

UNIVERSIDADE DE LISBOA
FACULDADE DE LETRAS



**PERCEPTION IN L2 IN A CLASSROOM ENVIRONMENT WITH
L2 PORTUGUESE CHINESE STUDENTS**

Catarina Maria Realinho Rafael

Dissertação orientada pela Prof.^a Doutora Helena Gorete Silva Moniz, e coorientada pela Prof.^a
Doutora Isabel Maria Martins Trancoso, especialmente elaborada para a obtenção do grau de
Mestre em Linguística

2022

0.1 Acknowledgements

As minhas primeiras palavras de agradecimento vão para as minhas orientadoras, Professora Doutora Helena Moniz e Professora Doutora Isabel Trancoso. Agradeço-lhes não só por terem aceite orientar o presente trabalho mas também pela paciência demonstrada ao longo dos anos como orientadoras, professoras e co-autoras.

Agradeço as orientações e indicações para construir um bom trabalho, bem como a revisão e o acompanhamento do mesmo. Aprecio muito o apoio que me foi também dado durante o ciclo de estudos, que exigiu o cruzamento de duas áreas do conhecimento bem distintas. Por último, agradeço a persistência, e o rigor exigido, mesmo nos momentos menos bons. Sem a vossa ajuda, o presente trabalho não teria sido realizado.

Deixo uma palavra de agradecimento ao ICLP, na FLUL, por me ter dado permissão para assistir às aulas do curso anual PLE (B1), contribuindo assim para a realização desta dissertação. Não posso também deixar de agradecer à Professora Rita Gonçalves pelo seu enorme espírito de cooperação durante fases fundamentais da realização desta dissertação, e pelas suas sugestões também como co-autora. Deixo também uma palavra de agradecimento à turma do curso anual B1, de 2018, pela sua participação fundamental.

Muito agradeço à minha família que sempre me apoiou e incentivou a lutar pelos meus objetivos (e a não desistir deles), mesmo quando isso exigisse muito de mim.

Quero deixar também um agradecimento ao Rodrigo, o sobrevivente, que ao longo dos anos sempre me acompanhou no meu percurso académico e profissional. Benditas horas dispendidas a discutir fonologia, enriquecidas pela tua paixão pela linguística. Também quero agradecer aos meus colegas de linguística que me proporcionaram excelentes momentos durante estes anos. André, Joana e Tony obrigada por partilharem comigo as várias horas de trabalho na Biblioteca, nos corredores da Faculdade, no cola e na Padaria Portuguesa. Moisés, Valéria, Ricardo e Nuno muito agradeço as sugestões dadas, e pela partilha da vossa experiência como mestrands.

Quero mencionar os meus professores do DLGR, que desde o início me porpocionaram uma

boa experiência na linguística, partilhando comigo métodos de trabalho e muito conhecimento.

Menciono também as minhas kussantas e os meus queridos amigos que me deram palavras e empurrões para chegar ao fim. Uma apreciação final aos com quem me cruzei (colegas, as minhas turmas multiculturais e professores) que, mesmo sem saberem, me deram motivação para atingir os meus objetivos.

Shukran jazeelan

Contents

0.1 Acknowledgements	3
List of Tables	9
List of Figures	13
0.2 Abstract	1
0.3 Resumo	3
1 Introduction	7
2 State of Art	11
2.1 Speech Perception	11
2.1.1 Native Listening	13
2.1.2 L2 Listening	16
2.2 EP Phonetic-Phonological Phenomena	18
2.2.1 Vowel Reduction	19
2.2.2 Connected Speech Processes	20
2.2.3 Impact of Vowel Reduction and Connected Speech Processes in L2 Perception	23
2.3 L2 Teaching	26

2.3.1	Classroom context	26
2.3.2	CALL	30
3	Methodology	33
3.1	Level adequacy and classes	33
3.2	B1 level classroom assistance	35
3.3	Perceptual experiments	37
3.3.1	Isolated word identification tasks	37
3.3.2	Cross word identification tasks with vowel reduction and connected speech processes	39
3.3.3	Control Group	41
4	Results	43
4.1	Results of perception experiments	43
4.1.1	Overall Results	44
4.1.2	Distribution of phonetic-phonological processes	47
4.1.3	Analysing Orthographic Transcriptions	51
4.2	Final Remarks	71
5	Discussion	73
5.1	Overall results	73
5.2	Distribution of phonetic-phonological processes	77
5.2.1	Vowel reduction	77
5.2.2	Connected speech processes	78
5.3	Analysing Orthographic Transcriptions	79

6	Proposed Didactic Sequence	81
6.1	Guidelines	81
6.2	Sets of exercises	84
6.2.1	Set 1	85
6.2.2	Set 2	87
6.2.3	Set 3	89
7	Conclusion	95
8	Appendixes	101
8.1	Single word identification	101

List of Tables

3.1	Examples of words targeted in task 1.	38
3.2	Examples of words targeted in task 2.	38
3.3	Examples of word sequences in task 3.	40
3.4	Examples of word sequences in task 4.	41
4.1	Percentage of correct answers of each experiment for Mandarin speakers.	45
4.2	Percentage of correct answers of each experiment for the control group.	45
4.3	Percentage of correct answers for each occurrence of vowel reduction for Mandarin speakers, in both tasks 1 and 2.	48
4.4	Percentage of correct answers for each occurrence of vowel reduction for native speakers, in both tasks 1 and 2.	49
4.5	Overall percentage of correct answers of each occurrence of connected speech processes for Mandarin speakers and the control group, middle and right columns, respectively.	50
4.6	Percentage of correct answers of each occurrence of voicing assimilation for Mandarin speakers and the control group, middle and right columns, respectively. . . .	50
4.7	Transcription of the answers of Mandarin speakers for words in the citation form of centralization and deletion of [i], in task 1	52

4.8	Transcription of the answers of Mandarin speakers for words in the citation form of deletion of [u], in task 1.	53
4.9	Transcription of the answers of Mandarin speakers for words in the citation form of deletion of [i], in task 1.	54
4.10	Transcription of the answers of Mandarin speakers for words produced with the centralization (and deletion) of [i], in task 2.	56
4.11	Transcription of the answers of Mandarin speakers for words produced with the deletion of [u], in task 2.	57
4.12	Transcription of the answers of Mandarin speakers for words produced with the deletion of [i], in task 2.	58
4.13	Transcription of the answers of Mandarin speakers for words with the production of [z] due to voicing assimilation, in task 3.	60
4.14	Transcription of the answers of Mandarin speakers for words with the production of [ʃ] due to voicing assimilation, in task 3.	61
4.15	Transcription of the answers of Mandarin speakers for words with the production of [ʒ] due to voicing assimilation, in task 3.	62
4.16	Transcription of the answers of Mandarin speakers for words with vowel encounters, in task 3.	64
4.17	Transcription of the answers of Mandarin speakers for words with vowel encounters, in task 4.	65
4.18	Transcription of the answers of Mandarin speakers for words with vowel encounters with clitics, in task 4.	66
4.19	Transcription of the answers of Mandarin speakers for words with production of [j], in task 4.	68
4.20	Transcription of the answers of Mandarin speakers for words with resyllabification with [l] and [r], in task 4.	69

4.21 Transcription of the answers of Mandarin speakers for words with haplologies, in task 4.	70
--	----

List of Figures

2.1	Spectrogram of the example sequence 'Eu gosto de aprender Português' ('I like to learn Portuguese')	12
6.1	Main page of the proposed 'O Jogo da Forca' ('The Hangman Game').	85
6.2	Example of the displayed topics that users can choose in the 'O Jogo da Forca' ('The Hangman Game'). The progress of the user is indicated, in percentage, in each topic.	85
6.3	Illustration of the environment of 'O Jogo da Forca' ('The Hangman Game')	86
6.4	Illustration of the Hangman game environment after the 'Ajuda' ('help') button is pressed. The gaps and some letters are displayed.	86
6.5	Illustration of the Hangman game environment when the user completes the given spaces. The parts of the body of a stickman are drawn in the figure with a gallows, displayed above the gaps.	87
6.6	Main page of the proposed 'Palavras Bomba-Relógio' ('The Time-Bomb Words') game.	87
6.7	Example of the displayed topics that users can choose in the 'Palavras Bomba-Relógio' ('The Time-Bomb Words') game. The progress of the user is indicated, in percentage, in each topic.	88
6.8	Example of the number of words that user want to include in a round, in the 'Palavras Bomba-Relógio' ('The Time-Bomb Words') game.	88
6.9	Illustration of the 'Palavras Bomba Relógio' ('Time-Bomb words') game environment when the user drags the letters, and drop them in the given spaces.	88

6.10	Illustration of the 'Palavras Bomba Relógio' ('Time-Bomb words') game environment when the game is over.	89
6.11	Main page of the proposed 'À Procura dos Limites' ('Searching for the Boundaries') game.	89
6.12	Example of the displayed options of the game 'À Procura dos Limites' ('Searching for the Boundaries'): words, texts and songs.	90
6.13	Example of the displayed topics (on words set) that users can choose in the 'À Procura dos Limites' ('Searching for the Boundaries') game. The progress of the user is indicated by a stars bar, in each topic.	90
6.14	Illustration of the 'À Procura dos Limites' ('Searching for the Boundaries') game environment when the user drags icons, and drop them in the empty squares, to mark whether the sounds are articulated (words set).	91
6.15	Example of the displayed topics (on texts set) that users can choose in the 'À Procura dos Limites' ('Searching for the Boundaries') game. The progress of the user is indicated by a stars bar, in each topic.	91
6.16	Illustration of the 'À Procura dos Limites' ('Searching for the Boundaries') game environment when the user drags icons, and drop them in the empty squares, to mark whether the sounds are articulated (texts set).	92
6.17	Example of the displayed songs that users can choose in the 'À Procura dos Limites' ('Searching for the Boundaries') game.	92
6.18	Example of the levels of difficulty of the songs set, in the 'À Procura dos Limites' ('Searching for the Boundaries') game.	92
6.19	Illustration of the 'À Procura dos Limites' ('Searching for the Boundaries') game environment when the user marks the word boundaries in the given lyrics of the easy level of difficulty (songs set).	93
6.20	Illustration of the 'À Procura dos Limites' ('Searching for the Boundaries') game environment when the user marks the word boundaries in the given lyrics of the hard level of difficulty (songs set).	93

0.2 Abstract

The purpose of this study is to contribute to the knowledge on the impact of common European Portuguese (EP) phonetic-phonological processes in second language (L2) learners. It is well established that L2 listening is a complex process, and that the most common difficulties among L2 learners are related to speech segmentation and word recognition. Due to the occurrence of connected speech processes, sounds are altered and the word boundaries can be hard to determine.

Vowel reduction within and across word boundaries is usually described as a very frequent process in EP. The reduction of vowels is even more evident in spontaneous speech, e.g. the word ‘telefone’ ([tɨlɨˈfõni] in the citation form, ‘telephone’) can be produced as [tˈfõn]. The interplay between these processes can be particularly impactful for L2 learners, in word recognition. Furthermore, this correlation is scarcely studied in a classroom setting.

The present study explores the impact of vowel reduction and connected speech processes in word recognition tasks from isolated words to continuous speech. Furthermore, it aims to understand the main difficulties that L2 learners, at the intermediate level B1, experience dealing with these phenomena. Lastly, it will contribute to understand not only the acquisition of vowel reduction and connected speech processes but also whether L2 learners could cope with them.

Therefore, it was designed a set of perception experiments involving these phenomena in increasing degrees of difficulty: single word identification without (i) and with vowel reduction (ii); word identification with simple (iii) and complex connected speech processes (iv). The experiments were conducted in an ecological setting of an intensive Portuguese course, of the intermediate level B1, at the University of Lisbon. A control group of EP native speakers also performed the experiments.

The overall scores revealed a decreasing tendency: (i) 94%; (ii) 65%; (iii) 31%; (iv) 16%. The results reveal that word recognition is compromised due to the connected speech processes. Vowel reduction and the consequent deletion of segments also affects the recognition of isolated spoken words even in read speech. The didactic outcomes of the experiments are relevant to contribute to the design of a proposal of a set of listening exercises focused on the practice of these

phonetic-phonological phenomena. The sequence is based on the use of Computer-Assisted Language Learning (CALL) technology, and includes two games and a set of perception exercises in which songs are also used as input.

Key words: Second Language; Perception; Vowel Reduction; Connected Speech Processes; CALL.

0.3 Resumo

Este estudo tem como objetivo contribuir para o conhecimento do impacto dos processos fonético-fonológicos do Português Europeu (PE), em aprendentes de PE Língua Segunda (L2). Pela nossa experiência como falantes nativos, sabemos que ouvir na própria língua nativa (L1) é um processo natural e intuitivo, ainda que complexo. No entanto, talvez por experiência própria também se saiba que, quando se aprende uma língua estrangeira, um dos maiores desafios está relacionado com a compreensão oral. Os mecanismos auditivos que intervêm na compreensão oral da língua nativa são os mesmos que intervêm na compreensão oral da L2.

Ouvir e compreender uma língua estrangeira é um processo muito complexo. Quando ouvimos, é necessário segmentar o discurso, e reconhecer as palavras no contínuo sonoro. No entanto, devido à ocorrência de processos fonéticos, próprios do discurso oral, os sons são alterados e coarticulados.

Ao processar a L1, os falantes nativos não têm quaisquer dificuldades. Pelo contrário, os aprendentes L2 podem ter várias dificuldades no que diz respeito ao reconhecimento de palavras no contínuo sonoro. Dada a ocorrência dos vários fenómenos de coarticulação, as fronteiras de palavra deixam de ser perceptíveis e evidentes.

Em PE, a redução vocálica está bem descrita na literatura, e é referida como um processo muito produtivo. Uma das suas características é promover a redução e o apagamento de vogais átonas. Como resultado, há uma frequente ocorrência de sequências fonéticas de consoantes nas produções orais. Interessa também salientar que a redução vocálica não só ocorre em fala espontânea como também é frequente em fala semi-espontânea e não espontânea.

O PE também é rico em processos de coarticulação, e muitos deles estão em direta relação com a redução e o apagamento das vogais átonas. Os processos de discurso oral contínuo muito bem estudados e descritos. Entre os processos de coarticulação estão os encontros vocálicos, os fenómenos de sândi, a assimilação do vozeamento e a inserção da iode [j]. Estes processos causam diversas alterações nos sons das palavras, sobretudo nos que estão presentes nas fronteiras de palavra. Além disso, promovem ainda a realização e a inserção de outros sons, que não têm uma representação gráfica. Outros são responsáveis pela reestruturação da estrutura silábica, causando a articulação

entre os sons das palavras vizinhas.

Do que é conhecido da literatura, a interação entre a redução vocálica e os fenómenos de sândi externo, coarticulação, assimilação, etc. , e o seu impacto na compreensão oral de aprendentes L2 ainda não está estudada. Além disso, ainda não existem estudos realizados em contexto ecológico de sala de aula, que se debrucem sobre esta correlação. Pelo que se sabe, também, o ensino destes processos fonético-fonológicos não é amplamente explorado nos diversos referenciais e programas de ensino de PLE. Como tal, existe a necessidade de consciencializar os alunos para a ocorrência destes processos na oralidade, e de como têm impacto na sua compreensão oral. Isso terá também uma relação direta com a aprendizagem e o desenvolvimento da proficiência dos estudantes.

Portanto, o presente estudo tem como objetivo explorar o impacto da ocorrência de processos fonético-fonológicos, tais como a redução vocálica e os processos da oralidade, em tarefas de reconhecimento e identificação de palavra. Estas tarefas foram realizadas quer em contextos de palavra isolada quer em contextos de discurso oral contínuo.

Além disso, este estudo procura entender, e descrever, as maiores dificuldades dos aprendentes L2, no nível intermédio B1, face à presença destes processos fonéticos na oralidade. Finalmente, este estudo será não só um contributo para um melhor conhecimento do processo de aquisição da redução vocálica e dos fenómenos de coarticulação, mas também para entender em que momento do processo de aprendizagem os estudantes precisam de treinar estas estruturas.

Portanto, foi criada uma sequência de testes percetivos que incluíram a ocorrência destes fenómenos em diferentes níveis de complexidade: tarefas de reconhecimento de palavra isolada sem (i) e com a produção de redução vocálica (ii); tarefas de reconhecimento de palavra com ocorrência dos processos de discurso oral em contextos simples (iii) e complexos (iv). Os testes de perceção foram aplicados em contexto ecológico de sala de aula de um curso intensivo anual de Português Para Estrangeiros (PLE), do nível intermédio (B1), na Universidade de Lisboa. Todos os participantes eram falantes nativos do Chinês Mandarim. Os testes também foram aplicados a um grupo de controlo, composto por falantes nativos do PE.

Nos resultados obtidos houve, em geral, uma tendência de decréscimo das percentagens para

as respostas corretas conforme o aumento da complexidade das estruturas linguísticas: (i) 94%; (ii) 65%; (iii) 31%; (iv) 16%. Estes resultados indicam que a ocorrência dos fenómenos do discurso oral comprometeram a tarefa de reconhecimento de palavra, nos diversos contextos. No entanto, este impacto foi bastante evidente nos contextos que testaram fala (semi-) espontânea. Além disso, a redução vocálica, e o conseqüente apagamento de vogais átonas, afetou ainda o reconhecimento de palavras isoladas.

Para complementar estes resultados, foi também realizada a análise das transcrições ortográficas dos alunos que permitiu a descrição dos erros ortográficos mais frequentes. Por outro lado, contribuíram também para entender a interação entre a oralidade e a escrita na sua aprendizagem do PE como L2.

As transcrições revelaram também que a presença da redução vocálica aumentou a produção de grupos consonânticos fonéticos, algo que se tornou evidente pelo facto de, frequentemente, apenas partes das palavras-alvo (ou apenas consoantes) terem sido transcritas. É importante também salientar que, aquando da realização das tarefas de identificação de palavras em contexto de fala (semi-) espontânea (com a presença de fenómenos de coarticulação), vários alunos transcreveram sequências de palavras articuladas, incluindo os sons articulados e alterados por processos de assimilação, por exemplo.

Do ponto de vista didático, os resultados obtidos foram relevantes para a construção de uma proposta de uma sequência didática com exercícios para a prática e para o treino da componente auditiva. Esta proposta de exercícios tem um especial ênfase no ensino dos processos fonético-fonológicos do PE. Esta sequência é baseada na tecnologia existente para a Aprendizagem de Línguas Assistida por Computador (ALAC).

Portanto, foram desenhados dois jogos, uma versão do clássico 'Jogo da Força' e o jogo das 'Palavras Bomba-Relógio'. Finalmente, a sequência também inclui uma série de exercícios de percepção oral, intitulada 'À Procura dos Limites' no qual se pretende praticar o reconhecimento das fronteiras das palavra em diferentes contextos de coarticulação. Além disso, a série inclui uma secção de exercícios lúdica, nos quais são utilizadas músicas portuguesas como input auditivo.

Palavras chave: Língua Segunda; Percepção; Redução Vocálica; Processos de discurso oral; ALAC.

Chapter 1

Introduction

The present dissertation attempts to understand the impact of common European Portuguese (EP) phonetic-phonological processes (Andrade and Viana, 1992; Frota, 2000; Mateus and Andrade, 2000; Mateus et al., 2005; Vigário, 2011) in word recognition with second language (L2) learners. Specifically, to understand the interplay of vowel reduction and connected speech processes (CSP) in isolated spoken words, and in continuous speech. This project also aims to contribute to teaching EP as a second language, providing materials and tools for teaching these processes.

Research in second language listening and teaching concluded that the main difficulties that L2 learners experience concerns word recognition in continuous speech (Henrichsen, 2006; Goh, 2000; Altenberg, 2005; Snijders et al., 2007; Cutler, 2012a; Mitterer and Tuinman, 2012; Grant, 2014; Alameen and Levis, 2015; Alammam, 2015; Graham and Santos, 2015; Goh et al., 2016; Reed, 2019). The alterations in the sounds (reductions, deletions, merges, assimilations, articulations, etc.) caused by the phonetic - phonological processes have a great impact in the students' perception, affecting their listening (Altenberg, 2005; Alameen and Levis, 2015; Alammam, 2015; Reed, 2019). Consequently, learners have difficulties spotting words and identifying its boundaries, and fail to recognize known words in the stream sound.

