial syllable stress and word length, where $r = -0.510$, n = 21, $p = 0.009$. This result is not surprising as the more syllables a word has, the greater the chance that the stress is placed on syllables other than the initial syllable. Also, there is a tendency in English for stress to shift rightwards as suffixes are added to a word and it becomes longer. The regression indicated there was no multi-collinearity.

	Percent correct responses	Frequency	Word length	Initial stressed syllable
Percent correct responses		0.433*	0.008	-0.005
Frequency	0.433*		-0.009	-0.022
Word length	0.008	-0.009		-0.510**
Initial stressed syllable	-0.005	-0.022	-0.510**	

Table 4.4: Recording 1 results of correlations between percent correct responses, frequency, word length, and initial stressed syllable.

The results of the regression for recording 1 indicated that when all three independent variables were entered together, there was no significant predictor of the number of correct responses to the target words. Regression Tables can be found in Appendix M.

In order to ascertain the individual contribution of each of the three variables to the number of correct responses to the targets, a stepwise multiple linear regression analysis was conducted. Word frequency was the only variable which explained a significant amount of the variance (14.5%) in the data for number of correct responses, $F(1,19) = 4.391$, $p = 0.050$).

Recording 2

The results, shown in Table 4.5 below, indicate that there is a weaker positive correlation between percent correct responses and frequency than in recording 1, where $r = 0.349$, n = 20, $p = 0.066$. As with recording1, and likely for the same reason, there was a robust negative correlation between initial syllable stress and word length, where $r = -0.663$, n = 20, $p = 0.001$.

	Percent correct responses	Frequency	Word length	Initial stressed syllable
Percent correct responses		0.349	-.060	0.062
Frequency	0.349		-0.368	0.348
Word length	-0.060	-0.368		-0.663**
Initial stressed syllable	0.062	0.348	-0.663	

Table 4.5: Recording 2 results of correlations between percent correct responses and frequency, word length, and initial stressed syllable.

The results of the regression for recording 2 indicated that when all three independent variables were entered, there was no significant predictor of the number of correct responses to the target words.

In contrast to the results for recording 1, the stepwise regression model for recording 2 was unable to predict the variance in the percent correct data when the three variables were added. In addition, no single variable emerged as a significant contributor to the ease with which words were recognised by participants. Regression Tables can be found in Appendix M.

Given that word frequency explained a significant amount of variance in the percent correct data in recording 1, but not in recording 2, an investigation into the differences between the distributions of frequency across recordings was undertaken to ascertain if this impacted the results. The total for each frequency band of one thousand was computed. These are presented in Figure 4.1 below.

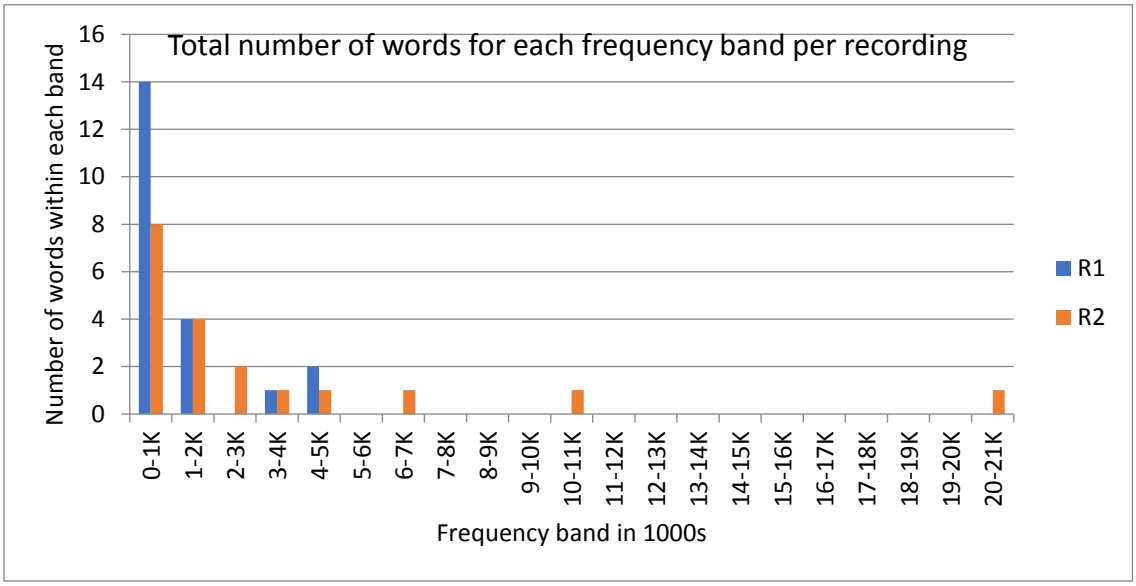


Figure 4.1: Total number of words for each frequency band per recording.

The analysis revealed that the frequency of recording 1 items ranged from 0-5,000, with 64% less than 1,000, and the frequency of recording 2 items ranged from 0-21,000, with 40% less than 1,000. The mean frequency of recording 1 items was 1055.19, and the mean frequency of recording 2 items was 3259.85. Therefore, it could be suggested that the effect of frequency explained a significant amount of variance in the percent correct data in recording 1 because the distribution was much more compact and regular overall, whereas for recording 2 it was much more expansive and less systematic. This illustrates the qualitative difference between the two recordings with respect to their frequency characteristics, and underscores the need for them to be analysed separately.

4.4. Discussion

It was pointed out in Section 2.4.5 that L2 lexical search processes in listening are under researched. Rost (2011:168) suggests, in the absence of evidence to the contrary, that L2 processes are similar to those identified in expert listening, however, given that the basis for competition in word recognition is entirely different in an L2, this may not be the case.

4.4.1. Word frequency

The results presented in Section 4.3 show that there was a relationship between the frequency of the target words in recording 1 and participants' successful word recognition, in that they were more likely to recognise more frequent words. This is unsurprising, given the nature of the listening process. Due to the time pressure of listening, there is a need for rapid mapping from acoustic input to words, and this mapping is likely to be more successful in the case of frequent words because it is more automatic. The finding reported here illustrates a similarity between L1 and L2 word recognition processes. It is generally accepted in L1 research that the frequency of a word is stored in the mind as part of its lexical entry (e.g. Aitchison, 2003; Marslen-Wilson, 1990; Morton, 1979), and so frequent words in spoken and written discourse are more easily processed, resulting in quicker and more successful recognition (Kirsner, 1994; Luce, 1986a).

In addition to automatisation (discussed in Section 2.4.2) being impacted by word frequency, L1 word recognition is linked to an awareness of the relative frequency of the word in speech. Competition in lexical search processes is supported by the frequency effects of the word in speech; more frequent words will be more highly activated than their lower frequency synonyms (Cutler, 2012). However, conclusions regarding this aspect of L2 listening processes cannot be drawn through the analysis reported here as it is not possible to know whether participants knew synonyms for any target words, or if they did, what the synonyms were and their relative frequency.

Responses to one target in recording 1 are surprising, however. In the target phrase '*pretty personal questions*', despite '*pretty'* being fairly frequent in spoken English in the sense of '*quite'* (1113 occurrences in the BNC), only 4% of participants gave correct

responses. This is likely due, to some extent, to the speaker heavily under-articulating the word, as the phrase was spoken fairly rapidly. This suggestion is in line with a proposal by Shockey (2003), who states that higher frequency words are articulated in a reduced form more often than lower frequency words, and this reduces their recognisability. Further support for this suggestion is that more than 81% of responses to this target were blank, and only 1.3% of responses (i.e., one participant) matched the initial stressed syllable. Therefore, participants may not have perceived enough acoustic evidence of the word to attempt to reproduce it.

The percentage of correct responses to several highly frequent targets in recording 2 are also surprising in terms of their low numbers, namely *'life'* (37%), '*same'* (39%), '*first'* (66%), and '*time'* (62%). In fact, '*first'* and '*time'* are the most frequent target words in both recordings. Similarly unexpected was that in the text none of these target words were followed by monosyllabic functors, which, due to their weak articulation, may have been incorrectly attached to the right of target words in the manner suggested in several theories of lexical segmentation (e.g., Cutler & Butterfield, 1992; Cutler et al., 1997a; Grosjean & Gee, 1987b), and thus causing mistaken word recognition. For example, several responses to the target phrase '*sick in the first place*' involved the functor '*in'* being attached to '*sick'*, i.e., '*second'* and '*sicking'*. However, the target words '*same'* and '*time'* appear in the phrase '*the same period of time*', in which it is evident that no weak functors appear after the two words.

It seems likely that the nature of the stimuli may have contributed to this result, as the recordings were sections of a lecture where connected speech features, such as elision, assimilation, and resyllabification, potentially reduce the recognisability of frequent words. It seems likely, therefore, that participants knew these target words, but did not identify them. This suggestion is in line with research by Pemberton (2004), who also required his EAP student participants to transcribe recordings using paused transcription tasks. He found that one in four of the 1000 most frequent English words were not recognised in connected speech after repeated listening, despite listeners knowing their orthographic forms. When considering the similarities between the current study and that of Pemberton (ibid.), it should be borne in mind that the results of the study reported here involved analysis of only content words. However, Pemberton's (ibid.) transcription tasks included function words, which are highly frequent and generally more affected by reduction in connected speech than content words, and therefore his results may have been amplified.

Overall, findings related to word frequency cues suggest that, although L2 listeners may be more likely to recognise more frequent words, the manner of articulation of words in connected speech seems to impact their recognisability. This factor is less likely to impact L1 listeners for two reasons. Firstly, they are more able to bring to bear top-down information, in the form of chunks and syntax, to aid the decoding of words affected by the non-standard phoneme articulation found in connected speech (Field, 2004a), whereas L2 listeners may not have the cognitive capacity to do so due to the pressures of time when listening (see Section 2.4.2). In addition, exposure to a range of phonemic variations, within and between words, that are produced in running speech enables listeners to lay down additional memory traces, and facilitates recognition of the variants in the speech stream when they occur. The more exposure to variants, the more efficient word recognition becomes (Goldinger, 1996; Johnson, 1996). L2 listeners are unlikely to be familiar with as many forms of phonemes in connected speech as L1 listeners, having been largely exposed to the citation forms of phonemes throughout their learning. Thus, they will not have stored memory traces of as many exemplars of the variation of phonemes as experts.

4.4.2. Word initial stress

Results from the analyses of both recordings revealed that word initial stress did not have a significant impact on the percent correct responses to target words. This finding appears to indicate a divergence between L2 and L1 listening processes; expert listeners of English rely heavily on stress in segmentation and, subsequently, word recognition. A word's stress pattern is stored in the mind as part of its lexical entry (see Section 2.3.2), and is highly influential in segmentation processes; over 85% of content words in English continuous speech are either monosyllabic or have initial syllable stress (Cutler, 1991). This expectation leads L1 listeners to segment words in continuous speech based on the likelihood that stressed syllables indicate new content words, whereas weak syllables are unlikely to be word initial (Cutler & Carter, 1987). The implication is that words with initial syllable stress should be easier to segment from the speech stream as they are fulfilling listeners' expectations.

The finding that word initial stress did not have a significant impact on the target words recognisability suggests that one of two types of behaviour may be operating in participants' listening processes. Either they did not recognise the saliency of stressed syllables, and so could not use them as indicators of new words in their lexical search process, or they recognised the saliency of stressed syllables, but were not familiar with the notion that the presence of stressed syllables can be used as an indicator of new words in English. The former suggestion seems somewhat unlikely, given the higher levels of clarity and longer duration of the full vowels in stressed syllables (Cutler, Dahan and van Donselaar1997a:172). In addition, even in reduced speech, prosodic features are likely to be maintained (Shockey, 2003). Therefore, given that English segmentation strategies are not universal across all languages (Laver, 1994), it seems that the latter behaviour is more likely, and L2 learners of a level similar to that of the participants may not be fully attuned to the markers of stress that prevail in English.

This suggestion appears to be in line with research undertaken by Cutler and her colleagues (e.g., Cutler, 2001; Cutler et al., 1997a; Cutler et al., 1986) which focussed on the importance of the rhythmic units in lexical segmentation (see Section 2.4.6). They largely found that listeners apply their L1 segmentation procedures to L2 spoken input and, given that errors in segmentation result in ineffectual lexical search processes, this behaviour impedes word recognition. However, similarities in the findings of the study reported here and the research by Cutler and her colleagues should be treated with caution as the research methods and stimuli were very different. The research by Cutler and her colleagues took place in laboratory conditions and used short segments of input of two to four syllables, some of which was in the form of non-words. Conversely, the study reported here used longer sections of natural speech where cues at all levels were available to aid word recognition and, thus, was more ecologically valid.

On the other hand, the findings contradict those of Field (2001), who explored how listeners insert boundaries between strong and weak syllables, and found that L2 listeners persistently chose the strong/weak pattern favoured by native speakers of English. In addition, research into native and non-native speakers' use of stress as a cue to segmentation by Sanders, Neville and Woldorff (2002), led them to conclude that L2 listeners are able to acquire new segmentation cues as a developmental process. However, the research method used by Sanders et al. (2002) was considerably different from the current study as, in line

with the work of Cutler and her colleagues, their stimuli involved sentences which had been manipulated in laboratory conditions to include non-words. It may, therefore, be imprudent to consider these findings alongside those of the current study.

4.4.3. Word length

Results from the analyses of both recordings revealed that word length did not have a significant impact on the percent correct responses to target words. This indicates that the increased perceptual evidence in the signal of longer words does not result in them being more easily recognisable. It was discussed in Section 2.3.2 that L1 word recognition operates at levels higher than syllabic, as evidenced by gating experiments (e.g., Grosjean, 1985; Shockey, 2003), and the finding reported here indicates that L2 listeners may also process speech at word or chunk level.

The findings add to the limited understanding of the possible effect of word length on L2 word recognition processes as, although much research into L2 word recognition has been conducted, none has been identified which addresses this issue. Also, in an L1 context, many investigations that were identified involved stimuli manipulated in laboratory conditions that were made up of non-words or individual syllables. The study reported here used longer sections of speech, which consequently included features of connected speech, and so the results could be considered more ecologically valid.

4.5. Summary

This chapter reported on the investigation into the lexical features of the target words which contributed to participants' correct responses. The chapter can be summarised as follows:

- Word frequency appeared to impact the word recognition of the participants, suggesting, predictably, more automatic mapping of acoustic input to words for higher frequency words. This operation is also apparent in L1 listening.

- The lack of relationship between initial syllable stress and percent correct responses indicates that participants were not making use of this feature in the same way that L1 listeners do, i.e., as a cue to lexical segmentation and word recognition.

- Word length did not have a significant impact on the percent correct responses to target words, indicating that both L1 and L2 word recognition processes may operate at word or chunk level.

The following chapter presents and discusses the results of the analysis of participants' erroneous responses, or miscues.

5. Data processing, results and discussion: miscues

5.1. Introduction

This chapter explores participants' incorrect responses, referred to as miscues, in order to explore the relationship between the breakdowns which occur in their understanding and the evidence available to them within the speech signal. This evidence is in the form of cues identified in the cognitive framework of analysis (see Appendix A). Specifically, this chapter addresses participants' use of: perceptual cues, and what are loosely referred to as 'contextual' cues; various types of perceptual cue; and word frequency cues, which compensate for misheard words. This chapter shows how the data were analysed, then presents the results of the investigation. Finally, it discusses the results with reference to the aims of the study, as seen in Chapter 1 – i.e., to contribute to the body of knowledge in L2 listening, and to inform the way listening is taught in both general English teaching contexts and on EAP courses

The results for the written responses to targets in the two recordings, see Tables 3.5 and 3.6 in Section 3.10, showed that 17.40% of responses to targets in recording 1 were miscues, and 14.35% of responses in recording 2 were miscues. The mean per participant of miscues in recording 1 was 3.65, and in recording 2 was 2.87. The results are presented separately by recording throughout this chapter, as each recording represents separate sets of phonological stimuli to which each group of participants responded.

5.2. Data processing

Before the data could be coded, it was necessary to establish the categories within which this would take place. The first stage of this process involved a broad classification in terms of whether participants were influenced by perceptual or contextual cues, and the second stage called for sub-categories of types of perceptual cues to be established.

5.2.1. Research Question 2: perceptual and non-perceptual miscues

Do L2 mid-level listeners in a British EAP context rely more on perceptual cues or the types of cue widely referred to as 'contextual' when compensating for words that have not been accurately recognised?

The L2 literature makes a general distinction between perceptual and contextual processing (e.g. Hasan, 2000; Rost, 2011; Vandergrift, 2007), and the division is similar to that made in Field's (2013) framework of analysis. He identifies two broad levels of processing, lower (ibid.:96) (referred to in this research as perceptual), and higher (ibid.:100) (referred to in this research as contextual). The development of these categories for use in the data analysis is outlined in this section.

Perceptual miscues

Perceptual miscues are instances of incorrect word recognition where responses were influenced by auditory cues provided by the speech signal. Examples of participant responses are provided in Table 5.1 below.

Target and its immediate co-text (target in bold)	**Example response**
the results are **disastrous**	"diseases"
Norway for instance has **witnessed** a drop	"business"

Table 5.1: Examples of perceptual miscues used by participants.

Contextual miscues

The criterion for classifying contextual miscues was that the miscue had been influenced by anything that could be described as knowledge of the situation under discussion, and/or of the text so far. Examples of participant responses are provided in Table 5.2 below. The text topic is health in the developed world.

Target and its immediate co-text (target in bold)	Example response
other countries are already doing more … [to reduce deaths from heart disease] Norway for instance has witnessed a **drop**	"drug"
we should be looking at how to prevent heart disease rather than concentrating only on how to **cure** it	"kill it"

Table 5.2: Examples of contextual miscues used by participants.

Co-textual miscues

Two issues became evident as the data were classified into the two broad categories of perceptual and contextual miscues. Firstly, a considerable number of miscue responses were syntactically appropriate within the clause, and it was felt that these could not be described accurately as either perceptual or contextual, as defined in Section 5.2.2. Examples of participant responses are provided in Table 5.3 below:

Immediate co-text (target in bold)	Example response
we have to get the **implicit** approval	**"recent**"
Norway for instance has **witnessed** a drop	**"fixed"**

Table 5.3: Examples of syntactically appropriate miscues used by participants.

Additionally, some miscues were lexically appropriate within the clause, but did not demonstrate use of contextual cues, i.e., knowledge of the situation under discussion, and/or of the text so far. Examples of participant responses are provided in Table 5.4 below:

Immediate co-text (target in bold)	Example response
to make sure that you get **optimum** response	"get **to know"** **"automatic** response"
what does that mean in terms of life **expectancy**	"life **experience**"

Table 5.4: Examples of lexically appropriate miscues used by participants.

Thus, *co-textual* was established as a third broad category of miscue. This category encompassed responses that were either or both syntactically or lexically appropriate within the immediate co-text. The lexically appropriate category of miscue also included lexical co-text in the form of collocation. A small number of miscues fitted this category, and examples of participants' responses are provided in Table 5.5 below:

Immediate co-text (target in bold)	**Example response**
asked some pretty **personal** questions	**"deep** questions" **"good** questions" **"detailed** questions" **"private** questions"

Table 5.5: Examples of miscues used by participants that collocate with co-text.

Types of co-textual miscue

Having established the relative importance of co-textual cues as types of what are often referred to as 'contextual' cues, it is informative to distinguish between types of co-textual miscues. On examination of the data, co-text was separated into syntactically and lexically appropriate miscues. It should be noted that, as with the types of perceptual cues, there is no one to one relationship between a cue and a participant response and any single participant response may show evidence of being syntactically and lexically appropriate. As such, it is not possible to statistically analyse this data.

Summary of broad categories

In summary, the definitions of the three broad categories are:

i) **Perceptual -** participants have attempted to represent the auditory cues in the speech signal either in the form of a similar word or in the form of a non-word (see example in Table 5.1 above).
ii) **Co-textual -** participants have made use of the immediately surrounding words of the utterance; they comprise two types:
    - **Syntactic appropriacy within current clause**: responses that fit the syntax of the clause (see example in Table 5.3 above)
    - **Lexical appropriacy within current clause:** responses where the lexis

used is appropriate within the clause (includes collocations and synonyms) (see example in Table 5.4 above).

iii) **Contextual -** participants have drawn upon their knowledge of the situation under discussion and/or the text so far (see example in Table 5.2 above).

Hybrid miscues

An additional observation, made during the categorisation of data for the investigation of Research Question 2, resulted in a third separate analysis being necessary. A proportion of miscues fell outside the three broad categories, perceptual, co-textual, and contextual, since participants' incorrect responses could be categorised under more than one single category, and these are referred to as 'hybrid miscues'. Thus, a participant may have been influenced by one of these cues or, given the highly interactive nature of listening (e.g., Field, 2013; McQueen & Cutler, 2010), may have been influenced by cues at two or three levels. These data have clearly contributed to the results and provide further elaboration of this research question. Examples of participants' hybrid responses are provided in Table 5.6 below:

Target	**Example response**	**Cues used**
the market research **process**	"the market research **protest"**	stress pattern match, initial syllable match, syntactically appropriate
asking various pretty **personal** questions	**"topical** questions"	stress pattern match, syntactically appropriate
in terms of life **expectancy**	"life **experience"**	stress pattern match, initial syllable match, syntactically appropriate, collocation

Table 5.6: Examples of hybrid perceptual and co-textual miscues used by participants.

In addition, a small number of miscues were influenced by a hybrid of both co-textual and contextual cues, and examples of participants' hybrid responses are provided in Table 5.7 below. The text topic is preparing effective questionnaires.

Target in bold	Example response	Cues used
any doubts about the length of the questionnaire or the style of **particular** questions	"style of **the precise** questions"	syntactically appropriate, relates to knowledge of the current discourse (lecture on designing questionnaires)
any doubts about the length of the questionnaire or the style of **particular** questions	**"specific** questions"	syntactically appropriate, relates to knowledge of the current discourse

Table 5.7: Examples of hybrid co-textual and contextual miscues used by participants.

Furthermore, a small number of miscues were influenced by both perceptual and contextual cues; an example of this in a participant's response is provided in Table 5.8 below. The text topic is health in the developed world.

Target in bold	Example response	Cues used
add to this inactivity an unhealthy diet and the results are **disastrous**	"results are **diseases"**	initial syllable match, stress pattern match, relates to knowledge of the current discourse

Table 5.8: Example of hybrid perceptual and contextual miscues used by participants.

5.2.2. Research Question 3: perceptual miscues

Which of the perceptual cues at syllable and word level are most heavily relied upon by L2 mid-level listeners in a British EAP context when compensating for words that have not been accurately recognised?

Until now, the perpetual miscues have been classified generally as a group. However, this broad classification clearly covers a variety of individual cues. Three perceptual categories were developed inductively by examining the data set to identify patterns of

participants' responses, and also with reference to empirical psycholinguistic literature relating to: i) how lexical information is stored and accessed; and, ii) perceptual salience.

Therefore, these categories were identified by a principled approach. The third research question explores these various types of perceptual cues used by L2 listeners. The criteria for classifying perceptual miscues are as follows:

- **Stress pattern of the target word match**

This refers to whether the stress pattern of the miscue matches the lexical stress of the target word. The literature relating to L1 storage of, and access to, lexical information suggests that a word's stress pattern is one of a number of cues stored in the mind as part of its lexical entry, and provides an important cue to word recognition (e.g., Aitchison, 2003; Grosjean, 1985) (see Section 2.3.2). Therefore, participants' successes at recognising the stress pattern of target words may indicate that they are storing the stress patterns of words in a manner similar to L1 listeners and using this cue to access words, even if a wrong word is recognised.

- **Initial syllable match**

This type of perceptual miscue enables analysis of the extent to which participants can reproduce the initial syllable of the target word. It is widely accepted (e.g. Aitchison, 2003; Bond & Shockey, 2014) that initial syllables serve as indicators of the presence of words and initiate a cohort of matches in the competition for L1 word recognition (e.g., Cutler, 2012; McQueen, 2007), discussed in Section 2.3.2. Therefore, participants' successes at recognising initial syllables of target words are an indication of their word recognition abilities, even if whole word recognition is not accurate.

An additional consideration within this category is the importance of initial stressed syllables, identified in 90% of English content words (Cutler, 1991), which initiate a lexical search (Grosjean & Gee, 1987b). Cutler and Butterfield (1992) refer to this as the MSS (also discussed in Section 2.3.2). In order to explore whether participants made use of initial stressed syllables as a cue, the *initial syllable* miscue was divided into initial *stressed* and initial *unstressed* syllable match. If participants were able to match initial stressed syllables with the target words, this may indicate that they were

relying on them to identify a cohort of words in the competition for word recognition, in the same way as L1 listeners.

It should be noted that Research Question 1 involved exploring successful use of initial stressed syllables in word recognition processes, whereby words were correctly recognised (Chapter 4).

- **Vocalic element of the stressed syllable match**

This type of perceptual miscue enables analysis of the extent to which participants can identify the peak, i.e., the vowel phoneme, of the initial stressed syllable of the target word. The perceptual salience of a word's stressed syllable is due to its loudness, duration, pitch, and quality, and the extended duration is carried largely by the vowel (Roach, 2009), as discussed in Section 2.3.2. As a result, vowels in stressed syllables are much more reliable as perceptual cues, not only because they last longer, but also because they are full, rather than reduced or weak, vowels. Evidence for the influence of stress in L1 word recognition processes is highlighted by slips of the ear data (e.g., Grosjean and Gee, 1987, Bond, 1999, Bond, 1979), which suggest that the vowels of stressed syllables are rarely mistaken.

If participants were able to recognise and match stressed vowels against a background of reduced or weak vowels, this may indicate sensitivity to the prosodic cues of English and, thus, the possibility that these important cues can be used in word recognition processes, even if it leads to a wrong word being recognised.

Examples of participant responses using the three categories of perceptual cue are provided in Table 5.9 below; these clearly show that it was possible that participants' choice of an incorrect word may reflect either one or any combination of perceptual miscue.

Target	Example response	Cues used
statistically **valid** sample	"value"	initial stressed syllable match, stress pattern of the target word match
in order to get **optimum** response	"often"	vocalic element of the stressed syllable match
the style of **particular** questions	"perlitual"	initial unstressed syllable match, vocalic element of the stressed syllable match, stress pattern of the target word match

Table 5.9: Examples of perceptual miscues used by participants.

5.2.3. Research Question 4: word frequency

To what extent do L2 listeners in a British EAP context rely on word frequency cues when compensating for words that have not been accurately decoded?

The data were examined for evidence of actual whole words (as opposed to non-words) which resembled the target. Their frequency levels, according to the BNC, were noted. The goal was to find whether frequency effects might, in some cases, have overruled the perceptual evidence processed by the listener, in other words, whether the word reported was more or less frequent than the target word. Responses influenced by co-text and context were excluded from this analysis due to the focus being at single word level.

Word frequency provides a vital cue to word recognition (e.g., Cutler, 2012; Grosjean, 1980). The literature relating to L1 storage of, and access to, lexical information suggests that knowledge of the frequency of a word is stored in the mind as part of its lexical entry, and the most frequent words are those most easily recognised (Kirsner, 1994; Luce, 1986a). The effect of word frequency on participants' correct responses was examined in Section 4.2, however this line of enquiry investigates whether word frequency cues may have influenced their incorrect responses.

5.2.4. Summary of all categories

The various categories of cue which influenced participants' incorrect responses are illustrated in the model below (Figure 5.1):

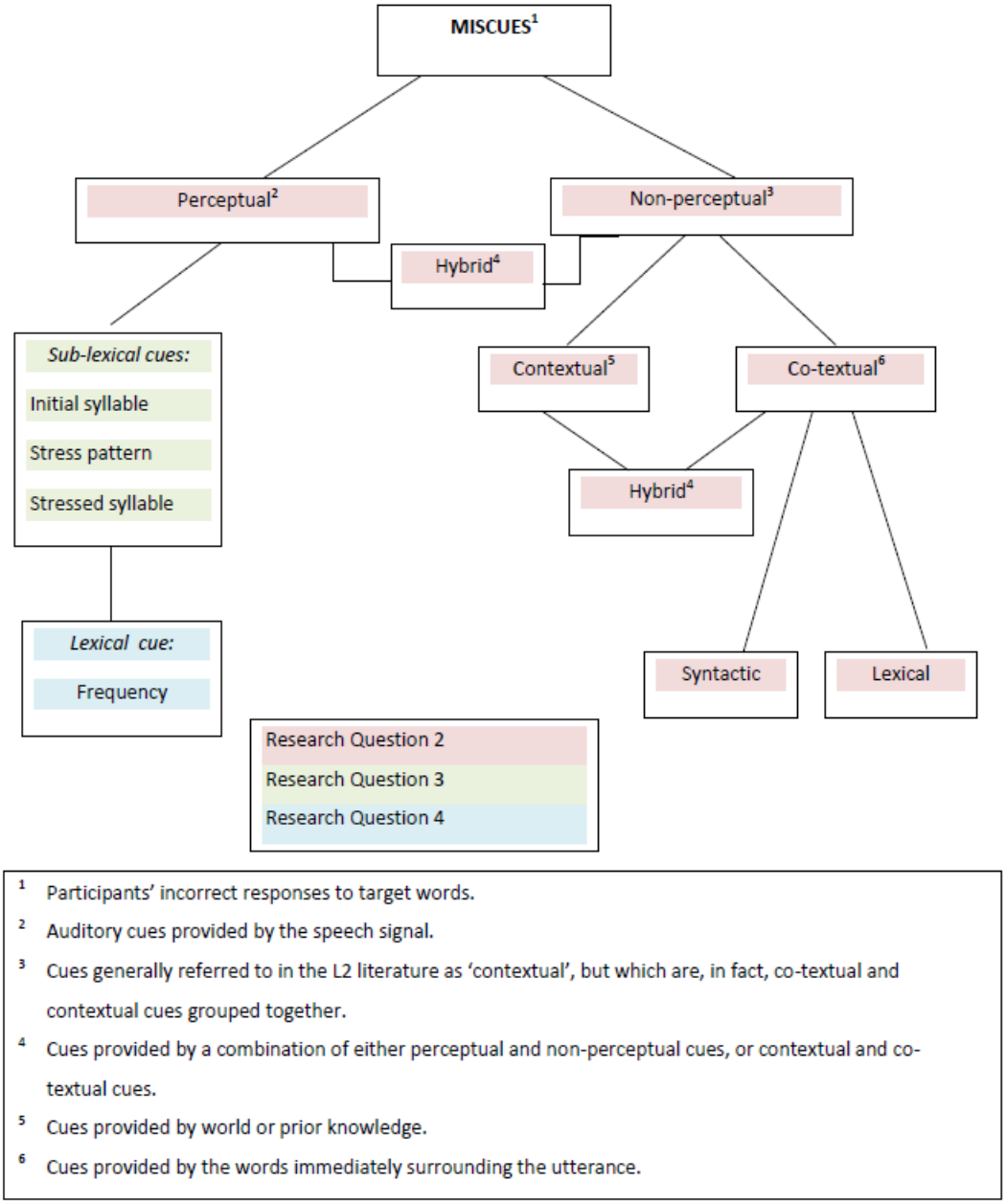


Figure 5.1: Model of categories of cue which influenced participants' incorrect responses.