Portuguese research on L2 teaching is focused on the discrimination of segments, and on the development of didactic materials to give the students perceptual training on segments to improve

their discrimination skills (Castelo and Santos, 2017; Castelo, 2018, 2019; Castelo and Freitas, 2019; Zhou et al., 2019, 2021).

Furthermore, EP research on natural language processing (NLP), and computer-assisted language learning (CALL) (Marujo, 2009; Marujo et al., 2009; Baptista et al., 2010; Correia, 2010; Correia et al., 2010; Ribeiro et al., 2010; Pellegrini et al., 2011; Silva et al., 2011; Pellegrini et al., 2012; Silva et al., 2012), contributed with exercises focused on the practice of perception and reading skills, grammar and vocabulary.

To our knowledge, research on the perception of vowel reduction and connected speech processes with L2 learners, at intermediate levels, is not explored yet in an ecological setting. Besides, there are no EP didactic materials that ensure the acquisition of vowel reduction and connected speech processes, and that give effective, and regular, perceptual practice of these processes so far.

Taking into account the linguistic literature review and the competencies for the B1 level, we put for the hypotheses of our work that:

(i) if connected speech processes and vowel reduction are present it will have a clear effect on the recognition of words in continuous speech; (ii) at the B1 level, if vowel reduction occur it will have a low effect on the recognition of spoken isolated words; (iii) if particular connected speech processes are present, it will differently affect the word recognition process.

Furthermore, these phonetic-phonological phenomena are more frequent in spontaneous than in read speech. Therefore, we put for the hypothesis that if different type of inputs are tested (read speech vs semi-spontaneous speech) it will have an effect on the perception of L2 learners.

To test our hypotheses, we designed two perceptual experiments with word identification tasks, increasing the difficulty from isolated words to continuous speech. A series of four tasks were designed and conducted to two groups: (i) 12 L2 students of EP and (ii) a control group composed of 12 EP native speakers. The first experiment encompassed words in citation form and words produced with vowel reduction, in read speech. Experiment two, concerned common connected speech processes, in semi-spontaneous and spontaneous speech, with vowel reduction naturally present in the spoken words.

The aim of this project is to contribute to a better understanding of the difficulties L2 learners, at the intermediate level B1, face dealing with vowel reduction and CSP whether with isolated words or continuous speech (in EP). Moreover, it will rely not only on the acquisition of these phenomena but also on the skills that students need to practise, and when is the right moment, in the learning process, to give that practice. This project aims to complement existing EP that aims to improve learners' L2 perception and pronunciation at the segmental level. Therefore, it proposes a didactic sequence of listening exercises, concerning these processes, with the existing methods and tools of Computer-Assisted Language Learning (CALL), described in the EP literature.

Chapter 2 provides a review of the recent research on the topic of speech perception, comparing both native and L2 listening, focusing in word recognition. It also describes vowel reduction and connected speech processes provided EP literature review. A discussion about the research in second language teaching focused on the effect of connected speech processes in L2 perception is also given. Lastly, it contains a review of the state of art including the EP research on the design of L2 teaching materials, for a classroom context, and on the use of CALL technology in the production of didactic materials.

Chapter 3 contains a detailed description of the methods used in the current study. The aim of this dissertation is to understand the impact of these phonetic-phonological processes in L2 learners, in a classroom setting. Therefore, the chapter provides a description of the classroom environment (since the researcher attended the classes) of a B1 course of Portuguese as a Foreign Language as well as the students, and the activities they performed during the course. The perceptual experiments applied to those students are also detailed.

Chapter 4 presents the overall results for the perceptual experiments as well as the results for each phenomenon. Finally, it also reports an analysis of the students' transcriptions. The discussion of the results with relevant literature, and a summary of the main findings of the experiments, are given in Chapter 5.

Chapter 6 briefly discusses the interplay of the main outcomes of the present study, and the previously presented criteria for the design of L2 teaching materials. It also describes the guidelines for the design of a proposed set of listening exercises focused on the practice of vowel reduction

and CSP.

The proposed didactic sequence includes two games - 'O Jogo da Força' ('The Hangman Game') and 'Palavras Bomba-Relógio' ('The Time-Bomb Words') - and a set of listening exercises focused on the detection of word boundaries in short sequences of words, and in real speech samples - 'À Procura dos Limites' ('Searching for the Boundaries'). This last set also uses Portuguese songs as input.

Finally, Chapter 7 summarises the main findings of the perceptual experiments described in Chapter 3 (reported in Chapter 4 and discussed in Chapter 5), and provides a review of the proposed didactic sequence, described in Chapter 6. Lastly, the limitations and issues identified in the current study are mentioned, and some lines of future work are therefore suggested.

Chapter 2

State of Art

In the present chapter it will be reviewed the body of research on the topic of the listening process. Therefore, section 2.1 will describe the state of art of the speech perception topic, exploring both native and L2 listening. It will also provide a discussion on the similarities, and the differences, between these both processes. Section 2.2 will describe vowel reduction, a frequent process in EP, and the EP connected speech processes. It will also include remarks on the interplay of these phonetic-phonological processes on spontaneous speech. This last section presents a further discussion on the impact of these processes in L2 listening.

Finally, section 2.3 will afford the major challenges in L2 teaching provided empiric data, and present research on the development of teaching strategies, and the design of didactic materials, for a classroom context. Furthermore, it will present EP state of art on the application of CALL technology on the design of didactic materials for Portuguese as a foreign language.

2.1 Speech Perception

The ability to communicate is a primary skill that distinguishes humans from non-human species. Listening skills are a key aspect of human communication. Commenting on spoken language processing, Sawusch (2005:7) argues that speech "is the starting point for the listening". Speech is a

continuous acoustic signal that our ear receives. Animals also deal with speech. Hear it, they will, but they are incapable of decode and parse the input. In other words, non-human species do not achieve the interplay of perception and listening processing.

Tatham and Morton (2006:201) define perception as "the active assignment of a symbolic representation to what is heard". Listeners map the acoustic input into linguistic representations in the brain. So, what makes speech perception so special is what happens in listeners' minds, and what abstract information, in the repertoire of listeners' native language, can be encouraged in the perception process (Cutler, 2012b; Heald and Nusbaum, 2014).

In summary, listening is a human-specific ability due to the human exclusive brain structures and the cognitive processes that support listening. For this reason, no native listener needs to be told how to process speech. Human perception skills are so intuitive as heartbeats. The Figure 2.1 below shows a spectrogram of the spoken sentence 'Eu gosto de aprender Português' ('I like to learn Portuguese'), produced by an European Portuguese native speaker.

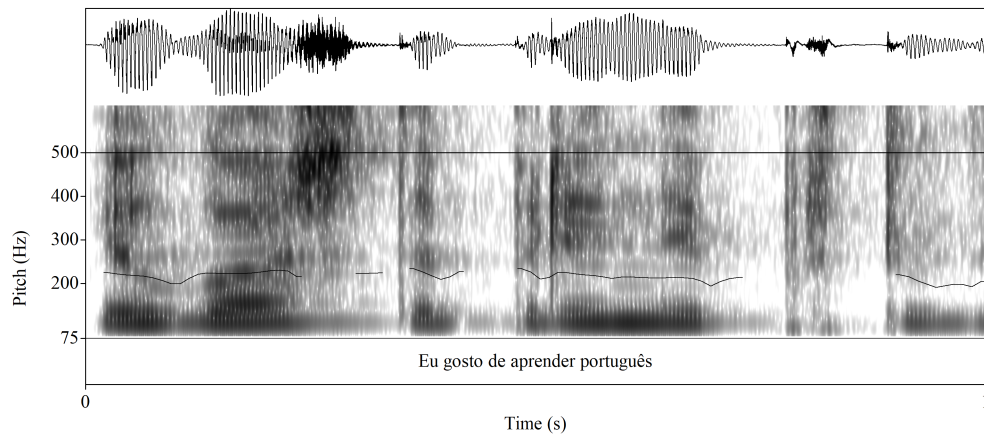


Figure 2.1: Spectrogram of the example sequence 'Eu gosto de aprender Português' ('I like to learn Portuguese')

In contrast to written pages, there are no spaces between the words since in continuous speech all sounds are linked and articulated. EP native speakers would map this physical input into discrete units when listening to the spoken sentence, determining which sounds and words were produced,

allocating the word boundaries. They would also do it in a matter of milliseconds with minor, and perhaps no effort.

Over the last decades, most research in speech perception has emphasised that listening occurs in modular stages. The mapping between the stream sound and comprehension is described in serial and static stages of processing (Swinney, 1979; Tanenhaus et al, 1979; Marslen-Wilson, 1987 apud in Magnuson (2016). Recently, significant advances in cognitive neurosciences have allowed a better understanding of the complex brain structures that support listening. Therefore, recent theories of speech perception (Heald and Nusbaum, 2014; Gaskell and Mirkovic, 2016; Magnuson, 2016) are discarding modular approaches towards assumptions that listening is an active process, and that cognitive factors such as attention, experience and plasticity are involved.

The following sections describe listening on the light of the dynamic interaction of the linguistic levels, and of what listeners deal with in word recognition without assuming a speech perception theory.

2.1.1 Native Listening

The key research question on this work is related to L2 listening. However, clearly understand how perception in second language happens, it is important to recognise the inherent complexity of native listening.

Human listening ability is unique. During perception, listeners segment speech, discriminate sounds and recognise small parts of speech (Cutler, 2012a; Mattys and Bortfeld, 2016). The brain decodes the continuous auditory signal and creates linguistic representations of that stream. Thus, listeners retain the perceptual information in memory to further assign it to abstract representations. Linguistic information, such as sound categories in the phonological repertoire, is already provided by the listener's L1 (Tatham and Morton, 2006; Weber and Cutler, 2006; Cutler, 2012b,a).

Speech perception involves recognising words and their boundaries from the stream speech. Listeners may use different cues that encourage word retrieval. Several lines of evidence suggest that the activation of potential word candidates, and its consequent competition, may be promoted

by low-level information, such as phonotactic cues (Mattys et al., 2007) and stress (Correia et al., 2015). Moreover, other independent categories of representation such as morphemes may also play an important role at early stages of processing even "before phonological representations are computed in detail" as Johnsrude and Buchsbaum (2016:7) points out. Cutler (2012a) also claims that word recognition is dependent on the temporal distribution of the sounds in the utterances. Therefore, listeners do not have to wait to hear the entire word to begin the lexical activation.

This is even more impressive given the complexity of the native mental lexicon and how fast listeners recognise words. In a lecture about comprehension of the written Portuguese, a teacher compared the native mental lexicon to the most extensive library worldwide, as if all libraries in the world were merged in just one giant library. The teacher underlined that finding a specific book in such a library would take a considerable time. The native vocabulary is similar to that vast library. During word identification, the brain retrieves a specific word from the 'mental library' which is structured with thousands of words, connections of meanings, and morphological and syntactic categories. The cognitive processes that support this process are so effective that word retrieval occurs in frames of milliseconds.

Hence, it is evident that the dynamic structure of the mental lexicon contributes to the complexity of the word recognition process. The native vocabulary is not an abstract list of words with forms and meanings. Cognitive processes allow an impressive semantic richness (Goh et al., 2016) and a dynamic interaction of complex semantic networks. In fact, Nenadic et al. (2019:97) suggested that when vocabulary is activated, "word meanings are accessed, not just their form". Therefore, in word recognition, there is a strong interplay of the complexity of the semantic connections in the vocabulary and the cognitive processes.

Existing research also recognises the important role played by the syntactic level in spoken word recognition. Mattys et al. (2007) tested ambiguous word sequences, such as 'takespins', and explored how the syntactic context would encourage the listeners to associate the sound [s] to the word 'takes', as an inflexion marker, or to the word 'spins', integrating an object noun phrase. In this study, the authors concluded that listeners make use of syntactic cues to segment and recognise spoken words.

Concerning other linguistic levels, there is evidence that prosody also plays a crucial role in speech perception. Lindfield et al. (1999) apud in Severino (2011) investigated the interplay of segmental information and prosodic features, in word recognition. The research indicated that listeners can recognise words even when the minimum of segmental information is presented if prosodic features are provided.

Native speakers are able to cope with speech variability due to their listening flexibility (Cutler, 2012a). For instance, the acoustic variability can be related to changes in sounds in the continuous stream. Spoken words do not retain their citation forms as in dictionaries: sounds are altered, linked, articulated or even deleted. Hence, the way a word sounds in isolation is different than when it is embedded in a sentence.

Listeners also deal with the speaker variability (Johnsrude and Buchsbaum, 2016), or even with pauses, repetitions and reformulations that speakers do in spontaneous speech (Moniz, 2006), as part of the parsing process. Moreover, the acoustic variability is encouraged by sociolinguistic factors, speech rate and even speech style (Lass, 1984; Mateus et al., 2005; Moniz et al., 2014; Mata and Moniz, 2016).

Listeners do not always face the ideal circumstances of listening, and perhaps there might exist noise in the environment. In these situations, the success in listening is related to the active role that the brain plays by masking the noise (Heald and Nusbaum, 2014). Therefore, the listening flexibility is proof of the crucial role that cognitive structures such as memory, attention and plasticity play in speech perception.

In conclusion, native listening is a complex process characterised by an interwoven of the linguistic levels and cognitive structures. Listeners also cope with extralinguistic variables such as sociolinguistic factors and speech variability. Despite the mentioned challenges in speech perception, speakers can achieve a successful listening.

It is impossible to decode speech, and achieve comprehension, unless the listener knows the language. Given this fact, it is interesting to address the questions: Would an L2 listener successfully decode speech as a native speaker? Do L2 listeners use other relevant cues in listening? Is L2

decoding easier, intuitive and automatic as native decoding is?

The current subsection presents native listening on the lights of the interaction of the different linguistic levels and the cognitive processes. The following subsection discusses L2 listening and the common difficulties learners experience in second language learning.

2.1.2 L2 Listening

A key research question in speech perception concerns how perception in second language happens. There are similarities between native and second language listening. Biological mechanisms such as the auditory system and the brain structures are the same whether processing a native or a second language. Decoding L2 speech also implies the interplay of linguistic (and extralinguistic) levels and cognitive processes (Heald and Nusbaum, 2014).

It has previously been observed that, to successfully listen, speakers have to segment speech and discriminate sounds (Cutler, 2012a; Mattys and Bortfeld, 2016), recognise words in stream speech and activate vocabulary from the memory (Magnuson, 2016). Data from several studies suggest that semantic richness, semantic networks (Goh et al., 2016; Nenadic et al., 2019), the language structure (Mattys et al., 2007) and prosodic information (Severino, 2011) play a vital role in word retrieval. Furthermore, the listening flexibility (Cutler, 2012b) is highlighted in the successful way listeners cope with an enormous acoustic variability, such as speaker variability (Lass, 1984; Cutler, 2012a) or noisy environments (Heald and Nusbaum, 2014).

However, it is well known, even from self experience, that listening in second language is harder than in first (Cutler, 2012a). Despite the similarities with native processing, L2 listening has considerably more issues involved turning it into a challenge to L2 learners. Their performance has also been shown to be related to their knowledge about the target language. In other words, it depends on the level of proficiency.

A frequent question on speech perception topic concerns whether the native language inhibits or promotes L2 perception. Learners transfer properties from their L1 to the L2 they are learning. In fact, recent research concluded that even in more advanced levels speakers transfer properties

from their L2 (Cutler, 2012a; Zhou et al., 2017), and experience difficulties especially in listening (Altenberg, 2005). Thus, in addition to the importance of the level of proficiency, an extensive experience with the language is also crucial (Gustafson and Bradlow, 2016).

The learners' L1 acts as a filter when learning a second language (Cutler, 2012a), and this 'filtering' affects listening. Speakers use language-specific strategies to segment speech (Alammar, 2015). Moreover, parsing the stream speech turns out to be more difficult when the range of native sounds does not match with the L2's range (Snijders et al., 2007), because learners may apply cues that are productive in their L1 but do not facilitate the segmentation of the L2 (Altenberg, 2005; Alammar, 2015; Reed, 2019). Consequently, the application of that segmentation strategy inhibits the recognition of words and its boundaries (Goh, 2000).

L2 learners may not be aware of how L2 sounds change in connected speech as well, and they may not be familiarised with these processes from the L1 (Weber and Cutler, 2006; Alameen and Levis, 2015). Hence, learners may recognise spoken isolated words but they may not when these same words occur in connected speech (Mitterer and Tuinman, 2012; Vandergrift and Goh, 2012).

A study conducted by Snijders et al. (2007) tested event-related potentials (ERP) techniques to understand the brain responses of native speakers, and L2 learners, when listening to known spoken words in isolation and embedded in sentences. The results revealed that the brain responses were slowed for the L2 learners, in comparison to native speakers. This corroborates the fact that decoding and segmenting continuous speech in a second language is a challenge. Furthermore, this reveals that the cognitive processes are affected during L2 processing.

The interaction of linguistic levels such as the vocabulary and semantics are known to influence L2 word recognition. The native vocabulary is much more developed and complex than L2 vocabulary (Cutler, 2012a). However, until a learner is an advanced and experienced user of the L2, the size of the vocabulary and the complexity of the semantic connections are conditioned. As a result, there might be a delay in the activation of words from the vocabulary during listening, compromising word recognition.

Other existing body of research on L2 perception indicates that syntactic cues promote the

recognition of spoken words and the vocabulary activation. For instance, Vandergrift and Goh (2012) suggests that content words are relevant cues during listening comprehension (when comparing to function words).

L2 listeners can also make use of suprasegmental information (Maastricht, 2018) such as prosodic features or rhythmic cues that encourage them to locate word boundaries. In addition, processing both linguistic and contextual information may promote the activation of pragmatic knowledge (Cook and Liddicoat, 2002, apud in Vandergrift and Goh, 2012). It is important, however, to highlight that a successful listening is influenced not only by the level of proficiency but also by the experience that speakers have using the language.

In sum, this section included a comparison between native and L2 listening processes, and described the main challenges L2 listeners experience. It was also discussed the role of the phonological processes in speech perception. As mentioned, some processes cause alterations in sounds, which may inhibit speech segmentation and the allocation of word boundaries. In addition, if these processes are not familiar to the learners, the impact of the linguistic transfer may be more evident.

In European Portuguese, the occurrence of phonetic-phonological phenomena is very productive. The following section 2.2 will describe how these processes occur in spoken EP, and how L2 learners' perception is affected.

2.2 EP Phonetic-Phonological Phenomena

Phonetic-phonological processes occur in every language and European Portuguese is not an exception. These processes are present in read speech but they are even more frequent in spontaneous speech. They promote alterations in sounds and, as a consequence, those sounds can be deleted, contracted, linked, assimilated, coarticulated, etc.

This section focuses on which phonetic-phonological phenomena occur in EP, according to the linguistic literature, and on their impact in L2 perception. Subsection 2.2.1 describes vowel reduction, a productive process in EP. Section 2.2.2 will describe the variety of connected speech

processes that occur in spoken EP: vowel encounters, voicing assimilation, production of [j], haplogophies, and external sandhi with [l] and [r].

2.2.1 Vowel Reduction

Vowel reduction is a phonological process in which the vowels, in unstressed positions, are altered. European Portuguese (Mateus and Andrade, 2000; Mateus et al., 2005; Vigário, 2011), has seven stressed vowels: /i, e, ε, o, ɔ, u/. The reduction of vowels consists in the raising of /e, ε, a, o, ɔ/ and the centralisation of /e, ε/. This process reduces the seven vowels to four phonetic vowels [i, i, u, ə] when produced in unstressed positions (pre- and post-tonic positions, with some exceptions).

A consequence of these alterations is the deletion of the reduced vowels. Portuguese research on this topic (Delgado Martins, 1975; Freitas, 1997; Fikkert and Freitas, 2006; Mateus et al., 2005; Vigário, 2011; Veloso, 2016) indicates that [u] and [i] are the most deleted vowels, and that this elision is influenced by syllable and word position, word frequency, and other features, such as speech rate and speech style.

The high deletion of vowels in unstressed syllables results in the production of sequences of phonetic consonant clusters. For instance, in word 'desprevenir' ([dɨʃprɨvɨ'nɨr], 'fail to provide') and word 'telefone' ([tɨlɨ'fɔni], 'telephone') (Mateus and Andrade, 2000:44) two to three vowels in the following syllables are deleted contributing to the production of sequences of three ([tɨ'fɔn]) and six ([dʃprɨv'nɨr]) consonants. Native speakers cope with the massive deletions of vowels, however, some parsing problems may arise. In a study investigating the perception of contrasts between lexical clusters and phonetic sequences, (Cunha, 2015) reported that the deletion of vowels and the consequent consonantal overlap can temporarily mislead EP native speakers. For instance, participants perceived words such as 'querer' ([k(i)'rer], 'to want') and 'crer' ([k'rer], 'to believe') as homophones. Thus, native speakers can rely on other cues such as context to recognise words. Nevertheless, more research on the frequency of occurrence of this process in spontaneous speech (and how it affects L2 perception) is needed.

2.2.2 Connected Speech Processes

European Portuguese has a variety of connected speech processes. The occurrence of these phonetic processes depends on features, such as the prosodic structure or stress (Frota, 2000; Mateus et al., 2005; Andrade, 2020). Vowel reduction, previously described, plays an important role across word boundaries. Hence, the consequent deletion of vowels also has an impact in connected speech processes (Mateus and Andrade, 2000). They can occur in both prepared and spontaneous speech (although with much more frequency in the former), and it appears that certain social and linguistic factors may affect the frequency, quality, and contexts of some CSP (Mateus et al., 2005; Oliveira et al., 2014; Oliveira, 2016; Paulino, 2016).

2.2.2.1 Vowel Encounters

The different occurrences of vowel encounters are well described in the EP linguistic literature (Andrade and Viana, 1992; Frota, 1995, 2000; Andrade, 2020). The reduction of vowels causes the deletion of sounds even within words. In vowel encounters across words, the deletion of vowels in word final position encourages an articulation process to the first vowel of the following word like Andrade and Viana (1992:61) illustrates: in the sequences 'triste amigo' ([ˈtristɐˈmigu], 'sad friend') and 'fosse obrigado' ([ˈfosɔbɾiˈgaðu], 'was obligated') the final vowel [i] of the word 'triste' and 'fosse' is deleted. Consequently, these words link to the following vowel resulting in the production of one articulated sequence. This is also true for the sequence 'deixe o livro' ([ˈdɛʃuˈlivɾu], 'leave the book') in which the last vowel of the word 'deixe' ([i]) is deleted and this word links to the following one, which is the article 'o' ([u]). Frota (2000:84) gives an example with a different final unstressed sound, e.g. 'oito amigas' ([ˈoitɐˈmigeʃ], 'eight friends'). In this case, the unstressed sound [u] is deleted and the first word links to the following sound, which is [ɐ].

Vowel encounters also stem from alterations in vowels that lead to the merge of both sounds. Andrade and Viana (1992:63) give examples such as the sequences 'diga Aldina' ([ˈdigaˈdine], 'say Aldina') and 'a abóbora' ([aˈbɔbɔˈɾɐ], 'the pumpkin'). The unstressed sound [ɐ], at the end of word 'diga', merges with the first sound of the word 'Aldina' ([a]) which is also at an unstressed position.

The same happens with the second sequence: the article 'a' ([ə]) merges with the first vowel, [ɐ], of the following word 'abóbora'. The merge of both sounds, in both previous examples, results in the production of [a].

It is interesting to highlight that prosodic features, prominence constraints, stress (Frota, 1995, 2000) and extralinguistic factors (Mata and Moniz, 2016; Paulino and Frota, 2016) affect vowel encounters by promoting or inhibiting them.

2.2.2.2 Production of Glides

In European Portuguese, some vowels are produced as a semi-vowel (or glide). Data from several studies suggest that the production of [j] is determined not only by prosodic features (Frota, 2000; Vigário, 2011; Mata and Moniz, 2016) but also by sociolinguistic features (Oliveira et al., 2014; Oliveira, 2016), e.g. dialects, age and speech style. (Andrade and Viana, 1992) illustrates some of these occurrences:

- disse-o ('said it', [ˈdisju], Andrade and Viana (1992:61))
- quinze armas ('fifteen guns', [ˈkizjˈarmɐʃ], Andrade and Viana (1992:63))
- onze horas ('eleven hours', [ˈõzjˈɔɾɐʃ], Andrade and Viana (1992:63))

According to the authors, some numerals are not produced with the production of [j], as with 'sete' ('seven') and 'dezassete' ('seventeen'). There are also cases of variation in the production of the glide as in (1) and (2).

1. It is possible to produce 'nove horas' ('nine hours') as [ˈnɔvˈɔɾɐʃ] or [ˈnɔvjˈɔɾɐʃ] (Andrade and Viana, 1992:63).
2. 'porque é que veio' ('why did it come') may be produced as [ˈpʷrkɐkˈvɐju] (Andrade and Viana, 1992:62) or [ˈpʷrkjɐkˈvɐju].

2.2.2.3 Voicing Assimilation

Voicing assimilation is very well explored in the EP literature (Andrade and Viana, 1992; Frota, 1995, 2000; Mateus and Andrade, 2000; Mateus et al., 2005), and it consists in the assimilation of voicing features from neighbour segments within and across words (Mateus et al., 2005). The fricative sound [f] is produced as [ʒ] (2) if the following sound is voiced, or as [z] (3) if it is succeeded by a vowel. When unvoiced segments occur, the sound of the fricative is not altered and it is produced as [f] (1).

In connected speech, the fricative assimilates the voicing features of the first sound of the following word. The examples in Frota (2000:53) clearly illustrates this point .

1. lápis preto ('blue pencil') - lápi[f]preto
2. lápis branco ('white pencil') - lápi[ʒ]branco
3. lápis azul ('blue pencil') - lápi[z]azul

The realisation of [z] and the consequent linking process imply a restructuration in the syllabic structure. Thus, [z] fills the onset position of the following syllable. The prosodic structure is a key factor in voicing assimilation (Frota, 2000, 2014). In the following sentence, the assimilation is blocked across IP boundaries: the final sound [f] in word 'jornalistas' ('journalists') is not produced as [z], as expected, because the following word does not belong to the same IP. Therefore, the sound of [f] is not altered.

- (Aos jornalistas,)IP (as angolanas ofereceram especiarias.)IP

(Frota, 2014:14)

2.2.2.4 External sandhi with [l] and [r]

Articulation processes are very productive in EP. External sandi with the liquids [l] and [r] is well documented in EP research (Mateus et al., 2005) and its occurrence is associated with a consequent

syllabic reestructuración. Mateus et al. (2005) points out that the grapheme <l> is produced as [ɫ] in coda position, as in words 'sa[ɫ]' ('salt') and 'sa[ɫ]gado' ('salted') (Mateus et al., 2005:229).

In contrast, if the following word starts with a vowel the final sound [ɫ] is altered and produced as [l] due to a linking process that causes a resyllabification. For instance, in the sequence 'sal amargo' ('sour salt'), [l] fills the onset position of the word 'amargo' thereby altering the syllabic structure ([ˈsa.lɐ.ˈmar.gʷ]). Furthermore, sandhi occurs with the sound [r], in coda position. For instance, in the sequence 'mar azul' ('blue sea'), the final sound [r] of 'mar' ('sea') links to the following word, and its syllabic structure is altered: [ˈma.rɐ.zuɫ] (Mateus et al., 2005:229).

2.2.2.5 Haplologies

The existing body of research on phonetic-phonological phenomena recognises haplologies as a frequent process in EP (Andrade and Viana, 1992; Frota, 2000). Haplologies occur when the final syllable of a word is similar, or identical, to the first syllable of the following word. By way of illustration, Andrade and Viana (1992) and Frota (2000) show data (see examples (1)-(3)) in which the vowel of the last syllable is deleted, the consonant segments geminate (2) or only a consonant is produced (1) and (3).

1. Faculdade de Letras ('Faculty of Arts', [fɐkʰɫˈdadˈletɾɐ]) (Andrade and Viana, 1992:65))
2. disse sim ('said yes', [ˈdiˈsi]) (Andrade and Viana, 1992:65))
3. campo poluído ('polluted field', [ˈcɐ̃pʷluˈidʷ]) (Frota, 2000:999))

2.2.3 Impact of Vowel Reduction and Connected Speech Processes in L2 Perception

As mentioned in section 2.1.2, L2 listening is a complex process (Cutler, 2012a). It is well established that speech segmentation and word recognition are far the major challenges in second language acquisition (Reed, 2019; Altenberg, 2005). Goh (2000) conducted a series of interviews with beginner and advanced learners of English to explore the difficulties students experienced in second language learning. The primary challenge students reported was recognising words in

continuous speech (even known words). This difficulty is a consequence of the presence of the phonetic-phonological processes in spoken speech (articulations, assimilations, reductions, deletions and insertions) that occur, in some form, in every language.

Graham and Santos (2015) interviewed intermediate L2 learners of English to understand the impact of these processes in listening. A common view amongst the students was that speech seemed unbroken, turning hard to know whether a word ended or started. Therefore, students, on the whole, reported difficulties in word boundaries allocation. Also, Henrichsen (2006) studied the impact of connected speech processes in L2 learners of English (of different proficiency levels). Participants were exposed to isolated sentences recorded with the absence and presence of sandhi phenomena. The results reveal that even the advanced students experienced difficulties in listening when continuous speech processes were present.

Therefore, the success in listening and segmenting speech depends on whether the occurrences are familiar from the listener's native language, and how/if learners are aware of these processes in the L2. Unfortunately, there is a lack of materials and effective perceptual training in L2 teaching concerning these phenomena (Cutler, 2012a; Alammari, 2015).

So far, there is little research on the impact of connected speech processes or reductions in L2 learners. Besides, most studies are lab-based, so they do not encompass, or replicate, spontaneous speech features such as disfluencies, and different prosodic and discursive patterns (Mitterer and Tuinman, 2012).

In European Portuguese, vowel reduction causes alterations in the sounds and promotes the deletion of vowels in unstressed positions. As a result, there can be a massive production of phonetic sequences of consonants. Native speakers cope with this deletion despite temporarily parsing problems (Cunha, 2015). However, L2 learners may have difficulties recognising reduced forms of words that they would recognise in written form or when spoken in isolation.

Connected speech processes are also responsible for various alterations in sounds, e.g. restructurings in the syllabic structure and merges of segments across words. Consequently, L2 listeners may have difficulties determining the word boundaries. Continuous speech processes and vowel re-

duction occur in both read and spontaneous speech. However, the interplay of these phenomena and the frequency of occurrence is even more evident in spontaneous speech than in carefully reading (Mateus et al., 2005; Cutler, 2012a; Mata and Moniz, 2016).

It is interesting to highlight that most foreign speakers and L2 learners of European Portuguese express a common opinion about EP speech. Listeners perceive the differences between a canonical pronunciation and a reduced form in spontaneous speech. The following citations were selected from two question and answer (Q&A) websites. The first is from Portuguese Language Stack ¹, a discussion forum about the Portuguese language, and the second citation is from Reddit website ².

“I have noticed that in European Portuguese, many native speakers don’t necessarily pronounce words the same way as the textbook would teach a foreign student to; in that many vowels are elided or even deleted. One example is in the Cidade FM slogan “só se quiseses”: when actually read out by the Portuguese announcer, it becomes something like “sósquisers”. I’m not really sure how this vowel deletion works, or how to specify it, except that Portuguese natives typically delete the last vowels of words?”

(in Portuguese Language Stack)

“Can someone explain disappearing vowels in European Portuguese? Is there a pattern to it? Does everyone in Portugal do it? Is there some sort of pattern to the disappearing vowels in Portuguese words? Is it a colloquial thing? Does everyone do it in casual speech or is it more related to an area? For instance, does this occur in news broadcasts?”

(in Reddit)

The citations reveal that both speakers were sensitive to the deletion of vowels in spontaneous speech. Interestingly, a speaker reported that EP native speakers do not produce words as the canonical form taught in textbooks. This indicates that there is also a lack of effective L2 teaching materials focused on connected speech processes and vowel reduction.

In general, research in second language speech perception is limited. Our knowledge of L2

¹<http://www.portuguese.stackexchange.com/questions/6420/are-there-any-rules-for-vowel-deletion-in-european-portuguese>

²http://www.reddit.com/r/Portuguese/comments/ao9g2x/can_someone_explain_disappearing_vowels_in/

perception in EP is even more so. So far, there is no research on how L2 learners of EP cope with phonetic-phonological processes in a real ecological context. This dissertation aims to explore the abilities of L2 learners of EP to segment speech and recognise words. Furthermore, it aims to understand the impact of vowel reduction and connected speech processes in word recognition.

2.3 L2 Teaching

The major goal of L2 learners is to have an active role in real communicative situations with native speakers. Therefore, recent research on second language teaching (Grant, 2014; Graham and Santos, 2015; Reed, 2019; Richards and Cotos, 2019) is exploring new teaching tips to help students improve their production and perception skills to prepare the students for real contexts.

The following section will present some strategies focusing on listening practise in a classroom context. Besides, some criteria for the design of teaching materials will also be discussed. Section 2.3.2 will mostly present EP state of art on the development of tools for teaching EP as a second language using Computer-Assisted Language Learning (CALL) technology.

2.3.1 Classroom context

Investigating the listening process is a continuing concern within the second language teaching field. Teachers are aware that listening in L2 is a challenge. Recently, considerable literature, e.g. Graham and Santos (2015); Castelo and Santos (2017); Castelo (2018); Maastricht (2018) and Reed (2019) has grown up around the theme of which are the most common difficulties L2 students experience during listening, and how it has implications in comprehension.

Segmenting continuous speech and detecting word boundaries appear to be the greatest challenges among L2 students (Cutler, 2012a; Vandergrift and Goh, 2012) and these are the determining factors of the success in word identification and recognition. In fact, students may recognise words in written form or when spoken in isolation. However, they may have difficulties recognising those same words when they are embedded in sentences.

Continuous speech processes and reductions in vowels occur, in some form or other, in every language. Therefore, it is clear that investing in listening practise is useful even at lower levels of proficiency. L2 teachers might be aware of what difficulties that students have in listening. However, teachers also have their own challenges selecting listening materials that are adequate to the students' needs.

Brito and Hernandez (2020) interviewed 95 teachers of Portuguese (Brazilian Portuguese) as a second language in the United States to understand the difficulties they have specially when teaching reduced forms, in the classroom. Interestingly, 65% of the teachers reported that the main challenge in teaching spoken reduced forms is the lack of adequate teaching materials. Some teachers pointed out that, despite they intend to expose lower level students to these forms, it is not easy to find attractive and interesting materials.

Furthermore, the authors concluded that curricular constraints and the lack of time to practise the listening skills in the classroom is another concern among the teachers. They find it essential to practise these reduced spoken forms since this is the type of input the students will deal with outside the classroom. Despite the use of videos, web media and recordings, some teachers reinforced the need for more materials to teach the reduced forms and the phonetic processes. Another common view among them is that teaching the phonetic-phonological phenomena (which cause alterations in sounds and the reduced forms of spoken words) is essential to be addressed in Portuguese courses for foreigners.

Cuicui (2012) also conducted a questionnaire to 127 teachers of Portuguese as a foreign language (PFL), and to 125 students of Portuguese (native speakers of Mandarin Chinese) about the use of teaching materials in PFL classes. 44% of the teachers reported that the lack of adequate materials is a challenge. Also, 81% of the participants (teachers and students) reported that the use of interesting materials is "extremely important" in the listening practise, and 79% (of the participants) pointed out that the use of varied listening materials is "extremely important".

These findings suggest that it is important to invest in adequate teaching materials focused on the mentioned phonetic phenomena. A key factor on the development of teaching materials is that they must be adapted to a particular teaching scenario, based on the students' needs. They also should

accomplish the learning goals pertaining to the respective levels of proficiency, and guarantee that the students receive that right skills at the right moment of the learning process (Azevedo, 2012).

Another important factor concerns the feedback and the evaluation that students can receive. Evaluating their progress will help the teachers to understand the students' needs, and the skills they need to practise. Giving feedback about their learning progression will motivate the students, encourage them to be autonomous and improve their self-evaluation ability (Castelo, 2018).

The majority of materials used in listening exercises are similar to lab-based input used in speech perception studies. Some of the listening materials are audio files with pre recorded sentences or texts, recorded by a native who is carefully reading (Mitterer and Tuinman, 2012). As a consequence, this type of materials do not include typical features of spontaneous speech, such as the variety of prosodic and discursive features or variations in speech rate. Besides, in casual speech alterations, deletions, articulations and changes in sounds are more frequent than in read speech (Altenberg, 2005). Authentic input also includes features such as disfluencies, e.g. pauses, breaks and repetitions which are natural in spontaneous speech (Moniz, 2006; Mata and Moniz, 2016).

Research in English as a Second Language (ESL) (Reed, 2019) also indicated that exposing learners to connected speech processes contributes to the development of their segmentation skills. Therefore, to design materials that rely on phonetic-phonological processes, teachers must take into account the type and the amount of input they expose the students to.

Using authentic speech materials containing real communicative situations will prepare the students for 'the outside classroom speech' (Castelo, 2018), and contributes to the development of their segmentation skills. Teachers may find various examples of authentic speech including different communicative contexts on youtube, ted talks, podcasts, interviews, radio, news, commercials, movies or other available resources. Barbosa (2014) also stated that the use of songs could be an efficient strategy, since it may result in interesting, dynamic, and attractive classes, and improve the students' memory.

Alves (2015) and Castelo (2018) also demonstrated that, despite the essential role of authentic speech materials, teachers at initial stages may use more controlled input and progressively integrate

spontaneous speech related to the topics taught in the classroom. Therefore, the teachers should always focus on the potential didactic outcomes of the selected input taking into account the type and the amount of that input.

Using authentic speech materials into the topics taught in the classroom could also be useful to integrate the vocabulary that students need to acquire in real contexts (Castelo, 2018). In this study, students had difficulties recognising spoken words they knew due to vowel reduction and the consequent deletions of segments. Therefore, including the target vocabulary in listening exercises may help students to be aware of how spoken words differ from their written form.

Research on second language teaching (Reed, 2019) has suggested methodological approaches to teach connected speech processes, specially in English. One of the most common suggested activities is the dictation task. Besides, researchers also indicated that including reflection questions will encourage the students to be sensitive to the occurrence of these processes and the changes they cause in the sounds. For instance, questions such as “What did you hear?”, “Does it make sense?” or “What was really said?” (Reed, 2019:407).

Other approaches (Graham and Santos, 2015) suggest the presentation of the vocabulary as a list of spoken isolated words and in connected speech contexts, e.g. a spoken sentence, and ask the students about the changes they perceived in the sounds. Further questions about the phonetic contexts for the occurrence of these processes are also useful. Another interesting suggestion to teach the connected speech processes is to present sentences or small texts with no space between words and then ask students to listen to its productions, mark the word boundaries and rewrite the text correctly.

Research in EP (Castelo and Santos, 2017; Castelo, 2018, 2019; Castelo and Freitas, 2019; Zhou et al., 2019, 2021) is mostly dedicated to the discrimination and production skills, and it is also focused on the perceptive training of segmental features of the EP repertoire of vowels and consonants (since this is also a common difficulty among L2 learners of EP). There is also interest on the development of activities and didactic materials to improve the students’ production and writing skills, in a classroom context. There is no evidence of EP research concerning the perceptual practise of vowel reduction and the occurrences of connected speech processes.

2.3.2 CALL

Nowadays, with the advances in technology for information systems much improvement has been done in the development of applications. One of the areas of these advances is the use of Computer-Assisted Language Learning (CALL). These systems aim to complement second language teaching and learning with the use of computer technology. This approach early aimed to supplement or replace the direct relation teacher-student in the classroom.

The design of these applications aims to offer the students interactive materials according to their learning necessities. Another advantage of these applications is that they motivate students and promote self-study.

The technology used in CALL includes NLP tools and even Text-To-Speech (TTS) synthesis and Automatic Speech Recognition (ASR) systems to develop materials targeting the practice of different skills. One of the aims of the use of this technology in L2 didactics field is the automatic generation of learning materials. Some of these generated exercises may be included in web pages with static designs, e.g. fill in the blanks. Moreover, recent approaches included the development of more interactive environments such as games with high quality graphics or audios, and 3D games. When performing these activities, the users can also receive feedback and scores so that the students can feel motivated.

One of the challenges in the development of these sources resides in maintaining a close connection between CALL activities and the real communicative contexts that students will face outside the classroom. Therefore, the input should reflect natural speech features in order to ensure the authenticity of the sources (Chapelle, 2001).