As a result of this model, eight broad categories of types of miscue were identified; they are summarised and defined in Table 5.10 below, with examples:

Category	Definition	Example target	Example response
1) Perceptual	Attempt to represent phonetic input, either in the form of a similar word or a non-word. Matches either one, or any combination, of: i) initial syllable (separated into stressed and unstressed) ii) vocalic element of the stressed syllable iii) stress pattern of the target word	statistically valid sample in order to get optimum response the style of particular questions	"value" "often" "perlitual"
2) Co-textual	Uses surrounding content of the utterance. Matches either one, or both, of: i) syntactic appropriacy within current clause ii) lexical appropriacy within current clause	to make sure that you get optimum response	"get to know" (syntactically and lexically appropriate)
3) Contextual	Draws upon knowledge of the situation under discussion and/or the text so far.		
4) Perceptual and co-textual	Hybrid response – influenced by both 1 and 2 above	asking various pretty personal questions	"topical questions"
5) Perceptual and contextual	Hybrid response - influenced by both 1 and 3 above	we are curing or controlling many illnesses	"controlling medicine"
6) Co-textual and contextual	Hybrid response - influenced by both 2 and 3 above	the style of particular questions	"specific questions"
Unclassified	Likely to be unsupported guess work	make sure that you get **optimum** response	"**can use** the sports"

Table 5.10: Summary of categories of miscues used by participants.

5.2.5. Unclassified miscues

A number of responses could not be classified and seemed to be the result of unsupported guess work. These were coded as neither perceptual, co-textual, nor contextual; examples of participants' responses are provided in Table 5.11 below:

Target	**Example response**
make sure that you get **optimum** response	**"can use** the sports"
Norway for instance has **witnessed** a drop	"**there** a drug"

Table 5.11: An example of unclassified miscues used by participants.

5.2.6. Categorising the data

In order to enable categorisation of which cues influenced each miscue, and to address Research Questions 2, 3 and 4, each target word was allocated a separate work sheet in the Excel spreadsheet discussed in Section 3.10. Next, the participants' responses to each word were listed in the first column, and each category of miscue was listed in the first row. Therefore, a cell was available which corresponded with each participant and each possible category of miscue. Responses were then coded to identify which category influenced participants' decoding decisions. Finally, the frequency, according to the spoken data of the BNC, of perceptual responses which were whole words was recorded in order to collate data relating to Research Question 4. A sample of fully coded data in respect of the target word '*trouble'* can be seen in Table 5.12 below, and the full set of data for every target word is available in Appendix F.

L1	Pl tst L	P #	*trouble*	about	it	Frequency of target 8826 and response	Initial syll matched	Vocalic component of stressed syllable matched	Stress pattern matched	Syntactic appropriacy within current clause	Collocation with preceding word	Semantic appropriacy within current clause	Synonym	Context: relates to knowledge of text so far	Context: relates to knowledge of situation under discussion	Null	Perceptual match	Co-textual match	Perceptual and co-textual match	Contextual match	Perceptual & contextual match	Co-textual & contextual match	Neither percep, syntactic-lexical nor concep match	Blank response	Correct response
TH	15	55	problem	1	1	28559			1	1	1	1		1							1				
CH	12	22	question	learn	1	25673										1							1		
CH	12	3	tr	ab	1																		1		
CH	11	11	travel	1	1	7221			1								1								
CH	13	30	travel	1	1	7221			1								1								
CH	14	37	travel	0	0	7221			1								1								
CH	13	24	travelled	1	1	2156			1								1								
CH	13	42	traveller	0	0	859			1								1								

Table 5.12: Sample of collated data for the target 'trouble'.
Key: L1 = first language, TH Thai, CH Chinese; Pl tst L = listening placement test score (out of x); P# = participant number; italicised word = target

5.3. Results

5.3.1. Research Question 2: perceptual and non-perceptual miscues

The wider L2 listening research refers to top-down processing variously as, context, semantic knowledge, pre-existing knowledge, world knowledge, or prior knowledge (e.g.Goh, 2002; Hansen & Jensen, 1994; Jensen & Hansen, 1995; Tsui & Fullilove, 1998; Vandergrift, 2002; Voss, 1984). No mention is made of co-textual cues. It is assumed that these studies have either not considered co-textual cues, or that co-textual and contextual cues may have been loosely grouped together and classified as contextual. However, the data in the current research illustrates the need to distinguish between co-textual and contextual processing in order to obtain a more accurate insight into L2 listeners' behaviour, as well as to enable comparison to that of expert listeners. Therefore, the term 'non-perceptual' is used to refer to both co-textual and contextual cues for the remainder of this thesis. In addition, this observation means that the analysis of Research Question 2 should be divided into to two separate analyses. The first is whether listeners rely more on perceptual or non-perceptual (i.e., co-text and context) cues, and the second is whether they rely more on co-textual or contextual cues.

Perceptual or non-perceptual miscues

Firstly, overall instances of the differences between the use of perceptual and non-perceptual cues were calculated. These instances included only miscues which fell into either of the two categories, perceptual and non-perceptual, and so did not include miscues which fell into the hybrid categories where more than one cue was used (Section 5.2.1). As a result of the separation of the hybrid categories, the instances of the use of perceptual and non-perceptual miscues are also expressed as i) a percentage of the **total** miscues, in Table 5.13 below, and ii) a percentage of the **total perceptual and non-perceptual miscues**, in Table 5.14 below. This is in order to provide a clear depiction of the overall instances of their use.

The total of all miscues reported by participants in recording 1 was 281, and in recording 2 the total was 270. In recording 1, 66 of 77 participants used cues in these categories, and in

recording 2, 76 participants did so. The remaining participants used only hybrid categories, and these miscues were not included in this analysis.

	Perceptual and non-perceptual miscues (% of all miscues by recording)	**Perceptual miscues (% of all miscues by recording)**	**Non-perceptual miscues (% of all miscues by recording)**
Recording 1 (Miscues=281) (Participants=66)	162 (57.65)	131 (46.62)	31 (11.03)
Recording 2 (Miscues =270) (Participants=76)	157 (58.15)	105 (38.89)	52 (19.26)

Table 5.13: Overall instances of the use of perceptual and non-perceptual miscues, and percentage of their use compared to all miscues by recording.

	Perceptual miscues (% of all perceptual and non-perceptual miscues by recording)	**Non-perceptual miscues (% of all perceptual and non-perceptual miscues by recording)**
Recording 1 (Perceptual and non-perceptual miscues=162) (Participants=66)	131(80.86)	31 (19.14)
Recording 2 (Perceptual and non-perceptual miscues =157) (Participants=76)	105 (66.88)	52 (33.12)

Table 5.14: Overall instances of the use of perceptual and non-perceptual miscues, and percentage of their use compared to all perceptual and non-perceptual miscues by recording.

The frequencies of miscues were then analysed by participant for each recording, and the use of these cues across participants was compared. Mean percentages per participant are presented in Table 5.15 below.

	Mean perceptual miscues (% of all perceptual and non-perceptual miscues)	**Mean non-perceptual miscues (% of all perceptual and non-perceptual miscues)**
Recording 1 (Perceptual and non-perceptual miscues=162) (Participants=66)	78.42	21.58
Recording 2 (Perceptual and non-perceptual miscues =157) (Participants=76)	66.88	33.12

Table 5.15: Mean percentage per participant of responses using only perceptual and non-perceptual miscues by recording.

It is evident, based on the tables presented above, that the majority of erroneous responses which relied on only one type of cue, i.e., perceptual or non-perceptual, relied on perceptual cues. This difference between reliance on perceptual and non-perceptual cues was explored using paired samples *t* tests, which revealed that the reliance on perceptual miscues was significantly higher than non-perceptual miscues for both recordings (Recording 1: t=8.144, df=65, $p<0.001$, Recording 2: t=3.937, df=75, $p<0.001$) (see Appendix N for *t* test output).

It is also clear that the level of cues used varied considerably across participants. The results were, therefore, analysed to establish how frequently each individual participant responded using perceptual, rather than non-perceptual, cues. The numbers of participants who responded with perceptual miscues are reported below in Figure 5.2, and as a percentage of total perceptual and non-perceptual miscues in Figure 5.3.

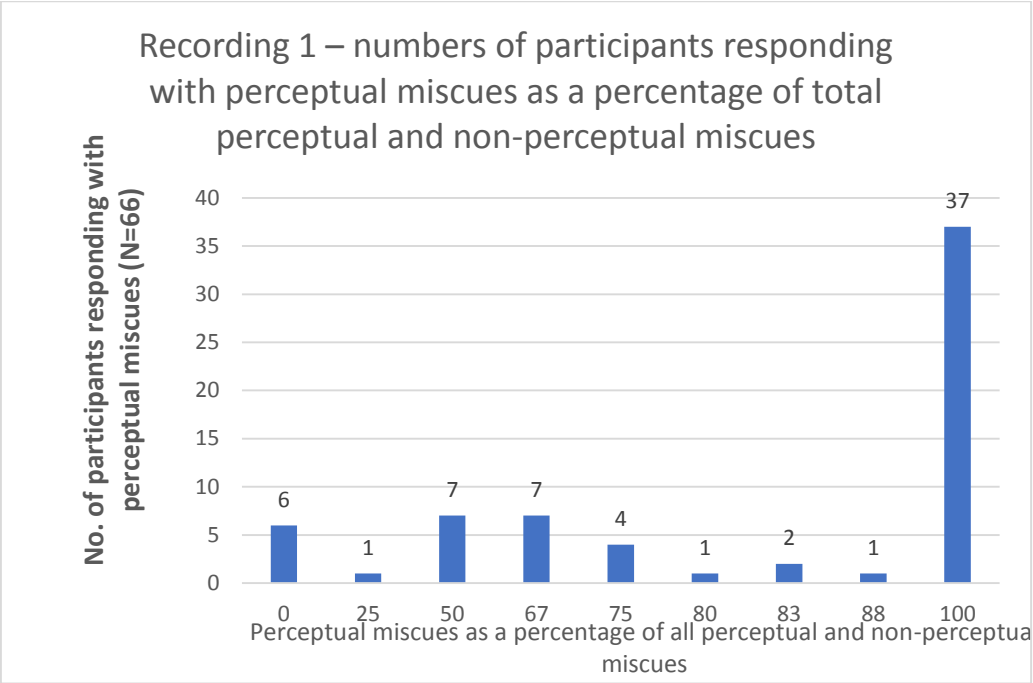


Figure 5.2: Recording 1 – numbers of participants responding with perceptual miscues as a percentage of total perceptual and non-perceptual miscues.

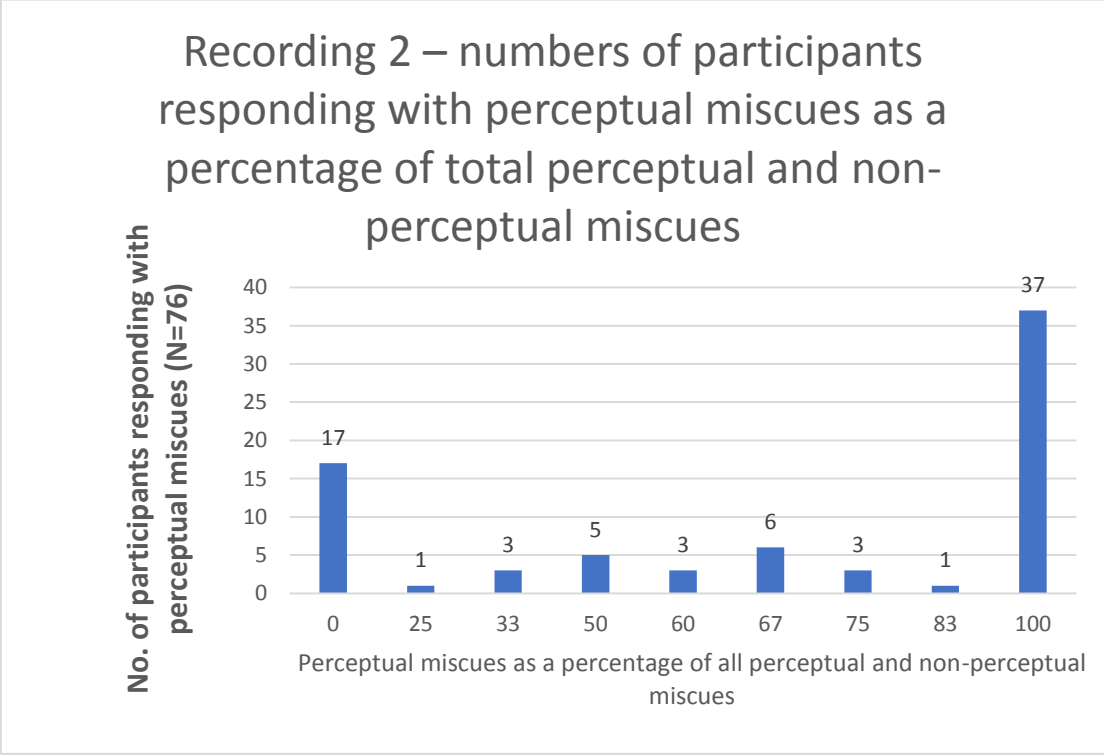


Figure 5.3: Recording 2 – numbers of participants responding with perceptual miscues as a percentage of total perceptual and non-perceptual miscues.

A striking feature of the bar charts is the very high number of participants whose miscues were entirely based upon perceptual information. They totalled 37 out of 66 in Recording 1 (56%), and 37 out of 76 (49%) in Recording 2. Thus, overall, a total of 74 out of 142 participants (52%) showed evidence of sole reliance on perceptual information in their miscues. A small number of participants reported only non-perceptual miscues, and these are represented in the figures as zero.

To enable a statistical comparison to be made, means of the perceptually and non-perceptually influenced responses were then calculated across participants and are presented in Table 5.16 below:

	Mean perceptual miscues (SD)	**Mean non-perceptual miscues (SD)**
Recording 1 (Participants=66)	1.98 (1.34)	0.47 (0.59)
Recording 2 (Participants=76)	1.38 (1.12)	0.68 (1.12)

Table 5.16: Mean per participant of responses using perceptual and non-perceptual miscues by recording.

The figures confirm an overall marked difference across individuals in their reliance on perceptual information as opposed to non-perceptual. However, the relatively high standard deviations indicate a degree of variation amongst participants.

Participants' perceptually influenced responses can generally be classified in two ways. The first type of frequent response was a non-word which was an approximate phonological representation of the target word using English grapheme-phoneme correspondence rules, e.g.,

	Target	**Response**
Recording 1	pre-test	protixt
Recording 1	valid	vallage, vanat, valant
Recording 2	expectancy	enffcience, emfassionsy
Recording 2	cure it	curiot, cuirt, crurite, curise
Recording 2	witness	vistance, vitnic

The second frequent type of response was a word which resembled the target word perceptually, but did not fit non-perceptual evidence in the signal. Several examples illustrate this point. Firstly, a section of Recording 1 refers to the ethics of administering questionnaires in research, and the target phrase, with the target word in bold, is, '*we got into a little bit of* ***trouble*** *about it*'. Of the eight miscues, five were forms of the verb '*travel'*, clearly perceptually similar to '*trouble'*, but not relevant to the topic of the text, and not part of the discourse so far. A further target in Recording 1 referred to, '*the* ***market*** *research process*', and a response was, '***muscles*** *product*'. The spelling of '*muscles'* was accurate, which implies the participant was not merely producing an approximation of the target word, but instead was relying on the perceptual similarities without attending to non-perceptual cues. Finally, two targets in Recording 2, which were selected as they lacked perceptual saliency, elicited interesting responses. The target, '*we are curing, or at least controlling many* ***illnesses***', elicited a response, '*controlling many* ***officers***'; and the target, '*stopping people getting sick in the* ***first place***', elicited a response '*sick in the* ***forest***'. Clearly, the participants providing these examples did not attend to non-perceptual cues. Evidence of this nature was also identified by Field (2008e) in response to a paused transcription task. It could be suggested that the nature of the task meant that participants felt compelled to write something on the response sheet, however, there are a great number of instances of blank responses, which actually suggests this may not be the case.

Perceptual, co-textual, or contextual miscues

The second analysis addressing Research Question 2 divides non-perceptual miscues into co-textual and contextual miscues. As such, it investigates to what extent 'contextual' cues, so-called in L2 literature, are in fact co-textual rather than contextual. Findings from the previous section regarding perceptual cues are carried forward to this section in order to provide a clear picture of the influence of all three types of cue on participants' erroneous responses.

The non-perceptual cues reported in Section 5.2.1 were categorised into co-textual and contextual miscues. Overall frequencies for the differences between the use of perceptual, co-textual, and contextual cues were calculated and are presented in Table 5.16 below. These figures include only miscues which fell into any of the three categories, perceptual, co-textual, and contextual, and so did not include those which fell into the hybrid categories. Due to the separation of the hybrid categories, the frequencies for the use of perceptual, co-textual, and contextual miscues were also expressed as a percentage of the total miscues, this was done in order to provide a clear depiction of the overall frequency of their use. The total of all miscues

reported by participants in Recording 1 was 281, and in Recording 2 the total was 270. In Recording 1, 66 of 77 participants used cues in these categories, and in Recording 2, 76 participants did so. The remaining participants used only hybrid categories, and these miscues were not included in this analysis. This is illustrated in Table 5.21 below.

	Perceptual, co-textual, and contextual miscues (% of all miscues by recording)	**Perceptual miscues (% of all miscues by recording)**	**Co-textual miscues (% of all miscues by recording)**	**Contextual miscues (% of all miscues by recording)**
Recording 1 (Miscues=281) (Participants=66)	162 (57.65)	131 (46.62)	28 (9.96)	3 (1.07)
Recording 2 (Miscues =270) (Participants=76)	157 (58.15)	105 (38.89)	48 (17.78)	4 (1.48)

Table 5.17: Overall frequencies for the use of perceptual, co-textual and contextual miscues and percentage of their use compared to all miscues by recording.

The frequencies of miscues were then analysed by participant for each recording, and the use of these cues across participants was compared. Mean percentages per participant are presented in Table 5.18 below.

	Perceptual miscues (% of all perceptual, co-textual, and contextual miscues)	**Co-textual miscues (% of all perceptual, co-textual, and contextual miscues)**	**Contextual miscues (% of all perceptual, co-textual, and contextual miscues)**
Recording 1 (Miscues=162) (Participants=66)	80.86	17.28	1.86
Recording 2 (Miscues =157) (Participants=76)	66.88	30.57	2.55

Table 5.18: Mean percentage per participant of responses using perceptual, co-textual and contextual miscues by recording.

Means per participant were then compared and are presented in Table 5.18 below:

	Perceptual miscues (SD)	**Co-textual miscues (SD)**	**Contextual miscues (SD)**
Recording 1 (Miscues=281) (Participants=66)	1.98 (1.34)	0.42 (0.58)	0.05 (0.21)
Recording 2 (Miscues =270) (Participants=76)	1.38 (1.12)	0.63 (0.75)	0.05 (0.22)

Table 5.19: Mean per participant of responses using perceptual, co-textual, and contextual categories of miscue by recording.

It is clear, from the tables presented above, that the majority of erroneous responses which relied on only one type of cue, i.e., co-textual or contextual, were largely reliant on co-textual cues. This difference between reliance on co-textual and contextual cues was explored using paired samples *t* tests, which revealed that the reliance on co-textual miscues was significantly higher than on contextual miscues for both recordings (Recording 1: t=4.728, df=65, $p<0.001$, Recording 2: t=6.539, df=75, $p<0.001$) (see Appendix M for *t* test output).

Therefore, the type of cue most relied upon is perceptual, followed by co-textual, and finally contextual, which is the type least relied upon. A further analysis was conducted to establish how frequently each individual participant who used non-perceptual cues responded by using co-textual cues. In Recording 1, out of 29 participants who used non-perceptual cues, 26 used 100% co-textual cues. In Recording 2, out of 39 participants who used non-perceptual cues, 35 used 100% co-textual cues.

Types of co-textual miscue

In order to investigate the proportion of co-textual miscues which were syntactically and lexically appropriate, overall instances of the use of these miscues across participants were calculated. These instances are made up of all miscues that have a co-textual element, and include those which fall into the hybrid categories. In Recording 1, 116 cues were co-textual, and

in Recording 2, 135 were co-textual. Instances are presented in Table 5.20 below. In Recording 1, 50 participants out of 77 used cues in these categories and in Recording 2, 66 participants out of 94 did so. The remaining participants did not use any co-textual miscues.

	Syntactically appropriate (% of total co-textual miscues)	**Lexically appropriate (% of total co-textual miscues)**
Recording 1 (co-textual miscues=116) (Participants=50)	89 (76.72)	27 (23.28)
Recording 2 (co-textual miscues=135) (Participants=66)	100 (74.07)	35 (25.93)

Table 5.20: Overall instances of the use of co-textual miscues across participants, and percentage of their use compared to all miscues by recording.

Results across both recordings indicate that, of the co-textual cues, syntactic cues exert more influence on participants' word recognition processes than lexically appropriate cues.

Hybrid miscues

The relative proportion of the hybrid cues, which fell into four distinct categories, are presented in Table 5.21 below. Due to the separation of the hybrid categories from responses using only one of perceptual, co-textual, or contextual miscues, the frequencies for these cues were also expressed as a percentage of the total miscues in order to provide a clear depiction of the overall frequency of their use. The total of all miscues reported by participants in Recording 1 was 281, and in Recording 2 it was 270. In Recording 1, 37 participants used cues in these categories, and in Recording 2, 55 participants did so. The remaining participants used only non-hybrid categories which were reported above.

	Perceptual and co-textual miscues (% of all miscues by recording)	**Perceptual and contextual miscues (% of all miscues by recording)**	**Co-textual and contextual miscues (% of all miscues by recording)**	**Perceptual, co-textual and contextual miscues (% of all miscues by recording)**
Recording 1 (Miscues=281) (Participants=37)	49 (17.44)	4 (1.42)	2 (0.71)	0
Recording 2 (Miscues =270) (Participants=55	45 (16.67)	1 (0.37)	2 (0.74)	25 (9.26)

Table 5.21: Overall figures for the uses of the mixed categories of miscue, and percentage of their use compared to all miscues by recording.

Next, the relative use made of hybrid miscues by participant was examined. Means per participant were compared and are presented in Table 5.22 below:

	Perceptual and co-textual miscues (SD)	**Perceptual and contextual miscues (SD)**	**Co-textual and contextual miscues (SD)**	**Perceptual, co-textual and contextual miscues (SD)**
Recording 1 (Miscues=281) (Participants=37)	1.32 (0.78)	0.11 (0.39)	0.05 (0.23)	0
Recording 2 (Miscues =270) (Participants=55)	0.82 (0.75)	0.02 (0.13)	0.04 (0.19)	0.45 (0.50)

Table 5.22: Means per participant of responses using hybrid categories of miscue by recording.

It is apparent from the tables presented above that the hybrid miscues used by participants across both recordings appear to be influenced more by perceptual than co-textual

and contextual cues, i.e., perceptual cues plus either co-text or context were more frequent than co-textual plus contextual cues. Furthermore, the cue identified in addition to perceptual miscues was largely co-textual, not contextual. It should also be pointed out that part of co-textual decoding involves perceptual processing as participants are identifying words within the text in order to decode co-textually. This analysis reinforces the evidence presented above concerning the importance of perceptual cues as, even though responses are hybrid, they are still partly perceptual, and only a minimal amount is wholly co-textual/contextual. The very high standard deviations in Table 5.20 show there is some variation between participants.

Responses to two particular targets are worthy of mention as they result in the noticeably high figure for a hybrid of all three categories: perceptual, co-textual, and contextual miscues. Thus, there is a possibility that participants may have been influenced by a cue at a single level, or that they were influenced at two or three levels. The interactive nature of listening (e.g., Field, 2013; McQueen & Cutler, 2010) supports this notion. Recording 2 is on the subject of health, and the speaker discusses aspects which include, illness prevention across various countries, drug use, and life expectancy. In the target phrase '*Norway has witnessed a* ***drop'***, with the target being '***drop'***, 14 participants responded with '*drug'*. In the phrase '*concentrating only on how to* ***cure*** *it*', with the target being '***cure'***,11 participants responded with '*kill'*. It was evident in both these cases that the words reported by a number of participants were perceptually similar, syntactically appropriate, and related to the topic of the lecture.

An interesting point is that, in Recording 1, 48.05% of participants reported responses using hybrid categories, and in Recording 2, 58.51% did so. This provides evidence that a reasonable proportion of participants were listening interactively, i.e., often relying on a range of cues, as do expert listeners.

5.3.2. Research Question 3: perceptual miscues

Section 5.2.2 indicated that a sub-division of the cues generally defined as 'perceptual' gave rise to three distinct sources of perceptual information that appeared to have influenced decisions. Responses showing examples of each are:

i) **stress pattern**: the stress pattern of the word reported matched that of the target word (i.e., '*topical'* was transcribed for the target *personal*)

iia) **initial syllable** (stressed): the target word carried lexical stress on the initial syllable and the word reported matched the stressed syllable (i.e., '*value'* was transcribed for the target *valid*)

iib) **initial syllable** (unstressed): the target word did not carry lexical stress on the initial syllable and the word reported matched the initial unstressed syllable (i.e., '*attaching'* was transcribed for the target *attacking*)

iii) **vocalic element of the stressed syllable**: the vocalic element of the stressed syllable of the word reported matched that of the target word (i.e., '*perlitual'* was transcribed for the target *particular*)

The data showed that the choice of an incorrect word may reflect a compound of types of perceptual cues. That is to say, any single participant response may show evidence of the influence of one or more perceptual cues, as exemplified in Section 5.3.1. Therefore, it is not possible to statistically analyse this data. Nonetheless, descriptive statistics are strongly indicative of some cues being more important than others.

The incidence of each of the four types of perceptual miscue is presented in Table 5.23 below. In addition, the use of each type of perceptual miscue is presented as a percentage of the total perceptual miscues. Perceptual miscues totalled 271 out of a total of 281 miscues in Recording 1, and 221 out of 270 in Recording 2. In Recording 1, 62 participants out of 77 used cues in these categories, and in Recording 2, 75 participants out of 94 did so. The remaining participants did not use any perceptual miscues.

	Initial stressed syllable match (% of all perceptual miscues)	**Initial unstressed syllable match (% of all perceptual miscues)**	**Vocalic element of stressed syllable match (% of all perceptual miscues)**	**Stress pattern match (% of all perceptual miscues)**
Recording 1 (Perceptual miscues=271) (Participants=62)	41 (15.13)	17 (6.28)	96 (35.42)	117 (43.17)
Recording 2 (Perceptual miscues=221) (Participants=75)	14 (6.33)	18 (8.14)	59 (26.70)	125 (56.56)

Table 5.23: Overall instances of types of perceptual miscue, and percentage of their use compared to all perceptual miscues by recording.

The use of each type of perceptual miscue as a percentage of the total perceptual miscues by recording is presented in Figure 5.4 below.

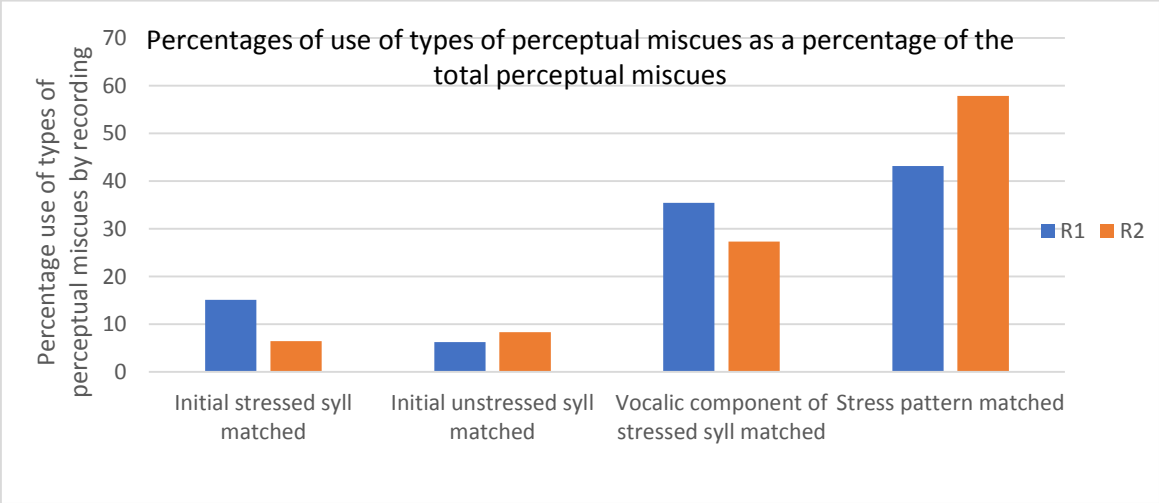


Figure 5.4: Recording 1 and 2 percentages of use of types of perceptual miscues as a percentage of the total perceptual miscues.

Descriptive statistics for both recordings indicated largely similar responses and strongly suggest that stress pattern match played an important role in participants' miscues for both recordings. Interestingly, a further prosodic cue, vocalic component of stressed syllable match, was also influential across both recordings. Looking at both stressed and unstressed syllables overall, it seems that they are less influential, although in Recording 1, initial stressed syllables were more influential than initial unstressed syllables. In Recording 2, however, the results show that initial stressed syllables were slightly less influential than initial unstressed syllables.

Summary of findings

Results extracted from these data can only be purely indicative as one response may have shown evidence of more than one type of perceptual miscue. However, results across both recordings are indicative that, across participants, a major proportion of miscues had a stress pattern which matched that of the target. Furthermore, the second largest proportion of miscues across both recordings identified the vocalic component of the stressed syllables of the target words. Finally, compared with the other perceptual cues, there was not a great deal of evidence that the initial syllable of a word was used as a cue to word recognition.

5.3.3. Research Question 4: word frequency

Descriptive statistics indicated that the number of responses which were more frequent than the target word was higher than those which were less frequent in both recordings. The data showing this are presented in Figure 5.5 below.