Recent resources have been developed mostly for English. For instance, Accent Perfect: American English Pronunciation App ³ is focused on the production skills but it also gives perceptual practise on the discrimination of American English sounds and includes a component containing sentences targeting some examples of the connected speech processes where the user can listen to

³<http://www.apps.apple.com/us/app/accent-perfect-american-english-pronunciation-app/id959744135>

the sentences. The webbased software English Accent Coach ⁴ aims to give perceptual training of English segments. Also, the purpose of the software Youglish (2019) ⁵ is to expose the user to natural stream speech. For instance, users can type in a word and listen to various productions of that word in context. Thus, they can train their perception skills concerning some connected speech processes.

EP research on this topic also contributed to the development of more useful tools and software to teach EP as a second language. For instance, REAP.PT (REAders-specific Practice PorTuguese) (Marujo, 2009; Marujo et al., 2009), a Portuguese version of REAP (Collins-Thompson and Callan, 2004; Heilman et al., 2006) (REAders-specific lexical Practice for improved reading comprehension), aims the practice of reading, listening, grammar and the learning of vocabulary, providing an interactive and individualised experience to the users (Baptista et al., 2010). The system automatically generates authentic study materials that are adequate to the needs and interests of the students. For instance, they are exposed to text-based exercises in which they can listen to the correspondent audio and learn new vocabulary. The system also generates questions about words included in the document (Correia et al., 2010; Correia, 2010).

An interesting application of the ASR technology on the development of this resource contains the automatic transcription of an EP daily broadcast news, contributing to the use of multimedia documents as learning materials. REAP.PT also includes exercises in a gaming context (Ribeiro et al., 2010; Silva et al., 2011, 2012), such as using 3D gaming environments, attractive to the students, motivating them.

Pellegrini et al. (2011) and Pellegrini et al. (2012) also describe the development of different listening comprehension games. For instance, in 'word-puzzle' games, the users can listen to sentences (automatically selected) and then reconstruct them (as in a puzzle) using a given list of words. Other exercises allowed the students to watch broadcast news and follow the automatic transcription of the audio. Interestingly, a karaoke feature added to these exercises allowed the user to follow the transcription as words were highlighted as they were produced.

⁴<http://www.englishaccentcoach.com/index.aspx>

⁵<https://youglish.com/>

Chapter 3

Methodology

This chapter describes the methodology used in this dissertation. Firstly, it will briefly characterise the attendance at an annual course of Portuguese as a Foreign Language, describing the students and some activities performed by them. Next, it will outline the design of a set of exercises handed out to those students and to a group of European Portuguese native speakers.

3.1 Level adequacy and classes

To understand the main difficulties that L2 learners experience in listening, including the researcher in the students' classroom context was considered to be an essential first step. Therefore, we were given permission to attend a course of Portuguese as a foreign language, taught at Instituto de Cultura e Língua Portuguesa (ICLP), at Faculdade de Letras da Universidade de Lisboa (FLUL).

The current study aims to understand the acquisition trajectory of vowel reduction and connected speech processes, and understand the right moment in the learning process to teach them. Thus, the learning goals and listening skills pertaining to different proficiency levels were taken into account. According to the Common European Framework of Reference for Languages (CEFR) ¹, the six

¹<https://www.coe.int/en/web/common-european-framework-reference-languages/the-cefr-descriptors>

levels of proficiency in a given language are divided into three main groups: 'Basic Users' (A levels), 'Independent users' (B levels) and 'Proficient users' (C levels).

Proficient users (C levels) are expected to have a native-like performance. As mentioned, even advanced L2 students have difficulties in speech segmentation (Altenberg, 2005). Despite these difficulties, exploring the segmentation abilities of advanced students may not contribute to understand the initial stages of the acquisition of the phonetic-phonological processes, and the students' learning trajectory.

Students at A levels (A1 and A2) have a fundamental knowledge. The size of the vocabulary of a beginner is much smaller than an intermediate or an advanced user (Cutler, 2012a). The semantic and syntactic knowledge is also limited. Regarding the overall listening skills, the CEFR indicates that students at A2 levels can understand general topics, expressions, keywords and main points related to immediate priority topics provided slowly articulated speech. Thus, students might recognise some words in connected speech if the speech is slowly articulated. However, this speech style may not encourage the natural interplay of vowel reduction and connected speech processes (as expected in spontaneous speech (Mateus et al., 2005)). Therefore, exploring the listening skills of A levels may not contribute to clearly understand the impact of these processes in continuous speech since students are not able to cope with standard speech.

Finally, the CEFR also indicates the 'Independent Users' (intermediate levels B1 and B2) can cope with standard spoken language. For instance, B2 students can understand native speakers on both familiar and unfamiliar topics. They can also understand TV news, plays, live interviews, films and radio documentaries, and follow lectures, talks and discussions. These learners can actively engage in conversations between native speakers. These students might experience difficulties in noisy environments, but they can use contextual clues to achieve a clear comprehension.

B1 students understand both general messages and main points of extended discussions, conversations, lectures, talks, news, TV programmes and audio broadcasts in clearly articulated standard speech. This input must be on familiar matters and related to topics of personal interest or everyday situations. However, they will sometimes need to ask the interlocutor to repeat particular words and phrases. The students can also recognise unknown words in continuous speech and deduce its

meaning from the context (if they are familiar with it).

The goal of this study is to understand what is the right moment, in the learning process, to aware students of how the phonetic-phonological phenomena affect listening, and give adequate perceptual training to improve the students' segmentation skills. B levels are a transitory stage in which learners cope with spontaneous speech and recognise words in the continuum stream. Thus, the B1 level may be the starting point of the development of strategies to deal with vowel reduction and connected speech processes, recognise words in spontaneous speech, and make use of other clues such as context to identify words. Therefore, we were given permission to attend an annual B1 course.

3.2 B1 level classroom assistance

The class consists of 12 students, all female, aged between 20 and 25. They were Mandarin Chinese native speakers, with the exception of a student who also spoke Cantonese. Previously to the course, the students learned Portuguese in China and then took the intensive course at ICLP. Among the course subjects, *Compreensão Oral e Escrita* ('Oral and Written Comprehension'), *Conversação* ('Conversation') and *Atividades de Imersão* ('Language Immersion Activities') were the mostly dedicated to listening, oral comprehension and spoken interaction.

The students performed listening exercises such as fill in the blanks, and listening comprehension with true or false questions, for instance. The teacher mostly used news, commercials, radio podcasts, interviews and songs as input. As homework tasks, for instance, the teacher asked the students to listen to interviews (with native speakers) and transcribe them. To train the production skills, the teacher asked the students to work on discussions, debates, oral Q&A, oral presentations and to perform role-play tasks. Also, the students frequently interviewed EP natives. With these exercises, students trained the spoken interaction among themselves, the teacher and other EP native speakers.

Overall, the students proved to have adequate skills pertaining to their proficiency level regarding both writing and reading components. They also had an adequate knowledge of the grammar

and vocabulary. The listening component revealed to be the one in which students experienced more difficulties and this was evident when the students tried to understand the general message and the main points of news, commercials, radio podcasts and broadcast audio materials delivered in standard speech. Besides, they revealed difficulties following the teacher's explanations, talks and questions (in oral Q&A) despite the delivery in clear articulated speech. The students reported that they only recognised some keywords. In those situations, the teacher repeated the questions and the explanations in slowly articulated speech to make herself understood. She also simplified the structure of her sentences and used synonyms of some expressions.

This highlights that at the B1 level students are developing the necessary skills to segment oral input in order to achieve a successful listening comprehension. It is interesting to underline that in discussions or spoken interactions among the students, they better understood each other's utterances than the teacher's productions. This is consistent with previous research by Van Engen et al. (2010), which concluded that L2 learners with shared L1 understand each other better than an L2 learner of a particular language understands a native speaker of that language.

Despite some difficulties, the students revealed to achieve the pertaining production skills for their level. However, it is interesting to highlight that it was noticed that students do not produce vowel reduction and connected speech processes. This resulted in the production of more vocalic segments (than expected in fluent EP), either in reading tasks or in (semi-)spontaneous speech. Even semi-vowels (present in diphthongs, Mateus et al. (2005)) were generally produced as their corresponding vowels. This is consistent with research on rhythmic patterns of L2 Chinese learners of EP (Zhou et al., 2017), which revealed that B1 students produced more vocalic segments than more advanced learners (C2 level).

Interestingly, the researcher also noticed that the students produce flat intonation patterns, and that they utter every word (even the unstressed words) as stressed applying a tone. Together these facts are evidence of linguist transfer (Cutler, 2012a; Maastricht, 2018), as expected for this proficiency level.

3.3 Perceptual experiments

This section describes two perceptual experiments applied to the B1 classroom (previously characterised) and to a control group of EP native speakers. Each experiment included two word identification tasks increasing the complexity of the structures - from single words to continuous speech. So that the exercises could be seen as a natural extension of the classroom dynamics, they were integrated into the topics taught in the course.

The design and the time frame of the experiments aim to understand the acquisition trajectory of the phonetic-phonological processes (concerning the competences for B1 levels), and underpin the didactic outcomes of the results. Before performing the experiments, both groups completed and signed an informed consent, and the researcher clarified the aim of the project.

3.3.1 Isolated word identification tasks

The first experiment took place around a month after the beginning of the course. It consisted of two single word identification tasks, similar to a dictation activity. Therefore, the students' transcriptions may give clues to understand not only the main difficulties they experience in listening, but also what they really perceived in the stream speech.

The first task tested stressed words targeting the frequent EP patterns of vowel reduction (Delgado Martins, 1975; Mateus et al., 2005). However, the words were produced as citation forms without applying vowel reduction. Thus, the vowels were not altered or deleted. To avoid tiring effects caused by an increase of the duration of the task, only ten words were selected. The target words were nouns with an average extension of 4 syllables.

Vocabulary also has a significant influence in speech perception. Therefore, to ensure students knew the words they were all selected from exercises performed in their classes within the same teaching unit or topic. The words were recorded by an EP native speaker, in a silent room, using the recording system from a Samsung Galaxy Note II smartphone. Recordings were then converted

Table 3.1: Examples of words targeted in task 1.

Word	Translation	Transcription
Identificação	Identification	[idẽtifike'sẽw]
Promoção	Promotion	[prumu'sẽw]
Telefone	Telephone	[teli'foni]

Table 3.2: Examples of words targeted in task 2.

Word	Translation	Transcription
Atividade	Activity	[etvi'dad]
Produtores	Producers	[prud'torf]
Entrevista	Interview	[ẽtr'viʃtẽ]

from m4a to WAV format with an online converter ² and edited in Praat software (Boersma and D, 2009). Each recording included 20 milliseconds (ms) of silence before and after the production of the word. The following Table 3.1 presents some of the words tested in this task.

The students were given a form with written instructions about the task. After reading it, the researcher resumed the instructions of the experiment. The students were instructed to write down the words, in their form, as soon as they identified them. The words were presented twice in a random order. Unfortunately, it was noticed that the acoustic conditions of the room were not the ideal ones, but a real scenario. The task took around 10 minutes to be completed.

In the second task, the students were exposed to words produced with vowel reduction and the consequent deletion of segments. The vowel reduction patterns followed the frequent ones described for EP (Delgado Martins, 1975; Mateus et al., 2005) such as the centralization of [i] (and its deletion), and the deletion of [u] and [i], in unstressed positions. The stimuli were again ten stressed words, all nouns, with an averaged extension of 4 syllables. The target words were also

²<https://onlineconvertfree.com/pt/convert-format/m4a-to-wav/>

selected from the teacher's exercises to ensure the students were familiar with the vocabulary. The recording and edition processes were replicated.

Before performing the task, the researcher briefly reviewed the instructions. The students heard the stimuli of each isolated word, which was presented twice, and wrote them down. The presentation of the stimuli was randomised. Table 3.2 includes some of the selected words for task 2.

Finally, the students were questioned about the difficulties they had when performing this experiment. Overall, the students reported that task 2 was more challenging than task 1, because the words were pronounced faster, and some letters and sounds were missing.

3.3.2 Cross word identification tasks with vowel reduction and connected speech processes

So far, L2 perception studies with European Portuguese have not addressed connected speech processes yet. These phenomena cause significant changes in the sounds, compromising listening. Consequently, even if learners know words in isolation, they may not recognise those same words in continuous speech (Vandergrift and Goh, 2012). This study aims to explore how word identification and perception are affected when words are embedded in connected speech processes.

Therefore, to increase the difficulty of the experiment, two more tasks were designed to test word identification in continuous speech combining vowel reduction (naturally present) with connected speech processes. This experiment replicates the fill in the blanks exercises performed in the course. The students' transcriptions will indicate their difficulties and how they perceived word boundaries when these phenomena are present.

For this reason, it was essential to use authentic speech material presenting real communicative contexts. Thus, for task 3 it was selected a news video ³, 01:29 minutes long, related to the topics the students were learning. The students were given an incomplete script of the audio with nine word pairs missing (totalling 18 different words). Table 3.3 illustrates the processes targeted in the sequences: vowel encounters (Andrade and Viana, 1992) (top row) and voicing assimilation

³<https://sicnoticias.pt/pais/2017-05-26-Ha-cada-vez-mais-jovens-a-sair-do-pais>

Table 3.3: Examples of word sequences in task 3.

Word	Translation	Transcription
nove anos	nine years	[ˈnɔvˈɐnuʃ]
já estava	was already	[ˈʒaʃˈtava]
jovens saem	youngster leave	[ˈʒovẽjʃˈsajẽj]
países preferidos	preferred countries	[pɐˈizʃprɐfˈridʃ]
dois mil	two thousand	[ˈdojʒˈmiɫ]
depois vêm	then come	[dˈpojʒˈvẽjẽj]
melhores oportunidades	better opportunities	[mˈɫɔrzɔprtuniˈdadʃ]
país onde	country where	[ˈpɛizˈɔd]

patterns such as the production of [ʃ], [ʒ] or [ʒ] (Mateus et al., 2005) (second, third and fourth rows, respectively).

To avoid an increase of the duration of the task, there were two occurrences for each category. The target sequences encompassed lexical words, all stressed, with three syllables on average. The words were familiar to the students since they were also taken from exercises they performed in their classes. Before performing the task, the students read the transcription in their forms, and they were told it was a fill in the blanks exercise. This task lasted around 8 minutes and only the audio of the video was presented (twice).

The design of task 4 was similar to the previous one but targeted more complex connected processes (Andrade and Viana, 1992; Mateus et al., 2005), increasing the difficulty of the task. It was selected an excerpt of a news video ⁴ (from Youtube), of 02:36 minutes, including real interviews. Therefore, the students were exposed to more spontaneous speech (Cutler, 2012a). The students were given a script of the audio with 11 word sequences (totaling 29 different words) missing. As Table 3.4 presents, the target sequences included vowel encounters (row 1), production of [j] (row

⁴<https://www.youtube.com/watch?v=9yWX0gqBdgo>

Table 3.4: Examples of word sequences in task 4.

Word	Translation	Transcription
vinte à hora	twenty per hour	[ˈvĩtaˈɔrɐ]
uma antiga colega	an old colleague	[ũmãˈtigekʷˈlɛgɐ]
sabe que o	knows that the	[ˈsabkju]
penso que a	I think that the	[ˈpẽskjɐ]
resultado do encontro	meeting result	[rzuɫˈtad:wẽˈkõtrʷ]
sede de trabalhar	thirst to work	[ˈsed:trɐbɐˈʎar]
tentar a sorte	try your luck	[tẽˈtarɐˈsɔrt]
mil euros	a thousand euros	[ˈmilˈewrɔʃ]

2), haplogologies (row 3), and resyllabification with [r] and [l] (row 4). Only two target sequences for each process were selected, and the interval time between these occurrences was also controlled to avoid tiring effects. The sequences included stressed and unstressed words (clitics, see Vigário (2011:179) from different lexical categories, ranging between 1 to 6 syllables.

The audio of the video was presented twice and this task lasted around 10 minutes. Both tasks 3 and 4 were performed on the same day, and implemented two weeks after the first experiment.

Finally, the researcher questioned about the difficulty of the exercises. The participants on the whole commented that the first task was easier than the second one (task 4). Also, they reported that they had difficulties understanding the interviewees in the audio of task 4, since they spoke too fast.

3.3.3 Control Group

To set up a baseline, the experiments were also applied to a group of 12 EP native speakers aged between 23 and 27. Participants were either undergraduate or postgraduate students at the University of Lisbon. These participants performed the same sequence of tasks in different moments. The

experiments were applied in silent rooms such as libraries and study or meeting rooms.

After carrying out the activities, the researcher also asked about the difficulty of the tasks. Regarding the first experiment, participants reported that words in task 1 were hyper articulated, but the pronunciation in task 2 was natural. They added they were not used to perform fill in the blanks exercises since they left high school. However, they had no difficulties performing the experiments.

Chapter 4

Results

This section includes the results of the perception experiments applied to the B1 classroom and the control group of EP native speakers. Firstly, it reports the overall results for the four tasks comparing both groups. Secondly, it describes the results per phenomena (vowel reduction and connected speech processes). Finally, an analysis of the student's transcriptions is also described to explore their main difficulties in listening and the frequent errors they commit.

4.1 Results of perception experiments

In these experiments students wrote their answers in worksheets. So, to facilitate the data analysis, the answers were transcribed to Word. When analyzing students' answers, several orthographic forms were used for the same word. However, it was important to ensure that words were identified despite the occurrence of wrong orthographic transcriptions. Therefore, to evaluate the accuracy of word identification, some criteria were carefully established to classify whether a word was identified, or not, despite the misleads in the transcriptions. These guidelines were established on the basis of the most common errors EP learners native speakers of Mandarin Chinese tend to commit (Zhou, 2017; Zhou et al., 2017, 2019, 2021; Nunes, 2015; Castelo and Santos, 2017; Castelo, 2018; Castelo and Freitas, 2019; Castelo, 2019), and on the fact that some alternations between graphemes

are common among EP speakers during the L1 acquisition (Freitas, 1997; Fikkert and Freitas, 2006; Freitas, 2003; Santos et al., 2014). Therefore, despite the occurrence of errors in the orthographic transcriptions it was possible to understand if the target words were identified.

Therefore, an answer was considered correct when:

- the word was orthographically correct;
- some graphemes could correspond to the phonetic transcription of such grapheme, for instance alternations between the grapheme 'i' and 'e', both corresponding to the sound [i], or alternations between 'u' and 'o' to express [u];
- the same grapheme is used to represent voiced and voiceless segments, for instance [t] and [d];
- changes in the order of the segments occurred as in metathesis.

The tasks in the second experiment targeted sequences of words. So, an answer was considered correct when all the words of a such sequence were identified (on the basis of the previous criteria).

4.1.1 Overall Results

The reduced dimension of the classroom (N=12) affected the amount of collected data and can only be seen as a representative sample. We calculated the rate of correct answers (in percentage), the mean, and the standard deviation. The overall results, the mean of correct answers, and the standard deviation of the four tasks are presented in Tables 4.1 (for Mandarin speakers) and 4.2 (for EP native speakers).

As mentioned, the first experiment involved single words identification. The first task targeted words in citation forms, and the second one words produced with vowel reduction. As shown in the Table 4.1, the percentage of correct answers was 94% in task 1 and decreased substantially to 65% in task 2. These results reveal that students identified the majority of the words in the first task (mean=9.4; sd= 0.8) while in the second one, students identified (on average) at least 6 words

Table 4.1: Percentage of correct answers of each experiment for Mandarin speakers.

Tasks	Correct Answers	Mean	s.d.
1 Single word without reduction	94%	9,4	0,8
2 Single word with reduction	65%	6,5	1,2
3 Simple connected speech	31%	2,5	1,7
4 Complex connected speech	16%	1,6	1,0

Table 4.2: Percentage of correct answers of each experiment for the control group.

Tasks	Correct Answers	Mean	s.d.
1 Single word without reduction	100%	10	0
2 Single word with reduction	100%	10	0
3 Simple connected speech	93%	7,4	0,8
4 Complex connected speech	98%	9,8	0,4

(mean= 6.5; sd=1.2). In both tasks, two outlier words were excluded: 'repórter', in task 1, and 'reportagem' in task 2.