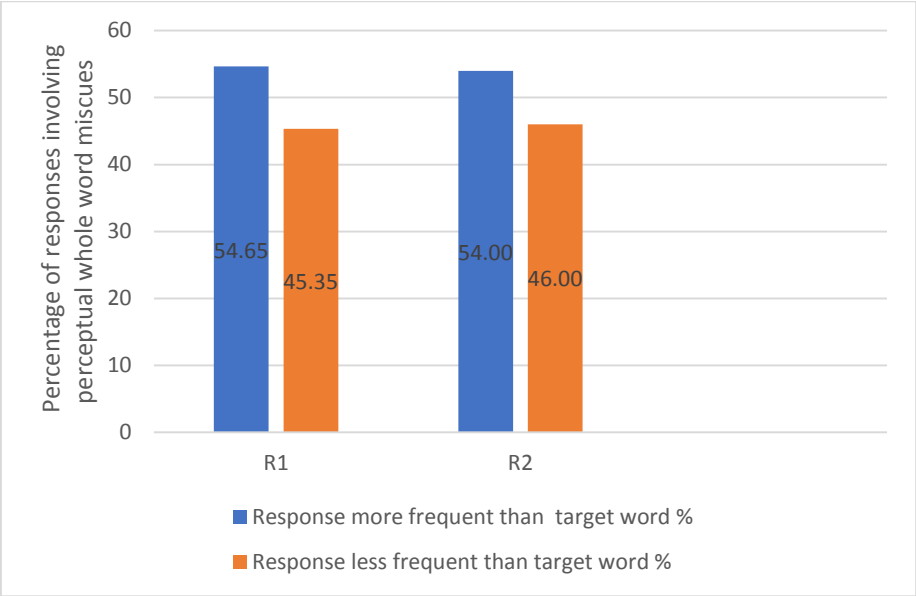


Figure 5.5: Frequency of participants' perceptual miscue responses compared to the target word by recording.

5.3.4. Summary

Section 5.3 has presented the findings of the investigation into the participants' miscue responses relating to Research Questions 2, 3, and 4. These are summarised as follows:

Research Question 2 ***(Section 5.2.1)***

Perceptual and non-perceptual miscues

Participants relied significantly more on perceptual cues than non-perceptual cues when compensating for word level and phrase level items that have not been accurately decoded

Perceptual, co-textual, and contextual miscues

Participants' reliance on contextual cues was very limited, and was greatly exceeded by their reliance on co-textual cues.

Types of co-textual miscue

Within the category of co-textual cues, those that appeared to be influenced by syntactic considerations were much more frequent than those which were influenced by lexical considerations

Hybrid categories of miscue

Some types of miscue responses appeared to reflect the influence of both perceptual and co-textual cues.

Research Question 3 ***(Section 5.2.2)***

Types of perceptual miscues

Three types of miscues were identified. Of the three, the most influential was prosody, primarily at word level, but also, to a lesser degree, at syllable level. Instances of initial syllable matching were few, although stressed initial syllables were slightly more likely to be matched than those which were unstressed.

Research Question 3 ***(Section 5.2.3)***

Perceptual miscues and word frequency

When actual words were reported instead of the targets, frequency effects appeared to account for the substitution in some cases, but by no means in all.

In the next sections, the results from Research Question 2 are discussed in the context of the aims specified for this research.

5.4. Discussion

This section discusses the results in order of the research questions.

5.4.1. Research Question 2: perceptual and non-perceptual miscues

Results presented in Section 5.2.1 show that participants appear to rely heavily upon perceptual cues in recognising words in connected speech, as illustrated by the types of error they make. The results of this study provide strong evidence to contradict the long-held tradition in the L2 listening literature that learners depend highly on, what is referred to very generally as, 'contextual' cues. For example, research by Voss (1984), using transcription tasks, pointed to high dependence on contextual cues, as did research by Long (1990), and Tsui and Fullilove (1998). However, the results of this study suggest that when L2 listeners are confronted with processing difficulties, it is in fact perceptual cues which are relied on to a much greater degree than contextual cues.

It could be suggested that when listening breaks down L2 listeners should *not* be relying on perceptual cues. They may well be aware of not having recognised the word correctly and should therefore also rely upon contextual cues. One possible reason for this 'misplaced' reliance could be related to a lingering effect of language learning. Learners begin by decoding at word level, and it may be that they fail to leave behind this early habitual behaviour as their learning progresses. This could cause them to fixate on the meanings of individual words, a type of behaviour noted by O'Malley, Chamot and Küpper (1989) to be present in less effective L2 listeners.

This argument is supported by the notion that the participants in this study, at approximately level B1 and 2, *should* be able to build meaning from wider context to some degree. According to the CEFR (Council of Europe, 2001:26), found in Appendix D, listeners at B1 level should be able to "understand the main points of clear standard speech on familiar matters regularly encountered", and at B2 level they should be able to "understand extended speech and lectures and follow even complex lines of argument provided the topic is reasonably familiar". This implies that the processes they employ at word level are becoming increasingly reliable and automatic, enabling sufficient cognitive capacity for a degree of meaning building from the wider context. Thus, it is likely that contextual cues are available to the participants,

however, when understanding breaks down they make an instinctive choice to rely on perceptual cues rather than contextual cues.

This choice seems rational, given the pressure of time in listening. After all, perceptual information is immediately available as 'hard evidence' from the speech signal, whereas contextual information must be interpreted and doing so imposes additional cognitive demands. Therefore, in instances where listening breaks down, listeners may be responding to the pressure of time by relying on perceptual cues, which is less challenging than making sure that the perceptually-led matches make complete sense in the current line of argument. The role of time constraints would appear to be supported by Kim (1995), who concludes that learners are able to shift their attention to top-down influences if they are allowed added time in the form of pauses at syntactic boundaries (see Section 2.4.7). Koster's (1987:56) findings also support this notion and he adds that if more processing time exists, top down processing is, in fact, much more influential for L2 listeners than for L1 listeners.

The findings resemble those which emerged from two other studies. Firstly, Field (2004a) conducted several experiments whereby his 31 participants (see Table 2.2) reported the final word of a sentence, or the final word in a group of words, and found that these listeners depended largely on perceptual, rather than on contextual, cues. In addition, Hansen and Jensen (1994) investigated L2 listeners behaviour in an academic context by analysing responses to a lecture comprehension test comprised of both detailed and global questions (see Table 2.2). Their material was of a similar nature to that used in the current study in that they re-recorded authentic academic lectures, the delivery of which, in their original form, were too fast for their listeners to comprehend (in the study reported here, the pilot study used an authentic natural lecture segment which was deemed too fast for participants' comprehensibility (see Section 3.5.1). Hansen and Jensen's (ibid.) findings showed that lower proficiency listeners were less able to answer questions requiring what they term 'top-down processing', despite asking the global questions after a second play of the text. Even though their research method did not accurately replicate real life listening, the similarity with the current study's findings in this area are noteworthy.

Finally, and also in an academic context, it is interesting to note that the findings seem to be in line with Airey and Linder's (2006) investigation of Swedish university students receiving physics instruction in both Swedish and English (see Section 2.4.3). Several examples of students' comments regarding their listening difficulties during English medium lectures

indicated that they focussed at word level, to the detriment of following the overall meaning. Given that these students had been studying English for over 10 years, contextual cues should have been available to them.

The findings also call into question the notion that compensatory strategies in reading are applicable to listening, as the processes of the two skills are frequently regarded as similar. Stanovich (1980:63), in his interactive compensatory model of reading, states that "a deficit in any particular process results in a greater reliance on other knowledge sources, regardless of their level in the processing hierarchy". Referring to a deficit in word level processing in reading, Perfetti and Roth (1980:45) state that "readers of lower skill are … more dependent on context, because of slow executing word level processes" . In relation to listening, this view is referred to by Field (2008c:28) as "the inverse correlation between an adequate interpretation of the signal and the amount of compensatory top-down processing that is brought to bear", and he posits that, in instances where they lack confidence in the accuracy of their decoding decisions, listeners may rely on high level cues (2008d; Field, 2011) (see Section 2.2). However, participants' extensive reliance on perceptual cues in this study suggests this may not be the case. It seems, therefore, that reading processes may differ from listening processes, possibly due to the differences between the two skills in terms of time constraints. Listening takes place in real time, however readers have time to engage in cognitively demanding compensatory strategies based on context.

The findings generate a further issue for discussion which relates to individual styles of listening. Calculations of means per participant showed that figures for perceptual miscues were noticeably higher than those for non-perceptual miscues. However, it was evident that there was a slight variation in behaviour between participants. Although a significant proportion of participants ignored non-perceptual cues and preferred a word with an approximate phonological match, despite the word not fitting the context of the text, or preferred a non-word with an approximate phonological match, there were still a small number of participants who ignored perceptual cues in what may have been an attempt to fit a word into the discourse representation they had built so far. This finding is in line with the views of Field (2008b) and Graham et al. (2008), who, when discussing strategic listening (see Section 2.4.11), posit that listeners' individual styles impact their listening behaviour. Thus, although there were significantly more perceptual than non-perceptual miscues, there are clearly some individual listeners who are more or less likely to rely on one type of cue than the other.

Perceptual, co-textual and contextual miscues

During classification of miscues, two crucial, related factors became evident regarding use of the term 'context' across the wider L2 literature. Firstly, L2 commentators use a variety of terms to refer to context, often interchangeably, and with varying definitions; and secondly, within their notion of 'context' they seem to include localised co-textual cues.

With regards to the first point, as noted in Section 5.3.1 the wider L2 listening research (with notable exceptions as discussed in Section 2.4.3) refers to top-down processing using various terms, e.g., context, semantic / pre-existing / world / or prior knowledge. Some intend to refer to these specific aspects, but many allow these categories to blur into what is loosely termed 'context', or use them interchangeably. For example Jensen and Hansen (1995:101) refer to expert listeners' use of "world knowledge … and background knowledge which is organized around scripts and schemata", and appear to associate this with the notion that "second language listeners use top-down processing". Tsui and Fullilove (1998:449) use several of the terms mentioned above interchangeably. For example, they refer to 'top-down processing' in a study examining only use of schema, and conclude that L2 listeners "need to learn to become less reliant on guessing from contextual *or* prior knowledge" (emphasis added], however they neither define nor differentiate between the two terms.

Turning to the second point, it seems that, generally, this 'catch all' notion of context subsumes cues provided by the co-text surrounding an utterance. However, it was evident from the data collected in this study that some responses appeared to draw upon local information in the input, i.e. syntactic and semantic cues within the immediately surrounding text, while others appeared to be related to general impressions of the topic or world knowledge, i.e. context. Hence, although in the initial investigation of Research Question two, the term 'non-perceptual' was coined to encompass both contextual and co-textual miscues, it was evident, from the results presented in Section 5.1.2, which shows that co-textual miscues were relied upon a great deal more the contextual miscues, that participants' use of co-textual miscues necessitated further investigation.

This notion of the existence of co-text and context is in line with psycholinguistic views of expert listening, notably the framework of analysis for this study (Field, 2013), but the distinction seems to have generally been overlooked in L2 listening research. Exceptions have been identified, in that a small number of researchers do indeed refer to co-text in their L2

studies, (e.g., Field, 2004a, 2008c, 2008d; Kim, 1995; Koster, 1987), and these are referred to in Chapter 0. Given that less proficient L2 listeners may not be able to use contextual cues to support or refute their decoding decisions due to pressure of time, and may lose track of the incoming words while doing so, it seems wiser for them to rely upon co-text, a cue that is phonologically present in the speech signal, and so less cognitively challenging to make use of. Indeed, processing in this manner is an integral part of parsing in expert listening processes, where incoming words, i.e., co-textual cues, are held in the mind until a syntactic structure can be imposed (Jarvella, 1971), and used as a cue to word recognition.

Not only have the findings of this study indicated that context is often referred to in terms that are too general, and rather loosely by many L2 researchers, but the results indicate that participants make significantly *more* use of co-textual than contextual cues, with use of specifically contextual miscues being extremely limited. This suggests that a substantial amount of L2 research which refers to top down and bottom up processing, and defines top down processing as use of context, prior knowledge etc., may indeed be misrepresenting L2 listening processes. Instead, it seems that co-text is a more highly influential cue to listening than context. In turn, this has critical implications in the manner listening is traditionally taught and materials are designed, and this is an issue discussed in Chapter 7.

Co-textual miscues

It was established in 0 5.2.1 that a major proportion of miscues were co-textually appropriate and this section discusses the results, see Tables 5.3 and 5.4, relating to the two types of co-text identified, namely syntactic and lexical. It should be noted that there is no one to one relationship between a co-textual cue and a participant response and any single participant response may show evidence of being syntactically and lexically appropriate.

Syntactically appropriate miscues

The results show that, of the co-textual miscues identified, instances of syntactically appropriate miscues were much higher than lexically appropriate miscues (76.72% of total co-textual miscues in Recording 1, and 74.07% of total co-textual miscues in Recording 2).

In addition to providing evidence of co-textual cue use, this finding reinforces the conclusion regarding the major influence of perceptual cues when listening breaks down. In order to make a match which is co-textually accurate, participants must have also decoded elements within the surrounding clause. The evidence that participants, in attempting to make a

lexical match, are using syntax as a cue, reflects psycholinguistic research. For instance, Grosjean and Gee (1987b:145-146) report that syntax is used to aid lexical segmentation, an element of word recognition (see Section 2.4.6). In addition, from a different perspective, Boland and Cutler (1996:314) state that "there is abundant evidence that syntactic decisions make use of detailed lexical information that is accessed as part of word recognition". The findings clearly illustrate the interactivity of the two types of information in L2 listening.

It can be assumed from the above that participants' ability to use syntactic evidence in the signal, as well as perceptual evidence, improves their rate of word recognition. This assumption is supported by Koster (1987), who investigated the use of co-text in lexical recognition and found that the lexical recognition of intermediate non-native listeners improved when co-textual cues were available (see Section 2.4.8).

Lexically appropriate miscues

Only a quarter of co-textual miscues identified were influenced by lexical co-text, illustrating that syntax is a far more constraining category of co-text. This reflects Bond's (1999:133) L1 listeners' slips of the ear data set, which illustrates that listeners are not constrained by lexical plausibility, but are constrained by syntax (see Section 2.3.4). Thus, it seems that, in terms of co-text, participants' behaviours are in line with that of L1 listeners, perhaps on the basis of dealing with the same issues in their own L1.

Of lexically appropriate miscues, few responses collocated with the surrounding co-text. This finding supports the view of Ellis et al. (2008), who report that L2 learners face enormous challenges acquiring a native-like range of collocations (see Section 2.4.5). Given that L1 vocabulary recognition draws upon a complex network to aid the lexical search (Aitchison, 2003), including frequent collocations, this is likely to impede the efficiency of participants' listening processes. This is especially concerning considering the academic context of the study reported here as, according to Conrad and Biber (2005), approximately 20% of academic prose is made up of three and four word collocations (see Section 2.3.2). Thus, the findings further suggest that academic L2 listeners of a similar proficiency to those in this study will face a higher cognitive load by decoding individual words rather than chunks.

An additional negative impact of difficulties with collocation on academic listening relates to the nature of lecture listening. It was discussed in Section 2.4.10 that comprehenders need to construct a mental representation of incoming information, and either map that

information onto current representations, or build new representations (Gernsbacher, 2013). In lectures, for instance, listeners need to identify a hierarchy of points and sub points in an argument. Less skilled listeners are more likely to face difficulties with this, and consequently, to view each point as being separate (Field, 2011). In lectures, this hierarchy of points and sub points is commonly indicated by chunks of frequently co-occurring words, such as '*on the other hand*', or '*in contrast*' (Conrad & Biber, 2005), and so it could be suggested from the results that L2 listeners' difficulties in recognising common collocations may impact their ability to recognise the relationship between points and sub points, and to build an accurate discourse representation.

An additional type of collocation typical in academic prose is referred to by Conrad and Biber (2005:66) as "specification of attributes", such as '*in the case of*' and '*the extent to which*'. An example of such a collocation in the study reported here is the target '*in terms of…'*, and it is alarming to note that only 9.6% of participants correctly identified this phrase. Given that the function of collocations in this "specification of attributes" class in the context of a lecture is "the conveying of precise information" (ibid.:69), it is clearly questionable how much of this information less skilled listeners may be able to comprehend.

Although evidence of the use of collocation was relatively rare, there were instances where a collocational link was established regressively despite the fact that the collocate did not resemble the target word. This illustrates that some participants were capable of a degree of regressive decoding, a behaviour noted in expert listeners (discussed in Section 2.3.3), whereby listeners may need to regress and reanalyse the speech stream as further cues become available (Shockey, 2003).

Hybrid categories of miscue

A striking characteristic of the data was that a considerable proportion of miscues fell outside the three broad categories (perceptual, co-textual, and contextual) as there were instances where more than one single category contributed to participants' decoding decisions. The results, presented in Section 5.3.1, are discussed here as a further elaboration of the second research question.

By far the largest part of participants' hybrid miscues had a perceptual component, combined with either a co-textual or a contextual one. The frequency of wholly co-textual/contextual was minimal. This finding is especially interesting for several reasons. Firstly,

the component mainly used in addition to perceptual miscues is co-textual rather than contextual. Thus, it would seem that the choice of a perceptually influenced response is sometimes strengthened by ensuring its goodness of fit within the current context or co-text. This finding underscores the suggestion above (see Section 5.4.1) that cues referred to in the wider L2 literature as 'contextual' are, in fact, largely co-textual. In order to use incoming co-text in listening processes, it is necessary to recognise words within the utterance, thus participants also decode perceptually. This reinforces the evidence of the importance of perceptual cues; even though responses are hybrid, they are still largely founded on the listener's ability to process perceptually. For example, in the recording about health in the developed world, responses to the target '*the number of* ***prescriptions***', included 'the number of *questions*' and '*the number of descriptions*'. In addition, the target '*market research* ***process*****',** elicited several responses of '*market research* ***projects***'. Whilst these are clearly perceptual matches, participants have also decoded the previous clause and made a match which is co-textually appropriate.

These findings further contradict the customary view, discussed in Section 2.4.3, that listeners rely heavily on top-down processing as compensation for perception difficulties. This notion derives from the work of several commentators (e.g., Rost, 2011; Vandergrift, 1997) who appear to have conflated co-textual cues with contextual. It seems that higher level processes are, to some extent, relied upon, but in the form of *co-text* not *context*, and it is sometimes relied upon *in conjunction with* perceptual, or bottom up, processing rather than as an alternative.

Finally, this finding is particularly interesting when comparing L2 listeners' behaviour to that of expert listeners; the extent to which those L2 listeners used hybrid miscues illustrates behaviour similar to that of expert listeners. Psycholinguistic research (e.g. Clark & Clark, 1977; Field, 2004b; Lieberman, 1967; McQueen & Cutler, 2010) views listening as highly interactive; listeners make use of multiple cues at various levels simultaneously. Evidence from the hybrid responses could suggest that even in an L2, listeners are capable of using multiple cues simultaneously.

5.4.2. Research Question 3: perceptual miscues

The findings discussed in Section 5.2.1 illustrated that participants are highly dependent on perceptual cues when listening breaks down. Due to the strength of this result, it was of interest to investigate the various types of participants' perceptual miscues to identify which they used most frequently as they attempted to resolve their listening breakdowns. The cues

investigated were lexical stress pattern, initial syllable (stressed and unstressed), and the vocalic element of the stressed syllable.

Lexical stress pattern

Although these results are purely indicative, bearing in mind that one response may have shown evidence of more than one perceptual miscue, it can be said, across the participants, that a large proportion (43.17% in Recording 1 and 56.56% in Recording 2) of the perceptual miscues identified matched the stress pattern of the target word. This finding supports the view that prosody is a major element within perception (e.g., Aitchison, 2003; Bond & Small, 1983; Cutler, 1984). Aitchison (2003:142) reports that the relationship between syllables and rhythmic patterns is "specified in the mental lexicon", and Bond (1999) concurs. Referring to evidence from her slips of the ear data set, Bond suggests that part of the representation of a word stored in the mind is its stress pattern, and this supports lexical retrieval (see Section 2.3.2). This view is supported by the notion that, in cases of malapropism, if the number of syllables is maintained, so is rhythmic structure (ibid.). This finding suggests that L2 listeners store some of the same information in their mental lexicon as experts.

The finding illustrates that participants were sensitive to the stress patterns of English and made use of prosodic elements of perceptual miscues; this behaviour is markedly similar to that of expert listeners. This is particularly noteworthy as the majority of participants were Chinese, and their L1 does not mark lexical stress in the same way as English. Whereas the metrical structure of English uses "all four dimensions of speech", i.e., pitch, loudness, duration, and quality, (Laver, 1994:512), Mandarin Chinese makes use only of pitch, with "each syllable associated with a characteristic pitch value". This is referred to by Laver as "syllable based use of lexical tone" (ibid.:465). Given that the peak of any strong syllable is a vowel phoneme due to its duration and quality (Roach, 2009:94), neither of which are dimensions of speech used in Mandarin Chinese, it could be assumed that these listeners have acquired prosodic knowledge of English incidentally throughout their language learning and recognise the cue it provides.

Initial syllable match

A words' initial syllable is one of several cues stored in the mind as part of its lexical entry which informs competition in lexical search processes (Grosjean & Gee, 1987a). Inaccurate matching of a word's initial syllable is likely to result in an inaccurate set of candidates being activated, which will hinder accurate word recognition. Viewing initial syllable

matches of both stressed and unstressed initial syllables together, findings indicate that, of the perceptual miscues identified, only a small proportion (21.41% in Recording 1 and 14.42% in Recording 2) matched the initial syllable of the target word. Although the findings presented in the previous subsection show that there are many instances of the lexical stress of the target word being reproduced in the reported word, i.e., prosodic matching, this indicates that there is less success in phonological matching. This finding is surprising, given that the onset of a word is usually more reliable as assimilation effects are more often regressive than progressive (Cruttenden, 2014:308), that is to say, word onsets alter in connected speech less often than word offsets (see Section 2.4.4).

In addition, this finding runs counter to Field's (2004a) results, which found that non-native listeners, like native listeners, seem to view word onsets as more reliable than their offsets. However, participants in the aforementioned study were reporting on individual words from a list, whereas in the current study they were reporting on phrases within a lecture. Hence, it could be suggested that the targets in the current study were more challenging to decode as they were subject to features of connected speech and words needed to be segmented from the speech stream. In addition, the study reported here is based on a much larger data set than are Field's results.

Initial stressed syllable match

Separating the initial stressed syllable statistics to examine initial stressed syllable matches provides mixed results. In Recording 1, initial stressed syllable matches, i.e., syllables which it is widely believed are likely to indicate the presence of a new word (e.g., Cutler & Butterfield, 1992; Cutler & Norris, 1988), were more frequent (15.13% of all perceptual miscues) than initial unstressed syllable matches (6.28% of all perceptual miscues). In contrast to Recording 1, there was no evidence to suggest that participants were more likely to match stressed initial syllables than those which were unstressed in Recording 2, with the figures for stressed initial syllable matches (6.33% of all perceptual miscues) being slightly lower than unstressed initial syllable matches (8.14% of all perceptual miscues).

Research into L1 listening, (e.g. Aitchison, 2003; Cutler, 1984; Grosjean & Gee, 1987b), discussed in Section 2.3.2, suggests that the prominence of stressed syllables aids their decoding and is an important cue to word recognition in expert listening. However, the results reported here do not provide compelling evidence that stressed syllables are influential in L2 word recognition processes.

As well as appearing to under-use this phonological cue to word recognition, a further consideration, in the context of the current study, is the impact that misidentification of initial syllables has on a scale larger than single word recognition. It was discussed in Section 2.3.2 that competition in a lexical search may be initiated by several factors, including stressed syllables. Grosjean and Gee (1987a:144) suggest that "a series of cohorts are activated where the stressed syllable in question is the first syllable of a subset of candidates", a notion supported by Cutler and Clifton (1984) and Gow and Gordon (1993). If listeners fail to match the initial syllable, the set of candidates activated will be inaccurate and lexical recognition will fail. This issue is exemplified by response to the target '*illnesses'* in Recording 2. Only 3.2% of participants (i.e., 3 participants) accurately transcribed this target, and, of the 49% of miscue responses (i.e., 46 participants) none accurately reported the initial stressed syllable. Thus, in all cases, an inaccurate cohort of possible lexical matches was likely activated, and likely led to confusion and breakdown in understanding of the utterance. This could also cause an inaccurate discourse representation to be carried forward, impacting more global understanding, a factor which, in an academic context, could seriously impact lecture comprehension, discussed in Section 2.4.10.

Although initial syllable matches were limited, it should be pointed out that this finding refers to the matching of the *whole* syllable, for example the miscue response '***val****ue'* to the target '***val****id'*. Hence, it is possible that participants may have recognised the presence of a word initial stressed syllable and, consequently, its likely indication of a new word, but simply not have matched the syllable accurately. For instance, of the 25 (26.6%) incorrect responses to the target word with initial syllable stress, '*witnessed'*, in Recording 2, only five (5.32%) matched the whole initial syllable. However, 15 (15.96%) matched the vocalic element of the stressed syllable. In addition, and reflecting the finding showing the successful matching of lexical stress discussed above, 16 responses (17.02%) to this target matched the lexical stress. Thus, despite only a small proportion of participants responding with an accurate whole initial stressed syllable, a much larger proportion of participants had, in fact, matched the lexical stress of the target word and segmented correctly. Given that a major proportion of the identified perceptual miscues matched the stress pattern of the target word, discussed above, it seems that the findings regarding lexical stress, vocalic component of the stressed syllable match, and initial syllable match should be considered together as all indicate participants' accuracy in segmentation. Nevertheless, accurate segmentation is only one stage in word recognition processes, and the extent to which stressed initial syllables were not matched is alarmingly low.

Initial unstressed syllable match

The findings in respect of initial weak syllable matches provide evidence of them largely being ignored or displaced. This is a result entirely expected given that weak vowels are less prominent in terms of saliency and duration, and are greatly reduced in connected speech (Roach, 2009). In addition, unstressed syllables do not generally indicate the presence of a new word (e.g. Cutler & Butterfield, 1992; Cutler & Norris, 1988) (Section 2.3.3). This finding is illustrated by several examples given in Table 5.24 below:

Targets with initial weak syllables (includes total number of miscues)	**% of total miscues where initial weak syllables were ignored (and examples)**	**% of total miscues where initial weak syllables were displaced or ignored (and examples)**
optimum **response** (23)	56.52 (sport, spot, spouse)	39.13 (the sports, you stop, to stop)
implicit approval (37)	43.24 (presses, plenty, recent)	0
implicit **approval** (35)	31.43 (prove, proval, prover)	17.14 (the prove, of proval, the prove)
we are not **attacking** (36)	66.67 (talking, typing, taken)	8.33 (the taking, ad taking)

Table 5.24: The percentage of initial weak syllables being ignored or displaced and examples.

This supports the view of Stemberger (1990:148), who states that an error at either of the levels of lexical stress or segmentation "automatically [leads] to an error at the other", hence, two breakdowns in decoding result from one error in perception. In these examples, although participants segmented incorrectly after listening had broken down, they were, in fact, behaving in a manner associated with expert listeners' strategies for segmentation, namely the MSS (e.g., Cutler & Butterfield, 1992; Cutler & Norris, 1988). This strategy refers to what has been demonstrated to be a crucial factor in lexical segmentation in English, i.e., the high frequency of content words that are monosyllabic or bear word-initial lexical stress (see Section 2.3.3). Thus, listeners recognise stressed initial syllables as likely indicators of new words. For example,

participants reported the two syllables of the target word 'response' as separate words, with the first unstressed syllable as a monosyllabic function word, and the second, as an approximation of the syllable. Furthermore, in the phrase 'we are not **attacking**', of 36 miscues (38.3% of all perceptual miscues), only three (3.19%) identified the schwa as the initial syllable of a three-syllable word, for instance they reported 'detecting', and 'protecting'. Responses to these targets indicate entirely expected behaviour as function words are most frequently pronounced in weak forms (Roach, 2009), and it is stressed syllables that generally indicate a new word.

Vocalic element of the stressed syllable

Although participants were only minimally successful at matching *whole* initial stressed syllables, they were far more successful at matching the perceptually salient vocalic element of the stressed syllable. Of the perceptual miscues identified, a sizeable proportion (35.42% of all perceptual miscues in Recording 1 and 26.70% in Recording 2 (see Table 5.23)) matched the vocalic element of the stressed syllable of the target word.

Participants' success at matching the vocalic element of the stressed syllable is unsurprising, given the prominence of stressed vowels discussed in Section 2.3.2. Vowel phonemes in strong syllables are articulated with more precision and with a longer duration than vowels in weak syllables, and are therefore more reliable (Roach, 2009). It seems that participants are sensitive to this reliability, and are perhaps using it as a strategy when they lack confidence in their decoding decisions. In addition, this evidence supports L1 research which finds that the vowels of stressed syllables are rarely mistaken in slips of the ear (Bond, 1979; Bond, 1999; Grosjean & Gee, 1987a) (see Section 2.3.2). Thus, it seems that the participants in this study are behaving in the same way as Bond and Small's (1983:473) native speaker participants, and "may consider stressed vowels as reliable phonetic information". It is also possible that participants are using duration, as well as stress, as a cue, a behaviour similar to that of expert listeners (Smith, Cutler, Butterfield, & Nimmo-Smith 1989).

Summary of perceptual cue discussion

The findings regarding perceptual miscues indicate three types of behaviour:

i) a high proportion of participants were sensitive to the prosody of whole words;
ii) participants were not generally successful at matching initial syllables, either stressed or unstressed; and,
iii) participants were fairly successful at matching stressed vowels.

This suggests that participants were successful at prosodic matching, but less successful at phonological matching. This finding should be viewed alongside one aspect of Research Question 1, namely the investigation into the effect of initial stressed syllables on participants' correct responses, discussed in Section 4.4.2. In both recordings, initial stressed syllables did not have a significant positive effect on the percent correct responses to the target words. Given that the prominence of stressed syllables is an important cue to word recognition in expert listening (e.g., Aitchison, 2003; Grosjean & Gee, 1987b), discussed in Section 2.3.2, it would be expected that words with initial syllable stress should be easier to recognise. The findings, in respect of both correct responses and miscues, suggest that L2 listeners, who are of a similar proficiency to the participants in the study reported here, are not making use of this cue in the same way as expert listeners. However, the evidence that the participants appear to be storing a word's stress pattern in the mind and using the pattern in lexical access processes implies that they are undertaking similar behaviour to that of L1 listeners.