The purpose of the second experiment was to test words embedded in connected speech processes with vowel reduction naturally present. Thus, the target sequences included word pairs, in task 3, and ranged between two or three words (stressed and unstressed), in task 4. What stands out in Table 4.1 is the overall rate of correct answers, which was less than 50%: the percentage of correct answers was 31% in task 3 and even lower in task 4 (16%). There can be noticed a variability in the responses. On average, students identified at least two sequences, in task 3, but some students identified even more sequences (mean=2.5; sd= 1.7). The number of identified sequences varied from one to three (mean= 1.6; sd = 1), in task 4.

An important note on the results: an outlier target sequence in task 3 was excluded ('cento afirma'), due to the fact that it is a mathematical expression written by extent, and these expressions are commonly considered in classes in its numeric form. In task 4, the sequence 'sede de trabalhar' was also excluded due to its metaphorical sense, which might have compromised word recognition.

In summary, what it can be clearly seen in the results for Mandarin Speakers is a decrease in the overall scores from task to task (see Table 4.1). For instance, there is a gap of 29% between task 1 (with words in citation form) and task 2 (words with vowel reduction). A notorious difference of 34% in the scores occurred when word identification tasks changed from isolated words to continuous speech. The results also show a slight decrease of 15% from task 3 to task 4.

Data from Table 4.1 can be compared with the data in Table 4.2, which shows that the control group had higher scores, as expected. The percentage of correct answers was 100% in the first Experiment (tasks 1 and 2). Participants scored 93% (mean=7.4; sd=0.8) in task 3, while in task 4 scored 98% (mean=9.8; sd=0.4). So, no greater differences in the scores for native speakers were evident despite temporarily parsing problems.

Together these results provide important insights into the impact of vowel reduction and connected speech processes in word recognition. Firstly, the differences in the scores for Experiment 1 and 2 reveal that recognising words in continuous speech (when connected speech processes and

vowel reduction are present) is more difficult than in isolation.

In fact, this corroborates **the initial hypothesis (i) about a clear effect of connected speech processes and vowel reduction on the recognition of words in continuous speech**. So, students at the intermediate level experience difficulties in the recognition of known words in the continuous stream due to the presence of these phonetic-phonological phenomena.

Besides, the results in Experiment 1 reveal that vowel reduction compromises word recognition even when words are produced in isolation. This is a surprising outcome since higher scores were expected to single words identification tasks according to the competences pertaining to B1 levels. In this case, **the hypothesis (ii) about the low effect of vowel reduction on the recognition of spoken isolated words is not confirmed since the results indicate that vowel reduction had a greater effect (than expected) on single word recognition**.

Moreover, **the differences in the scores also confirm the stated hypothesis (iv) about the effect of the type of input that students are exposed to**. In this research, the Experiment 1 tested read input while Experiment 2 tested authentic materials. The scores in the Table 4.1 reveal that the tasks with (semi-) spontaneous speech had lower scores than the tasks with read speech.

Finally, when questioned about the difficulty of the experiments, the participants were unanimous in the view that the speech in Experiment 1 was easier to understand than speech in the second experiment. The students also argued that speech in task 4 seemed faster than in task 3. Together these facts also indicate that the type of input also affects the perception of L2 learners.

4.1.2 Distribution of phonetic-phonological processes

This section details the results of both groups of participants for vowel reduction (and its absence) tested in Experiment 1, and for connected speech processes, tested in Experiment 2.

Table 4.3: Percentage of correct answers for each occurrence of vowel reduction for Mandarin speakers, in both tasks 1 and 2.

Vowel Reduction Patterns	Correct answers (citation form)	Correct answers (reduced form)
Centralization (and deletion) of [i]	97%	61%
Deletion of [u]	92%	58%
Deletion of [i]	94%	78%

4.1.2.1 Citation Form vs. Vowel Reduction

The words tested in task 1 (in the citation form) targeted contexts for centralization of [i] and its deletion, deletion of [u] and deletion of [i] in unstressed positions. The second task encompassed words produced with vowel reduction in the same target contexts. The following Table 4.3 shows the percentage of correct answers of each vowel reduction pattern (in both tasks) for Mandarin speakers: centralization of [i] and its deletion (row 1), deletion of [u] (row 2) and deletion of [i] (row 3).

It can be seen from the data in Table 4.3 that no great differences were found in the scores for citation forms in each context. Centralization and deletion of [i], as in word 'discurso', had a rate of 97%, deletions of [u], as in word 'comunicação' rated 92%, and deletion of [i] in unstressed positions, as in word 'telefone', had 94%.

The most interesting aspect of this table is the variability in the scores for words produced with vowel reduction. The pattern of deletion of [i] in unstressed positions had the highest percentage (78%), but the rates for deletion of [u] as in words 'computador' and 'produtores, and centralization (and deletion) of [i], as in words 'atividade' and 'exibição', decreased to 58% and 61% respectively.

It is interesting to compare the results of both citation and reduced forms of words, in each target context. For instance, there is a difference of 16% between the percentages for the citation form and the reduced form of words targeting the deletion of [i], in unstressed positions. However, the difference in the scores of these both forms increased in the contexts targeting the deletion of

Table 4.4: Percentage of correct answers for each occurrence of vowel reduction for native speakers, in both tasks 1 and 2.

Vowel Reduction Patterns	Correct answers (citation form)	Correct answers (reduced form)
Centralization (and deletion) of [i]	100%	100%
Deletion of [u]	100%	100%
Deletion of [i]	100%	100%

[u] (34%) and centralization (and deletion) of [i] (36%).

Finally, the Table 4.4 presents the results for the control group for each target context, in both tasks. It is notable that the percentage of correct answers was 100% in all contexts, either in citation or reduced forms.

4.1.2.2 Connected Speech Processes

The second experiment targeted word sequences with connected speech processes and vowel reduction naturally included. Voicing assimilation and vowel encounters were tested in task 3, whereas task 4 tested production of [j], resyllabification with [r] and [l], haplogogies and other occurrences of vowel encounters. Table 4.5 provides an overview of the correct answers, in percentage, of these phonetic-phonological processes for both test and control groups (middle and right columns, respectively): voicing assimilation (row 1), vowel encounters (row 2) production of [j] (row 3), resyllabification with [r] and [l] (row 4), and haplogogies (row 5).

It can be seen from the data in the Table 4.5 that the results for the control group were above 90% whereas the results for Mandarin speakers were under 30%. Concerning this last group (see middle column), the rate of correct answers for vowel encounters was 26%, and 24% for the occurrences of voicing assimilation. production of [j] and resyllabification with [r] and [l] had identical scores (21%). What is striking about the scores for L2 learners in this table is the absence of correct answers when words targeted haplogogies (0%).

Table 4.5: Overall percentage of correct answers of each occurrence of connected speech processes for Mandarin speakers and the control group, middle and right columns, respectively.

Processes	Correct answers	Correct answers
	Mandarin Speakers	Control Group
Voicing Assimilation	24%	93%
Vowel Encounters	26%	96%
Insertion of [j]	21%	100%
Resyllabification with [r] and [l]	21%	100%
Haplologies	0%	100%

Table 4.6: Percentage of correct answers of each occurrence of voicing assimilation for Mandarin speakers and the control group, middle and right columns, respectively.

Processes	Correct answers	Correct answers
	Mandarin Speakers	Control Group
Assimilation [z]	13%	92%
Assimilation [ʃ]	29%	92%
Assimilation [ʒ]	29%	96%

The results set out in the Table 4.5 (see right column) show that native speakers rated 100% in the occurrences of production of [j] and haplologies, 96% in vowel encounters and resyllabification with [r] and [l], and 93% in the occurrences of voicing assimilation. In sum, no greater differences in the scores for each process were observed for this group for each target process.

It is also interesting to include a further analysis on the results for voicing assimilation patterns. In task 3, the target sequences included assimilation of [z] (row 1), assimilation of [ʃ] (row 2) and assimilation of [ʒ] (row 3). The Table 4.6 below provides the scores of each occurrence for both test and control groups (on the middle and right columns, respectively).

From the data for Mandarin speakers, in the Table 4.6, it can be seen that the occurrences

of assimilation of [z] (as in the sequence 'país onde') resulted in the lowest percentage of correct answers (13%). The scores for assimilation of [ʃ] and [ʒ] were 29%. In contrast, and as expected, the control group achieved higher scores: 92% for assimilation of [z] and [ʃ], and 96% for assimilation of [ʒ].

Together these results provide important preliminary insights into the impact of connected speech processes in L2 perception, at level B1. According to hypothesis (iii) it was expected that some connected speech processes had more impact than others in word recognition. In fact, the results reveal that haplogogies, for instance, might compromise word recognition, since there were no correct responses. The detailed analysis of the scores for the occurrences of voicing assimilation suggests that some patterns can affect the allocation of word boundaries. **Despite the small amount of data, these preliminary results may indicate that some connected speech processes may have a more evident impact in L2 perception than others confirming hypothesis (iii).**

4.1.3 Analysing Orthographic Transcriptions

Gap filling exercises and dictation tasks allow the students to write what they perceived in the stream speech. For instance, students can transcribe the identified words and their transcriptions will indicate the accuracy of both listening and writing components. Furthermore, when learners do not recognise the words, their transcriptions will reveal which sounds and segments they actually perceived.

In this research, the analysis of the transcriptions may give important insights on how vowel reduction, and its consequent deletion, affects word perception (even in read speech) and how the alterations in the sounds caused by connected speech processes affect the allocation of word boundaries. Therefore, this section presents the transcriptions of the students by task.

4.1.3.1 Experiment 1

Tables 4.7, 4.8 and 4.9 present the students' transcriptions of the words tested in task 1 (words produced without vowel reduction): centralization and deletion of [i], deletion of [u] and deletion

Table 4.7: Transcription of the answers of Mandarin speakers for words in the citation form of centralization and deletion of [i], in task 1

Identificação	Discurso	Utilização
Indentificação	Descurso	Utilização
Identificação	Descurso	Utilização
Identificação	Descurso	Utilização
Identificação	Discurso	Utilização
Identificação	Discurso	Utilização
Identificação	Discurso	Utilização
Identificação	Discurso	Utilização
Identificação	Discurso	Utilização
Identificação	Discurso	Utilização
Identificação	Discurso	Utilização
Identificação	Descurso	Utilização
Ecenfiticação	Discurso	Utilização

of [i], respectively.

From the data in the Tables 4.7, 4.8, 4.9, it is apparent that the students tend to commit the same pattern of errors. For instance, Table 4.8 shows examples of an alternation between 'o' and 'u' to represent the sound [u] in words 'documento' and 'promoção', which were written as 'ducomento' and 'promução'. The students also alternated between 'i' and 'e' to represent the sound [i], as in 'discurso' and 'descurso' (see Table 4.7). Thus, the learners used different graphemes to represent the same sound.

Furthermore, it is also common the occurrence of alternations between 't' and 'd' to represent the voiced and voiceless sounds [d] and [t], as in word 'consumitor' (to represent 'consumi[d]or', Table 4.8), and 'demefora' (to represent '[t]elefone', Table 4.9). This indicates that the discrimination of this pair of voiced and voiceless consonants is difficult for these students.

Table 4.8: Transcription of the answers of Mandarin speakers for words in the citation form of deletion of [u], in task 1.

Comunicação	Promoção	Documento	Consumidor
Comunicação	Promoção	Ducomento	Consumido
Comunicação	Promoção	Documento	Consumidor
Communicaçãõ	Promuçãõ	Ducumento	Consumitor
Comunicação	Promoção	Documento	Consumidor
Comunicação	Promuçãõ	Ducomento	Consumidor
Comunicação	Promoção	Documento	Consumidor
Comunicação	Promoção	Documento	Consumidor
Comunicação	Parmação	Ducomento	Consumitor
Comunicação	Promoção	Documento	Consumidor
Comunicação	Durmusão	Ducomento	Consumidor
Comunicação	Ruçãõ	Documento	Consumidor
Comunicação	Propuçãõ	Documento	Consumidor

Table 4.9: Transcription of the answers of Mandarin speakers for words in the citation form of deletion of [i], in task 1.

Telefone	Representação	Conhecimento
Telefone	Representação	Conhecimento
Telefone	Representação	Conhecimento
Telefone	Representação	Conhecimento
Telefone	Representação	Conhecimento
Telefone	Representação	Conhecimento
Demefora	Representação	Conhecimento
Telefone	Representação	Conhecimento
Teleforne	Repersentação	Conhecimento
Telefone	Representação	Conhecimento
Telefone	Representação	Conhecimento
Telefone	Representação	Conhecimento
Telefone	Representação	Conhecimento
Telefone	Representação	Conhecimento

From these data it is also interesting to note that the students tend to simplify the syllabic structure of the words, avoiding consonant sequences on the onset position. Thus, a common used strategy was metathesis, and it can be seen in the transcriptions of the word 'representação', e.g. 'repersentação' (see Table 4.9), and of the word 'promoção', e.g. 'parmação' and 'durmusão' (see Table 4.8). Some students also deleted the segments in coda position, as in the word 'consumidor', presented in the Table 4.8, which was written as 'consumido'.

Interestingly, some transcriptions may reveal which are the prominent cues even when words are not recognised. As an example, the transcriptions of the words 'promoção' and 'telefone' reveal that students mostly perceived the stressed syllable, e.g. 'parmação', 'durmusão', 'rução' and 'propução' (see Table 4.8, for word 'promoção'), and 'demefora' (see Table 4.9, for word 'telefone').

Turning now to the second task, Tables 4.10, 4.11 and 4.12 present the students' transcriptions of the words targeting the same patterns, but produced with vowel reduction: centralization and deletion of [i], deletion of [u] and deletion of [i], respectively.

As previously established, students commit similar patterns of errors. For instance:

1. alternations between 'i' and 'e', as in 'esvisão' and 'isovisão' to represent '[i]xibição' (see Table 4.10), and in 'intrevista' to represent '[ẽ]ntrevista' (see Table 4.12). Also, students represent the sound [i], as in the word 'acontecimento', by writing 'aconticimento' (see Table 4.12).
2. alternations between 'o' and 'u' to represent sound [u], as in 'compotador', presented in the Table 4.11, to represent 'computador';
3. alternations between 't' and 'd' to represent [t] or [d], as presented in the Table 4.12, for word 'acontecimento', e.g. 'acondicimento'.

In addition, Table 4.12 shows examples of metathesis (a strategy to simplify the syllabic structure), as in the word 'entrevista', which was transcribed as 'intervista'.

Another important aspect of the transcriptions can be seen in Table 4.10, when students transcribed the word 'exibição' as 'exvisão' and 'esvisão', and in Table 4.12, in which the word 'tele-

Table 4.10: Transcription of the answers of Mandarin speakers for words produced with the centralization (and deletion) of [i], in task 2.

Participação	Atividade	Exibição
Participação	Atividade	Exibição
Paricipação	Atividade	Edivição
Participação	Atividade	Etuição
Participação	Atividade	Efinação
Participação	Atividade	Esvisão
Participação	Atividade	Exibição
Participação	Atividade	Isuvição
Parcitação	Atividade	Etição
Partipação	Atividade	Isovisão
Partipação	Atividade	Edivição
Participação	Atividade	Exibição
Partipação	Atividade	Exvisão

Table 4.11: Transcription of the answers of Mandarin speakers for words produced with the deletion of [u], in task 2.

Comunidade	Produtores	Computador	Educação
Comunidade	Prontos	Computador	Educação
Comunidade	Prodores	Computador	Emocação
Comidade	Prodores	Computador	Educação
Comunidade	Produtos	Computador	Educação
Comidade	Quintos	Computador	Emigração
Comunidar	Produz	Computador	Educação
Comunidar	Produz	Computador	Educação
Comunidaden	Quintos	Compotador	Equeção
Comunidade	Produz	Computador	Educação
Comunidade	Quintos	Computador	Educação
Comunidade	Frutos	Computador	Educação
Conidade	Produz	Computador	Educação

Table 4.12: Transcription of the answers of Mandarin speakers for words produced with the deletion of [i], in task 2.

Televisão	Entrevista	Acontecimento
Televisão	Entrevista	Acondicimento
Televisão	Entrevista	Acomtimento
Televisão	Intrevista	Aconticimento
Televisão	Entrevista	Aconhecimento
Televisão	Entrevista	Aguarcimento
Tervisão	Intrevista	Acontecimento
Televisão	Entrevista	Acontecimento
Televisão	Intrevista	Acontecimento
Televisão	Entrevista	Acontinuito
Televisão	Entrevista	Acontimento
Televisão	Entrevista	Reconhecimento
Televisão	Entrevista	Aquandotimento

visão' was transcribed as 'tervisão'. In these cases, it is evident that students perceived the deletion of the segment [i].

Furthermore, the deletion of [u] (see Table 4.11) resulted in transcriptions such as 'prontos', 'prodores' and 'produz' for word 'produtores', and 'comidade' and 'conidade' for word 'comunidade'. Also, some students transcribed the word 'participação' (see Table 4.10) as 'partipação', 'parcipação' and 'parcitação', and the word 'acontecimento' (see Table 4.12) as 'acontimento'. In these examples, the students simplified the phonetic sequence of consonants that resulted from the deletion of the vowels.

Despite they perceived the deletion of vowels, some students intended to represent the deleted segments with a grapheme, even without identifying the word. For instance, some students represented the deleted segment [i] (due to the centralization of [i]) in the word 'exibição' by writing 'edivição', 'efinição', 'isovisão' and 'isuvição'. The Table 4.11 also provides examples of transcriptions including the representation of the deleted segment [u], as in 'comunidar' (to represent 'comunidade'), and 'emocação' (to represent 'educação').

It is also interesting to highlight that students transcribed parts of the words, similar words and non words with the same (or similar) segments, representing the stressed syllable. As an example, the word 'acontecimento' (presented in Table 4.12) was transcribed as 'aconnhecimento', 'aguarci-mento', 'reconhecimento' and 'aquandotimento'. Some students wrote 'equeção' and 'emigração' to represent the word 'educação', and wrote 'quintos' and 'frutos' to represent the word 'produtores' (both from Table 4.11). Besides, there are also occurrences of other non words, such as 'etuição' and 'etição' to represent the word 'exibição', in the Table 4.10.

4.1.3.2 Experiment 2

Turning now to Experiment 2, the transcriptions of the sequences of words are set out (per phenomena) in the following Tables: voicing assimilation (Tables 4.13, 4.14 and 4.15), vowel encounters (Tables 4.16, 4.17 and 4.18), production of [j] (Table 4.19), resyllabification with [r] and [l] (Table 4.20), and haplogologies (Table 4.21).

Table 4.13: Transcription of the answers of Mandarin speakers for words with the production of [z] due to voicing assimilation, in task 3.

Melhores Oportunidades	País Onde
Condições melhores	País
Procurar	País
Melhor	-
Melhor oportunidade	País que
Melhorar	País
Melhor	País
Melhores oportunidades	Países
Melhor	País
Cada vez mais oportunidades	País onde
Melhor oportunidades	-
Melhores oportunidades	País
Melhorar	Países

The first set presents the transcriptions of the word pairs targeting voicing assimilation, tested in task 3: Table 4.13 presents assimilation of [z], Table 4.14 assimilation of [ʃ], and Table 4.15 assimilation of [ʒ].

As these data confirms, overall, the students correctly transcribed at least one of the words, or transcribed parts (or segments) of both words. For example, the sequences targeting the realization of [z] ('melhor oportunidades' and 'país onde', in Table 4.13), were transcribed as 'melhor', 'melhorar', 'melhor oportunidade', 'melhor oportunidades', 'procurar' (which contains segments of word 'oportunidades'), and as 'país' (the word "onde" was rarely transcribed).

In addition, the sequences 'jovens saem' and 'países preferidos', targeting the realization of [ʃ] (see Table 4.14), were mostly transcribed as 'jovens', 'países' or 'preferidos'. Also, there are transcriptions of parts of both words, as in 'país preferidos', 'país preferem', 'jovens sai', 'alguns

Table 4.14: Transcription of the answers of Mandarin speakers for words with the production of [ʃ] due to voicing assimilation, in task 3.

Jovens Saem	Países Preferidos
Alguns jovens	Países muito felizes
Jovens saem	País preferidos
Jovens sai	País preferem
Jovens saem	Países
Jovens	País preferidos
-	Países preferidos
Jovens têm	Países preferidos
Jovens	Países
Jovens saem	Países preferidos
Jovens	Países preferidos
Jovens	Países
Jovens	Países

Table 4.15: Transcription of the answers of Mandarin speakers for words with the production of [ʒ] due to voicing assimilation, in task 3.

Depois Vêm	Dois Mil
Depois	imigna imgrimos
Depois	trimidor
pois que	dois mils
Depois veem	dois mil
Depois	dois mil
Depois	2 mil
Depois	dois mil
Depois	duzentos
Depois vai	dois mil
Depois	dois mil
Pois	pricar
Que	trinta

jovens' and 'jovens têm'.

The sequences in the Table 4.15 show that students mostly transcribed one of the words of the sequence, such as the word 'pois' of the word pair 'depois vêm', and 'dois' of the sequence 'dois mil'. Some students also wrote parts (or segments) of the words, as in 'depois vai', 'pois que', 'pois' and 'que' to represent the sequence 'depois vêm', and there are also transcriptions including non words, such as 'trimidor', 'trinta' and 'pricar' to represent the sequence 'dois mil'. These examples illustrate that when the words are not recognised, the segments that students transcribed included the stressed syllable, which was already identified as a relevant cue.

Closer inspection of the Tables 4.13 and 4.14 shows that students perceived the assimilation process, and the resulting production of the sounds [z] and [ʒ]. Thus, the transcriptions included examples of the use of graphemes to represent those sounds.

For instance, the sequence 'paí[z]onde' was transcribed as 'países', in which the grapheme <s> represents the sound [z]. Besides, there is a transcription of non words, such as 'imigna imgrimos', in which the grapheme <g> represents the sound [ʒ] produced in the word pair 'doi[ʒ]mil'.

Finally, it was also noticed in the transcriptions the use of synonymous, or related words, of the target sequences, such as 'condições melhores' and 'cada vez mais oportunidades' to represent the pair 'melhores oportunidades', as presented in Table 4.13. Also, there are occurrences of 'duzentos' to represent 'dois mil', in Table 4.14, and 'países muito felizes' to represent 'países preferidos', as Table 4.15 illustrates.

If we now turn to vowel encounters, Tables 4.16, 4.17 and 4.18 will show the transcriptions of these occurrences. Table 4.16 provides the transcriptions of the word pairs targeted in task 3. Tables 4.17 and 4.18 show the transcriptions of the sequences tested in task 4: vowel encounters with stressed words, in Table 4.17, and with clitics (Vigário, 2011), in Table 4.18.

As already mentioned, vowel encounters had the highest rate of correct answers (26%, see Table 4.5). Overall, students maintained a pattern in the transcriptions: recognising and writing one of the words of the sequence. Regarding the sequences with prosodic words, Table 4.16 includes examples in which the students mostly transcribed the word 'nove', of the sequence 'nove anos', and

Table 4.16: Transcription of the answers of Mandarin speakers for words with vowel encounters, in task 3.

Nove Anos	já Estava
Nove Anos	Estava
Nove Anos	Estava
Noves	Já estava
Nove Anos	Estava
Nove Anos	Estava
Nove	Já estava
Nove Anos	Já estável
Nove Anos	Estável
Nove Anos	Já estava
Nove Anos	Já estava
Nove Anos	Estável
Nove	Estava

Table 4.17: Transcription of the answers of Mandarin speakers for words with vowel encounters, in task 4.

Já Amanhã	Vinte à Hora	Cadeia Hoteleira
Comer	-	Redeira palavra
-	Oitente	Cabela prenta
Amanhã	Vinte	Cabela
-	-	Cabela
Passado	Vinte	Capeira
Na amanhã	Bo hora	Cateia Mdeia
Amanhã	Oitenta	Cadeira
-	20	Praia
-	35 horas	Cadeira branca
-	Oito	Cabela clara
Amanhã	Horas	Cadeira
-	-	Colega

Table 4.18: Transcription of the answers of Mandarin speakers for words with vowel encounters with clitics, in task 4.

De Encontrar	Uma Antiga Colega
Encontrar	Marquetina direito
De Encontrar	Uma colega
De Encontrar	Vai diga colega
De Encontrar	A colega
De Encontrar	A colega
Dem contrado	Matica corena
De Encontrar	Uma colega
De com certeza	Diga
De Encontrar	Antiga colega
De Encontrar	Uma colega
De Encontrar	Uma colega
De Encontrar	Uma colega

the word 'estava', of the sequence 'já estava'. Also, the word 'amanhã' was the mostly recognised and transcribed word of the sequence 'já amanhã', presented in the Table 4.17. In this Table, it is also clear that the sequence 'vinte à hora' was partly identified. For instance, students either wrote the word 'vinte' or 'hora', and the word 'horas'.

Data from these Tables can be compared with the data in Table 4.18 which show that when the sequences targeted vowel encounters with unstressed words, the same pattern is present: students mostly identified one of the words of the sequence as, for example, the word 'encontrar' in the sequence 'de encontrar', and the words 'uma colega' and 'antiga colega', in the sequence 'uma antiga colega'. These results indicate that the deletion of the vowels and the articulation process, and the assimilation and the consequent merge of vowels compromised the recognition of the target sequences.

Furthermore, there are transcriptions of parts and segments of the target words, and transcriptions of similar words, e.g. 'estável' to represent 'já estava' (see Table 4.16), and 'vai diga colega' to represent 'uma antiga colega' (see Table 4.18). Besides, there are occurrences of non words, and pseudo words, with similar sounds, e.g. 'matica corena' and 'marquetina direito', to represent 'uma antiga colega', and 'oitente' to represent the sequence 'vinte à hora' (see Table 4.17).

What stands out in Table 4.17 is the occurrence of several transcriptions of the sequence 'cadeia hoteleira', e.g. 'cadeira', 'cateia', 'cabela', 'cadeira branca', 'cabela clara', 'cabela preta', 'redeira palavra', 'capeira', 'correira', 'redeia', 'praia' and 'colega'. In these cases, the students also perceived some segments of the words and the stressed syllable (without identifying and recognising the target words).

Finally, another interesting aspect of the transcriptions of vowel encounters is that some of them indicate that students were sensitive to the articulation and assimilation processes. They transcribed a sequence with similar sounds, indicating they perceived a merged sequence without identifying the both words of the target sequence. For instance, the sequence 'nove anos', in Table 4.16, was transcribed as 'noves', and the sequence 'de encontrar', in Table 4.18, was transcribed as 'tem contrado' and 'de com certeza'.

Table 4.19: Transcription of the answers of Mandarin speakers for words with production of [j], in task 4.

Sabe Que o	Penso Que a
Será que	-
Sabe Que o	-
Sabe um	-
Sabe Que o	Quere
Sabe o	Ctou
Sabe Que o	Teve que
Sabe Que o	Uma
Sa	-
Sabia um	Q.b
Seja	Prime que
Sai do	-
Sabe que o	-

The transcriptions of the sequences targeting the production of the semi-vowel[j] are set out in the Table 4.19.

From these data (Table 4.19), it is clear that the students recognised one or two words of each sequence, e.g. the students transcribed the word 'que', in the sequence 'penso que o', and words such as 'sabe um', 'sabe o' and 'sabia um' to represent the sequence 'sabe que o'. Besides, there are occurrences of related words or even parts of the sequence such as 'quere', 'ctou', 'teve que', 'prime que', 'q.b' and 'uma' (to 'penso que a'), and 'será que', 'seja', 'sai do' and 'sa' (to 'sabe que o').

Turning to the occurrences of resyllabification, the Table 4.20 below presents the collected data for this process. The left column presents the examples for resyllabification with [l], and the right column for examples with [r].

Table 4.20: Transcription of the answers of Mandarin speakers for words with resyllabification with [l] and [r], in task 4.

Mil euros	Tentar a Sorte
Meio ouros	-
Uma milhão	-
Miho	Tentar
Mil euros	Tentar
Million	Tentar
Milhares	-
Mil euros	Tentar
Milhões	-
Mil euros	Tentar
Mil euros	Entrevista
Mil euros	Voltar
Milhões	-

Table 4.21: Transcription of the answers of Mandarin speakers for words with haplogologies, in task 4.

Resultado do Encontro
Recionista
Prerrecutar concordo
Resutar encontro
Resultado encontro
Resultar
Resultado
Resultar por encontrar
Contrário
Resultador por encontro
Resudado
-
-

The interesting aspect of this Table 4.20 is that students perceived the pair ‘mil euros’ as an articulated sequence, without determining the word boundaries. This is evident in some transcriptions of this sequence, e.g. ‘milhão’, ‘milhares’, ‘milhões’ and “million”, which are similar to the merged sequence. Besides, the students wrote one of the words of the sequence (or parts of it) such as ‘tentar’ and ‘voltar’ to represent ‘tentar a sorte’, and other words with similar sounds as ‘ouros’, representing ‘mil euros’.

Finally, the Table 4.21 presents the transcriptions for the occurrences of haplogologies which is just a sequence ‘resultado do encontro’, since the previous deletion of an outlier sequence, as already mentioned.

It is apparent from this Table 4.21 that the students mostly transcribed words such as ‘resultado’ and ‘encontro’. There are also occurrences of parts and segments of the target sequences, similar

words and non words with similar sounds, e.g. 'prerrecutar concordo', 'recionista', 'resultar', 'resulta', 'resuta', 'contrário' and 'resudado'. It is also interesting to highlight occurrences such as 'resultado encontro', 'resutar encontro', 'resultador por encontro' and 'resultar por encontrar', in which it is clear that this process affected the allocation of word boundaries and its recognition, and some students tried to represent what they perceived by including the word 'por'.

4.2 Final Remarks

Overall, these results indicate that L2 learners had more difficulties recognising words in continuous speech than in isolation, which is related to the occurrence of the connected speech processes and vowel reduction. Besides, the identification of spoken isolated words was compromised due to vowel reduction and its consequent deletion of segments.

Further analysis also reveals that some connected speech processes may bring more difficulties in word boundaries allocation. Preliminary results indicate that some patterns of vowel reduction may have a greater impact in word recognition than others.

The analysis of the transcriptions reveal that learners have difficulties in the perceptual discrimination of some segments, and that they apply strategies such as metathesis and coda deletion to simplify the syllabic structure of the words. It was also noticed that learners deleted consonants to cope with the production of phonetic sequences of consonants due to the deletion of vowels.

In addition, some students transcribed parts or segments of the words, including the stressed syllable. There are also occurrences of non-words as an attempt to represent the perceived sounds. Despite sometimes the students did not recognise the words, their transcriptions reveal that they perceived the deletion of the vowels, and transcribed the absence of segments.

Surprisingly, some students tried to represent the missing sounds with other graphemes. Regarding the recognition of sequences targeting connected speech processes, the transcriptions reveal that students identified whether one of the words and parts or segments of both words. Some students also wrote similar words, sometimes synonyms, and non words (or pseudo words) including the

segments they perceived. Some transcriptions indicate that students perceived the sequences as a block, without determining the word boundaries, and wrote merged sequences representing the altered sounds with graphemes.

Taken together, these results suggest that vowel reduction also had influence across word boundaries because students perceived the deletion of vowels and the consequent presence of the connected speech processes.

Chapter 5

Discussion

5.1 Overall results

The results of these perceptual experiments indicate that the students, at the B1 level, have great difficulties facing continuous speech due to the interaction of vowel reduction and the connected speech processes. Experiment 1, with single word identification had higher rates than Experiment 2, which involved word identification in continuous speech. As mentioned, the rates of correct answers in the tasks of the first experiment were 94% and 65% (task 1 and task 2, respectively).

Previous studies (Goh, 2000; Altenberg, 2005; Snijders et al., 2007; Cutler, 2012a; Graham and Santos, 2015; Gaskell and Mirkovic, 2016; Reed, 2019) have demonstrated that the major challenges in second language acquisition are related to speech segmentation and word recognition.

On the question of connected speech processes, the results for Experiment 2 indicate that word recognition and speech segmentation can be heavily affected, due to the presence of these processes and vowel reduction (naturally present). The rate of 31% (mean=2.5; sd=1.7) in task 3 means that word identification is compromised when vowel reduction, voicing assimilation and vowel encounters occurred across word boundaries.

In addition, the increased complexity of task 4 brought more difficulties to the students. In this task, the target word pairs encompassed vowel encounters, haplogogies, production of [j] and

resyllabification with liquids [r] and [l]. This means students dealt with stressed and unstressed words, and that some sequences targeted three words. A decrease to 16% (mean= 1.6; sd= 1) in the scores reveal that the occurrence of these connected speech processes (which implied an increase of the complexity of the structures) and vowel reduction (always present within words and across word boundaries) also compromised word identification. In this case, the students identified (on average) at least one target sequence.

In accordance with the present results, research on L2 perception (Cutler, 2012a,b; Vandergrift and Goh, 2012; Alammari, 2015) argued that word recognition in continuous speech is a challenge, due to the occurrence of the phonetic-phonological processes (which are present in every language).

The alterations in the sounds promoted by these phenomena compromise the recognition of word boundaries. Consequently, L2 learners can successfully recognise isolated spoken words, but they may not recognise those same words (or other known words) in continuous speech (Henrichsen, 2006; Goh, 2000; Snijders et al., 2007; Mitterer and Tuinman, 2012; Vandergrift and Goh, 2012). These findings are in line with the results for the Experiment 2, confirming that the occurrence of connected speech processes and vowel reduction (very productive in EP) compromises the recognition of known words in continuous speech (at this level).

Surprisingly, vowel reduction was found to have also a great impact in word recognition even in single word identification tasks. As mentioned, the results revealed that students successfully identified the words in citation form (94%), when all segments were present. However, when words were produced with vowel reduction the rate of correct answers decreased substantially to 65%.

These findings suggest that the occurrence of all the phonetic segments was useful to listeners to recognize and identify most words. However, in task 2, vowel reduction caused alterations in the vowels promoting its deletion and, as a result, several sequences of consonants were produced. So, the deletion of segments and the production of consonants compromised the recognition of known words, even in read speech. Therefore, vowel reduction affects word recognition even when they are produced in isolation or in read speech.

These results can be explained by the fact that vowel reduction (and the CSP) are not productive

in Mandarin Chinese (Duanmu, 2007; Lin, 2007). Thus, learners transfer properties from their L1 to the L2 they are learning (Cutler, 2012a; Maastricht, 2018). As a result, they might be applying segmentation strategies that are productive in Mandarin Chinese but not in EP, when listening to EP utterances (Goh, 2000; Alammari, 2015; Altenberg, 2005; Reed, 2019).

This finding is consistent with Cutler (2012a) that argues that successful listening depends on efficient decoding skills which are determined by whether learners are familiar with these phonetic processes from their L1. In addition, this transfer process is even more evident at lower and intermediate levels (Zhou et al., 2017).

This is also in agreement with the overall opinion of the students about the difficulty of the experiment. They reported that words in the second task were pronounced faster than in task 1, and that some letters or sounds were missing. Thus, task 2 was more difficult than task 1. Other comments, from web forums, mentioned in the subsection 2.2.3 also reveal that L2 learners can perceive the deletion of segments, especially the deletion of vowels. It can thus be suggested that L2 learners perceive vowel reduction even within words, and that it has impact even in the identification of spoken isolated words.

Furthermore, it is important to take into account the correlation between the presence of vowel reduction and single word identification on the light of the competences pertaining to the B1 levels. According to the CEFR, B1 levels are able to deal with connected speech, since they can understand the general message and the main points of conversations, discussions, news etc. Thus, it is expected that students, at this level, recognise spoken isolated words produced with vowel reduction, since they are able to cope with more complex structures and continuous speech (in which vowel reduction is naturally present). Therefore, higher scores were expected for single word identification tasks. Moreover, the previous results for tasks with continuous speech also reveal that the occurrence of these processes inhibited word recognition and compromised oral comprehension, since students do not achieve the meaning of most target words.

In Experiment 2, the students were exposed to authentic speech materials, e.g. (semi-) spontaneous speech. They reported that the speech in both tasks was fast but, in task 4, some parts of the interview were even faster, compromising their listening. In accordance with this unanimous

opinion, previous studies (Cutler, 2012a) have demonstrated that all L2 learners might share the same opinion that foreign languages are spoken faster, overestimating speech rates.

However, it is encouraging to compare these results with previous observations by Mitterer and Tuinman (2012) that indicate that L2 learners should not limitate their exposure to L2 speech to classroom or language tapes, since this type of input is close to a canonical pronunciation, and it is far from spontaneous speech that students face outside the classroom (Moniz, 2006; Johnsrude and Buchsbaum, 2016; Mata and Moniz, 2016). In general, therefore, it seems that it is essential to expose students to authentic speech materials and real communicative contexts to improve their listening skills.

Nevertheless, despite the native language of the classroom, these results are also likely to be related to the lack of awareness of these processes (Cutler, 2012a; Alammr, 2015). It can thus be suggested that there is a strong need of teaching and practising these phenomena in the classroom even at lower levels.

Offering this effective perceptual training will encourage the development of the students' segmentation abilities, and promote the vocabulary activation. If students are aware of vowel reduction, and how the sounds change in connected speech, they will then be better prepared to deal with native speakers outside the classroom. This effort will have a positive effect on the learning trajectory since it contributes to the achievement of the learning goals. Finally, it also encourages the students' L2 proficiency (Alammr, 2015; Reed, 2019).

The results of the control group reveal that natives do not have problems recognizing isolated words with or without vowel reduction. In fact, the native speakers reported that the stimuli in the first task was hyperarticulated and similar to low quality synthesised speech. In contrast, the words in task 2 seemed more natural. This might indicate the presence of vowel reduction, and the deletion of vowels, could be an important feature to define the nativeness of an utterance in EP, even in read speech.

Furthermore, the scores for task 3 (93%) and task 4 (98%) indicate that natives did not have greater difficulties processing connected speech, in comparison with the test group. However, some

parsing problems may arise (Cunha, 2015). Some participants reported that they were not used to perform fill in the blanks exercises, since high school. Thus, they argued that it was difficult to listen and to write down all words correctly, and quickly, because they were focused on the comprehension of the audio.

A limitation of this study is the small amount of collected data (10 words per task, for instance in both tasks 1 and 2). However, it is also important to remember that testing a big amount of data would increase the duration of the tasks and therefore cause tiring effects. Since this study aimed to be applied in a classroom setting, the experiments were designed to be seen as a natural extension of the materials used in the classroom.

In addition, the size of the sample is reduced ($N=12$). Despite that, it was essential to select a group of students taught by the same teacher to ensure all the students followed the same teaching program. Testing classrooms taught by different teachers would increase the variability of the classroom and the results. Nevertheless, it would not be possible to ensure that different classrooms performed similar exercises and learned the same structures at the exact moment the experiments were applied, since the teaching program is flexible.

5.2 Distribution of phonetic-phonological processes

5.2.1 Vowel reduction

With respect to the influence of vowel reduction patterns in L2 learners' perception, it was analysed the rate of correct answers of each pattern, in Experiment 1. The most relevant finding (and unanticipated one) was that some vowel deletion patterns might have a greater impact than others in word recognition.

These results may be somewhat limited by the small amount of tested data. However, this finding, while preliminary, suggests that the alterations in the vowels, due to the centralization process (and its consequent deletion) and the deletion of [u] in unstressed positions may bring difficulties to L2 students in word identification. EP research (Delgado Martins, 1975; Mateus

et al., 2005) on vowel reduction patterns concluded that [u] and [ɪ] are the most deleted vowels in EP utterances. However, more research on this topic is needed to understand the impact of the different vowel reduction patterns in word recognition (in isolated words and in continuous speech) with L2 learners of EP.

In addition, a comparison between the scores of words in citation form (targeting the same patterns) showed an evident gap of 36% between the citation form and the reduced form of the words targeting the centralization of [ɪ] and its deletion. A possible explanation for the highest gap in the scores between the two tasks might be that this pattern of vowel reduction not only implies the deletion of the vowel but also a centralization process in which the vowel [ɪ] is produced as [ɪ̞] resulting in a spoken form distant from its canonical pronunciation (as in citation forms) taught in the classroom or in textbooks.

5.2.2 Connected speech processes

On the question of the effect of connected speech processes, this study found that some processes may have more impact in word recognition than others. As described in Table 4.5, the similar scores for the different processes might indicate that they are equally difficult to the learners. However, the absence of correct answers for the sequences targeting haplogogies may suggest that the alterations in the sounds caused by this process may heavily inhibit the allocation of word boundaries (and its recognition).