5.4.3. Research Question 4: perceptual miscues and word frequency

The findings presented in Section 5.2.3 indicate that the participants' whole word responses which resembled the target word perceptually were likely to be of higher frequency (according to the BNC) than the target word. This is further evidence that L2 listening processes may be similar to expert listening, as L1 research shows that information about the frequency of a word is one of several cues stored as part of its lexical entry, and the most frequent words are those most easily processed (Morton, 1979) . Thus, expert listeners are known to recognise high frequency words more quickly than those of low frequency as increased exposure to traces of words (see Section 2.3.2) enables listeners to build an awareness of their relative frequency (e.g., Kirsner, 1994; Luce, 1986a) . In an L2 context, the participants in this study seem to have recognised probabilities of the occurrence of words, and so it is more frequent words that they prefer. This supports the suggestion that L2 listeners store some of the same information in their mental lexicon as experts. Given that L2 listeners operate in conditions of uncertainty and under pressure of time, as discussed in Section 2.4.2, it seems understandable that their miscues are mainly perceptual and influenced by frequency.

Findings in respect of one aspect of Research Question 1, namely the investigation into the effect of frequency on participants' correct responses (Section 4.3), support the proposal that higher frequency words are accessed more easily by L2 listeners. In Recording 1, frequency had a significant positive effect on the percent correct responses to the target words.

5.5. Summary

This chapter provided details of the data analysis, and presented and discussed the results in respect of Research Questions 2, 3, and 4. Analysis of participants' miscue responses has shed light on L2 listening behaviour, specifically the cues within the listening process on which they rely most.

The findings call into question a number of customary beliefs in previous L2 listening research. Firstly, L2 listening is generally discussed in the binary terms of either top down, implying contextual, or bottom up, implying perceptual, processes. However, findings illustrate that not only are localised co-textual cues part of L2 listening processing, but they are, in fact, more important in L2 listening that contextual cues. Additionally, the general view in L2 listening research, that contextual cues are used by L2 listeners to compensate for perceptual difficulties, is also refuted by the findings. Evidence shows that top down processing may indeed be used as compensation, but largely in the form of co-textual and not contextual cues, and often in combination with perceptual cues. Thus, it seems that L2 listening processes may be generally misconstrued. This, of course, has consequences in the field of L2 listening teaching.

The next chapter presents and discusses the results of the investigation into the extent of participants' strategy use by examining their confidence in their decoding decisions.

6. Data processing, results and discussion: participant confidence - a follow up investigation

6.1. Introduction

Chapters 4 and 5 analysed participants' transcriptions of short pieces of recorded material in order to investigate their accurate responses to establish lexical level cues relied on (Chapter 4), and their inaccurate ones to establish which perceptual and non-perceptual cues could have led participants astray (Chapter 5). It is of interest to discover how certain the participants were of the responses that they gave in order to identify:

i) how confident participants were of their word recognition when it was accurate, thus indicating their self-awareness as listeners during this task;
ii) to what extent participants' incorrect responses derived from mishearings using normal automatic word recognition processes, and to what extent their incorrect responses derived from a problem of recognition of which they were aware, thus indicating the use of compensatory strategies; and,
iii) whether participants' oral responses were similar to those obtained in writing in the main task reported in Chapter 5.

This small-scale qualitative investigation sought to consider these issues by questioning participants as to their confidence in their word recognition decisions immediately after they had reported the words. As such, it addresses Research Question 5:

To what extent are L2 mid-level listeners in a British EAP context aware of their errors of word recognition?

It should be noted that the findings reported in this chapter have no implications in terms of the quantitative studies reported in Chapters 4 and 5 investigating the relative importance of lexical, perceptual, and non-perceptual cues. Instead, they shed some light on the extent to which those cues were used automatically as part of lexical search processes, and the extent to which they were part of a compensatory strategy.

6.2. Data processing

The oral paused transcription task, described fully in Section 3.6.2, took place approximately one week after the written task. There were 22 volunteer participants, 11 of whom had listened to Recording 1 in the written task and so listened to Recording 2 in this task, and 11 of whom had listened to Recording 2 in the written task and so listened to Recording 1 in this task. The task consisted of two stages, both of which were recorded. Firstly, the participants responded to each pause in an oral paused transcription task with no interruption from the researcher. Immediately after this, the stimulated recall task took place, this involved playing back the recording of the first stage as a prompt and pausing the recording at each target to ask the participants whether they were sure that their responses had been correct.

After the task, the recordings of the participants' responses were transcribed, including their remarks regarding their confidence, and these can be seen in Appendix C. Next, their responses to the target words were transcribed in the same manner as the written paused transcription task and analysed. An example is shown in Table 6.1 below, and the fully coded data is in Appendix F:

L1	Pl tst L	P	Phrase to report (targets in bold) and responses			
			flow	of	the	**questionnaire**
CH	11	1	words	and	0	1

Table 6.1: An example of a coded participant response to one phrase.
Key: L1 = first language; CH = Chinese; Pl tst L = listening placement test score; P = participant number; 1 = correct response; 0 = no response

With regards to coding participants' responses relating to their confidence, there were four possible categories of response depending on whether the participant was sure or not sure, and the response was correct or incorrect. Each category was interpreted by the researcher as shown in Table 6.2, below.

Category	Interpretation
confident when response correct	awareness of accurate word recognition
confident when response incorrect	example of mishearing (i.e., failure of perceptual processes)
not confident when response correct	possible evidence of strategic behaviour
not confident when response incorrect	possible evidence of strategic behaviour

Table 6.2: Categories of participants' responses about confidence of their decoding decisions, and the researcher's interpretation.

Participants' lack of confidence was interpreted as possible evidence of strategic behaviour due to the view put forward in Section 2.4.11 that listening strategies are "compensatory techniques that are used to fill gaps in word recognition or in understanding" Field (2008b:2). Thus, listening strategies are used when listeners are unsure of their decoding decisions and rely on information from levels of input other than perceptual.

Although participants were commenting on their certainty about more than one word in the target phrase, they generally made it clear to which words they were referring. For example, in response to the target phrase '*presentation is very important*', a participant repeated the phrase correctly and, when questioned about her confidence, responded "the first maybe not correct exactly, the other I think is right". This response was coded, for '*presentation'*: not confident when response correct, and for '*important*': confident when response correct. If it was not possible to attribute a response to a particular word, it was not coded. For example, in response to the target phrase '*controlling many illnesses*', a participant responded '*controlling menu illnesses*'. When questioned about her confidence, she replied "*er I just can manage the key words the firsterly is erm what*", and this response was not coded. Samples that illustrate response types are presented below in Table 6.3. Miscues were typed according to the researcher's judgement based on the acoustic evidence from the participants.

Target (bold) and participant response	Verbal report	Category of confidence
approval 1	sure of approval	Confident when response correct
optimum optimus	not very sure. optimum, optimal, I don't know the word	Not confident when response incorrect
pretty personal questions clearly 1 1	very sure of all the words	Confident when response incorrect (*pretty*), confident when response correct (*personal questions*)
flow of the questionnaire 1 0 0 1	not sure of flow	Not confident when response correct (*flow)*, confident when response correct (*questionnaire*)

Table 6.3: Samples illustrating participant verbal report responses.

Total responses to each of the four categories were then calculated on the same spreadsheet, along with the number of words on which participants had reported. This can be found in Appendix L.

6.3. Results

Descriptive statistics were calculated for the total number of responses to each category and as a percentage of overall responses for each recording. These are presented in Tables 6.4 and 6.5, below. In Recording 1, 204 words were reported on, and in Recording 2 the total reported on was 197.

Category of confidence	Total responses per category	Percent of overall responses (N=204)
Confident when response correct	112	54.90
Confident when response incorrect	36	17.65
Not confident when response correct	13	6.37
Not confident when response incorrect	43	21.08

Table 6.4: Participants' responses in respect of confidence in their decoding decisions for Recording 1.

Category of confidence	Total responses per category	Percent of overall responses (N=197)
Confident when response correct	108	54.82
Confident when response incorrect	20	10.15
Not confident when response correct	14	7.11
Not confident when response incorrect	55	27.92

Table 6.5: Participants' responses in respect of confidence in their decoding decisions for Recording 2.

The trend across both recordings is essentially similar, and is illustrated in Figure 6.1, below.

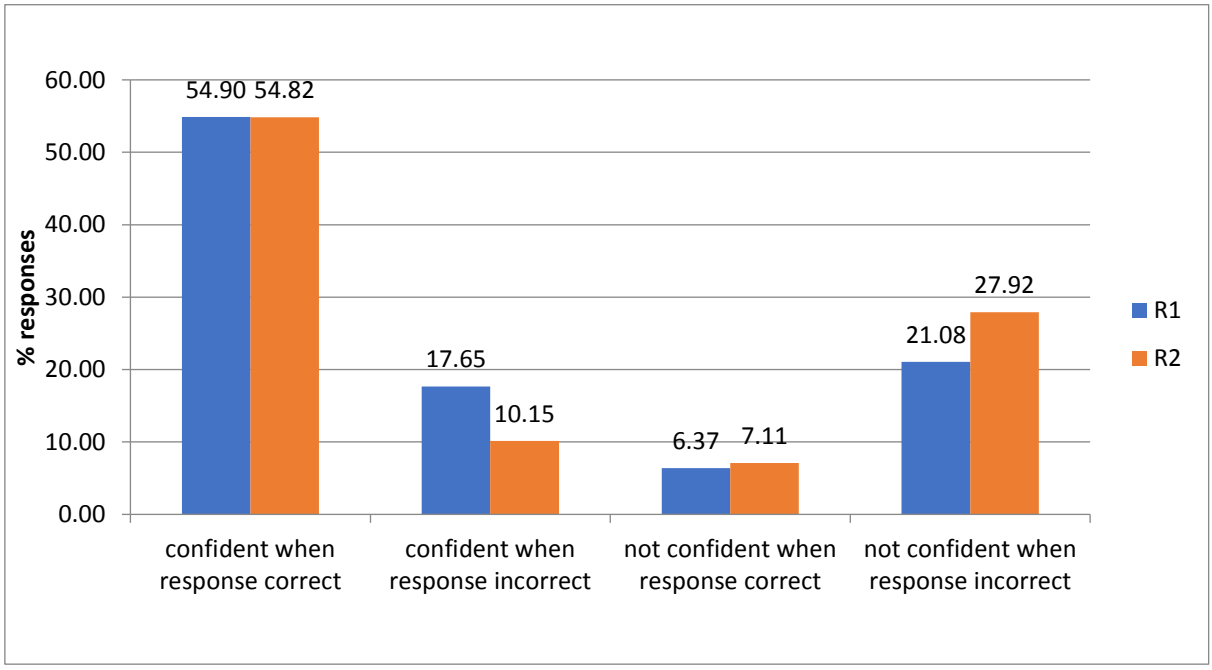


Figure 6.1: Percent of participant responses in respect of confidence in their decoding decisions by recording

Participants' self-awareness

With regards to the first point identified in Section 6.1, relating to listeners' self-awareness, the figures across both recordings indicate that participants were predominantly confident of their correct responses, and few lacked confidence in their correct responses. This indicates that they are showing a high degree of self-awareness as listeners.

Strategic behaviour

The second point identified in Section 6.1 related to the extent of the use of compensatory strategies, i.e., instances of uncertainty of word recognition. In circumstances where the uncertainty related to both correct and incorrect responses, the total for Recording 1 was 27.45% and in Recording 2 was 35.03%. With regards to incorrect responses, the results indicate that there were a greater number of incorrect responses to target words in which participants were not confident, than those in which they were confident. This indicates that there was more strategic behaviour taking place than mishearings using normal automatic word recognition processes.

Category of cues most associated with strategic behaviour

Participants' incorrect responses in which they were not confident, i.e., behaviour interpreted as strategic, were explored further to identify possible patterns with regards to the cues relied upon in this strategic behaviour. The number of responses in this category to each word was calculated. The minimum number of incorrect responses to any word was one and the maximum was six, as not all participants responded to all words. Target words to which there were three or more incorrect and uncertain responses were included in the analysis as patterns in cues used could then be noted. The cues on which the participants relied in their strategic behaviour were identified using the same process developed for the main analysis of miscues in Section 5.2.1, i.e:

i) **Perceptual -** participants have attempted to represent the auditory cues in the speech signal either in the form of a similar word or in the form of a non-word.
ii) **Co-textual -** participants have made use of the immediately surrounding words of the utterance; they comprise two types:
    - **Syntactic appropriacy within current clause**: responses that fit the syntax of the clause
    - **Lexical appropriacy within current clause**: responses where the lexis used is appropriate within the clause (includes collocations and synonyms)
iii) **Contextual -** participants have drawn upon their knowledge of the situation under discussion and/or the text so far.

The responses to target words to which there were three or more incorrect and uncertain responses are shown in Tables 6.6 and 6.7 below.

Target	flow	Cues	implicit	Cues	approval	Cues	statistically	Cues
Participant responses	floor	p,co-	impricit	p	sample	co-	sasisterly	p
	float	p,co-	presit	p	improval	p	this thee	
	follow	p	resistance	p	of proval	p	existing	p
			a present	p			stickule	
			impricit	p				
			impleasant	p				

Table 6.6: Participants' incorrect responses in which they were not confident and cues relied on for Recording 1
Key: p = perceptual, co- = co-textual. Target words in grey.

Target	illnesses	Cues	prescriptions	Cues	disastrous	Cues	witnessed	Cues
Participant responses	orifist	p	presepction	p	disaster	p	this is	p
	annoratives		prespiction	p	grayassess		insistent	
	of this is	p	prosciption	p	distaterous	p	written	p, co-
	undersis	p	craptions		dairy		been	
	of this	co-	cresbitions	p	laxitous			
			crab					

Table 6.7: Participants' incorrect responses in which they were not confident and cues relied on for Recording 2
Key: p = perceptual, co- = co-textual. Target words in grey.

It is evident across both recordings that, where participants have perceived a problem of recognition, and are thus behaving strategically, they have drawn largely on perceptual cues.

Category of cues most associated with mishearings

It was of interest to investigate participants' incorrect responses in which they were confident, i.e., those identified as mishearings. An analysis was undertaken with this category of responses using the same procedure as established in the previous section. The responses to target words to which there were three or more incorrect and certain responses are shown in Tables 6.8 and 6.9 below.

Target	optimum	Cues	approval	Cues	sample	Cues	particular	Cues
Participant responses	optenser	p	prue	p	example	p	the	co-
	optimus	p	of proval	p	example	p	the	co-
	optimember	p	improval	p	example	p	the	co-
					example	p	the	co-

Table 6.8: Participants' incorrect responses in which they were confident and cues relied on for Recording 1.
Key: p = perceptual, co- = co-textual

Target	prescriptions	Cues	expectancy	Cues	drop	Cues
Participant responses	persquiptions	p	expecancy	p	drug	p, con
	perscription	p	agency	co-	jot	p
	questions	co-	expence it	p	draf	p

Table 6.9: Participants' incorrect responses in which they were confident and cues relied on for Recording 2.
Key: p = perceptual, co- = co-textual, con = contextual

Clearly, the results are comparable to those relating to participants' strategic behaviour in the previous section, when reporting unfamiliar words. Thus, it seems that responses in which compensatory strategies were employed, and those which were identified as simple mishearings, both made use of with largely perceptual cues.

Comparison of responses to written and oral tasks

A further line of enquiry was to identify the extent to which participants responses when reporting orally were similar to those when reporting in writing in the main, quantitative task (reported in Chapter 5). However, given the small number of participants who took part in the oral task (see Section 3.6.2), this comparison is purely indicative.

The findings reported above reflect those of the main analysis of miscues reported in Section 5.3.1, i.e., miscues were largely perceptual, followed by those which were co-textual, while minimal contextual cues were seen.

Closer analysis of the erroneous responses to individual words which were frequently not recognised in both tasks revealed that there were some instances of responses to the oral task being the same as those reported for the written task. Instances are shown in Table 6.10 below.

Target	Oral response	Written response
response (R1)	3 occurrences of 'sponse'	6 occurrences of sponce, sponse, spones
approval (R1)	6 occurrences 'improve, improver, improval'	15 of improve, improver, improval/el
illnesses (R2)	3 occurrences of 'of this'	13 occurrences of 'of this'
attacking (R2)	3 occurrences of 'talking/taking'	6 occurrences of talking/taking
drop (R2)	2 occurrences of 'drug/s'	14 occurrences of drug/s

Table 6.10: Instances of responses in the oral task being the same as in the written task.

In addition, in both modes of reporting, a number of participants' responses appeared to be a result of attempts to parrot the sounds heard, i.e., reporting a non-word which was an approximate phonological representation of the target word. For example, eight written and three oral responses parroted 'optimum'. This behaviour was particularly noticeable when participants were able to recognise and report the preceding word/s, as they appeared to isolate these and then parrot the remaining group of unrecognised sounds. For example, in the target phrase '*the number of prescriptions*' participants recognised and reported '*number'* and parroted their interpretation of the word '*prescriptions'*. There were two occurrences of this behaviour in the oral data and 15 occurrences in the written data.

There is also evidence in this orally reported data that stressed syllables were influential. Instances of incorrect segmentation took place where target words' initial syllables were unstressed, whereby participants segmented at the stressed syllable and reported the initial weak vowel syllable as a function word. This behaviour is also evident in the written data. For example, three oral responses to the target '*implicit'* were '*a pressit*', '*a pricit*', and '*a present*'. In the written data, responses of '*a present*', '*a presit*', and '*no places*' were seen. For the target '*approval*', two oral responses were '*and proval',* and '*of proval'*. In the written data, variations on this response occurred six times.

6.4. Discussion

Participants' self-awareness

The finding that participants lacked confidence in their incorrect responses may indicate a degree of successful self-monitoring by participants and suggests that they were aware that the perceptual information on which they were building their understanding was approximate. This may imply awareness that their understanding may need to be reconsidered as the utterance continues and further cues become available. However, despite this proposed awareness, L2 listeners of a similar level to the participants may lack the abilities to reconsider and successfully revise their understanding due to the time pressures inherent in listening. In research by Graham et al. (2008), comprising of case studies of two listeners at lower and higher levels, the lower level listener, who was a similar level to the participants in the current study, reported being generally unable to revise her understanding upon realising that she had misunderstood areas of the text. Similar findings have also been reported by Field (2008e) and Tsui and Fullilove (1998) in relation to listeners of lower levels of proficiency. Thus, although listeners of B1 to B2 level may be aware that their word recognition has been unsuccessful, they may be unable to repair the damage by reconsidering their interpretation of the input.

Strategy use and associated cues

The results of this subsidiary investigation, reported in Section 6.3, indicate that a minimal number of erroneous responses were likely to have been the result of mishearing and a greater number were likely to be the result of strategic behaviour. These results inform those of Field (2011), who, on finding that his 47 B2 level participants' were able to recognise far fewer words than expected in a listening task, concluded that they must rely to a great degree on compensatory strategies to augment areas of the text they could not decode. Field (ibid.) had not differentiated between mishearing due to automatic word recognition processes and attempts at decoding using compensatory strategies, however, the results obtained in the current study support his conclusion.

Furthermore, participants were largely confident in their correct decoding decisions (percent overall responses Recording 1=54.85, Recording 2=53.74) and lacked confidence in their incorrect decoding decisions (percent overall responses Recording 1=21.52, Recording 2=31.72). This suggests that participants were generally able to distinguish between these two situations. Firstly, they were aware when they had accurately understood the input, and secondly they were aware when their word recognition was unsuccessful. In the second instance, they

appeared to be operating strategically by attempting a lexical match using the cues in the input available to them to support or refute their uncertain decoding decisions. Thus, the participants seem not to have misheard the input, but were simply trying to make sense of it. This suggestion is supported by psycholinguistic views of listening which propose that listeners draw on various cues in the speech signal in instances of uncertainty (e.g., Bond 1999, Field 2006).

However, a difference with regards to the behaviour of expert listeners and the participants of the study reported here is noticeable. Expert listeners have available to them cues at higher levels than perceptual, thus in circumstances where word recognition is impaired, perhaps as a result of a noisy environment, they listen strategically by drawing on co-textual or contextual cues to supplement the incomplete auditory input. In contrast, in situations of uncertainty, the participants in this study relied largely on various types of perceptual cue, such as stressed syllables with their more salient vocalic elements, and this suggests that co-textual and contextual cues may not be available to them. This finding relating to the small-scale oral task should be treated with caution due to the small numbers of participants. Nonetheless, findings relating to the qualitative task, where miscue responses were largely perceptual, support this suggestion and were discussed in detail in section 5.4.1. It was suggested that listeners at approximately level B1 and 2, i.e., the level of the participants, should be able to use higher level cues to build meaning to some degree. However, O'Malley et al. (1989) report that less proficient L2 listeners may focus more heavily on individual word recognition than those of a higher level, and so this focus may leave insufficient cognitive capacity to consider higher level cues. Furthermore, the time pressure of listening may cause listeners to rely on instantly available perceptual cues due to the more cognitively challenging demands of interpreting and weighing up evidence from higher level cues. Evidence from both Kim (1995) and Koster (1987) supports this, as they noted that additional processing time enabled L2 listeners to make use of higher level cues. Therefore, the strategic behaviour of the participants towards both the oral and written task, namely their reliance on perceptual cues, may simply be because cues at higher levels are not available to them.

It is of interest to note that the findings established through this small-scale oral task can inform those of the main analysis of incorrect responses, addressing Research Question 2 and reported in Section 5.3.1. During the main analysis, it was not possible to consider to what extent participants' decoding decisions were strategic, that is to say, they may have been aware of their misinterpretation of perceptual information and so were acting strategically, or whether they were, in fact, under the impression that their decoding decisions were accurate. However, the

oral task was designed specifically to investigate this and it was found that more participants lacked confidence in their incorrect responses, suggesting that they may have been acting strategically, than were confident of them. Their strategic behaviour led them to rely largely on perceptual cues when listening broke down, followed by syntactic co-textual cues, as well as cues which were a hybrid of both these types. Given that the findings relating to the quantitative written task revealed similar types of cue reliance, it is possible that a comparison can be made. Essentially, despite the small numbers of participants who took part in this qualitative task, the results strengthen the quantitative findings reported in Chapter 5; they indicate that the quantitative findings may have been the result of strategic behaviour and not simply the result of mishearing.

Comparison of responses to written and oral tasks

A final consideration relates to the task design, specifically the mode of reporting. A possible objection to the way in which responses were elicited in the written paused task, discussed in Section 3.3.1, was that participants reported in writing what they heard, as in previous research using this method (Field, 2008a, 2008c; Pemberton, 2004). It is possible, therefore, that participants' ability to transfer information from spoken to written form, along with their general writing and spelling proficiency, may have compromised the evidence they reported. However, although the oral task reported in this chapter was small-scale, there is a noticeable similarity in the patterns of behaviour between the responses elicited orally and those elicited during the written task reported in Chapter 5. In fact, several of the miscue responses were identical in both tasks. This may go some way to mitigating the limitation which may exist in the validity of data elicited in writing.

6.5. Summary

Whilst the results of this follow-up investigation were established through a small-scale task, they are nonetheless informative for several reasons. The quantitative findings discussed in Chapter 5 established the relative importance of the cues that were used in participants' erroneous decoding decisions, however it was not possible, at that time, to consider the extent to which those cues were used automatically, as part of lexical search, or the extent to which they were used as compensatory strategies. The data analysed in the current chapter suggest that more responses were likely to have been a result of strategic behaviour, and this indicates the cues were used in compensatory strategies. In addition, the notion that the participants were more likely to be aware that their listening had broken down, and so their interpretations of utterances

were tenuous, are also of interest, both for L2 listening researchers and those involved in listening pedagogy. Finally, the similarity between data elicited by written and oral means may help to inform research design using the paused transcription method.

The topic of the following chapter turns to the original motivation for this study, and discusses the overall findings in the context of L2 listening teaching.

7. Applications to practice

7.1. Introduction

Although this study is not primarily a pedagogical one, its fundamental motivation was entirely pedagogical, due to the author's conviction, itself a result of several years of experience as an English language teacher, that general L2 course books seemed not to be efficient at teaching listening. This study was viewed as a means of exploring that conviction. It was undertaken with the ultimate aim of increasing knowledge in the under-researched field of L2 listening, specifically word recognition processes, which could then be utilised to improve L2 listening training. In a pre-sessional EAP context, in which the study is set, more effective listening training would better prepare learners for academic lecture listening in their degree courses. This chapter discusses how the findings can be applied to the field of English language listening teaching.

Firstly, the traditional comprehension approach to L2 listening instruction is discussed, illustrating the limitations of this approach in relation to the major findings of the study. Following this, various alternative teaching approaches that the research findings support are put forward, which should be generalisable to listeners of a profile similar to the participants of this study. It is important to note that these alternative approaches are not presented as methods to be administered separately, with one precluding others; there is a noticeable relationship and overlap between them. Then, the discussion turns to the way in which technology enhanced learning (TEL) can be used to support these alternative approaches, and finally, the specific benefits of rethinking listening training in an academic context are put forward.

7.2. The comprehension approach

The generally accepted comprehension approach to teaching listening, espoused in course books and pre-service training courses around the world, involves learners being motivated to listen to a text, listen to it again, and answer questions related to it (Field, 2008d). The learners' ability to answer those questions indicates whether they have been successful. This approach seems to have been growing less popular in recent years and scholars in the field have expressed the view that the comprehension approach provides listening practice, i.e., opportunities to listen, but does not teach the skill (Field, 2008d; Graham, 2011; Graham et al.,

2008; Lynch, 2004). In addition, evidence has been put forward which illustrates that teachers of listening are aware of, and dissatisfied with, this situation (e.g. Graham, 2011; Graham et al., 2008). The study reported here has explored the word recognition processes of L2 listeners of approximately B1 to B2 level, and it is evident that the comprehension approach, in its typical format, is not a suitable means to improve word recognition, and so does not improve comprehension.

7.2.1. Limitations of the comprehension approach

L2 listeners' use of the main categories of miscue: perceptual, co-textual, and contextual

One of the major findings of this research, discovered during the investigation into the second research question, was that the types of cues widely referred to as 'contextual' should be viewed as two separate categories: co-textual and contextual. As mentioned in Section 5.4.1, this conflicts with the long-held convention, both in the literature (e.g. Hasan, 2000; Tsui & Fullilove, 1998; Vandergrift, 1997; Vandergrift & Baker, 2015) and in teaching materials (e.g. De Chazal & Rogers, 2013; Soars & Soars, 2008), that views listening as two distinct processes, namely, perceptual and contextual. For instance, tasks in listening course books generally comprise of local and global questions, what Field (2008d:14) refers to as "extensive listening followed by general questions in context", and "intensive listening followed by detailed comprehension questions". Questions which call for syntactic parsing, or the use of syntax in inferencing, are generally not seen. Of course, comprehension questions directed at perceptual and contextual processes implicitly require use of co-textual cues. However, it seems that co-textual cues are often conflated with contextual cues in the traditional approach to teaching listening as it has been defined in Sections 1.2 and 7.2 (e.g. De Chazal & Rogers, 2013; Soars & Soars, 2008), or conflated with vocabulary, in the instance of syntactic co-text (Vandergrift & Baker, 2015), rather than being addressed in their own right.

An additional major finding related to the second research question is that, in viewing 'contextual' cues as two separate categories, it is in fact co-textual and not contextual cues which are more heavily relied upon by participants, after perceptual cues. More specifically, of the two types of co-textual cue investigated, i.e., syntactic and lexical, it is syntactic cues which are more influential in participants' decoding decisions. Although general prediction-type tasks, common in course books which use the comprehension approach, could be said to activate listeners' sematic networks, i.e., lexical co-text, tasks which involve syntactic processing are less

common, as previously stated. In fact, no teaching material could be identified during the course of this research which addressed the parsing of incoming speech using syntactic cues despite calls from scholars, such as Brown (1977) who has pointed out its importance for several decades. By paying little attention to the use of syntactic co-text in L2 listening processes, course books seem to be overlooking a cue identified in this research as much more influential than context.

Types of perceptual miscues

Results investigating the second research question found that perceptual cues were relied upon by participants significantly more than co-textual and contextual cues in listening breakdowns. However, this reliance is not necessarily profitable for listeners as their ability to decode initial syllables in listening breakdowns was largely unsuccessful. This greatly impacts their word recognition processes as inaccurate cohorts of lexical matches are likely to be activated in the competition process (Broersma & Cutler, 2008), as discussed in Section 2.4.5. However, course books using the comprehension approach do not generally promote low level decoding skills in connected speech, other than the placement of lexical stress in isolated words (e.g. Campbell & Smith, 2009; Soars & Soars, 2008) .

Results of this study show that even stressed initial syllables were, on the whole, not decoded successfully, despite their prominence, and results related to the first research question show that stressed initial syllables did not appear to influence participants' successful word recognition. Given the importance of this cue in the segmentation and word recognition processes of expert listeners, i.e., the likelihood that stressed initial syllables indicate new words (Cutler & Butterfield, 1992; Cutler & Norris, 1988), it is worrying that traditional listening course books do not promote the role of lexical stress in longer segments of connected speech.

A point in favour of the comprehension approach to listening teaching relates to the results which indicate that participants need to make more use of contextual cues in their listening processes. If listeners of a similar profile to the participants of this study are relying heavily on perceptual cues, general comprehension tasks could encourage them to listen more strategically by checking their decoding decisions against higher level cues, such as co-text and context. However, it should be considered that contextual cues require an element of deduction from listeners, and so impose additional cognitive demands (Field, 2008d). As a result, and due to time pressure, listeners may ultimately be forced to rely on perceptual cues without checking

whether matches make complete sense (Koster, 1987). This complication would be mitigated if word recognition were more automatic.

7.3. Possible alternative approaches

It is evident from the previous section that the comprehension approach, in its traditional format (discussed in Section 1.2), chiefly provides listening practice, and Field (2008d), recognising this, puts forward alternative methods for teaching L2 listening. The findings of the study reported here support several of these methods and they are discussed in relation to the methods outlined below:

- *A diagnostic approach*: subsequent to general comprehension tasks, wrong answers are examined to ascertain what element of listening processes caused the breakdowns. Practice exercises are then used to address these elements.
- *A process approach*: instruction based systematically on the listening processes of expert listeners, as established by psycholinguistic literature.
- *Awareness raising*: low level decoding exercises to highlight aspects of connected speech.

(Field, 2008d:327 - 8)

In addition, two further approaches are put forward in light of the findings of the study reported here; these are also discussed below:

- *Pronunciation to promote listening*
- *Spoken vocabulary training*

A diagnostic approach

The aim of a diagnostic approach to teaching listening is to provide an outcome, in the form of responses which identify listening breakdowns, which is used to guide remedial tasks. It requires teachers to analyse the outcome in order to pinpoint reasons for the breakdowns, and to identify or design tasks to address them (Field, 2008d). This is similar to Wilson's (2003) 'discovery listening' in which listeners identify their own listening breakdowns by comparing

their task responses to the recording transcription. However, the final, and arguably the most useful, step is missing in Wilson's (ibid.) seemingly truncated approach; he does not suggest remedial exercises, but posits that listeners' 'noticing' of the causes of breakdowns is sufficient to improve their listening skills.

A limitation of a diagnostic approach is that it calls for a degree of expertise in listening processes with which not all teachers are equipped, not only to identify why listening broke down, but also to design remedial tasks. Thus, perhaps the diagnostic approach is not suitable in all contexts. As an alternative, remedial tasks of this nature could be designed based on generalisable results of research into listening breakdowns across various proficiencies. These tasks could be added as a final step to traditional comprehension-type material, as suggested by Cauldwell (2002).

By way of illustration, the findings of this study provide a diagnostic outcome generalisable to listeners of a similar profile to the participants (approximately B1 to B2 level). Evidence regarding Research Questions 1 and 2, for instance, suggests that listeners of this proficiency may not be making use of stressed initial syllables as a cue to lexical segmentation. To address this, post-comprehension remedial tasks could feature exercises highlighting how this cue can be used. More proficient listeners may require fewer decoding tasks, but instead they may need to practise making inferences or disambiguating word senses. On the other hand, lower levels may need decoding tasks based on short, non-authentic, graded texts (Field, 2008d). These could be established more precisely through research similar to this study, but which explores the issue with listeners of varying proficiencies. Ultimately, further research could enable teachers to anticipate problems likely to be encountered with specific texts, e.g., chunks of speech likely to present segmentation difficulties due to the features of connected speech, and the use of pre-designed post-comprehension remedial tasks, rather than having to diagnose and address breakdowns spontaneously.

One such chunk in the study reported here was, '*sick in the first place*', where several participants had difficulty segmenting the first two words and reported '*second'* or '*sicking'*. In this instance, armed with the prior knowledge that listeners would likely attach the weak functor '*in'* to the right of the strong syllable '*sick',* remedial tasks could be designed, before the lesson, which address resyllabification.

A process approach

A process approach centres around the division of listening processes into separate elements which can be taught and practised in the classroom (Field, 2008d). These processes are often referred to in the L2 literature as 'sub-skills' (e.g. Ridgway, 2000; Rost, 2011), however, following psycholinguistic literature which underpins this research, they are referred to here as processes. A framework of these processes formed the departure point for the analysis in this research (see Appendix A), and the findings provide evidence of the processes which cause listeners difficulties, or are underused by listeners. Thus, a process approach to the teaching of listeners of a similar proficiency to the participants can be facilitated by the findings. For instance, participants' lexical search processes appeared to be restricted by difficulties segmenting words with weak initial syllables. Thus, syllable level decoding processes could be addressed in the classroom by exercises involving awareness-raising of weak forms, and other aspects of connected speech, a suggestion which Cauldwell (2002) has put forward and for which he has designed materials.

An additional example from the findings of the current study relates to evidence that there were fewer instances of participants using co-textual and contextual cues, and this illustrates that listeners do not seem to be checking their decoding decisions against the co-text, or against their previously established knowledge of the context of the text. It was suggested in Section 5.4.1 that participants at this level should have contextual cues available to them, according to the CEFR, but may be choosing to rely on perceptual evidence in the signal. Thus, tasks are called for which utilise meaning-building processes to confirm decoding decisions. Several such tasks are outlined by Field (2008d) which involve processes such as: predicting, by using knowledge of the world; activating schema; and, noticing word sense. In addition, the process of self-monitoring could be addressed using exercises to monitor for inconsistencies (e.g. Field, 2008d:253), although texts would need to be carefully graded to provide listeners with time to undertake such cognitively complex processes.

Finally, the results of this study have highlighted a training need that is rarely addressed in pedagogical materials, i.e., the use L2 listeners make of syntactic cues. Psycholinguistic research has established that cues of this nature are a key element of the listening process (e.g. Cutler, 2012; McQueen & Cutler, 2010), so it seems appropriate that tasks supporting a process approach should reflect this. Field (2008a) puts forward several suggestions for the design of teaching materials based on recognising syntactic cues, online parsing, and inferencing from syntactic functions. Yet, there seems to be a lack of exercises of this nature in listening course

books. Ultimately, raising learners' awareness of the benefit of using co-textual cues and meaning-building processes to support perceptual evidence in the signal may promote behaviour more in line with that of expert listeners; a view supported by many L2 listening researchers (Goh, 2002; Vandergrift & Goh, 2012).

In sum, the findings of this study can be used to provide a source for the design of tasks which support a process approach, and which are aimed at listeners of a similar profile to the participants. Further, research of a similar nature to the current study, but across various proficiencies of listener, could provide evidence on which to base tasks designed for other levels of listener.

Raising awareness of aspects of connected speech

Pronunciation of words in the speech stream varies considerably due to aspects of connected speech, such as assimilation (Roach, 2009), and L2 listeners face the challenge of words often sounding very different in the speech stream compared to their citation form. This may have impacted the finding that participants in the study were largely unsuccessful at decoding initial syllables. This notion is commented on by Vandergrift and Goh (2012), who state that tasks aimed specifically at decoding help listeners to notice the irregularities and modifications of connected speech. Listeners can be made aware of examples of aspects of connected speech by the use of micro dictation tasks (Field, 2003), which the author has successfully used in listening instruction to highlight how words deviate from their citation form. For instance, a section of speech is played in small segments at a natural speed, which contains words likely to be known by the listeners, but which may be unrecognisable due to aspects of connected speech. After the listeners have transcribed their understanding of the utterance and compared their transcription to a partner's, the written form of the text is revealed and the teacher uses the listeners' transcriptions to diagnose which aspects of natural speech hindered their word recognition. This diagnosis can inform future remedial instruction.

Research in the field of phonetics provides a foundation for tasks of this nature, and short tasks can be designed to address the various features discussed in the literature. For instance, Brown (1977:68) points out that the /t/ and /d/ of regular past tense forms of verbs are more often elided between consonants than they are pronounced, hence inflections are often unreliable syntactic cues as they are regularly of weak quality or are omitted. A task designed by the author addresses the issue, and involves dictating sentences where the surrounding co-text supports

recognition of the temporal nature of the sentence, despite the past tense ending being omitted. For example, in the sentence '*my colleagues and I walked to the meeting yesterday despite the rain*', the '*ed*' ending of '*walk'* is elided with the subsequent word '*to'*. If the spoken sentence is cut between those two words and only the first half is presented, learners invariably report the event as being in the present, likely a habitual action. Once they hear the whole sentence and recognise the temporal, co-textual cue 'yesterday', they alter their responses and report a past event. Exercises such as this raise listeners' awareness of the importance of using other cues to overcome the inconsistency of the signal.

Support for this approach to training in aspects of connected speech can be found in two areas of L2 listening research. Firstly, evidence from a number of studies reported in Section 2.4.12. (e.g., Linebaugh & Roche, 2015; Siegel & Siegel, 2013) illustrates that such training can improve listeners' decoding abilities. Furthermore, Khaghaninezhad and Jafarzadeh (2014), discussed in the same section, have also indicated connections between lower-level perception training and listening comprehension.

Published material is also available which addresses the issue of connected speech, highlighted by the results of this study, although it may be less suitable for lower proficiency listeners as it seems largely to use fully authentic speech. For instance, Cauldwell (e.g., 2002; 1996; 2016) has long espoused the need for decoding training, including for prosodic features, and has created a great deal of material with this direct aim, largely in electronic format which allows easy access to sound files. Also, Thorn (2012) uses authentic texts to teach listening, and her books (e.g., 2013) do so using low level decoding exercises designed to target areas where listening is likely to break down. However, Lynch (2004) uses graded, non-authentic lectures to teach features of connected speech as follow up exercises to comprehension and note-taking tasks, which may be accessible to learners of a similar proficiency to the participants of the current study. Despite this evidence that authors recognise the need for decoding training, the transition of these ideas into activities within a widely available series of course books does not appear to have been established.

***Pronunciation to promote listening*:**

Pronunciation training implicitly involves listening skills, however, teaching pronunciation of individual words and phonemes, rather than words in connected speech, limits the impact of this type of pronunciation training on listening processes. Words spoken individually are not liable to the variation which results from features of connected speech, and

so, when they practise pronunciation of words in isolation, learners are not exposed to the type of speech they will hear in real-life listening (see Section 2.1.1), and which this study finds is required. On the contrary, it has been established that the production of longer sections of connected speech aids decoding. For example, in a study by Linebaugh and Roche (2015:5), mentioned in Section 2.4.12 and Table 2.3, training took place using words of minimal pairs within 'carrier sentences' (e.g., "The man said ___ again"). Results showed that the experimental group performed better than the control group in listening tests after training in pronunciation. Similarly, in a study by Ahangari, Rahbar and Maleki (2015), also in Table 2.3, their experimental group received a total of 30 hours extra pronunciation practice in one term, and their findings showed that the experimental group performed significantly better in the listening post-test than the control group, who received no treatment. Therefore, they conclude that pronunciation practice can improve listening comprehension.

Pronunciation to promote listening is chiefly an extension of the tasks mentioned in the previous section aimed at raising awareness of aspects of connected speech. The author has used such tasks for several purposes based on the findings of the current study and recommendations from Field (2008d). For instance, learners produce phrase and clause length utterances illustrating: i) aspects of connected speech; ii) prosodic features at word and clause level; and, iii) lexical stress as a cue to segmentation. Materials designed for training of this nature are widely available online (e.g. Cauldwell, 2016; Hancock, n.d.), but are severely lacking in traditional course books for teaching listening.

Spoken vocabulary training

Vocabulary course books tend to focus on learning words in their written form (e.g. Campbell, 2012; McCarthy & O'Dell, 2001), however, knowing the orthographic form of words does not guarantee they will be recognised in natural connected speech (e.g. Field, 2008e; Milton & Hopkins, 2006) where features, such as resyllabification and assimilation, may affect their pronunciation. Evidence from this study supports this view. For instance, the target word '*sick'* was likely to be known to participants in its orthographic form, however only 11 out of 94 participants (8.5%) correctly recognised the word. This may be because resyllabification took place with the subsequent word '*in*', and thus resulted in segmentation being more challenging. Further evidence presented in Section 4.4.1 shows that recognition rates in speech of several highly frequent targets were surprisingly low, namely *'life'* (37%), '*same'* (39%), '*first'* (66%), and '*time'* (62%), and it seems reasonable to suggest that they would largely have been recognised in their orthographic form. Other research presents similar findings, for example,

Pemberton (2004) found that around three out of four frequent words in authentic connected speech were not recognised by his participants (see Section 2.4.5).

These findings support the proposal that vocabulary teaching material should be developed to teach vocabulary presented orally and embedded in phrase and clause length utterances which illustrate its various possible deviations from citation forms in natural connected speech. This proposal is linked to the previous section which suggests that pronunciation materials should involve longer chunks of speech for the same reason.

Finally, the findings regarding participants' lack of use of lexical co-text indicate a need to teach vocabulary in the form of collocations and frequent formulaic chunks. Course books teach this in written form (e.g., O'Dell & McCarthy, 2017) but, as discussed above, there is a need for oral material. Ellis et al. (2008) point out the challenges L2 learners face in the acquisition of collocations, and the view that L2 learning must involve acquiring common collocations and idioms (Ellis, 1997) should surely include acquisition in spoken form.

The role of the listening teacher

A final point crucial in the discussion of teaching listening, and alluded to above, is the need for teachers to be conversant with a basic understanding of the listening process, and able to notice and explain aspects of connected speech to learners. After all, the traditional approach towards teaching listening does not call for knowledge of this kind as addressing reasons for incorrect answers to comprehension questions is not generally part of the approach. This view is in accordance with Cauldwell (2002), who adds that a post-listening phase within a teaching exercise should be added to improve students decoding skills; one which includes extracts from the previously heard listening text and which highlights connected speech. In view of this, materials designed to incorporate the approaches discussed above should include thorough guidance notes for teachers.

7.4. Technology enhanced learning

TEL and decoding training

Technology enhanced learning (TEL) forms an important strand in general educational practice, and is ideal for use in the type of L2 listening training suggested above, and based on the findings of this study. Small-scale, technology-based decoding exercises, such as those recommended by Field (2008d), can be used as part of all approaches suggested in Section 7.3.

However, teachers using TEL to design listening tasks should not only be aware of the psycholinguistic processes of listening, as mentioned above, but, according to Kessler (2007), should also be formally taught how to manipulate the software to suit their needs, rather than rely on their informal IT knowledge, often gained through personal experience. Thus, TEL should form an integral part of teacher training in order for teachers to evaluate the effectiveness of the technological tools available to them (Rost, 2007).

A method of teaching listening using TEL is put forward by Hulstijn (2003), who discusses software which provides a series of stages to be followed by learners working independently. The stages are:

i) listen to the recording;
ii) think for themselves whether they have understood everything that was said;
iii) replay the recording as often as they deem necessary;
iv) reveal the text … to read what they have just heard;
v) realize what they should have understood;
vi) replay the recording as many times as is necessary to be able to understand everything that was spoken without the aid of the written text.

(Hulstijn, 2003:421)

Hulstijn (2003) posits that, in this way, learners can identify and work on improving their own difficulties. However, it could be suggested that learners may only identify where their listening broke down, and not *why* it broke down. It could be suggested that after stage v), further stages be added. A teacher with knowledge of the listening process could anticipate which areas of the text were likely to cause difficulties, as discussed above, and analyse why this was the case. Then, small-scale exercises which address the problems could be designed (Field, 2008d), e.g., segments of the sound files could be manipulated to slow the speech down at points where difficulties occurred.

This illustrates the need for TEL listening exercises to be designed based on empirically attested processes of listening; IT must be employed appropriately and skilfully if it is to have any value in language acquisition. An illustration of such listening material is Cauldwell's (2013) iPad application, Streams of Speech. The application comprises several features, such as downloadable sound files and exercises aimed at improving decoding.

TEL and raising awareness of aspects of connected speech

An approach to teaching listening was put forward above which involved bringing aspects of connected speech to learners' attention, i.e., awareness raising (Field, 2008d). This can be achieved using sound file manipulation technology to adapt existing course book sound files. For instance, traditional listening lessons can be enhanced by adding decoding activities, as suggested by Cauldwell (2002), and mentioned above. After completing comprehension type listening practice, sections of the text which provide examples of speech degraded by features such as resyllabification and assimilation, can be spliced from the existing recording and used separately, using sound file editing software. These can be used to illustrate why listening may have broken down, and also for pronunciation tasks, which, according to research (e.g. Linebaugh & Roche, 2015), facilitate perception; these were discussed in Section 7.3 above.

Additionally, numerous websites exist which provide illustrations of the practice of aspects of connected speech. An example is *Aspects of Connected Speech* (Brett, n.d.), which appears to be highly regarded and has been used in L2 listening research (Ahmadian & Matour, 2014), and by the author for decoding training. The website provides examples of a wide range of connected speech features, as well as production exercises.

TEL and pronunciation to aid listening

Numerous software packages exist to teach pronunciation, however, they should be carefully evaluated before being used in the classroom. Evidence from the study reported here indicates that being able to recognise words spoken in isolation does not guarantee they will be recognised in longer chunks of speech. Therefore, software which presents single words may not be as useful as software which presents words within phrases and sentences.

TEL pronunciation materials are a suitable means of raising learners' awareness of lexical stress, and exercises of this nature are now being incorporated in material, such as, Speech in Action (Cauldwell, 2016) and Pronunciation Power (English Computerised Learning Inc., n.d.). However, no TEL pronunciation materials were identified which specifically address initial stressed syllables as a cue crucial to segmentation processes (Cutler, 1995), and the results of the current research indicate a need for this type of awareness raising.

7.5. Teaching academic listening

This study is situated within an academic context, and so it is appropriate to discuss the specific nature of EAP listening training. The fundamental aim of university pre-sessional courses is to prepare students for their degree courses, during which lectures are the central method for delivering information to groups of students. As discussed in Section 2.4.10, lectures are usually densely packed with information and are of a considerable duration (Rost, 2011). Thus, they present a major challenge to L2 listeners who must not only decode the speech, but also concurrently perform many other tasks, such as: ascertain the organisation and structure of a lecture; identify main and supporting points and their relationship; integrate new information into the existing discourse; and, recognise the lecturer's attitude and inferences (Field, in press). What is more, they have to contend with aspects of natural speech, such as hesitations, fillers, and false starts, as well as lecturers' jokes, asides, and metaphors. Finally, in addition to the myriad challenges mentioned above, listeners must also take notes (Flowerdew & Miller, 2005). It is evident that decoding a lecturer's speech is only one of many such processes involved in lecture listening, however it is decoding which is fundamental to all other processes (e.g. Field, 2008d; 2011; Goh, 2000; 2008). If decoding takes place with a higher degree of automaticity, then greater cognitive capacity is available to the listener for higher level processes (e.g. Field, 2013; Goh, 2000). Despite this, EAP listening material generally focuses on the comprehension approach and top-down strategy instruction (Phillips, 2006; Thaine & McCarthy, 2012). If the current pedagogical approach taken in EAP listening course books is viewed alongside the results of the study reported in Section 3.10, which showed that only 53.37% of target words were decoded accurately in Recording 1 and 42.40% in Recording 2, it could be suggested that leaners' needs are not being met. L2 listeners of a level common in pre-sessional courses, i.e., B1 to B2, evidently face great challenges in word recognition, and so are unlikely to be able to utilise higher level cues to fully comprehend lectures.

This lack in EAP listening materials also applies to EAP teachers' guides, which generally do not seem to promote tasks addressing lower level decoding. An example of this situation is illustrated in *EAP Essentials: a teacher's guide to principles and practice* (Alexander, Argent & Spencer 2008). The recommendation for how teachers can develop listeners' competence revolves around the design of material which retains the authenticity of the target situation, i.e., a lecture, but with adjustments made to the material to reduce the cognitive load, for instance, by reducing the speed of lecture delivery or its duration. This is, of course, a logical piece of advice, however, the tasks Alexander et. al. (ibid.) recommend while using this

type of material generally involve strategies related to high level processing, such as activating prior knowledge. They advocate dictation tasks, which are an ideal tool for teaching aspects of connected speech, however, instead their tasks are aimed only at pre-teaching vocabulary items, rather than as post-listening tasks. Unfortunately, the text does not refer to the notion that cognitive load can be reduced by improving decoding skills.

Fortunately, of late, increasing numbers of authors are producing pedagogical materials reflecting the listening processes identified by psycholinguistic research, although little seems to refer directly to lecture listening and EAP. In one EAP listening course book identified, the second edition of *Study Listening* (Lynch, 2004), the author anticipates where listening may break down in the lecture texts he uses and includes what he refers to as "troubleshooting" tasks involving decoding connected speech and understanding infrequent language, such as idioms. He provides transcripts and exercises, using them to improve prosodic aspects of speech. This is in line with the suggestion made in Section 7.3 regarding a possible drawback of the diagnostic approach, i.e., that it requires teachers to be knowledgeable of expert listening processes and features of connected speech, as this course book addresses that issue. Lynch's (ibid.) course book could be viewed as providing the aforementioned pre-designed post-comprehension remedial tasks. An additional course book, one which has moved on from the traditional approach to teaching listening, is *Contemporary Topics* (Solorzano & Frazier, 2016). The material is organised around college lectures and includes traditional comprehension activities augmented by exercises in pronunciation and vocabulary using words embedded in sentences, as suggested in Section 7.3. Tutors can then, if they are trained to do so, exploit the sentences to include training on aspects of connected speech and prosodic features using TEL in the manner discussed in Section 7.4.

7.6. Summary

This chapter has considered the application of the findings of this study to listening teaching. It is evident that the comprehension approach cannot sufficiently fulfil the needs of L2 learners, largely because it is based on tradition, rather than theory. Hence, a re-evaluation of pedagogical methods and materials is called for which aligns more closely with current cognitive frameworks of listening (e.g. Field, 2013), and L2 listening research based on these frameworks. In an academic context, the linguistic requirements which pre-sessional students must meet in order to progress to their degree programmes are clearly outlined by their receiving Schools and

Departments, and it is the responsibility of EAP course providers to equip them as efficiently as possible with the skills needed to meet these needs.

In addition, the rapid progress of technology, and emergence of innovative material designers, such as Cauldwell (2016), bode well for the future of TEL for listening. Virtual and augmented reality and 3D imaging developments will no doubt find their way into the classroom. The challenge will be to ensure that these new pedagogical paradigms, supported by theory, replace traditional ones sooner, rather than later.

8. Conclusion

8.1. Introduction

This chapter briefly summarises the study and the main findings, which it reviews in the context of the research questions. It then reviews the limitations of the study in terms of theoretical and methodological factors. Suggestions are made as to how the extensive data set gathered in the current study could be used for further research, and how new studies using the framework of analysis developed through the theoretical model of listening could enhance the understanding of L2 listening across all proficiencies.

8.2. Summary of the study

This study was motivated by the author's desire to better understand the behaviour of L2 listeners of English, to ascertain why current teaching approaches seemed ineffectual, and to identify ways of re-evaluating them. In order to achieve these aims, it investigated lexical, perceptual, and contextual cues to word recognition by analysing participants' correct and incorrect responses to a listening task. To this end, a psycholinguistic framework of listening was identified to form a benchmark against which L2 listening behaviour could be compared, and to inform the set of coding categories used in the analysis. Quantitative data were gathered using a paused transcription task, the design of which was informed by empirical psycholinguistic research, and thus considered suitable for eliciting participants' natural listening processes. A subsidiary qualitative task added insight into the strategic behaviour of L2 listeners when listening breaks down.

Given that the proficiency level of the participants was controlled, and so relatively homogenous, the participants can be considered to be representative of the wider population of L2 listeners of a similar level. Therefore, the results are viewed as generalisable across this population. Furthermore, the study took place in an academic context and the specific demands of academic listeners were considered throughout, so generalisations across similar contexts are also appropriate. The extension of the findings across the wider population means that the implications of the results on the suitability of general and EAP listening instruction can be

discussed, and recommendations for a re-evaluation of listening instruction methods and materials can be put forward.

8.3. Summary of the major findings

8.3.1. Research Question 1

The review of the literature identified little research investigating the lexical characteristics of words that influence L2 word recognition processes, and so the first line of enquiry for the current study was to investigate the impact of word frequency, word length, and initial syllable stress on participants' correct responses to the task. The research question was identified as:

> *What word level information contributes to the word recognition of L2 listeners in a British EAP context?*

With regards to word frequency, the findings indicated that participants were more likely to recognise more frequent words in Recording 1, illustrating a similarity between L1 and L2 word recognition processes. Quicker and more successful recognition of frequent words is due to mapping being more automatic (Kirsner, 1994; Luce, 1986a), and this finding illustrates that the frequency of a word seems to be stored in L2 listeners' minds, as with L1 listeners (Aitchison, 2003). However, it was evident that features of connected speech reduced the recognisability of even highly frequent words.

A major finding was that word initial stress did not have a significant impact on the likelihood of words being accurately recognised, despite this cue being very influential in L1 segmentation processes (Cutler, 2012). Thus, L2 listeners not fully attuned to the markers of stress that prevail in English may be less successful in their word recognition processes.

Finally, this study contributed to the limited understanding of the effect of word length on L2 word recognition processes by finding that this feature did not impact the likelihood of successful word recognition. This result suggests that L2 listeners process speech at word or chunk level, in a manner similar to L1 listeners (e.g. Grosjean, 1985; Shockey, 2003).

It was suggested in Chapter 7 that listening teaching should include training using small-scale decoding tasks involving longer sections of speech. Teachers conversant with the cognitive processes of listening could teach their learners how features of connected speech degrade the speech signal, and how initial syllable stress is a cue to segmentation. Chapter 7 discussed how regular practice could aid word recognition and how trained teachers could provide this.

8.3.2. Research Question 2

It has been illustrated in previous research that L2 listeners' ability to decode connected speech is not as efficient as is generally assumed (e.g. Field, 2011; Pemberton, 2004). To compensate for this difficulty, they make strategic decisions, and it was the purpose of this study to explore this behaviour. Hence the cues relied on by L2 listeners were investigated in relation to three research questions. The first is shown below:

> *Do L2 listeners in a British EAP context rely more on perceptual cues or the types of cue widely referred to as 'contextual' when compensating for words that have not been accurately recognised?*

With regards to listeners' reliance on perceptual cues versus cues widely referred to as 'contextual', perceptual cues were more heavily relied upon. This major finding contradicts the assumption made by many L2 commentators over many years that, where processing of the perceptual signal is unreliable, listeners fall back on contextual information, what Field (2008d:127) refers to as 'context saves the day'.

Participants' reliance on perceptual, rather than contextual, cues when listening breaks down demonstrates a behaviour which is understandable, given the pressure of time imposed on listeners. Perceptual evidence in the signal is instant, hard evidence, whereas contextual information requires interpretation, and is thus time-consuming and cognitively challenging for L2 listeners. Literature which has investigated L2 listening seems to vary in its outcomes, however a great deal of the research has not been informed by full, psycholinguistic models of listening processes, it has been relatively small scale, has not used longer sections of connected speech, or has been based on listener self-reports (e.g., Goh, 2000; Graham, 2006; Jensen & Hansen, 1995; Tsui & Fullilove, 1998). The limitations of each of these points were discussed in Chapter 2. Thus, the results of the current study fill a gap in the literature as they were obtained

using a quantitative analysis, underpinned by a widely accepted psycholinguistic model of listening, i.e., Field's 2013 model detailed in Section 2.2.

Two important findings, which have generally not been addressed in the literature, are: firstly, that co-textual cues should be viewed separately from contextual cues; and secondly, that co-text is an important source of information for L2 listeners. It is possible that the notion of co-text may have been implicitly included within a loosely defined view of context. As such, very few studies have considered the use L2 listeners make of co-text, hence this study contributes considerably to the literature.

The reliance on perceptual cues is supported by the finding that sometimes L2 listeners appeared to draw on information that involved two types of cue, referred to as hybrid. When this behaviour took place, a large proportion of these hybrid cues showed signs of the influence of perceptual evidence in the form of co-text. This is considered as perceptual evidence as co-textual decoding requires recognition of words within the surrounding utterance. Thus, this reinforces the finding that listeners attach great importance to perceptual evidence. Perhaps the greatest significance of this finding is that it illustrates that L2 listeners are capable of behaving like L1 listeners in terms of drawing upon multiple cues. Furthermore, when they do so, co-textual evidence seems to be more important than contextual. Neither of these notions seems to have been addressed in the literature, although Alderson, Brunfaut, and Harding (2014:329) identify a need to investigate what they refer to as "faulty interactions between different levels of processing working together", a concept which this finding may inform.

The findings also add to the literature by illustrating that compensatory strategies in listening may not be in line with those in reading, despite the processes of the two skills commonly being viewed as similar (discussed in Section 5.4.1. It has been argued that lower level readers rely on context in cases of difficulties with decoding at word level, however the participants of this study did not seem to be behaving similarly, relying instead on perceptual cues. Indeed, this seems intuitively likely given that the timing of reading is under the control of the reader who can pause to reflect on wider issues of consistency.

In relation to co-textual influences, it was suggested that L2 listeners would make limited use of collocation due to the challenges L2 learners face in acquiring them (e.g., Ellis et al., 2008). This was supported by the findings, and even the use of lexically appropriate miscues which were not collocates was limited. In addition, a further line of enquiry in respect of co-

textual cues referred to the use of syntax when listening breaks down. Although no research was identified which has investigated this area, an absence noted by Brown (2006), it was thought that L2 listeners' use of syntactic cues would be limited due to lack of mastery of the syntax of the language. This is refuted; on the contrary, findings indicate that participants' miscues were indeed influenced by syntactic co-text. Thus, a major gap in the literature relating to the process of parsing has been addressed.

The investigation which addressed this research question highlighted gaps in the knowledge of L2 listening processes when their processes are considered against the framework of L1 processes. This lack of a full understanding may go some way to explaining why traditional pedagogical approaches have not been more widely challenged. The findings support the call for a review of current teaching practice to allow the inclusion of a focus on word recognition training; extensive suggestions for which were put forward in Chapter 7. More automatic word recognition would free cognitive capacity for L2 listeners to make use of higher level cues to support, or refute, their decoding decisions, and to build meaning.

8.3.3. Research Question 3

This research question investigated the extent to which participants' responses that resembled the target word were more or less frequent than the target word. In doing so, it addressed an issue not previously identified in L2 listening research. The research question is as follows:

> *To what extent do L2 listeners in a British EAP context rely on word frequency cues when compensating for words that have not been accurately recognised?*

The finding that participants' responses were likely to be words more frequent than the target word is a further indication that L1 and L2 listeners store similar lexical information in the mind, namely the probabilities of the occurrence of words (e.g., Kirsner, 1994; Luce, 1986a). In addition, this shows that at B1 to B2 level, learners appear able to use criteria based on frequency in order to propose possible matches for spoken words that they have failed to recognise.

8.3.4. Research Question 4

L2 listeners' perceptual errors, and the cues they used as they attempted to resolve them, were explored and this is addressed in the fourth research question, which is as follows:

> *Which of the identified perceptual cues at syllable and word level are most heavily relied upon by L2 listeners in a British EAP context when compensating for words that have not been accurately recognised?*

The findings relating to the first and second research questions indicated that L2 listeners may be storing similar information in the mind as part of the representation of a word in a manner similar to L1 listeners. L1 listeners store information regarding the relationship between frequency, syllables, and rhythmic patterns (Aitchison, 2003; Bond, 1999), and this informs lexical retrieval. A finding from Research Question 4, i.e., that participants appeared to be storing lexical stress patterns, suggests that L2 listeners store some of the same information in their mental lexicon as experts. A further point worthy of note is that the majority of participants were Chinese, and their L1 does not mark lexical stress in the same way as English. Thus, participants with an L1 based on prosodic characteristics, other than stress, appeared not be disadvantaged.

The findings that participants were largely unable to match initial syllables, both stressed and unstressed, is evidence that their difficulties in phonological decoding are likely to impact their segmentation and word recognition processes. This view is supported by psycholinguistic literature which illustrates that initial syllables initiate a cohort of possible lexical matches in competition for word recognition (see Section 2.3.2). Despite limited success at matching initial stressed syllables, participants appeared to be sensitive to the reliability of stressed vowels, in a manner similar to L1 listeners (Bond, 1999; Grosjean & Gee, 1987a). It is also possible that participants used duration as a cue, a behaviour similar to that of L1 listeners (Smith, Cutler, Butterfield, & Nimmo-Smith 1989). The impact of the competition process for word recognition going awry due to initial syllable mismatching is huge, and L2 listeners are likely to lack the cognitive capacity to revise their decoding of initial syllables and activate a further set of competitors. This highlights a striking need for decoding training, suggested in Chapter 7, and supported by research findings discussed in Section 2.3.2.