In the same line of what was previously stated, a preliminary analysis of the scores for the different patterns of voicing assimilation also might indicate that they can differently affect word recognition. A possible explanation for the differences might be related to the fact that the realization of [z], due to voicing assimilation, causes a restructuration in the syllabic structure (Mateus et al., 2005; Frota, 2000) bringing difficulties in the detection of word boundaries. These results may be somewhat limited by the small amount of tested data. However, as already mentioned, increasing the target words would also increase the duration of the tasks, and the ecological classroom setting would be compromised.

A final note on the impact of the different connected speech processes. There is so far little research on word recognition or on the perception of these phenomena with L2 speakers of European Portuguese. Thus, these results cannot be compared with similar research. Future studies on the perception of connected speech processes and vowel reduction patterns with L2 learners are therefore recommended.

5.3 Analysing Orthographic Transcriptions

To complement the findings of this study, the transcriptions of the students were carefully analysed and described in the previous section 4.1.3. Therefore, this analysis focused on understanding which cues were prominent even when words were not identified. Also, a deep look into the students' transcriptions contributes to understand how the alterations in the sounds, e.g. articulations, merges, deletions etc., inhibited the detection of word boundaries since students wrote what they actually perceived when the target words were not identified.

One interesting finding is that students tend to commit the same patterns of errors. It is clear that the L1 of the students is acting as a filter conditioning their perception (Cutler, 2012a). Since these facts were well established in literature (Zhou, 2017; Zhou et al., 2017; Castelo, 2019; Castelo and Freitas, 2019; Zhou et al., 2019, 2021) they were taken into account in the criteria for the evaluation of the answers (as described in the Section 4.1).

There are several possible explanations for these results, especially taking into account the native language of the students. In accordance with the present findings, previous studies have demonstrated that Mandarin Chinese speakers have difficulties discriminating between the pair of voiced and voiceless consonants [t] and [d] (Nunes, 2015). In addition, Mandarin Chinese (Duanmu, 2007; Lin, 2007) has a simple syllabic structure allowing just a segment in the onset position and, in coda position, there is a small set of consonants allowed to occur in this position. Also, the segment [r] is not included in the repertoire of Mandarin Chinese consonants. Therefore, the occurrence of metathesis is a very common strategy.

The most surprising finding concerns how the transcriptions indicate that the students perceived

the deletion of vowels and the consequent production of consonant clusters (see Tables 4.10, 4.12 and 4.11). From the data in Table 4.17 is also evident that students tend to simplify the resulting sequences of consonants, caused by the deletion of vowels, by deleting one (or more) of them.

The analysis of the transcriptions of words targeting connected speech processes was useful to understand if the students noticed the changes in the sounds, and if they determined the word boundaries. This consideration would also give insight into how the recognition of words was affected.

The results of this analysis indicate that the students mostly recognised one of the words of the target sequence. On the same line of the previous results, when the words were not identified the students transcribed parts, or segments, of both words including representations of the stressed syllable. This is consistent with (Correia et al., 2013) who indicated that stress is a prominent cues in word identification to EP native speakers. However, more research on the impact of EP features such as stress and suprasegmental information, for L2 learners, is therefore recommended.

Another interesting finding was that in some transcriptions it was clear that the students did not allocate the word boundaries, and they did a direct transcription of what they really perceived. Thus, the transcriptions reflected articulations, merges, deletions and other changes in the sounds (see Tables 4.18, 4.16 and 4.20, in Chapter 4).

For instance, for the occurrences of voicing assimilation, the students wrote graphemes with similar phonetic realisations as an intend to represent the altered sounds, e.g. the sequence 'paí[z]onde' was transcribed as 'países', in which the grapheme <s> has the same phonetic realisation as the assimilated sound [z] (see Table 4.13).

Chapter 6

Proposed Didactic Sequence

The section 2.3 reviewed some criteria discussed in the linguistic literature to the design of teaching materials, and the EP research that has been done concerning perceptual training for classroom contexts. In addition, it was briefly discussed the use of CALL technology in the production of materials to teach EP as a foreign language.

The current study aimed to understand not only the acquisition trajectory of vowel reduction and connected speech processes, but also whether L2 learners could cope with them. Therefore, this chapter will review the didactic outcomes of the present study in order to set up guidelines for the design of listening exercises using CALL tools which will, therefore, be presented as a proposal of a didactic sequence in section 6.2.

6.1 Guidelines

The purpose of this study is to contribute with a proposal of a didactic sequence of listening exercises that aims to give practice on the perception of the EP phonetic-phonological phenomena. Therefore, the outcomes of the perceptual experiments are valuable to understand whether, in the learning process, an effective listening practise targeting these phenomena should be given to the students.

The findings of the perceptual experiments applied to the students gave an insight into how

vowel reduction and connected speech processes affected students' perception. As mentioned, the main outcome of this study is that word recognition is compromised due to the occurrence of vowel reduction and connected speech processes in continuous speech. Moreover, the presence of vowel reduction, and the consequent deletion of segments, affected the recognition of spoken isolated words whether in read speech or in (semi-) spontaneous speech.

The analysis of the orthographic transcriptions complemented these results. Due to the deletion of segments, some students recognised and transcribed only parts or segments of the target words in the tasks where vowel reduction occurred. The absence of some graphemes in the transcriptions also indicated that students perceived the deletion of segments, which compromised the recognition of some target words.

The same happened when the sequences targeted connected speech processes. Due to the alterations in the sounds, the students mostly recognised one of the words of the sequence. Some transcriptions were also similar to a phonetic transcription of the merged sequence of sounds, indicating that the students perceived an articulated sequence, without determining the word boundaries. In these last cases, some students represented the altered sounds with graphemes that have similar phonetic realisations.

These results indicate that, at this level, students are developing their segmentation skills. Therefore, considering the competencies for B1 levels¹ (already described in chapter 3, and the given outcomes of the experiments, it is possible to conclude that the intermediate level B1 may be the right moment in the learning process for students to receive an effective listening training.

Therefore, the proposal of the didactic sequence includes exercises that will rely on the pertained competencies for the B1 level. Moreover, the exercises were designed considering the difficulties learners experienced when dealing with vowel reduction and connected speech processes in word identification tasks (previously reviewed).

However, as mentioned, vowel reduction had a greater impact in the recognition of isolated

¹<http://www.coe.int/en/web/common-european-framework-reference-languages/the-cefr-descriptors>

words than expected for this level. So, this practise should rely not only on the perception of connected speech processes, but also on vowel reduction in real speech samples.

Furthermore, this perceptual practice should be considered even at lower levels of proficiency to ensure the students are aware of the impact of vowel reduction in isolated spoken words. This will also contribute to the development of segmentation skills, preparing the students for the occurrence of this process in continuous speech. Moreover, this practise will contribute to the achievement of the mentioned learning goals for the B1 level.

Therefore, the proposal of the didactic sequence also includes exercises for the practise of vowel reduction for A2 levels. According to the CEFR, at the A2 level, students can understand expressions and keywords related to topics of immediate priority provided slowly articulated speech. Students can also retain essential information from TV news reporting events and can follow the main points of short recorded materials about everyday topics. Besides, they can also understand the general topic and exchange relevant information provided speech is slowly articulated during a conversation, or discussion, about familiar matters.

One of the goals of the proposal of listening exercises is to encourage the progression of the proficiency of the students. So, for the design of these exercises, the learning goals and the listening competencies pertaining for the B2 level were also considered. B2 students can understand native speech on both familiar and unfamiliar topics. They can also understand TV news, plays, live interviews, films, radio documentaries and follow lectures, talks and discussions. B2 students should be able to actively engage in conversations between native speakers. Despite some difficulties they may have in noisy environments, they can use contextual clues, or main points spotting, to achieve a clear comprehension.

To achieve that goals, students should receive practise on real speech contexts. Therefore, the exercises included in the didactic sequence will also explore in different ways the use of real speech samples to give perceptual training concerning vowel reduction and connected speech processes. This will, therefore, encourage the progression and the proficiency of the students, targeting the competencies for future levels of proficiency.

A further consideration of the referencial Camões ², developed by Camões I.P. ³ to teaching EP as a foreign language, reveals that the practice of these phonetic-phonological processes is not included in the recommended teaching program of the different proficiency levels. However, as well as the CEFR, it contributes with an extensive description of the pertaining listening competencies to all levels.

The sequence was designed considering the topics taught in the classroom so that it can complement the classes and the curricular program. Given the previous summarised competencies, designing exercises increasing the difficulty degree will also be essential to promote the students' progression.

The design of the sets were based on the discussion (present in the section 2.3 about the guidelines for the production of teaching materials to apply in a classroom context, and for the ecological use of CALL technology.

6.2 Sets of exercises

The present section describes the proposal of a didactic sequence of exercises targeting the perceptual practise of vowel reduction and connected speech processes. The set of exercises were designed by the researcher, relying on the current uses of CALL technology for the production of listening materials.

The first set is related to the classic hangman game and it aims to give practise on vowel reduction at lower levels. The Set 2 is also dedicated to the perceptual practise of vowel reduction but in a different learning environment. This second set is named as 'Palavras Bomba-Relógio' ('The Time-Bomb Words'). The Set 3 is dedicated to the practise of vowel reduction and connected speech processes in real speech samples. This set has different levels of complexity and it also includes songs as input.

²http://www.instituto-camoes.pt/images/REFERENCIAL_ebook.pdf

³<http://www.instituto-camoes.pt/activity/centro-virtual/referencial-camoes-ple>

6.2.1 Set 1

The first set of exercises, 'O Jogo da Forca', aims to introduce and practise vowel reduction within words, at lower proficiency levels (A levels). This exercise replicates the classic hangman game, as presented in Figure 6.1. The purpose of this game is to practise the listening skills and, at the same time, the writing component.



Figure 6.1: Main page of the proposed 'O Jogo da Forca' ('The Hangman Game').

This game is divided into thematic units according to the *Referencial Camões*, and the user can select which unit he/she wants to practise (see Figures 6.2 and ??).



Figure 6.2: Example of the displayed topics that users can choose in the 'O Jogo da Forca' ('The Hangman Game'). The progress of the user is indicated, in percentage, in each topic.

In contrast to the classic hangman games, in which the players have to guess the intended word and fill in the blanks, this game focuses on listening skills. Therefore, the players have to listen to words and fill in the given blanks. As shown above in Figure 6.3, the user has to press the 'play' button to listen to the audio of a word. A slowed down speech button on the right side is also available.

In this exercise, the player has to write the letters in a unique empty gap (see Figure 6.3). However, the user can display the blanks for each letter, and some letters, by pressing the 'Ajuda'



Figure 6.3: Illustration of the environment of 'O Jogo da Forca' ('The Hangman Game')

('help') button. For instance, when the user requests help for the first time, the gaps of each letter are displayed. Then, each time the user presses the help button, a letter is shown (to a maximum of three letters). Since the purpose of this game is to practise vowel reduction, only consonants are displayed, as Figure 6.4 illustrates.

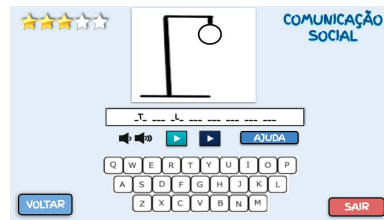


Figure 6.4: Illustration of the Hangman game environment after the 'Ajuda' ('help') button is pressed. The gaps and some letters are displayed.

The user can type the letter using a virtual keyboard displayed in the page. When a letter is correctly selected, it is written in the gap. As the user guesses the correct letters, the word will be written without marking the blanks of the following letters (unless the gaps were already displayed). If the user picks a wrong letter, the color of the key changes, and a song is played so that the user understands that a wrong letter was selected. Furthermore, each time a wrong letter is selected, the parts of the body of a stickman are drawn in the figure with a gallows (above the gaps), as Figure 6.5 presents.

When the man is completely drawn (after six errors), he is hung and the user loses the game. Misleads in orthographic accents will not be considered as errors since the main goal of the game is to practise listening skills. However, when the word is fully written, the orthographic accents are replaced.

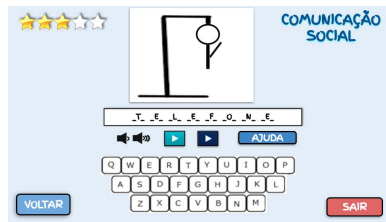


Figure 6.5: Illustration of the Hangman game environment when the user completes the given spaces. The parts of the body of a stickman are drawn in the figure with a gallows, displayed above the gaps.

The users can control the level of difficulty by displaying the gaps and the letters, or by pressing the slowed down speech button. In addition to the immediate feedback given to the user during the selection of letters, the game also contains a progress bar on the left side. When selecting a topic, the user can also see how much of each topic he/she already completed (see Figure 6.2).

6.2.2 Set 2

The second set is the game 'Palavras Bomba-Relógio' ('The Time-Bomb Words'), as illustrated in 6.6, and it is also focused on the perceptual practice of vowel reduction patterns. Similar to the previous game, the users can select the topic they want to work on, in each game, as Figure 6.7 shows. In each game the user can choose the amount of words he/she wants to include: sets of 5, 10 or 15 words (see Figure 6.8).

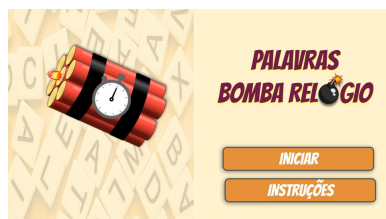


Figure 6.6: Main page of the proposed 'Palavras Bomba-Relógio' ('The Time-Bomb Words') game.

In this game, the user has to listen to the words and write them down as fast as possible. Figure 6.9 below illustrates the environment of the game: after playing the audio of the word (pressing the 'play' button) a timer starts, and the user has to fill in the blanks with the right letters before the



Figure 6.7: Example of the displayed topics that users can choose in the 'Palavras Bomba-Relógio' ('The Time-Bomb Words') game. The progress of the user is indicated, in percentage, in each topic.



Figure 6.8: Example of the number of words that user want to include in a round, in the 'Palavras Bomba-Relógio' ('The Time-Bomb Words') game.

time finishes. The user has to drag the letters from the given virtual keyboard, and drop them in the empty blanks. He/She can also play the audio many times as needed.

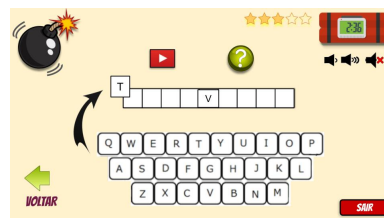


Figure 6.9: Illustration of the 'Palavras Bomba Relógio' ('Time-Bomb words') game environment when the user drags the letters, and drop them in the given spaces.

A help button, represented by a '?' icon, is also present and it can be pressed a maximum of three times. Each time the user presses it, a consonant is given. A countdown timer and a picture of a shaking bomb are also displayed in the right corner. During the game, a background music is also played but the user can control the volume or turn it off. When the words are correctly identified the user earns points and, if the word is correctly transcribed before the time finishes, extra points are

given. If the user wants to leave the game, he/she can press the button 'sair' ('leave'), but punctuation is not given.

As presented in the Figure 6.10, when the user does not complete the task and the time finishes, the shaking bomb blows. The user receives a message that the game is over, and two options for leave or play again are displayed. If the user selects the play again button ('jogar outra vez'), the system runs the same game again (with the same set of topics and words).

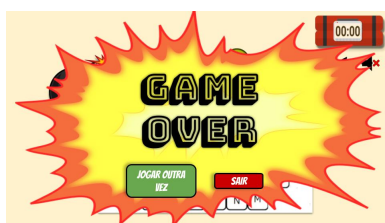


Figure 6.10: Illustration of the 'Palavras Bomba Relógio' ('Time-Bomb words') game environment when the game is over.

6.2.3 Set 3

The following proposal of exercises is focused on the allocation of word boundaries since this a common difficulty among L2 learners. Thus, this set targets the occurrences of connected speech processes and vowel reduction across word boundaries. Figure 6.11 illustrates the third proposed game: 'À Procura dos Limites' ('Searching for the Boundaries').

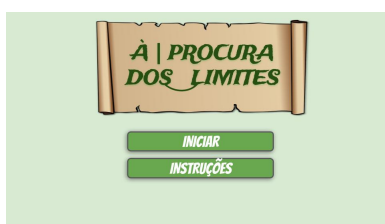


Figure 6.11: Main page of the proposed 'À Procura dos Limites' ('Searching for the Boundaries') game.

As presented in Figure 6.12, these exercises are divided into three main sets: word pairs, texts and songs. The exercises in the first set aim to present sequences of words, in which a phonetic

process is present, and contrasting sequences, in which the phonetic process does not occur. This set targets linking processes and merges of sounds in vowel encounters, resyllabification with [l] and [r], voicing assimilation and haplogogies. The difficulty of the exercises depends not only on the phonetic process but also on the type of input and the complexity of the structures.

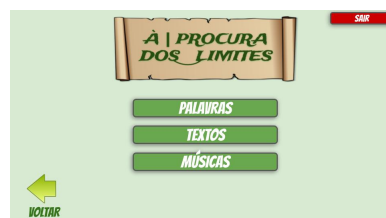


Figure 6.12: Example of the displayed options of the game 'À Procura dos Limites' ('Searching for the Boundaries'): words, texts and songs.

After selecting a topic (see Figure 6.13), some word pairs are displayed in the screen, and the user has to listen to the audio of the sequence, pressing a play button, and mark whether he/she noticed were 'mingled' or articulated. To do so, the user has to drag and drop boxes with symbols to mark the absence and the occurrence of these processes, as illustrated in the Figure 6.14. These boxes can be dropped in other empty boxes between the target words.



Figure 6.13: Example of the displayed topics (on words set) that users can choose in the 'À Procura dos Limites' ('Searching for the Boundaries') game. The progress of the user is indicated by a stars bar, in each topic.

Immediate feedback is also given when a marker is dropped in the wrong place: the box turns to red and a short song, indicating a wrong answer, is played. In contrast, when the marker is dropped in the right place, it turns green and also a short song is played. Also, a progress bar is filled as the user is completing the topic.

The second set of exercises aims to present more contexts for the occurrences of connected



Figure 6.14: Illustration of the 'À Procura dos Limites' ('Searching for the Boundaries') game environment when the user drags icons, and drop them in the empty squares, to mark whether the sounds are articulated (words set).

speech processes and vowel reduction, with real speech samples. In the main page, as shown in Figure 6.15, the user can select the topic he/she wants to work on, and the exercises rely on the presentation of news, reports or podcasts related to the chosen topic.

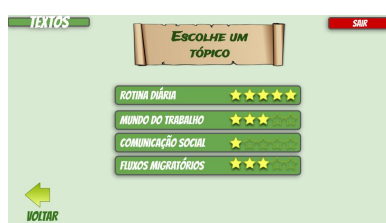


Figure 6.15: Example of the displayed topics (on texts set) that users can choose in the 'À Procura dos Limites' ('Searching for the Boundaries') game. The progress of the user is indicated by a stars bar, in each topic.

As Figure 6.16 shows, it is presented the orthographic transcription of the audio with some empty squares between the word boundaries. The user can play a video file, or just listen to the audio when podcasts are presented. Then, the user has to drag and drop the boxes containing the markers already used in the previous set of exercises to select whether the sounds were mingled, or linked, between words. The feedback procedure is the same as the previous set of exercises. The audios and the videos can be played many times as the user wants. When all the boxes are correctly filled the exercise is completed.

The third set of exercises aims to practise the allocation of word boundaries, using songs. In the initial page, the user can pick a song from a predefined playlist, as indicated in the figure 6.17. The playlist contains music from different genres and artists, and the songs and the lyrics were selected taking into account its didactic potential.

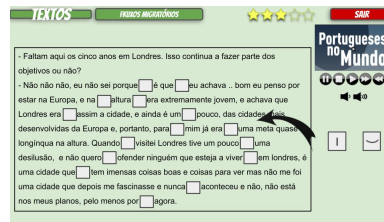


Figure 6.16: Illustration of the 'À Procura dos Limites' ('Searching for the Boundaries') game environment when the user drags icons, and drop them in the empty squares, to mark whether the sounds are articulated (texts set).