8.3.5. Research Question 5

The degree of L2 listeners' confidence in their word recognition formed a subsidiary qualitative investigation to provide an indication of whether they made strategic decisions regarding cues relied upon in order to compensate for their word recognition difficulties, or whether their word recognition errors were the result of mishearings. No previous research was identified which has explored this area of L2 listening. The research question is as follows:

To what extent are L2 listeners in a British EAP context aware of their errors of word recognition?

The results showed that participants were generally aware when they had accurately understood the input, and when listening had broken down. In the second instance, they appeared to be operating strategically by attempting a lexical match using perceptual cues in the input to support or refute their uncertain decoding decisions. Thus, the participants seem to be trying to make sense of the input rather than mishearing it. This informs the overall findings of Research Questions 2, 3, and 4, as participants' miscue responses were more likely to be the result of strategic behaviour than the result of normal word recognition processes. In addition, the types of responses elicited orally were comparable to those elicited in writing, suggesting that written paused transcription tasks are useful for obtaining insights into listening behaviour.

8.4. Limitations of the study

8.4.1. Participants

The participants were drawn from a cohort of students on summer pre-sessional courses at the University of Reading, and this particular cohort was largely made up of Mandarin Chinese speakers (77%). Therefore, it was not possible for the sample to include a balanced population in terms of L1. The result was that the data did not allow for comparisons across L2s with a variety of metrical forms. With regards to their L1, it was discussed in Chapter 5 that Mandarin Chinese does not mark lexical stress in the same way as English, i.e., of "the four dimensions of speech", pitch, loudness, duration and quality, it makes use only of pitch, (Laver, 1994:512). Despite this, the findings indicate that participants perceived lexical stress and it was assumed that they had acquired prosodic knowledge of English throughout their language learning. However, they appeared not to recognise the cue it provides to segmentation, and this

study has not been able to identify whether this is a result of their L1, or if it is common across all L1s. Quantitative data from other L1 groups would inform this issue.

It is also possible that other variables may impact the cognitive listening processes within L1 groups, such as educational background, and indeed other variables may also exist. Hence, replicating this study with participants from specific native language backgrounds may be informative.

8.4.2. Task

The written paused transcription task was necessarily administered by teachers of the pre-sessional course. Every effort was made to ensure the task was administered in exactly the same way to each class, but, of course, it was not possible to monitor this. It was also not possible, due to logistics and time constraints, to provide the teachers with face-to-face training beforehand, although both verbal and written instructions were piloted and positive feedback from teachers was received, which should go some considerable way to mitigating this limitation.

Every endeavour was made to design the task so that it replicated a classroom exercise in the hope that participants would behave naturally, rather than viewing the task as some kind of test and, thereby altering their behaviour in some way; that the task took place in their usual class room with their usual teacher helped in this aim. However, there were two features of the task which would have stood out as novel. Firstly, participants had to fill out biodata and sign consent on their response sheet, and secondly, it was unlikely that they had experienced a task of this design in their previous English language learning experience. Nevertheless, listening is a subconscious cognitive process, and so the participants' responses to the task were likely to reflect natural listening processes.

With regards to the oral reports, 14 participants volunteered to take part. A small sample size such as this is not uncommon in verbal report tasks (Gass & Mackey, 2000), and was unsurprising, given that participants undertook the task in their own time when their focus was, naturally, on their pre-sessional studies. Clearly, a larger sample size would have strengthened the results, and therefore, despite the participants being representative of the larger population, the results of the oral task should be viewed as indicative.

Finally, establishing categories for coding the data using the tools of grounded theory was necessary. This was because no previous research had been identified where an attempt had been made to analyse quantitative data supported by psycholinguistic theory to provide a reliable framework. Although establishing categories in this way was time-consuming, with the support of literature in the field of L1 listening processes and the psycholinguistic framework selected to underpin the study, identification of which cues should be included in the coding categories was achievable. For example, research into L1 listening processes supports the view that phonemes are unlikely to be used in word recognition, with the exception of stressed vowels, and so participants' responses were not coded to show phoneme-by-phoneme matches, however coding included stressed vowel matches. Once the categories had been identified, coding the data became straightforward. Having undertaken the extensive task of coding quantitative data using the categories I developed, I have no reason to suggest any amendments to the categories should take place, and would recommend them for use in further research with a similar aim.

Despite the limitations reported here, I would undertake further research using the same methods, and have no hesitation in suggesting future researchers also consider these methods, as shown in the next section. Some of the limitations could, in fact, provide a departure point for future research, in that research could be designed specifically to mitigate them. For instance, participants of more varied L1 backgrounds would enable comparisons of results across L1s, and a large sample size for the verbal reports task could provide more reliable results. It becomes clear in the following section that a major reason I would undertake further research using these methods, in spite of the limitations, is that the data set generated provides ecologically valid data which could be used to address various other research aims.

8.5. Recommendations for future research

A great advantage of collecting such an extensive data set is that it can be used for further research focusing on other areas of L2 listening. Several possibilities come to mind. Firstly, Pemberton (2004), whose data collection method was very similar to that of the current study, suggests that an unknown lexical item appearing in a phrase is likely to result in listeners being unable to recognise known words surrounding it. This was discussed in Section 2.4.5. Pemberton (ibid.) highlights the need for future research to investigate this subject and, given this tendency was also noticed in the data generated in this study, it would be entirely possible to analyse it specifically in order to address this issue. This is because a selection of the paused transcription targets were chosen specifically because they were likely to be unknown by the participants.

Hence, analysis of these specific targets could shed light on the impact of unknown words on the decoding of surrounding known words.

Additionally, the working memory capacity of L2 listeners has been discussed by several scholars (see Section 2.4.2), and further understanding could be achieved using this study's data set. It has been reported extensively that decoding difficulties place increased demand on L2 listeners' working memory, and that this may impact the number of words L2 listeners are able to retain in working memory until they are parsed (e.g., Call, 1985; Cook, 1979; Meara, 1980). Baddeley (1999) suggests that working memory functions in such a way that the most recent items are more easily reported, and thus it could be suggested that L2 listeners will have difficulty parsing longer sections of an utterance. This view is evidenced by Pemberton (2004) in his study using paused transcription tasks. He found that several sections of his material were too long for his participants to retain in working memory, whereas they managed to report on the shorter sections with less difficulty. Field (2011), in a similarly designed paused transcription task, also found evidence supporting this view as his participants were generally able to report the final word in an utterance with more accuracy than earlier words. It is worth noting that these participants were of a similar proficiency to the participants in the study reported here. In both the Pemberton (2004) and the Field (2011) studies, participant numbers were relatively small, 27 and 47 respectively, and this aspect of L2 working memory capacity could be investigated using the quantitative data of the current study, thereby broadening the participant numbers to a generalisable level.

A further area of L2 listening which could be explored using the data set generated here, relates to segmentation cues. In Section 2.3.3, a model of segmentation was put forward by Mattys et al. (2005:488) relating to L1 listening, which proposed that segmentation takes place in a hierarchical form, with some cues being more influential than others. The current study investigated all cues in this hierarchy, except the phonotactic constraints of English, which the author posits are more influential to L1 listeners than lexical stress, a view with which McQueen (1998) concurs. It would add to the knowledge of L2 listeners' behaviour if the responses of the participants in this study were analysed to investigate the impact of the phonotactic constraints of English.

It would be of great interest to use the data set of the current study as a bench mark and perform the task again with participants of a higher proficiency level in order to ascertain whether more proficient listeners' reliance on perceptual cues is as substantial. Hansen and Jensen (1994) suggest higher proficiency listeners may have more contextual cues available to

them, but they provide no evidence for this. An alternative to administering a further task could be to reorganise the data in the current study to address this. When establishing participant criteria, a number of participant responses were excluded in order to analyse data from mid-level listeners only. Exclusions included participants who scored under 11, out of a possible 23, in the listening placement test, and those who scored 18 or over. The response set could be reorganised to exclude a mid-range of responses, leaving only responses from higher and lower levels to be analysed. In this way, behaviours of the two opposing proficiency groups could then be compared.

With regards to the mode of reporting, the paused transcription task in the study required participants to report in writing. In contrast, the qualitative data was elicited in response to a task requiring verbal reporting. The results of the oral task were very similar to those of the written task, but were not reported due to the small sample size (N=14). It would be of interest to explore the impact of the mode of reporting in paused transcription tasks with a comparative study of the results obtained through oral and written reporting.

8.6. Concluding remarks

This study has contributed to the body of knowledge in SLA and challenged traditional views, both theoretical and pedagogical. The fundamental strength of this study is that it was informed, in all aspects, by empirically attested psycholinguistic theory, a feature which strengthens the reliability of its findings. The outcome of the study leads me to argue that L2 learners' needs are not being met by current teaching methods. The impact of this is wide-ranging, given that listening ability impacts learners' success in training in all skills. In all contexts, I argue that materials should be revised to include tasks teaching, not practising, listening, and teachers should be trained to use them. In an academic context, I maintain that EAP teachers should be conversant with the specific demands of academic listening as this will impact the way students are trained on their pre-sessional courses, and subsequently, their success in their future academic studies.

The implication of these arguments is that English language materials writers and teacher trainers should re-examine their understanding of listening processes in order to disseminate techniques for teaching listening, based on theory rather than tradition. This recommendation will take time and determination to put into practice. Perhaps highly regarded institutions, such as Higher Education EAP providers and The British Council, could instigate this pedagogical paradigm shift, after which others may follow.

REFERENCES

Adam Just, M., & Carpenter, P. (2004). A theory of reading: from eye fixations to comprehension. In N. Unrau & R. B. Ruddell (Eds.), *Theoretical Models and Processes of Reading (6th Edition)* (pp. 748-783). Newark, DE: International Reading Association.

Ahangari, S., Rahbar, S., & Maleki, S. E. (2015). Pronunciation or listening enhancement: Two birds with one stone. *International Journal of Language and Applied Linguistics, 1*(2), 13-19.

Ahmadian, M., & Matour, R. (2014). The Effect of explicit instruction of connected speech features on Iranian EFL learners. *International Journal of Applied Linguistics and English Literature, 3*(2), 227-236.

Ahmed, R. (2015). Five essential listening skills for English learners Retrieved 01/02/2018, from https://www.britishcouncil.org/voices-magazine/five-essential-listening-skills-english-learners

Airey, J., & Linder, C. (2006). Language and the experience of learning university physics in Sweden. *European Journal of Physics, 27*(3), 553.

Aitchison, J. (2003). *Words in the Mind: An Introduction to the Mental Lexicon*: John Wiley & Sons.

Al-jasser, F. (2008). The effect of teaching English phonotactics on the lexical segmentation of English as a foreign language. *System, 36*, 94–106.

Alameen, G. (2014). *The effectiveness of linking instruction on NNS speech perception and production.* Doctor of Philosophy, Iowa State University, Ames, Iowa. USA.

Alderson, J. C., Brunfaut, T., & Harding, L. (2014). Towards a theory of diagnosis in second and foreign language assessment: Insights from professional practice across diverse fields. *Applied Linguistics, 36/2*(2), 236–260.

Alexander, O., Argent, S., & Spencer, J. (2008). *EAP Essentials: a Teacher's Guide to Principles and Practice*. Reading: Garnet Publishing Ltd.

Anderson, J. (1995). *Cognitive Psychology and its Implications* (4th ed.). New York: Freeman.

Anderson, J. (1996). *The Architecture of Cognition*. Mahwah, N.J.: Lawrence Erlbaum Associates.

Baddeley, A. (1992). Working memory. *Science, 255*(5044), 556-559.

Baddeley, A. D. (1999). *Essentials of Human Memory*: Taylor & Francis.

Baghrahi, A. K., Shariati, M., & Tajadini, M. (2014). The effects of assimilation and elision teaching on listening comprehension of EFL junior high school students. *International Journal of Language Learning and Applied Linguistics World, 5*, 264-273.

Barnbrook, G., Mason, O., & Krishnamurthy, R. (2013). *Collocation: Applications and Implications*. London: Palgrave Macmillan.

Bekleyen, N. (2009). Helping teachers become better English students: Causes, effects, and coping strategies for foreign language listening anxiety. *System, 37*(4), 664-675.

Benson, M. J. (1994). Lecture listening in an ethnographic perspective. *Academic listening: Research perspectives*, 181-198.

Bloom, B. S. (1954). The thought process of students in discussion. *Accent on teaching: Experiments in general education*, 23-46.

Boland, J. E., & Cutler, A. (1996). Interaction with autonomy: Multiple output models and the inadequacy of the Great Divide. *Cognition, 58*(3), 309-320.

Bond, Z. (1979). Listening to elliptic speech: Pay attention to stressed vowels.

Bond, Z. (1999). *Slips of the Ear: Errors in the Perception of Casual Conversation*. San Diego, Calif. ; London: Academic.

Bond, Z., & Shockey, L. (2014). What slips of the ear reveal about speech perception. *Linguistica Lettica, 22*(107), 107-113.

Bond, Z., & Small, L. (1983). Voicing, vowel, and stress mispronunciations in continuous speech. *Perception & Psychophysics, 34*(5), 470-474.
Brett, D. (n.d.). Aspects of Connected Speech Retrieved 24 February 2017, from http://davidbrett.uniss.it/phonology/aspects_of_connected_speech_inde.htm
Broersma, M., & Cutler, A. (2008). Phantom word activation in L2. *System, 36*(1), 22-34.
Brown, G. (1977). *Listening to Spoken English* (2nd ed.). London: Longman.
Brown, G. (1986). Investigating listening comprehension in context. *Applied Linguistics, 7*(3), 284-302.
Brown, G. (2006). Second language listening. In K. Brown (Ed.), *Encyclopaedia of Language and Linguistics* (Vol. 11, pp. 81-88). Oxford: Elsevier.
Brunfaut, T., & Révész, A. (2015). The role of task and listener characteristics in second language listening. *TESOL Quarterly, 49*(1), 141-168.
Buck, G. (1988). Testing listening comprehension in Japanese university entrance examinations. *JALT journal, 10*(1), 15-42.
Bybee, J. (2000). The phonology of the lexicon; evidence from lexical diffusion. In S. K. M. Barlow (Ed.), *Usage-Based Models of Language* (pp. 65-85). Stanford: CSLI.
Call, E. (1985). Auditory short term memory, listening comprehension and the Input Hypothesis. *TESOL Quarterly, 19*(4), 765-781.
Campbell, C. (2012). *English for Academic Study: Vocabulary : Study Book*: Garnet Education.
Campbell, C., & Smith, J. (2009). *English for Academic Study: Listening*. Reading: Garnet Publishing, Limited.
Carbonell, J. G. (1982). Learning by Analogy: Formulating and Generalizing Plans from Past Experience. In R. S. Michalski, J. G. Carbonell & T. M. Mitchell (Eds.), *Machine Learning: An Artificial Intelligence Approach* (137-161). Palo Alto, Calif.: Tioga Press.
Cauldwell, R. (2002). Grasping the nettle: The importance of perception work in listening comprehension Retrieved 22 Febuary, 2016, from www.developingteachers.com
Cauldwell, R. (2013). *Phonology for Listening: Teaching the Stream of Speech*. Birmingham: speechinaction.
Cauldwell, R. T. (1996). Direct encounters with fast speech on CD-audio to teach listening. *System, 24*(4), 521-528.
Cauldwell, R. T. (2016). Speechinaction Retrieved 24 February 2017, from www.speechinaction.com
Charmaz, K. (2014). *Constructing grounded theory*: Sage.
Clark, H. H., & Clark, E. V. (1977). *Psychology and Language: An Introduction to Psycholinguistics*: Harcourt Brace Jovanovich.
Cohen, A. D. (1998). *Strategies in Learning and Using a Second Language*: Routledge.
Cohen, L., Manion, L., & Morrison, K. (2007). *Research Methods in Education* (6th ed. ed.). London: Routledge.
Connine, C., & Pinnow, E. (2006). Phonological variation in spoken word recognition: Episodes and abstractions. *The Linguistic Review, 23*(3), 235-245.
Conrad, S. M., & Biber, D. (2005). The frequency and use of lexical bundles in conversation and academic prose. *Lexicographica*.
Content, A., Kearns, R., & Frauenfelder, U. (2001). Boundaries versus onsets in syllabic segmentation. *Journal of Memory and Language, 45*(2), 177-199. doi: http://dx.doi.org/10.1006/jmla.2000.2775
Cook, V. J. (1979). Aspects of memory in secondary school language learners. *Interlanguage Studies Bulletin Utrecht, 4*(2), 161-172.
Council of Europe. (2001). *Title Common European framework of reference for languages : learning, teaching, assessment*. Cambridge Cambridge University Press.
Creswell, J. W., & Clark, V. L. P. (2007). *Designing and Conducting Mixed Methods Research*: SAGE Publications.
Cruttenden, A. (2014). *Gimson's Pronunciation of English* (8 ed.). London: Taylor and Francis.

Cutler, A. (1984). Stress and accent in language production and understanding. *Intonation, accent, and rhythm: Studies in discourse phonology* (8) 77-90.
Cutler, A. (1986). Forbear is a homophone: Lexical prosody does not constrain lexical access. *Language and Speech, 29*(3), 201-220.
Cutler, A. (1990). Exploiting prosodic probabilities in speech segmentation. In G. T. M. Altmann (Ed.), *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives.* (pp. 105-121). Cambridge, MA, US: The MIT Press.
Cutler, A. (1991). Linguistic rhythm and speech segmentation. In J. Sundberg, L. Nord & R. Carlson (Eds.), *Music, Language, Speech and Brain : Proceedings of an international symposium held at Wenner-Gren Center, Stockholm* (pp. 157-166). London Macmillan.
Cutler, A. (1995). Exploiting prosodic probabilities in speech segmentation. In G. Altmann (Ed.), *Cognitive Models of Speech Processing* (pp. 105-121). Cambridge, MA: MIT Press.
Cutler, A. (2001). Listening to a second language through the ears of a first. *Interpreting 5* (1), 1-23.
Cutler, A. (2012). *Native Listening: Language Experience and the Recognition of Spoken Words*: MIT Press.
Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language, 31*(2), 218-236
Cutler, A., & Carter, D. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech & Language, 2*(3-4), 133-142
Cutler, A., & Clifton, C. (1984). The use of prosodic information in word recognition. In H. Bouma & D. Bouwhuis (Eds.), *Attention and Performance: Control of Language Processes* (pp. 183-196). Hillsdale, NJ: Erlbaum.
Cutler, A., & Clifton, C. (1999). Comprehending spoken language: A blueprint for the listener. In C. Brown & P. Hagoort (Eds.), *The Neurocognition of Language* (pp. 123-166). Oxford: Oxford University Press.
Cutler, A., Dahan, D., & van Donselaar, W. (1997a). Prosody in the comprehension of spoken language: a literature review. *Language and Speech 40*(2), 141–201.
Cutler, A., Dahan, D., & Van Donselaar, W. (1997b). Prosody in the comprehension of spoken language: A literature review. *Language and Speech, 40*(2), 141-201.
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*(4), 385-400.
Cutler, A., Mehler, J., Norris, D., & Segui, J. (1987). Phoneme identification and the lexicon. *Cognitive Psychology, 19*, 141-177.
Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 113-121.
Cutler, A., & Norris, D. (1979). *Monitoring sentence comprehension.* Hillsdale, N.J: Erlbaum.
Davis, M. H., Marslen-Wilson, W. D., & Gaskell, M. G. (2002). Leading up the lexical garden path: Segmentation and ambiguity in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 28*(1), 218.
De Chazal, E., & Rogers, L. (2013). *Oxford EAP: A Course in English for Academic Purposes*: Oxford University Press New Zealand.
De Cock, S. (1998). A Recurrent Word Combination Approach to the Study of Formulae in the Speech of Native and Non-Native Speakers of English. *International Journal of Corpus Linguistics, 3*(1), 59-80.
Delattre, P. C., Liberman, A. M., & Cooper, F. S. (1955). Acoustic loci and transitional cues for consonants. *The Journal of the Acoustical Society of America, 27*(4), 769-773.
Dörnyei, Z. (2007). *Research Methods in Applied Linguistics : Quantitative, Qualitative, and Mixed Methodologies*. Oxford: Oxford University Press.
Ellis, N. (2003). Constructions, chunking, and connectionism: The emergence of second language structure. *The handbook of second language acquisition*, 63-103.

Ellis, N., Simpson-Vlach, R., & Maynard, C. (2008). Formulaic language in native and second language speakers: Psycholinguistics, corpus linguistics, and TESOL. *TESOL Quarterly, 42*(3), 375-396.
Ellis, N. C. (1997). Vocabulary acquisition: Word structure, collocation, word-class, and meaning. *Vocabulary: Description, acquisition and pedagogy*, 122-139.
English Computerised Learning Inc. (n.d.) Retrieved 24 February 2017, from http://www.englishlearning.com/
Ericsson, K. A., & Simon, H. A. (1987). Verbal reports on thinking.
Field, J. (1997). Monitoring for breakdown of understanding. In J. Field, A. Graham, E. Griffiths & K. Head (Eds.), *Teachers Develop Teachers Research 2.* Whitstable,: IATEFL.
Field, J. (1998). Skills and strategies: towards a new methodology for listening. *ELT Journal Volume, 52*(2), 110-118.
Field, J. (2001). *Lexical segmentation in first and foreign language listening.* PhD thesis. Cambridge University. Cambridge
Field, J. (2003). Promoting perception: lexical segmentation in second language listening. . *ELT Journal 57*, 325–334.
Field, J. (2004a). An insight into listeners' problems: too much bottom-up or too much top-down? *System, 32*, 363-377.
Field, J. (2004b). *Psycholinguistics: The Key Concepts*. London: Routledge.
Field, J. (2008a). Bricks or mortar: Which parts of the input does a second language listener rely on? *TESOL Quarterly, 42*(3), 411-432.
Field, J. (2008b). Guest editor's introduction emergent and divergent: A view of second language listening research. *System, 36*(1), 2-9.
Field, J. (2008c). The L2 listener: Type or individual? *RCEAL Working papers in English and Applied Linguistics in Honour of Gillian Brown* (Vol. 12, pp. 13-32). Cambridge: University of Cambridge.
Field, J. (2008d). *Listening in the Language Classroom*. Cambridge: Cambridge University Press.
Field, J. (2008e). Revising segmentation hypotheses in first and second language listening. *System, 36*(1), 35-51. doi: http://dx.doi.org/10.1016/j.system.2007.10.003
Field, J. (2011). Into the mind of the academic listener. *Journal of English for Academic Purposes, 10*(2), 102-112.
Field, J. (2012). The cognitive validity of the lecture-based question in the IELTS listening paper *IELTS Research Reports* (Vol. 2, pp. 391-453).
Field, J. (2013). Cognitive validity. In A. Geranpayeh & L. Taylor (Eds.), *Examining Listening: Research and Practice in Assessing Second Language Listening* (pp. 77-151). Cambridge: Cambridge University Press.
Field, J. (in press). *Rethinking the Listening Test*. Sheffield: Equinox.
Flege, J. (1995). Second language speech learning theory, findings, and problems. In W. Strange (Ed.), *Speech Perception and Linguistic Experience : Issues in Cross-Language Research* (pp. 233-277). Baltimore: York Press.
Flowerdew, J. (1994). Research of relevance to second language lecture comprehension: An overview. *Academic listening: Research perspectives*, 7-29.
Flowerdew, J., & Miller, L. (1992). Student perceptions, problems and strategies in second language lecture comprehension. *RELC Journal, 23*(2), 60-80.
Flowerdew, J., & Miller, L. (2005). *Second Language Listening : Theory and Practice*. Cambridge: Cambridge University Press.
Fodor, J. A. (1983). *The Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, Massachusetts: MIT Press.
Gass, S. M., & Mackey, A. (2000). *Stimulated Recall Methodology in Second Language Research*. New York: Routledge.

Gellatly, A. (1986). *The Skilful mind: An Introduction to Cognitive Psychology*. Milton Keynes: Open University.
Gernsbacher, M. A. (2013). Language comprehension as structure building. Oregon: University of Oregon.
Goh, C. (2000). A cognitive perspective on language learners' listening comprehension problems. *System, 28* 55-75.
Goh, C. (2002). Exploring listening comprehension tactics and their interaction patterns. *system, 30,* 185-206.
Goh, C. (2008). Metacognitive instruction for second language listening development: Theory, practice and research implications. *RELC Journal, 39*(2), 188-213. Retrieved from
Golato, P. (2002). Word parsing by late-learning French–English bilinguals. *Applied Psycholinguistics, 23*(3), 417-446.
Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1166-1183.
Gow, D. W., & Gordon, P. C. (1993). Coming to terms with stress: Effects of stress location in sentence processing. [journal article]. *Journal of Psycholinguistic Research, 22*(6), 545-578. doi: 10.1007/bf01072936
Graham, S. (2006). Listening comprehension: The learners' perspective. *System, 34*(2), 165-182.
Graham, S. (2011). Self-efficacy and academic listening. *Journal of English for Academic Purposes, 10*(2), 113-117. doi: http://dx.doi.org/10.1016/j.jeap.2011.04.001
Graham, S., Santos, D., & Vanderplank, R. (2008). Listening comprehension and strategy use:A longitudinal exploration. *System, 36* 52–68.
Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception and Psychophysics, 28*, 267–283.
Grosjean, F. (1985). The recognition of words after their acoustic offset: Evidence and implications. *Perception & Psychophysics, 38*(4), 299-310.
Grosjean, F., & Gee, J. P. (1987a). Prosodic structure and spoken word recognition. *Cognition, 25*(1-2), 135-155.
Grosjean, F., & Gee, J. P. (1987b). Prosodic structure and spoken word recognition.
Halliday, M. A. (1966). Lexis as a linguistic level. *In memory of JR Firth*, 148-162.
Hancock, M. (n.d.). Pronunciation as a listening skill, from http://hancockmcdonald.com/ideas/pronunciation-listening-skill
Hansen, C., & Jensen, C. (1994). Evaluating lecture comprehension. In J. Flowerdew (Ed.), *Academic Listening: Research Perspectives* (pp. 241). Cambridge: Cambridge University Press.
Harding, L., Alderson, J. C., & Brunfaut, T. (2015). Diagnostic assessment of reading and listening in a second or foreign language: Elaborating on diagnostic principles. *Language Testing, 32*(3), 317-336.
Harley, B. (2000). Listening Strategies in ESL: Do Age and L1 Make a Difference? *TESOL Quarterly, 34*(4), 769-777. doi: 10.2307/3587790
Hasan, A. S. (2000). Learners' perceptions of listening comprehension problems. *Language, Culture and Curriculum, 13*(2), 137-153.
Hulstijn, J. (2003). Connectionist models of language processing and the training of listening skills with the aid of multimedia software. *CALL 16*(5), 413-425.
Hyland, K. (2006). *English for Academic Purposes: An Advanced Resource Book*. Abingdon: Routledge.
Jarvella, R. (1971). Syntactic processing of connected speech. *Journal of Verbal Learning and Verbal Behavior, 10*(4), 409-416.
Jensen, C., & Hansen, C. (1995). The effect of prior knowledge on EAP listening-test performance. *Language Testing, 12*(1), 99-119.

Johnson-Laird, P. N. (1983). *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Cambridge M.A.: Harvard University Press.
Johnson, K. (1996). Speech perception without speaker normalization. In K. a. M. Johnson, J. (Ed.), *Talker Variability in Speech Processing* (pp. 145-165). San Diego: Academic Press.
Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher, 33*(7), 14-26.
Jusczyk, P. W., Cutler, A., & Redanz, N. J. (1993). Infants' preference for the predominant stress patterns of English words. *Child development, 64*(3), 675-687.
Kelly, P. (1991). Lexical ignorance: The main obstacle to listening comprehension with advanced foreign language learners. *International Review of Applied Linguistics in Language Teaching, 29*(2), 135–149.
Kessler, G. (2007). Formal and informal CALL preparation and teacher attitude toward technology. *Computer Assisted Language Learning, 20*(2), 173-188.
Khaghaninezhad, M. S., & Jafarzadeh, G. (2014). Investigating the Effect of Reduced Forms Instruction on EFL Learners' Listening and Speaking Abilities. *English Language Teaching, 7*(1), 159.
Kim, H. (1995). Intake from the speech stream: speech elements that L2 learners attend to. In R. Schmidt (Ed.), *Attention and Awareness in Foreign Language Learning* (pp. 65-84). Hawai'i University of Hawai'i Press.
Kirsner, K. (1994). Implicit processes in second language learning. *Implicit and explicit learning of languages*, 283-312.
Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *The Journal of the Acoustical Society of America, 59*(5), 1208-1221.
Klatt, D. H., & Cooper, W. E. (1975). Perception of segment duration in sentence contexts *Structure and Process in Speech Perception.* (pp. 69-89). Berlin, Heidelberg: Springer.
Klatzky, R. L. (1984). *Memory and Awareness: An Information-Processing Perspective*: WH Freeman.
Koster, C. (1987). *Word Recognition in Foreign and Native Language: Effects of Context and Assimilation*. USA: Foris Publications.
Kuo, Y. (2010). Using partial dictation of an English teaching radio program to enhance EFL learners' listening comprehension. *Asian EFL Journal Professional Teaching Articles, 47*, 4-29.
Laver, J. (1994). *Principles of Phonetics*: Cambridge University Press.
Leeman, J. (1999). *Recasts in Spanish as a second language: an empirical study of negative evidence and enhanced salience.* Paper presented at the Second Language Research Forum, Minneapolis.
Lehmann, H., Hoffman, S., & Schneider, P. (2002, 2007). BNCweb Retrieved 01/01/18, 2018, from http://corpora.lancs.ac.uk/BNCweb/
Liberman, A., Cooper, F., Shankweiler, D., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review, 74*(6), 431.
Liberman, A. M. (1957). Some results of research on speech perception. *The Journal of the Acoustical Society of America, 29*(1), 117-123.
Lieberman, P. (1967). *Intonation, perception, and language*: M.I.T. Press.
Linebaugh, G., & Roche, T. B. (2015). Evidence that L2 production training can enhance perception. *Journal of Academic Language and Learning, 9*(1), A1-A17.
Locastro, V. (1994). Learning Strategies and Learning Environments. *TESOL Quarterly*, *28*(2), 409-414. doi: 10.2307/3587445
Logie, R. H. (1999). Working memory. *The Psychologist, 12*(4), 174-178.
Long, D. (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.

Luce, P. (1986a). A computational analysis of uniqueness points in auditory word recognition. *Perception & Psychophysics, 39*(3), 155-158.
Luce, P. A. (1986b). Neighborhoods of words in the mental lexicon. *Research on Speech Perception, Technical Report No. 6*.
Lynch, T. (2004). *Study Listening: A Course in Listening to Lectures and Note Taking*: Cambridge University Press.
Lynch, T. (2006). Academic listening: Marrying top and bottom. In E. a. A. M.-F. Uso-Juan (Ed.), *Current Trends in the Development and Teaching of the Four Language Skills* (pp. 91-110). Berlin: Mouton de Gruyter.
Lynch, T. (2011). Academic listening in the 21st century: Reviewing a decade of research. *Journal of English for Academic Purposes, 10*, 79-88.
Macaro, E. (2001). Learning strategies in second and foreign language classrooms. *London: Continuum, 282*.
Mackey, A., Gass, S., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition, 22*(04), 471-497.
Marslen-Wilson, W. (1990). *Activation, competition, and frequency in lexical access.* Cambridge, MA: The MIT Press.
Marslen-Wilson, W., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology, 10*(29-63).
Marzban, A., & Abdollahi, M. (2013). The effect of partial dictation on the listening comprehension ability of Iranian intermediate EFL learners. *International Research Journal of Applied and Basic Sciences, 5*(2), 238-244.
Mattys, S., White, L., & Melhorn, J. (2005). Integration of multiple speech segmentation cues: A hierarchical framework. *Journal of Experimental Psychology, 134*(4), 477-500.
McCarthy, M., & O'Dell, F. (2001). *English Vocabulary in Use Upper-Intermediate with Answers*. Cambridge: Cambridge University Press.
McQueen, J. (1998). Segmentation of continuous speech using phonotactics. *Journal of Memory and Language, 39*, 21-46.
McQueen, J. (2004). Speech perception. In K. Lamberts & R. Goldstone (Eds.), *The Handbook of Cognition* (pp. 255-275). London: Sage.
McQueen, J. (2007). Eight questions about spoken-word recognition. In M. G. Gaskell (Ed.), *The Oxford Handbook of Psycholinguistics* (pp. 37-53). Oxford: Oxford University Press.
McQueen, J., Norris, D., & Cutler, A. (1994a). Competition in spoken word recognition: Spotting words in other words. . *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 621-638.
McQueen, J. M., & Cutler, A. (2010). Cognitive processes in speech perception *The Handbook of Phonetic Sciences* (pp. 489-520). Chichester: Blackwell.
McQueen, J. M., Norris, D., & Cutler, A. (1994b). Competition in spoken word recognition: Spotting words in other words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(3), 621.
Meara, P. (1980). Syntactic structure and memory span in second language learners. *Interlanguage Studies Bulletin, 5*(2), 31-54.
Meara, P. M. (2009). *Connected Words: Word Associations and Second Language Vocabulary Acquisition.* Amsterdam: John Benjamins.
Miller, G. (1956). The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information, from http://www.musanim.com/miller1956/
Miller, J. L. (1981). Effects of speaking rate on segmental distinctions. *Perspectives on the study of speech*, 39-74.
Milton, J. (2009). *Measuring Second Language Vocabulary Acquisition*. Bristol: Multilingual Matters.

Milton, J., & Hopkins, N. (2006). Comparing phonological and orthographic vocabulary size: Do vocabulary tests underestimate the knowledge of some learners? *Canadian Modern Language Review, 63*(1), 127-147.
Morton, J. (1979). Word recognition. *Psycholinguistics, 2*, 107-156.
Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics, 24*(2), 223-242.
Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*(3), 231.
Norris, D. (1994). Shortlist: a connectionist model of continuous speech recognition. *Cognition, 52*(3), 189-234
Norris, D., McQueen, J., & Cutler, A. (1995). Competition and segmentation in spoken-word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(5), 1209-1228.
Norris, D., McQueen, J. M., Cutler, A., & Butterfield, S. (1997). The possible-word constraint in the segmentation of continuous speech. *Cognitive Psychology, 34*(3), 191-243.
O'Dell, F., & McCarthy, M. (2017). *English Collocations in Use Advanced Book with Answers: How Words Work Together for Fluent and Natural English*. Cambridge: Cambridge University Press.
O'Malley, J. M., Chamot, A., & Kupper, L. (1990). *Learning Strategies in Second Language Acquisition*. Cambridge: Cambridge University Press.
O'Malley, J. M., Chamot, A. U., & Küpper, L. (1989). Listening comprehension strategies in second language acquisition. *Applied Linguistics, 10*(4), 418-437.
Oxford, R. L. (1990). *Language Learning Strategies: What Every Teacher Should Know*. New York: Newbury House.
Pemberton, R. (2004). *A preliminary investigation of difficulties faced by EAP listeners: the relative importance of lexical knowledge*. Paper presented at the Studies in lexis. Proceedings of a seminar on lexis., Hong Kong http://repository.ust.hk/ir/Record/1783.1-1084
Perfetti, C. A., & Roth, S. (1980). Some of the Interactive Processes in Reading and Their Role in Reading Skill. Pittsburgh P.A.: University of Pittsburgh
Phillips, T. (2006). *Listening Skills in English*. Reading: Garnet Publishing, Colleges of Technology Oman.
Pisoni, D. B. (1997). Some thoughts on "normalization" in speech perception. *Talker variability in speech processing*, 9-32.
Pisoni, D. B., & Luce, P. A. (1987). Acoustic-phonetic representations in word recognition. *Cognition, 25*(1-2), 21-52.
Pollack, I., & Pickett, J. M. (1964). Intelligibility of excerpts from fluent speech: Auditory vs. structural context. *Journal of Verbal Learning and Verbal Behavior, 3*, 79–84.
Posner, M. I., & Snyder, C. (1975). Attention and cognitive control. In R. Solso (Ed.), Information processing and cognition: The Loyola symposium (pp. 55-85). New Jersey: Lawrence Erlbaum.
Price, P. J., Ostendorf, M., Shattuck-Hufnagel, S., & Fong, C. (1991). The use of prosody in syntactic disambiguation. *The Journal of the Acoustical Society of America, 90*(6), 2956-2970.
Ridgway, T. (2000). Listening strategies - I beg your pardon? *ELT Journal, 54*(2), 179-185.
Roach, P. (2009). *English Phonetics and Phonology: A Practical Course*. Cambridge: Cambridge University Press.
Rost, M. (2007). I'm only trying to help: A role for interventions in teaching listening. *Language Learning & Technology, 11*(1), 102-108.
Rost, M. (2011). *Teaching and Researching Listening*: Longman.

Rumelhart, D. E. (2004). Toward an interactive model of reading. In N. J. Unrau & R. B. Ruddell (Eds.), *Theoretical Models and Processes of Reading (6th Edition)* (pp. 719-747). Neward, DE: International Reading Association.

Sanders, L., Neville, H., & Woldorff, M. (2002). Speech segmentation by native and non-native speakers: The use of lexical, syntactic, and stress-pattern cues. *Journal of Speech, Language, and Hearing Research Online., 45*(3), 519-530

Schafer, A. (1995). *The role of optional prosodic boundaries.* Paper presented at the Eight Annual CUNY Conference on Human Sentence Processing Tucson, Arizona.

Schmidt, R. (1992). Awareness and second language acquisition. *Annual Review of Applied Linguistics, 13*, 206-226.

Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129-158.

Sebastian-Galles, N. (2005). Cross-language speech perception. In D. Pisoni & R. Remez (Eds.), *The Handbook of Speech Perception* (pp. 546-561). Malden, Massachusetts ; Oxford: Blackwell.

Segalowitz, N., Trofimovich, P., Gatbonton, E., & Sokolovskaya, A. (2008). Feeling affect in a second language: The role of word recognition automaticity. *The Mental Lexicon, 3*(1), 47-71.

Segui, J., Frauenfelder, U., & Halle, P. (2001). Phonotactic constraints shape speech perception: Implications for sublexical and lexical processing. In E. Dupoux (Ed.), *Language, Brain and Cognitive Development: Essays in Honor of Jacques Mehler* (pp. 195-208). MA: MIT Press.

Shockey, L. (2003). *Sound Patterns of Spoken English*. Oxford: Blackwell.

Siegel, J., & Siegel, A. (2013). Empirical and attitudinal effects of bottom-up listening activities in the L2 classroom. *ELT World Online, 5*, 1-25.

Sinclair, J. (1991). *Corpus, Concordance, Collocation*: Oxford University Press.

Sinclair, J. M. (1987). Collocation: a progress report. In R. Steele & T. Threadgold (Eds.), *Language Topics* (Vol. 1, pp. 319-332). Amsterdam: John Benjamins.

Siyanova, A., & Schmitt, N. (2008). L2 learner production and processing of collocation: A multi-study perspective. *Canadian Modern Language Review, 64*(3), 429-458.

Smith, M. R., Cutler, A., Butterfield, S., & Nimmo-Smith, I. (1989). The perception of rhythm and word boundaries in noise-masked speech. *Journal of Speech, Language, and Hearing Research, 32*(4), 912-920.

Soars, L., & Soars, J. (2008). *New Headway English Course: Intermediate. Student's book*. Oxford: Oxford University Press.

Solorzano, H., & Frazier, L. (2016). *Contemporary Topics 1*. Upper Saddle River, United States.

Stanovich, K. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly, 16*(1), 32-71

Stemberger, J. P. (1990). Wordshape errors in language production. *Cognition, 35*(2), 123-157. doi: http://dx.doi.org/10.1016/0010-0277(90)90012-9

Strange, W. (1995). Cross-language studies of speech perception: A historical review. In W. Strange (Ed.), *Speech Perception and Linguistic Experience : Issues in Cross-Language Research* (pp. 3-45). Baltimore: York Press.

Strauss, A. L., & Corbin, J. (1994). Grounded theory methodology: an overview. In N. Denzin & Y. Lincoln (Eds.), *Handbook of Qualitative Research* (pp. 273-285). Thousand Oaks, California: Sage.

Tauroza, S. (1993). Recognizing words in continuous speech: how important are word-final consonants? *ELT Journal, 47*(3), 211-218. doi: 10.1093/elt/47.3.211

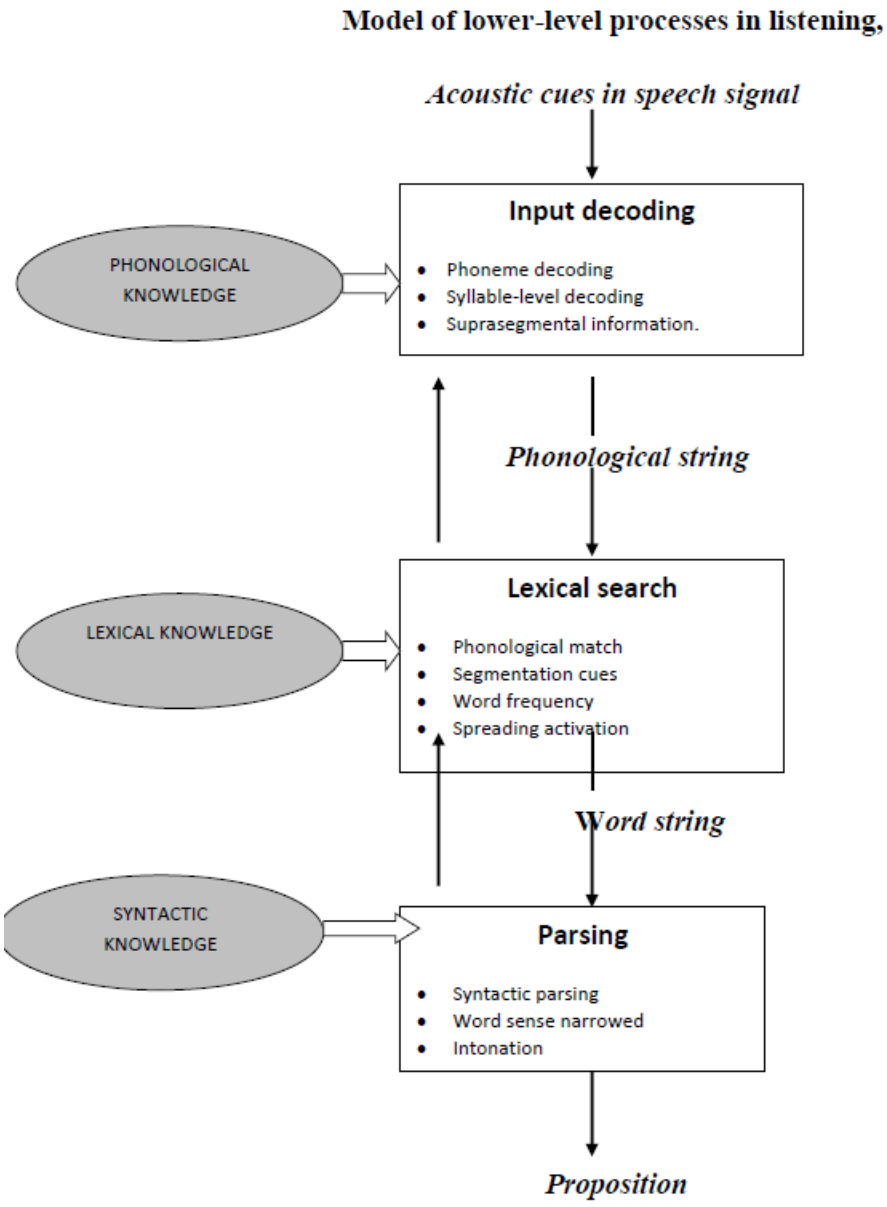
Teddlie, C., & Tashakkori, A. (2010). Overview of contemporary issues in mixed methods research. *Handbook of mixed methods in social and behavioral research*, 1-41.

Thaine, C., & McCarthy, M. (2012). *Cambridge Academic English B1+ Intermediate Class Audio CD and DVD Pack: An Integrated Skills Course for EAP*. Cambridge: Cambridge University Press.
Thorn, S. (2012). Debunking authentic listening. *Modern English Teacher, 21*(2), 65-69.
Thorn, S. (2013). *Real Lives, Real Listening, Elementary*: HarperCollins Publishers Limited.
Tsui, A., & Fullilove, J. (1998). Bottom-up or top-down processing as a discriminator of L2 listening performance. *Applied Linguistics, 19*, 432–451.
Tyler, L. K. (1984). The structure of the initial cohort: Evidence from gating. *Perception & Psychophysics, 36*(5), 417-427.
van Dijk, T. A., & Kintsch, W. (1983). *Strategies of Discourse Comprehension*. New York: Academic Press.
van Donselaar, W., Koster, M., & Cutler, A. (2005). Exploring the role of lexical stress in lexical recognition. *The Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 58A*(2), 251-273.
Vandergrift, L. (1997). The strategies of second language (French) listeners. *Foreign Language Annals, 30*, 387–409.
Vandergrift, L. (2002). Listening: theory and practice in modern foreign language competence. *Subject Centre for Languages, Linguistics and Area Studies*, 2009, from http://www.llas.ac.uk/resources/gpg/67#
Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language Teaching, 40*(3), 191-210.
Vandergrift, L., & Baker, S. (2015). Learner variables in second language listening comprehension: An exploratory path analysis. *Language Learning, 65*(2), 390-416.
Vandergrift, L., & Goh, C. C. M. (2012). *Teaching and Learning Second Language Listening: Metacognition in Action*: Routledge.
Voss, B. (1984). *Slips of the Ear*. Tübinger Guntar Narr Verlag.
Warren, R. (1970). Perceptual restoration of missing speech sounds. *Science, 167*(3917), 392-393. doi: 10.1126/science.167.3917.392
Weber, A., & Cutler, A. (2006). First-language phonotactics in second-language listening. *The Journal of the Acoustical Society of America, 119*(1), 597-607.
Wesche, M. B., & Paribakht, T. (2009). *Lexical inferencing in a first and second language: Cross-linguistic dimensions* (Vol. 46). Bristol: Multilingual Matters.
Wilson, M. (2003). Discovery listening - improving perceptual processing. *ELT Journal, 57*(4), 335-343.
Wray, A. (2002). *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press
Yeldham, M. (2009). *Approaches to second language listening instruction: Investigating the top-down/bottom-up debate*. Unpublished doctoral dissertation. University of Melbourne. Melbourne.
Yule, G. (1988). Highly confident wrong answering—and how to detect it. *ELT Journal, 42*(2), 84-88.

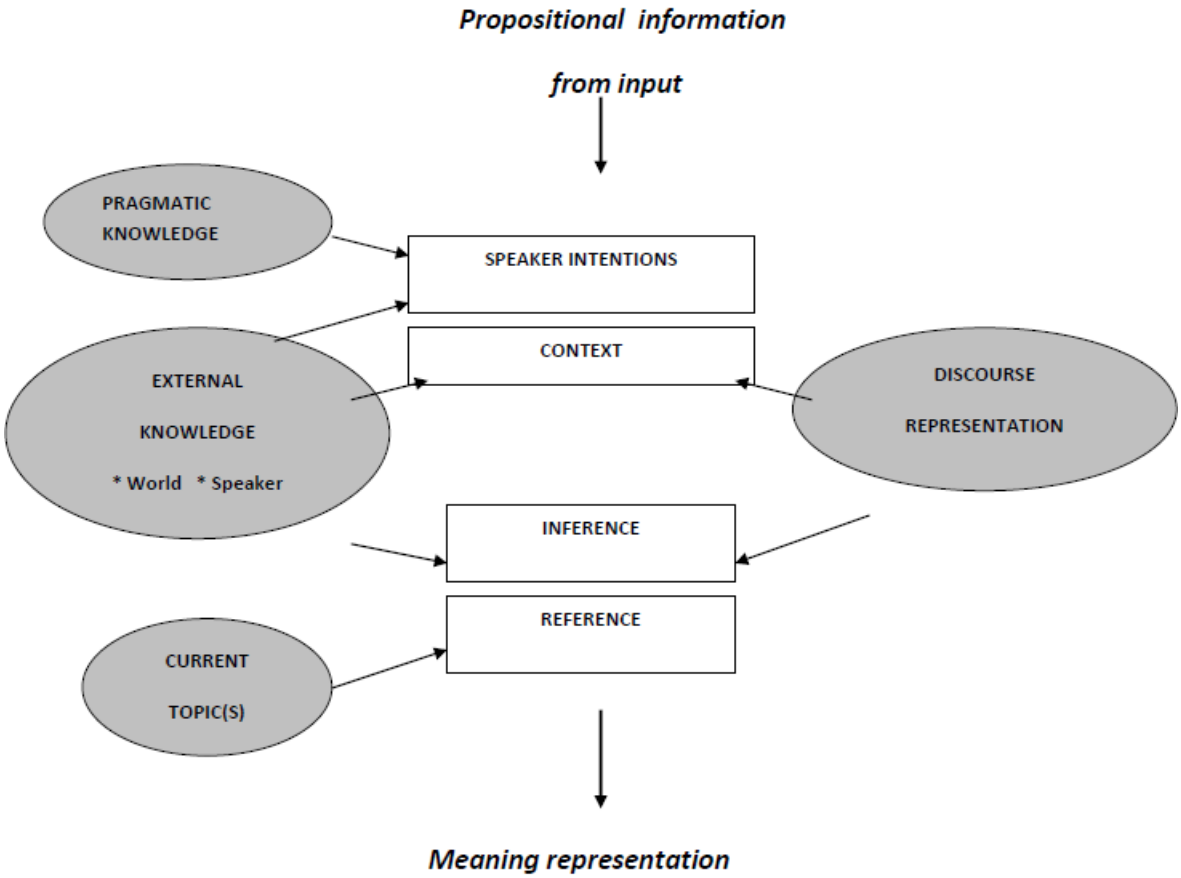
APPENDICES

APPENDIX A

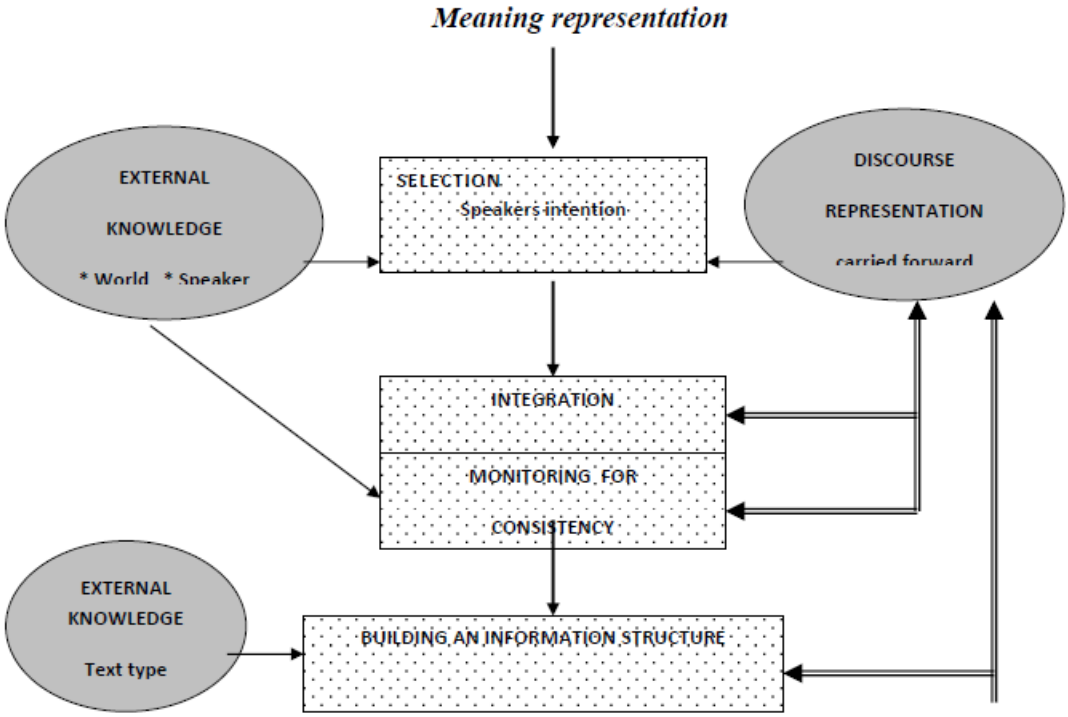
Model of listening processes
(drawing upon Cutler & Carter 1999 and Field 2008)



Model of meaning enrichment



Model of information handling in listening



(Field 2013)

Appendix B
Alphabetical list of common features of connected speech

Feature	Definition	Example
Assimilation	This is the process whereby one sound becomes like another neighbouring sound. It should be noted that coarticulation is a phenomenon closely related to assimilation; the major difference is that assimilation is used as a name for the process whereby one sound becomes like another neighbouring sound, while coarticulation, though it refers to a similar process, is concerned with articulatory explanations for why the assimilation occurs, and considers cases where the changes may occur over a number of segments.	*voice*: /hæv/ (have) becomes /I hæf tu/ (I have to) *place*: /ðæt/ (that) becomes /ðæk gɜ:l/ (that girl), /ðæt/ (that) becomes /ðæp bɔɪ/ (that boy) *manner:* /t/ /get sʌm əv ðæt səʊp/ (get some of that soap) becomes/s/ in two words /ges sʌm əv ðæs səʊp/
Cliticisation	Occurs at the morphological level. It is a process which combines two morphemes, but does not create a new word.	I am -> I'm cannot -> can't going to -> gunna
Coarticulation	This is concerned with articulatory explanations for why assimilation occurs, and considers cases where the changes may occur over a number of segments. It is the overlap of articulations and shows how surrounding phonemes are influenced by the target phoneme.	/ðæt/ (that) becomes /ðæk gɜ:l/ (that girl) /ðæt/ (that) becomes /ðæp bɔɪ/ (that boy)
Elision	Elision of vowels in English usually happens when a short, unstressed vowel occurs between voiceless consonants. In some cases we find a weak voiceless sound in place of the normally voiced vowel that would have been expected[1]. Elision also occurs when a vowel occurs between an obstruent consonant and a sonorant consonant such as a nasal or a lateral: this process leads to syllabic consonants[2]. Elision of consonants in English happens most commonly when a speaker "simplifies" a complex consonant cluster[3].	1. /pə'hæps/ (perhaps) /pə'teɪtəʊ/ (potato) /'baɪsɪkəl/ (bicycle) /fɪ'lɒsəfɪ/ (philosophy) 2. /sʌdn̩/ (sudden)/ /ɔ:fl/ (awful) 3. /ˈkʌpˌbɔrd/ (cupboard) becomes /ˈkʌbərd /ækts/ (acts) becomes /æks/ /twelfθnaɪt/ (twelfth night) becomes /twelθnaɪt/ or

		/twelfnaɪt/
Reduction	When a syllable in English is unstressed, it frequently happens that it is pronounced differently from the "same" syllable when stressed; the process is one of weakening, where vowels tend to become more schwa-like (i.e., they are centralised), and plosives tend to become fricatives.	'photograph' /ˈfəʊtəgrɑːf/, 'photography' /fəˈtɒgrəfi/, 'photographic' /ˌfəʊtəˈgræfɪk/ – when one of the three syllables does not receive stress its vowel is reduced to /ə/.
Resyllabification	This is a phonological process in which consonants are attached to syllables other than those from which they originally came.	went in -> wen tin tried out -> try dout

Definitions and examples are adopted from: Roach, P. (2009) English phonetics and phonology glossary (A little encyclopaedia of phonetics)**.** Available at: www.***cambridge.org/elt/peterroach***/resources/Glossary.

Appendix C

Verbal report: transcriptions of participants' responses to the question: "are you sure of those words?"

Recording 1
P = Participant

P 5

- So, the representation is important... sure
- order your response. Respond… Yeah, I'm sure.
- By the market presearch… No; I remembered there has another word, but I didn't grab it.
- We didn't pru… I just sure about the last one.
- I will go the details, and this is per personal (pause) detaily... no it is not that. I think from the first word to the except the last word. I think I heard them.
- Got to trouble of that. Got trouble for that… Yeah
- It is really important the present thing… I heard it in the metre of the sentence and er, yeah.
- (inaudible.)… No.
- The styles of the questions and I also hear two hundred; two percent people… Yeah. I think.

P 15

- Exactly presentation is very important... The first maybe not correct exactly, the other I think is right.
- Optimum misponse... yes I don't know what it is mainly about optimum misponse, I think I don't know.
- Flows the questionnaire... maybe flow, follow, follow the questionnaire; I think is this.
- By the market research process... maybe. I guess; yes, all right yeah.
- Impricit approval... lots of problems, I don't know; every word I don't know exactly what words they are.
- Pretty personal question really... large part; maybe the last word.
- You can imagine got a little trouble bout it... of course I'm not sure, but I think the larger meanings is trouble about student.
- You need to pretest there... some problem I think, I don't know what the meaning, I think it's not the whole sentences, I think the comma is the, interrupts the whole sentences, so I don't know what the talk about.

- Statistically exam; example... I'm not sure, large problem, I don't know the words meaning.
- The style of the particular questions... A large part I'm sure, particular maybe is not a…
- Pretesting is very important... maybe.

P 17

- Presentation is very important… sure
- Up to the shmultz… No. I can't follow it, this term.
- An the flow questionnaire… Most are correct. I think there may be some word in flow and questionnaire.
- By the market procer... research process. Mostly sure.
- Impersistant of approval… Not very sure. The word before proover. Or some word the nexis. Happy with second word
- Very private questions… Be sure, yeah.
- The trouble bout it... Mostly sure.
- You need to pretest thing… I'm sure.
- If you want a sasisterly.... I missed it. Not sure.
- The style of these questions…. One or two word not sure. I think the word before question.

P 18

- Presentation is very important… sure
- Optimum response… Not very sure. Optimum, optimal, I don't know the word.
- Flows the questionnaire... No, not sure of flows.
- Market process…Yeah. Yeah, the market process.
- Presit approval. No, presit… I don't know what that is. A preset. Approval I'm sure.
- Very personal questions… sure
- How to look trouble about it… No. The word before trouble.
- You need to pretest a thing…. Not sure.
- Statixicly valance examper… yeah. Maybe example, sample, I'm not sure.
- The style of the questions… Very sure. Yeah.

P 26

- Presentation is very important... yes.
- Optenser response... yes I'm sure.
- Flows the questions... I'm not sure of the word... flow.

- Called the market researches... I'm so not sure... before the research, that word.
- Implist ofproval, improve... before the improve, that word is not exactly.
- Very personal questions...yes...ok.
- Lot of troubles... I'm not sure... something troubles, I'm not sure before the trouble, sorry.
- You need the protesting... yeah.
- Statifically... I'm not sure... yes… certifically.
- The style of the questions... I'm sure.

P 28

- Presentation is very important... No problems.
- Optimus response... yes.
- The float of question air... The first word I didn't hear correctly.
- The mark research programme... Probably correct.
- A present improval... Yeah, I just heard the last word, the first two not I'm not sure.
- Very clear personal question... The second word I don't know... the rest is ok.
- Get trouble about it... No it's correct.
- You need to pretest things... the last two words is not sure.
- But this thee was (pause) examples... not sure the whole sentence I got right... Example is right...
- The style of particular questions... no problems.

P 29

- The presentation is very important... I think this is correct.
- Octem response... Yeah, this time I had a problem on the first word.
- The flow of the questionnaire... Yeah I think.
- Coller down by the market research (pause) something… Yeah the last word I can't. The name I don't know exactly... colled down.
- Impricit approval.... The first word I can't pronounce exactly.
- Pre; pre testing reading... The first word, I think I can't hear. I just remember it as three words.
- Its got trouble about it... The words before the trouble I can't hear either.
- You need to pretesting... The pretesting I don't know how to spell it, I don't know the word.
- Existing example... The first word no, example.
- Style of the questions... I think this is sure.

P 30

- Presentation is very important... I think it's correct.
- Optimember response... optimember response... no. yeah obitum, (*spells out*) o.b.t.i.a.m.
- The flow of the questionnaire... yeah.
- By the market research process... yeah, yeah.
- To do impleasant approval... impleasant maybe... approval, I think I know, yeah.
- Wears of questions really... no, I think, yeah.
- A globe of trouble about it… word of trouble, I'm missing the word... before trouble.
- You need to protest the thing... no problems.
- Statistically valid sample... (*says words quietly to themself*) yeah.
- The stickule of questions; I'm missing the word... the; and the questions between... between the and questions I mean.

P 32

- The presentation is very important... yeah.
- Optimum sponse... Optimum, I know this word but I'm not sure how to spell it. The other word no.
- He wants to follow the questionnaire... I'm sure about question yeah.
- By the market research process...very sure.
- Implicit improval... I'm not sure. the second word.
- Very clearly personal questions... maybe. yeah, all of them.
- Trouble bout it... yeah, I'm sure these two words but; others I'm not sure.
- You need to pretest a... maybe. the last word I'm not sure.
- Statistically valid example... I'm not sure the middle words.
- The style of particular questions... I'm very sure.

P 34

- Presentation is very important... yes.
- In order to get some response... yes. miss words at the beginning I think.
- The flow of the questionnaire... correct I think.
- By the market research precess... I'm not sure about the last word, precess.
- To get approval... I think some words missed before approval. yeah but I'm not sure, yeah.
- But it was asking very pre, pre-sessional course questions, sorry. No, I thought it were too long at the beginning and I can't catch the key words.

- A little bit trouble bout it... yeah.
- You need to protest the thing... yeah.
- Statistically balid sample... yeah.
- sties of particular questions... yeah.

Recording 2

P 1

- Controlling many of this… Yes sure of controlling
- Stop people getting sick… Yes
- Life in the UK is true… missing one word or (pause). is true, yeah, yeah.
- The number of prespection… I'm not sure of the last word.
- We are not attacking… Maybe attacking, I'm not sure.
- During the same period of time… Er (pause) ok.
- Disastrous… ok. (researcher: And before that?) Er are I didn't...
- Expecancy… No, not sure.
- Written a drop… not sure
- Concentrate on how to cure it… Yeah.

P 3.

- controlling menu illnesses… Er (pause) I just can manage the key words. The firsterly is erm, what. The secondly and the, the drug without the secur, security of this.
- The second, er first place… Er, a, no, not very correct.
- increasing that is true… Yes.
- The number of per, er, persquiptions… Yes.
- We are not talking… Erm I'm not; I'm not sure.
- A period of time… Er yeah, it's a difference between er, female and males with…
- Unhealthy er um unhealthy, er dairy… (Pause) Er sorry I'm not sure.
- I just catched last word agency… Yeah
- Have vigness a drug… Yes
- How to cure it… Yes

P4

- Or at least controlling many orifist… No, the last word is difficult.
- If stopping people, getting, in the first place… The preposition before the first place or

something.

- Increasing that is true… Yes, I'm sure.
- We can see the number of presciptions… The last word again presciption. Preskiptions.
- We are not attacking… Yes.
- Increase 70 percent by over period of time…I'm quite sure.
- The faster… Yes. The faster.
- And what does it mean in term of life expectancy… Yes. I can hear all them.
- Norway for instant has witnessed a drop… I'm quite sure.
- We should concentrate only how, not only how to cure it… I'm not sure concentrate, at the beginning of this.

P 8

- Controlling many annoratives... I'm not sure about the last word.
- Is stop people getting in the first place... maybe the word is single or plural, does it have s or no.
- The life in the UK does is increasing that is true... yeah.
- We can see the number prosciption... the last word. prosciption something like that.
- We are not attacking... yeah.
- Over the same period of terms... I'm happy.
- Disaster... I'm not sure… the word before but I can't remember it.
- Like affectancy... the last word... I hear but I'm not sure about the word.
- Has this is the drop... the middle word is like a word and I can't catch.
- How to kill it... yeah. yeah.

P11

- Controlling many of this is… I'm not sure actually. Some of this is.
- Stopping people, to take this place… sure of this place.
- Is increasing, slat true… I'm sure.
- We can see the numbers of craptions… The last word I'm not sure.
- tacking…yeah I'm sure.
- same period of time…No difficulties.
- Is laxitous. I think it's difficult.
- this is …The last word. I don't know.
- For instance erm, the drop…I missed a word. After for instance.

- how to cure it … sure

P 19

- Controlling many inaudible… I didn't catch the last word. I think that's the most important word.
- Is stop people from getting the first place… difficulties? Only one word. Getting what.
- increasing and that is true… yes
- The number of crespeshons… I don't know the meaning of the last word.
- To emphasise that we are not, we are not… one word missing.
- same period of time… sure
- Unhealthy diet is a disaster. Yeah, many problems.
- in terms of life expectancy… sure
- witnessed a jot… yes
- how to cure it… The last two... cure it. I didn't catch it cure, is the meaning.

P 20

- Controlling many... I don't think the last word is correct.
- Getting the first place… Yes, I think so.
- inaudible
- number of prescriptions… sure
- we are not attacking… I'm not sure about the last words.
- Over the same periods of the time… sure
- Healthy and unhealthy are disaster, problem has disastrous… I'm not quite sure in the middle.
- What terms of means was this expectancy… Yes I'm sure.
- Has witnessed a drug. I'm not quite sure. The last word.
- Concentrate … (no response)

P21

- (Did not understand what to do and did not answer)
- Getting sick on first place… Er, yes I think so.
- Is that true… No problems
- Number of prescription… sure
- We are not talking… I think so.

- same period of time… it might be
- And disastrous… Before the disastrous there were some words, but I couldn't catch it.
- The meaning of (inaudible) expectancy… No problems.
- Norway had drop… I couldn't catch one word.
- how to cure it… sure

P 24

- Controlling many undersis… Not sure. About heart disease.
- Sick in the first place… Not sure.
- The life in the UK increasing, that true… true yes, but I missed some words.
- The number of perscription… Yes, yes.
- We are not attacking… sure
- Increase 70% in the same period of term…Yes. I think it's all the words.
- (inaudible) gray assess…This I'm not sure.
- In terms of life expectancy… sure
- Norway insistent a drop… I'm not sure. The last words. I think I missed some words.
- How to prevent rather than how to kill it… Maybe, yes.

P 27

- Controlling many of this... I'm sure... yes.
- The first place...yes... the first place, yeah.
- That is true... yeah.
- The number crab; I don't know the last word... yeah...
- We are not tacking... No. the last one.
- Over the same period of time... yeah.
- And a what disaterous... I don't know what it means... the whole.
- Last… the last word I can't catch it. the final one I forgot.
- Has witness a draft... no. the last word.

- How to cure it... sure.

P 31

- Many ullnesses... I maybe have some problem. may be sometimes I can listen this word, but I can't understand the whole sentence meaning.
- In the first place...sure.
- Is increasing, that is true...may be sure.
- The number of er, questions...no, the last word, I'm not sure.
- We are not attacking...yeah.
- In the… of time, I'm not sure of the last three words.
- The disastrous...yeah. yeah, I just remember the last two words.
- Expense it (*pause*)… no (*pause*)… sorry.
- Has been a draf... something a draf.yes
- How to cure it...yeah maybe.

Appendix D

Common European Framework of Reference for Languages: Listening descriptors

Common Reference Levels: global scale

B2	Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialisation. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. Can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and Independent disadvantages of various options.
B1	Can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc. Can deal with most situations likely to arise whilst travelling in an area where the language is spoken. Can produce simple connected text on topics which are familiar or of personal interest. Can describe experiences and events, dreams, hopes and ambitions and briefly give reasons and explanations for opinions and plans.

Common Reference Levels: self-assessment grid – the listening skill

	B2	I can understand extended speech and lectures and follow even complex lines of argument provided the topic is reasonably familiar. I can understand most TV news and current affairs programmes. I can understand the majority of films in standard dialect.
	B1	I can understand the main points of clear standard speech on familiar matters regularly encountered in work, school, leisure, etc. I can understand the main point of many radio or TV programmes on current affairs or topics of personal or professional interest when the delivery is relatively slow and clear.

https://www.coe.int/en/web/common-european-framework-reference-languages/home

Appendix E

Participant response sheets

Task 1

This exercise is designed to practise your listening skills as part of your ISLC programme. I am a listening researcher and would like to use the answers you give for my research project.

Your answers will be treated as anonymous. You will not be referred to by name, only by a number.

If you **are happy** for me to use your answers, please sign here _____

If you **are not happy** for me to use your answers, please sign here _____

Thank you, Jane Ward

Family name: _____ First name: _____

First language: _____

Is this your first time in an English-speaking country: YES / NO

If it is **not** your first time in an English-speaking country, how long in total have you spent in an English-speaking country: _____

You will hear a section of a recorded lecture which has 10 pauses within it. When you hear a pause you should write the last 4 or 5 words you heard. You **don't** have to write the whole sentence you heard, just the **last 4 or 5 words**. Don't worry about your spelling. The pauses are quite short, so you should start to write as soon as you hear the pause. You have 1 line for each pause. **Do not change your answers after the recording**. You will hear the recording once only.

---

1) _____

2) _____

3) _____

4) _____

5) _____

6) _____

7) _____

8) _____

9) _____

10) _____

PLEASE DO NOT CHANGE YOUR ANSWERS

Task 2

Student Consent Form

You have kindly agreed to take part in a research project which aims to improve the teaching of listening.

You will be asked to complete a short exercise which is designed to practise your listening skills. I am a listening researcher and would like to record and use the answers you give for my research.

If you are happy for me to use your answers, please sign here

_____

Thank you,

Jane Ward

Family name: _____ First name: _____

First language: _____

Is this your first time in an English-speaking country: **YES / NO**

If it is **not** your first time in an English-speaking country, how long in total have you spent in an English-speaking country: _____

You will hear a section of a recorded lecture which has 10 pauses within it.

When you hear a pause you should repeat the last 4 or 5 words you heard. You **don't** have to repeat the whole sentence you heard.

The pauses are quite short, so you should start to speak as soon as you hear the pause.

You will hear the recording once only.

APPENDIX F

Coded data for written task

See accompanying USB

Appendix G
Transcriptions of recordings

Recording 1 (pauses in italics)

now I'm gonna show you lots of examples of different types of questions that you can ask and here are some very general design issues though your questions they need to be precise as you'll see in a moment they need to be well ordered and I'm gonna show you some examples of good and bad questions in a moment you need to decide very carefully I think on the ordering of I think there's not really any excuse for it these days in a sense for getting this part of it wrong and certainly *presentation is very important*

so will talk a little bit about presentation and how you are able to order your questions to make sure that *you get optimum response*

so you've got a set very clear objectives as to what your questionnaire is designed to achieve you need to say something about how you're going to collect the data er the sorts of questions that you can have the way that you word them *the flow of the questionnaire*

and so on obtaining approval is very important in the University we have a body known as the ethics committee technically speaking if you go out if you go outside the university to research anything you need to get the approval of the ethics committee and the ethics committee is in many ways a very good idea the University and indeed any market research body that doesn't want its name pulled down *by the market research process*

so we have two if we get a go out and indeed if students are doing projects we have *to get the implicit approval*

of the ethics committee sometimes that can come from the head of the department but two or three years ago just for your information the group er this group was actually a group of undergraduate students erm they decided to do market research project which er was part of the assessment for the course and they were given a free choice as to what subject they wanted to ask people about and the explicit instruction was that the people they researched should only be members of the course and this group came and said we want to do a kind of a personal survey and what this was it was it was actually fairly innocent although I did say you know this must be kept strictly within the group and it sort of I won't go into details but it was asking *various pretty personal questions really*

and the next thing I heard the next thing I heard of it was that somebody called up from I can't remember where but they had actually been accosted by one of the students somewhere downtown and been asked these questions and as you can imagine we got into *a little bit of trouble about it*

and we hadn't cleared it I hadn't cleared its basically with the committee because it wasn't well I didn't believe it was going out I didn't think it was going outside so there are there are obviously very potential problems in that but you do need to obtain approval there is a market research society code on the practice of asking questions on how to do research you need to pre-test the questionnaire this is really important those of you some of you will be doing this your dissertation and some of you I know collecting primary data *you need to pre-test the thing*

because you're the researcher you very close to the subject you know what you're talking about you've got to check that other people do as well and if you want *a statistically valid sample*

of one hundred or 200 people then you've got a make sure that your collecting the data properly and it's here that these pre-tests or pilots they are gonna tell you whether it's gonna work on not so make sure whether you do pilots and this can be half a dozen different people that you question I mean you'll soon find out whether you've got any potential or any doubts about the length of the questionnaire *or the style of particular questions* or whether the sort of questions that you're asking are valid you'll soon find out from that so piloting or pretesting is really important

Recording 2 (pauses in italics)

what I would like to do now is look at the health situation in the developed world with particular reference to the United Kingdom I think the situation can be summarised briefly like this firstly life expectancy how long people are living is increasing secondly we are taking more and more drugs and as a result of this we are curing or *at least controlling many illnesses*

however what we are not doing as well as we should is stopping people *getting sick in the first place*

let me just illustrate this point with some statistics I've said that life expectancy in the UK is increasing *and that is true*

for example let's look at men aged between 35 and 74 the number of men in this age group who died dropped by 42% between 1990 and 2000 now that is a huge fall 42% fewer deaths in this age group over a ten-year period now it is clear to me that much of this fall has been due to the amount of drugs we take now to cure problems if we look at heart disease for example and the drugs we take to regulate or cure it we can see that *the number of prescriptions*

issued by doctors has almost quadrupled increased by just under 400% in the last 20 years this includes drugs to lower blood pressure and to reduce cholesterol so we really are becoming a nation of pill takers but and this is the point I want to emphasise *we are not attacking*

the underlying causes of heart disease one major cause of heart disease is physical inactivity and in the UK we are becoming more and more inactive we are taking less physical exercise if you look at the statistics on your handout you will see these illustrate that since the 1970s the average number of miles travel on foot has dropped by around a quarter to just about 23% and the number of miles travelled by bike has dropped by one third in other words we are walking less we are cycling less by contrast the number of miles people drive has increased by 70% over *the same period of time*

so more use of the car and less physical exercise is the overall picture add to this inactivity an unhealthy diet and *the results are disastrous*

look at the figures for obesity in the UK the percentages of obese adults has almost doubled in the last 12 years a rise of about 92% so as a nation we are becoming more open piece as a result of poor diet and a lack of regular physical activity and what does this mean *in terms of life expectancy*

Well over 100,000 people die every year as a result of heart disease and a third of these deaths so more than 33,000 deaths according to the British Heart Foundation are premature in other words people are dying earlier than they should do so here in the UK we could do more other countries are already doing more Norway for instance *has witnessed a drop*

of 54% in the number of deaths in men aged between 35 and 74 in the last 10 years of the 20th century and as we saw earlier in that same age group Britain has a figure of 42% so although that seems good we could and we should be doing more and we should be looking at how to prevent heart disease rather than concentrating only on *how to cure it*

Appendix H
Instructions for teachers

Tell students they are doing an exercise based on a segment from a lecture about *the design of questionnaires for research/health in the developed world*.

Tell them I would like to use their answers for my research. If they agree, they should **sign** in the first space at the top of the form and if they don't want me to use them, they should sign in the second space at the top of the form. Hand the students the **response sheet.**

There is a box for them to complete with **personal info** but please remind the students their responses will be anonymous (as it says at the top of the form).

Under the personal info box is **instructions for the task**. Please ensure they understand and emphasise that they shouldn't change their answers (it's not a test!) and **play the example pause** so they know what to expect.

Play the recording.

At the end, please collect the papers and check they have signed the top.

Many thanks

Jane

j.e.ward@student.reading.ac.uk

Appendix I
Application to Ethics Committee and consent

School of Languages and European Studies
Department of Applied Linguistics

University of Reading

ETHICS COMMITTEE

Project description

Context

Little is known about second language (L2) listening despite its importance in academic performance. This interdisciplinary study applies insights from empirical psycholinguistic research in order to increase our understanding of the listening processes of L2 learners in academic contexts. Its goal is to identify which aspects of spoken input are most salient to L2 listeners and to construct a taxonomy of areas of difficulty to be targeted by instructors.

Subjects

The Subjects will be EAP students taking part in the pre-sessional courses in the International Study and Language Centre (ISLC) in 2010. In 2009 there were approximately 270 students and a similar number is anticipated in 2010. All the students will be approached for consent.

Method of approaching subjects

ISLC Course Directors have agreed that the research purposes can be explained and information/consent forms distributed to the subjects during the course induction.

Procedures – Task 1

Using an established procedure, subjects listen to an authentic recorded lecture into which pauses have been inserted at irregular intervals. Whenever a pause occurs, they transcribe the last four or five words heard. The recording will be approximately five minutes long and ten pauses will be inserted in each.

Procedures – Task 2

Approximately 15 subjects will be asked to perform a second task. A third recording will be played to a single participant who will repeat the last four or five words heard before the

pause. The researcher will discuss the answers given with the participant. The task will be recorded.

Pedagogical value of tasks

The Director of the ISLC supports this research and it was agreed that, as the tasks have pedagogical value, they would be included within the curriculum and thus would not detrimentally impact the summer courses.

Data storage and access

A data set of the words transcribed will be recorded by the researcher using an Excel spreadsheet on a home PC. The data will be accessible to two people via USB drive: the researcher and the researcher's supervisor. In addition, the analysed data will be subject to inspection by a judge to ensure reliability of analysis (University of Reading staff member). The USBs will be returned after use.

Anonymity is not possible as the researcher will need access to the subjects' identity in order to be able to consider implications of nationality and length of time spent in an English speaking country. However confidentiality will be guaranteed.

The data will be retained indefinitely by the researcher.

Pilot study

A pilot study will take place during week 7-8 of the summer 2010 term. The subjects will number approximately 12 and will be of an identical nature to those of the main study.

Memo

University of Reading
School of Languages and European Studies
Department of Applied Linguistics

To Jane Ward

Please reply to: Vicky Chondrogianni

From Vicky Chondrogianni

Copy John ~~Filed~~ Field
Alan Tonkyn

Date 31st August 2010

Your application for Ethical Approval

Your project entitled **'A study of second language listeners in an academic context'** has been considered by the School Ethics Committee, and we are pleased to report that the Committee has raised no ethical objections, and subject to your undertaking to store of the consent forms in the Department Office in the normal way, it has accordingly given permission for the project to proceed under the exceptions procedure as outlined in paragraph 6 of the *University's Ethics Guidance to Schools*.

Signed

Alan Tonkyn

Vicky Chondrogianni

On behalf of the School Ethics Committee,
Prof. Mary Bryden, School Director of Research,
Prof. Christopher Duggan, Head of School

Appendix J

Participant ethics information sheets and consent forms

School of Languages and European Studies
Department of Applied Linguistics

Participant Consent Form

Project title: **A study of second language listeners in an academic context**

1. I have read the Information Sheet relating to this project.

2. I understand the purposes of the project and what will be required of me, and any questions have been answered to my satisfaction.

3. I understand the arrangements described in the Information Sheet in so far as they relate to my participation.

4. I understand that my participation is entirely voluntary and that I have the right to withdraw from the project any time.

5. I have signed my response sheet to indicate my agreement to participate in this project.

Department of Applied Linguistics

School of Languages and European Studies
HumSS Building
The University of Reading
Whiteknights, PO Box 241
Reading RG6 6AA

Phone 01183788141
Email appling@reading.ac.uk

Researcher:
Jane Ward
Email: j.e.ward@student.reading.ac.uk

Supervisor:
Dr.John Field
Email: j.c.field@reading.ac.uk

p

INFORMATION SHEET TASK 1

All International Study and Language Centre (ISLC) students are being asked to take part in a research project which aims to improve the teaching of listening.

As part of your Listening and Speaking programme you will be given an exercise to practise your listening skills. I am a Listening researcher and would like to use the answers you give for my research.

If you agree for your answer papers to be used for the research, please sign the form on the next page.

If you sign the form but then decide you DO NOT want your answer papers to be used for the research, please tell your Listening and Speaking teacher as soon as possible.

Your answer papers will be treated confidentially and will be stored securely in the School of Languages and European Studies after the project.

If this research is published, students' names will NOT be given. If you wish, you may see the results of the research when it is completed and should contact me by email.

This project has been subject to ethical review by the University Ethics and Research Committee, and has been allowed to proceed.

If you have any queries or wish to clarify anything about the study, please feel free to contact my supervisor at the address above or by email at j.c.field@reading.ac.uk

Researcher:
Jane Ward
Email: j.e.ward@student.reading.ac.uk

Supervisor:
Dr.John Field
Email: j.c.field@reading.ac.uk

INFORMATION SHEET TASK 2

A small group of International Study and Language Centre (ISLC) students are being asked to take part in a research project which aims to improve the teaching of listening.

I am a Listening researcher investigating students' abilities to understand natural speech. I would like you to listen to a short real-life recording and at certain times you will be asked to repeat what you have heard. I will then briefly discuss your answers with you. This exercise will be recorded.

In return for your help, I am offering a specialised listening lesson which will help you to improve your listening abilities.

If you agree to assist in this research, please sign the form on the next page. If you agree and sign the form but then decide you DO NOT want to take part in the research, please tell me or your ISLC teacher as soon as possible.

The recordings will be stored securely in the School of Languages and European Studies after the project.

If this research is published, students' names will NOT be given. If you wish, you may see the results of the research when it is completed and should contact me by email.

This project has been subject to ethical review by the University Ethics and Research Committee, and has been allowed to proceed.

If you have any queries or wish to clarify anything about the study, please feel free to contact my supervisor at the address above or by email at j.c.field@reading.ac.uk

APPENDIX K

Inter and intra rater coded data

See accompanying USB

APPENDIX L

Coded data for oral task

See accompanying USB

Appendix M

Regression output: lexical features of correct responses

Recording 1 Multiple Regression

	Percent correct responses	Frequency	Word length	Initial stressed syllable
Percent correct responses	1.00	.433	.008	-.005
Frequency	.433	1.00	-.009	-.022
Word length	.008	-.009	1.00	-.510
Initial stressed syllable	-.005	-.022	-.510	1.00

Correlations				
		Percent correct responses	Frequency	Word length
Pearson Correlation	Percent correct responses	1.000	.433	.008
	Frequency	.433	1.000	-.009
	Word length	.008	-.009	1.000
	Initial stressed syllable	-.005	-.022	-.510
Sig. (1-tailed)	Percent correct responses	.	.025	.487
	Frequency	.025	.	.485
	Word length	.487	.485	.
	Initial stressed syllable	.491	.462	.009
N	Percent correct responses	21	21	21
	Frequency	21	21	21
	Word length	21	21	21
	Initial stressed syllable	21	21	21

Correlations		
		Initial stressed syllable
Pearson Correlation	Percent correct responses	-.005
	Frequency	-.022
	Word length	-.510
	Initial stressed syllable	1.000
Sig. (1-tailed)	Percent correct responses	.491
	Frequency	.462
	Word length	.009
	Initial stressed syllable	.
N	Percent correct responses	21
	Frequency	21
	Word length	21
	Initial stressed syllable	21

Variables Entered/Removed[a]			
Model	Variables Entered	Variables Removed	Method
1	Initial stressed syllable, Frequency, Word length[b]	.	Enter

a. Dependent Variable: Percent correct responses

b. All requested variables entered.

Model Summary[b]						
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics	
					R Square Change	F Change
1	.434[a]	.188	.045	25.697	.188	1.312

Model Summary[b]			
Model	Change Statistics		
	df1	df2	Sig. F Change
1	3[a]	17	.303

a. Predictors: (Constant), Initial stressed syllable, Frequency, Word length

b. Dependent Variable: Percent correct responses

Coefficients[a]						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	41.663	14.262		2.921	.010
	Frequency	.008	.004	.434	1.984	.064
	Word length	.958	13.173	.018	.073	.943
	Initial stressed syllable	.733	13.426	.014	.055	.957

Coefficients[a]				
Model		Correlations		
		Zero-order	Partial	Part
1	(Constant)			
	Frequency	.433	.434	.434
	Word length	.008	.018	.016
	Initial stressed syllable	-.005	.013	.012

a. Dependent Variable: Percent correct responses

Recording 1 Stepwise Multiple Regression

Correlations				
		Percent correct responses	Frequency	Word length
Pearson Correlation	Percent correct responses	1.000	.433	.008
	Frequency	.433	1.000	-.009
	Word length	.008	-.009	1.000
	Initial stressed syllable	-.005	-.022	-.510
Sig. (1-tailed)	Percent correct responses	.	.025	.487
	Frequency	.025	.	.485
	Word length	.487	.485	.
	Initial stressed syllable	.491	.462	.009
N	Percent correct responses	21	21	21
	Frequency	21	21	21
	Word length	21	21	21
	Initial stressed syllable	21	21	21

Correlations		
		Initial stressed syllable
Pearson Correlation	Percent correct responses	-.005
	Frequency	-.022
	Word length	-.510
	Initial stressed syllable	1.000
Sig. (1-tailed)	Percent correct responses	.491
	Frequency	.462
	Word length	.009
	Initial stressed syllable	.
N	Percent correct responses	21
	Frequency	21
	Word length	21
	Initial stressed syllable	21

Variables Entered/Removed[a]			
Model	Variables Entered	Variables Removed	Method
1	Frequency	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).

a. Dependent Variable: Percent correct responses

Model Summary[b]						
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics	
					R Square Change	F Change
1	.433[a]	.188	.145	24.311	.188	4.391

Model Summary[b]			
Model	Change Statistics		
	df1	df2	Sig. F Change
1	1[a]	19	.050

a. Predictors: (Constant), Frequency
b. Dependent Variable: Percent correct responses

Coefficients[a]						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	42.537	6.765		6.288	.000
	Frequency	.008	.004	.433	2.095	.050

Coefficients[a]						
Model		Correlations			Collinearity Statistics	
		Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)					
	Frequency	.433	.433	.433	1.000	1.000

a. Dependent Variable: Percent correct responses

Excluded Variables[a]						
Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics
						Tolerance
1	Word length	.011[b]	.054	.958	.013	1.000
	Initial stressed syllable	.004[b]	.021	.984	.005	1.000

Excluded Variables[a]			
Model		Collinearity Statistics	
		VIF	Minimum Tolerance
1	Word length	1.000[b]	1.000
	Initial stressed syllable	1.000[b]	1.000

a. Dependent Variable: Percent correct responses
b. Predictors in the Model: (Constant), Frequency

Collinearity Diagnostics[a]					
Model	Dimension	Eigenvalue	Condition Index	Variance Proportions	
				(Constant)	Frequency
1	1	1.621	1.000	.19	.19
	2	.379	2.067	.81	.81

a. Dependent Variable: Percent correct responses

Residuals Statistics[a]					
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	42.54	83.19	51.33	11.391	21
Residual	-47.815	38.445	.000	23.696	21
Std. Predicted Value	-.772	2.796	.000	1.000	21
Std. Residual	-1.967	1.581	.000	.975	21

a. Dependent Variable: Percent correct responses

Recording 2 Multiple Regression

Correlations				
		Percent correct response	Frequency	Word length
Pearson Correlation	Percent correct response	1.000	.349	-.060
	Frequency	.349	1.000	-.368
	Word length	-.060	-.368	1.000
	Initial stressed syllable	.062	.348	-.663
Sig. (1-tailed)	Percent correct response	.	.066	.401
	Frequency	.066	.	.055
	Word length	.401	.055	.
	Initial stressed syllable	.398	.066	.001
N	Percent correct response	20	20	20
	Frequency	20	20	20
	Word length	20	20	20
	Initial stressed syllable	20	20	20

Correlations		
		Initial stressed syllable
Pearson Correlation	Percent correct response	.062
	Frequency	.348
	Word length	-.663
	Initial stressed syllable	1.000
Sig. (1-tailed)	Percent correct response	.398
	Frequency	.066
	Word length	.001
	Initial stressed syllable	.
N	Percent correct response	20
	Frequency	20
	Word length	20
	Initial stressed syllable	20

Variables Entered/Removed[a]			
Model	Variables Entered	Variables Removed	Method
1	Initial stressed syllable, Frequency, Word length[b]	.	Enter

a. Dependent Variable: Percent correct response
b. All requested variables entered.

Model Summary[b]						
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics	
					R Square Change	F Change
1	.357[a]	.128	-.036	25.718	.128	.781

Model Summary[b]			
Model	Change Statistics		
	df1	df2	Sig. F Change
1	3[a]	16	.522

a. Predictors: (Constant), Initial stressed syllable, Frequency, Word length
b. Dependent Variable: Percent correct response

Coefficients[a]						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	37.975	17.380		2.185	.044
	Frequency	.002	.001	.382	1.504	.152
	Word length	3.144	16.432	.061	.191	.851
	Initial stressed syllable	-1.635	16.959	-.030	-.096	.924

Coefficients[a]				
Model		Correlations		
		Zero-order	Partial	Part
1	(Constant)			
	Frequency	.349	.352	.351
	Word length	-.060	.048	.045
	Initial stressed syllable	.062	-.024	-.023

a. Dependent Variable: Percent correct response

Residuals Statistics[a]					
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	36.44	77.14	44.00	9.030	20
Residual	-39.282	34.011	.000	23.600	20
Std. Predicted Value	-.837	3.670	.000	1.000	20
Std. Residual	-1.527	1.322	.000	.918	20

a. Dependent Variable: Percent correct response

Recording 2 Stepwise Multiple Regression

No variables were entered into the equation.

Correlations				
		Percent correct response	Frequency	Word length
Pearson Correlation	Percent correct response	1.000	.349	-.060
	Frequency	.349	1.000	-.368
	Word length	-.060	-.368	1.000
	Initial stressed syllable	.062	.348	-.663
Sig. (1-tailed)	Percent correct response	.	.066	.401
	Frequency	.066	.	.055
	Word length	.401	.055	.
	Initial stressed syllable	.398	.066	.001
N	Percent correct response	20	20	20
	Frequency	20	20	20
	Word length	20	20	20
	Initial stressed syllable	20	20	20

Correlations		
		Initial stressed syllable
Pearson Correlation	Percent correct response	.062
	Frequency	.348
	Word length	-.663
	Initial stressed syllable	1.000
Sig. (1-tailed)	Percent correct response	.398
	Frequency	.066
	Word length	.001
	Initial stressed syllable	.
N	Percent correct response	20
	Frequency	20
	Word length	20
	Initial stressed syllable	20

Appendix N

Paired samples t tests: participants' use of perceptual, non-perceptual, co-textual and contextual cues Recordings 1 and 2

Recording 1 Wilcoxon Signed Ranks Test

Perceptual v non-perceptual

	t	df	Sig. (2-tailed)	
Pair 1	PerceptualR1 - NonperceptualR1	8.144	65	.000

Co-textual v contextual

	t	df	Sig. (2-tailed)	
Pair 1	CotextR1 - ContextR1	4.728	65	.000

Recording 2 Wilcoxon Signed Ranks Test

Perceptual v non-perceptual

	t	df	Sig. (2-tailed)	
Pair 1	PerceptualR2 - NonperceptualR2	3.937	75	.000

Co-textual v contextual

	t	df	Sig. (2-tailed)	
Pair 1	CotextR2 - ContextR2	6.539	75	.000