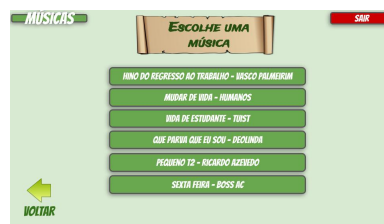


Figure 6.17: Example of the displayed songs that users can choose in the 'À Procura dos Limites' ('Searching for the Boundaries') game.

After selecting a song, the user needs to select the level of difficulty (see Figure 6.18. Figure 6.19 illustrates the exercise when the easy level ('fácil') is selected. As it can be seen, the corresponding lyrics are presented, and some sentences have no spaces between words. Also, features such as punctuation and capitalization are absent.



Figure 6.18: Example of the levels of difficulty of the songs set, in the 'À Procura dos Limites' ('Searching for the Boundaries') game.

The user must use the mouse to select whether he/she wants to mark the boundary between words. Thus, he/she has to determine the word boundaries based on orthographic cues, and when listening to the song. To clean the answer, the user can select the icon with 'x' to delete the mark. As

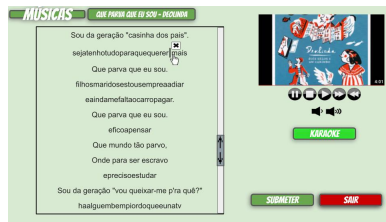


Figure 6.19: Illustration of the 'À Procura dos Limites' ('Searching for the Boundaries') game environment when the user marks the word boundaries in the given lyrics of the easy level of difficulty (songs set).

figure 6.20 illustrates, when the hard level ('difícil') is selected, all the lyrics have no punctuation, capitalization or white spaces between the words.

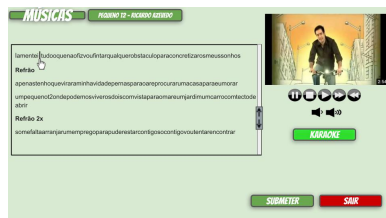


Figure 6.20: Illustration of the 'À Procura dos Limites' ('Searching for the Boundaries') game environment when the user marks the word boundaries in the given lyrics of the hard level of difficulty (songs set).

The target sentences cannot have punctuation or capitalization so that the users can not use this type of information as cues to locate the word boundaries. The user can play the song many times as needed. After marking the word boundaries, the user can submit his/her answers. The wrong answers will be marked in red. When the exercise is finished, a karaoke option is available so that the student can listen to the song, as the produced words are highlighted in the lyrics.

Chapter 7

Conclusion

As mentioned in the literature review, listening in a second language is much more challenging than in a native language (Cutler, 2012a,b; Vandergrift and Goh, 2012; Heald and Nusbaum, 2014; Maastricht, 2018). The major challenges L2 listeners experience are related to speech segmentation and word recognition (Altenberg, 2005; Snijders et al., 2007; Cutler, 2012a; Mitterer and Tuinman, 2012; Alameen and Levis, 2015; Alammam, 2015; Reed, 2019).

The occurrence of phonetic-phonological processes in the stream speech can contribute to the mentioned difficulties (Altenberg, 2005; Alameen and Levis, 2015; Alammam, 2015; Reed, 2019). These processes occur in some form or other in every language and they are responsible for alterations in the sounds. Therefore, L2 learners have difficulties in segmenting speech and detecting word boundaries, as research in second language teaching also confirmed (Henrichsen, 2006; Goh, 2000; Grant, 2014; Graham and Santos, 2015; Goh et al., 2016).

Vowel reduction is very productive in EP (Mateus and Andrade, 2000; Mateus et al., 2005), and it promotes alterations in the vowels and its consequent deletion (Delgado Martins, 1975; Cunha, 2015). In addition, it has been reported in the EP literature that connected speech processes (Mateus and Andrade, 2000; Mateus et al., 2005; Andrade and Viana, 1992; Frota, 2000) have a strong relationship with vowel reduction (Andrade and Viana, 1992; Frota, 1995, 2000). However no data was found on the association between these phonetic-phonological phenomena

and word recognition in EP with L2 learners. Moreover, there is no evidence of research about the emergence of abilities to deal with these processes in L2 listening. To our knowledge, no L2 perception studies were conducted in a classroom setting, in EP, with L2 intermediate students.

For those reasons, the purpose of the current project was to identify the main difficulties that L2 learners of EP experience in word recognition due to the presence of vowel reduction and the connected speech processes, in the stream speech. This project also aimed to understand how these processes affect the recognition of words not only in continuous speech but also in isolation. Finally, it aimed to understand the acquisition trajectory of these processes. From the didactic point of view, this fact is interesting to ensure that the right skills are trained at the right moment of the learning process in order to encourage a progression of the students' proficiency.

The present study was applied to a group of students of EP at the B1 level, and it was conducted in a classroom setting. It was set out to determine the effect of vowel reduction and connected speech processes, in word recognition, from isolated words to continuous speech. Therefore, a series of perceptual experiments with four word recognition tasks was designed.

These experiments increased the complexity of the structures, starting from isolated word identification including words in citation forms (task 1) and words produced with vowel reduction (task 2). Word recognition was also tested in semi-spontaneous and casual speech, with authentic input, in which the target sequences had contexts for the occurrence of common EP connected speech processes (task 3 and 4). Vowel reduction was naturally present in the stimuli of tasks 3 and 4, since authentic speech materials were used as input.

As mentioned, the first experiment involved two tasks of single word identification, and it was similar to dictation tasks. The second one, which included two word identification tasks in continuous speech, replicated fill in the blanks exercises.

Taking into account the review of the state of art about perception in L2 and the impact of connected speech processes and vowel reduction in L2 students, **the following hypotheses were stated: (i) a clear effect of connected speech processes and vowel reduction on the recognition of words in continuous speech; (ii) a low effect of vowel reduction on the recognition of spoken**

isolated words, at the B1 level; (iii) the presence of particular connected speech processes will differently affect word recognition; (iv) an effect of the type of input (read speech vs semi-spontaneous speech) on the perception of L2 learners.

The higher percentages of correct answers for the first experiment (see table 4.1.1: 94% in task 1, and 65%, in task 2) reveals that students, at this level, successfully identify isolated words. However, the difference of 34% between both tasks indicates that the production of vowel reduction (and the deletion of segments) affected the students' perception, and the recognition of the spoken isolated words in task 2.

The gap between the scores of task 2 (single word identification with vowel reduction) and task 3 (word identification with CSP) suggests that recognising words in continuous speech is harder than in isolation (the rates for tasks 3 and 4 were 31% and 16%, respectively, according to table 4.1). The students had difficulties identifying known vocabulary when it was in continuous speech or, in other words, embedded in CSP.

In sum, the overall results highlight that the accuracy in word identification depends on the complexity of the structures. For instance, word recognition is much more challenging in continuous speech than when words are produced in isolation. Furthermore, the interplay of vowel reduction and connected speech processes have a great impact on the perception of the word boundaries. Therefore, the occurrence of these phonetic-phonological affects the recognition of spoken words. Lastly, the deletion of segments and the production of phonetic clusters of consonants (due to vowel reduction) compromised the recognition of isolated known words, even in read speech.

It is also important to highlight that Mandarin Chinese is phonologically distant from EP. Therefore, some features such as the syllabic structure of words or the inventory of sounds are different between these two languages (Duanmu, 2007; Lin, 2007). Moreover, the vowel reduction process is not frequent in Mandarin Chinese. Consequently, the students may apply segmentation and decoding strategies that are productive in Mandarin Chinese but not in EP (Goh, 2000; Altenberg, 2005; Alammar, 2015; Zhou, 2017; Reed, 2019; Zhou et al., 2019, 2021). This is evidence of a linguistic transfer, a common process in a L2 learning process (Cutler, 2012a; Castelo and Santos, 2017; Zhou et al., 2017; Castelo, 2018; Maastricht, 2018; Castelo, 2019; Castelo and Freitas, 2019).

On the basis of these conclusions, **the hypothesis (i) about a clear effect of connected speech processes and vowel reduction on the recognition of words in continuous speech was confirmed.** As mentioned, the scores for the recognition of spoken isolated words in the first experiment were lower than expected (considering what is required for the listening competences for the B1 level, according to the CEFR). Therefore, it can be concluded that **the hypothesis (ii) about a low effect of vowel reduction on the recognition of spoken isolated words, at the B1 level, is not confirmed.**

Furthermore, the results also indicate that the type of input also affected the perception of the students, since the interplay of vowel reduction and the CSP, and the presence of speech features, such as disfluencies, was more evident in the (semi-) spontaneous speech samples presented in tasks 3 and 4. Therefore, **the hypothesis (iv) about the effect of the type of input (read speech vs semi-spontaneous speech) on the perception of L2 learners is also confirmed.**

Further analysis on the effects of vowel reduction patterns (see table 4.3) revealed a greater gap between the scores of the citation and reduced forms of the words targeting the centralization of [i] and its deletion, and the deletion of [u]. This finding may indicate that these patterns of vowel reduction may have a greater impact in the identification of words.

The overall results for connected speech processes (see table 4.5) reveal that vowel encounters had the highest rates (31%), followed by voicing assimilation with 24% and, lastly, both resyllabification with [l] and [r] and production of [j] with 21%. In contrast, the occurrences of haplogogies had 0%. These preliminary results may suggest that CSP have different impacts on the student's perception.

The different scores for each pattern of voicing assimilation (see table 4.6) may also reveal, for instance, that learners may be more sensitive to the production of sound [z] than the alternations between [ʃ] and [ʒ]. With this in mind, **hypothesis (iii) that stated that the presence of particular connected speech processes would differently affect word recognition is also confirmed.**

The analysis of the students' transcriptions (see tables 4.7 to 4.21) reveal that, at this level, students still transfer properties from their L1. Learners have difficulties in the perceptual discrimina-

tion of some segments, and they apply strategies, such as metathesis and coda deletion, to simplify the syllabic structure of the words. It was also noticed that learners deleted consonants to cope with the massive production of sequences of consonants due to the deletion of vowels. When the words were not identified, the students transcribed parts or segments of both words, also including representations of the stressed syllable.

Another interesting finding was that in some transcriptions it was clear that the students did not allocate the word boundaries, and they did a close transcription of what they really perceived, reflecting linking processes, merges, deletions and other changes in the speech.

A final note on the didactic outcomes of the experiment. The findings of the perceptual experiments were also useful to understand whether in the learning process an effective listening practice, concerning these phenomena, should be given to the students.

Despite the fact that the phonetic-phonological processes are well described in the literature, there is a lack of teaching these processes, and giving practise on them, in L2 classes. This is an important factor, since the knowledge (or the lack of it) about these phonetic-phonological processes can promote, or inhibit, the development of L2 speech segmentation abilities.

It can thus be suggested that there is a strong need for awareness on such topics even at lower levels, and therefore offer an effective perceptual training to the students. Consequently, it will encourage the accuracy of their segmentation abilities, the recognition of spoken words, and the activation of the vocabulary. Finally, it will have a positive effect on the didactic side, since it will contribute to the achievement of the learning goals, encouraging the proficiency of the students.

This dissertation gives a contribution to the methods of teaching EP as a second language. The aim of the proposed set of listening exercises is to give practise on the perception of phonetic-phonological phenomena in study, based on the existent CALL technology tools, well explored in EP literature (Marujo et al., 2009; Baptista et al., 2010; Correia et al., 2010; Correia, 2010; Ribeiro et al., 2010; Pellegrini et al., 2011; Silva et al., 2011; Pellegrini et al., 2012; Silva et al., 2012).

One of the goals of the proposed didactic sequence is to promote the acquisition of these processes which is a strong need, as suggested by the findings of the perceptual experiments in this

study. The exercises were designed considering the learning goals, and the listening competencies, pertaining to the A2, B1 and B2 levels (according to the CEFR) to, therefore, encourage the progression of the students.

Furthermore, the difficulties that the participants in this study experienced dealing with vowel reduction and connected speech processes, in word identification tasks (previously described), were also considered in the design of the exercises.

The sequence of exercises includes two games: 'O Jogo da Forca' ('The Hangman Game') and 'Palavras Bomba-Relógio' ('The Time-Bomb Words'). It was also designed a set of listening exercises named 'À Procura dos Limites' ('Searching for the Boundaries'). The focus of the sequence is to give practise on the detection of word boundaries whether in short sequences of words or in real speech samples. This set also includes a component with songs in which the users must mark the boundaries of the words in the lyrics, and there is also a karaoke feature to encourage the users to sing along the songs, to practise both perception and production skills.

Future work will tackle students of more levels of proficiency to compare the acquisition trajectory of these processes in long term. More improvements in the methods applied in this study are therefore recommended: (i) targeting more stimuli to ensure more data are collected, (ii) increase the number of participants in both test and control groups, (iii) design more exercises to target connected speech processes, (iv) use different types of input taking into account the speech rates. In addition, the proposed set of exercises and games should be improved, completed, tested, and evaluated.

Bibliography

- Alameen, G. and Levis, J. (2015). Connected speech. In Reed, M. and Levis, J. M., editors, *The Handbook of English Pronunciation*, pages 159–174. Wiley Blackwell.
- Alammar, A. (2015). Second language perception of word segmentation. In Lipani, L., Olsen, M., and Merchant, D., editors, *Proceedings of LSUGA's Second Interdisciplinary Conference in Linguistics*. The Linguistics Society at the University of Georgia.
- Altenberg, E. (2005). The perception of word boundaries in a second language. *Second Language Research*, 21(4):325–358.
- Alves, U. K. (2015). Ensino de pronúncia na sala de aula de língua estrangeira: questões de discussão a partir de uma concepção de língua como sistema adaptativo e complexo. *Revista Versalete*, 3(5):392–413.
- Andrade, A. (2020). Vocalismo. In Raposo, E., Nascimento, M., Mota, M. A., Segura, L., Mendes, A., and Andrade, A., editors, *Gramática do Português*, volume 3, pages 2589–2613. Fundação Calouste Gulbenkian, Lisboa.
- Andrade, E. and Viana, M. C. (1992). Que horas são às (1)3 e 15? *Actas do VIII Encontro Nacional da Associação Portuguesa de Linguística Lisboa: APL*, 8:59–66.
- Azevedo, M. L. (2012). Seleção e produção de materiais didáticos para a aula de PLE/PLS. Master's thesis, Faculdade de Letras, Universidade do Porto, Porto.
- Baptista, J., Costa, N., Guerra, J., Zampieri, M., Cabral, M., and Mamede, N. (2010). P-awl:

- academic word list for portuguese. In Pardo, T., Branco, A., Klautau, A., Vieira, R., and Lima, V., editors, *Computational Processing of the Portuguese Language*, volume 6000 of LNCS, pages 120–123. Springer, Berlin, Heidelberg.
- Barbosa, V. (2014). A canção nas aulas de português língua estrangeira aliada à compreensão e produção orais. Internship report, Faculdade de Letras, Universidade do Porto, Porto.
- Boersma, P. and D, W. (2009). Praat: Doing phonetics by computer. Retrieved from <http://www.praat.org>.
- Brito, E. and Hernandez, R. (2020). Perspectives on reduced forms instruction in portuguese. *Journal of the National Council of Less Commonly Taught Languages*, 28:51–80.
- Castelo, A. (2018). Ensino da componente fonético-fonológica: uma síntese e um exemplo de português para estrangeiros. *Linguística: Revista de Estudos Linguísticos da Universidade do Porto*, 12:41–72.
- Castelo, A. (2019). A produção de vogais em contexto de processo de redução vocálica em aprendentes chineses. In Alonso, C., Russo, V., Vecchi, R., and André, C., editors, *De Oriente a Ocidente: estudos da Associação Internacional de Lusitanistas*, volume 5, pages 7–34. Angelus Novus.
- Castelo, A. and Freitas, M. J. (2019). Produção de vogais orais tónicas do PLE por falantes nativos de chinês mandarim. *Orientes do Português*, 1:47–58.
- Castelo, A. and Santos, R. (2017). As vogais do português entre aprendentes chineses e suas implicações no desenvolvimento de um programa de português. In Gonçalves, L., editor, *Português como língua estrangeira, de herança e materna: abordagens, contextos e práticas*, pages 123–136. Boavista Press, Roosevelt, New Jersey.
- Chapelle, C. A. (2001). *Computer applications in second language acquisition: foundations for teaching, testing and research*. Cambridge Applied Linguistics. Cambridge University Press, Cambridge.

- Collins-Thompson, K. and Callan, J. (2004). Information retrieval for language tutoring: An overview of the reap project. In *Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, number 2 in SIGIR '04, page 544–545, New York, USA. Association for Computing Machinery.
- Correia, R. (2010). Automatic question generation for reap. pt tutoring system. Master's thesis, Instituto Superior Técnico, Universidade de Lisboa, Lisboa.
- Correia, R., Baptista, J., Mamede, N., Trancoso, I., and Eskenazi, M. (2010). Automatic generation of cloze question distractors. In *Second Language Studies: Acquisition, Learning, Education and Technology*, Tokyo, Japan. ISCA Archive.
- Correia, S., Butler, J., Vigário, M., and Frota, S. (2015). A stress “deafness” effect in european portuguese. *Language and speech*, 58(1):48–67.
- Correia, S., Frota, S., Butler, J., and Vigário, M. (2013). Word stress perception in european portuguese. In *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, pages 267–271.
- Cuicui, C. (2012). A seleção e a produção de materiais didáticos no processo do ensino do português aos alunos chineses. Master's thesis, Faculdade de Ciências Sociais e Humanas, Universidade Nova de Lisboa.
- Cunha, C. (2015). Portuguese lexical clusters and CVC sequences in speech perception and production. *Phonetica*, 72(2-3):138–161.
- Cutler, A. (2012a). *Native Listening: Language Experience and the Recognition of Spoken Word*. MIT PRESS, Cambridge, Massachusetts.
- Cutler, A. (2012b). Native listening: The flexibility dimension. *Dutch Journal of Applied Linguistics*, 1(2):169–187.
- Delgado Martins, M. R. (1975). Vogais e consoantes do português: estatística de ocorrência, duração e intensidade. *Boletim de Filologia*, 24(1-4):1–11.

- Duanmu, S. (2007). *The Phonology of Standard Chinese*. Oxford University Press, Oxford.
- Fikkert, P. and Freitas, M. J. (2006). Allophony and allomorphy cue phonological acquisition: Evidence from the european portuguese vowel system. *Catalan journal of linguistics*, 5:83–108.
- Freitas, M. J. (1997). *Aquisição da estrutura silábica do Português Europeu*. PhD thesis, Faculdade de Letras, Universidade de Lisboa, Lisboa.
- Freitas, M. J. (2003). The acquisition of onset clusters in european portuguese. *Probus*, 15(1):27–46.
- Frota, S. (1995). Os domínios prosódicos e o português europeu: fenómenos de sandhi. *Actas do X Encontro Nacional da Associação Portuguesa de Linguística*, pages 221–237.
- Frota, S. (2000). *Prosody and focus in European Portuguese: Phonological phrasing and intonation*. Garland Publishing, New York.
- Frota, S. (2014). The intonational phonology of european portuguese. In Sun-Ah, J., editor, *Prosodic typology II: The phonology of intonation and phrasing*, pages 6–42. Oxford University Press, Oxford.
- Gaskell, M. and Mirkovic, J. (2016). *Speech perception and spoken word recognition*. Psychology Press, London, England, 1st edition.
- Goh, C. (2000). A cognitive perspective on language learners’ listening comprehension problems. *System*, 28(1):55–75.
- Goh, D., Yap, J., Lau, C., Ng, M., and Tan, L.-C. (2016). Semantic richness effects in spoken word recognition: A lexical decision and semantic categorization megastudy. *Frontiers in psychology*, 7:Article 976.
- Graham, S. and Santos, D. (2015). *Strategies for second language listening: Current scenarios and improved pedagogy*. Palgrave Macmillan.
- Grant, L. (2014). Prologue to the myths: What teachers need to know. In Grant, L., editor, *Pronunciation myths: Applying second language research to classroom teaching*, pages 1–33. University of Michigan Press, Michigan.

- Gustafson, E. N. and Bradlow, A. R. (2016). French speech segmentation in liaison contexts by L1 and L2 listeners. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, 7(1):1–16.
- Heald, S. and Nusbaum, H. (2014). Speech perception as an active cognitive process. *Frontiers in Systems Neuroscience*, 8:Article 35.
- Heilman, M., Collins-Thompson, K., Callan, J., and Eskenazi, M. (2006). Classroom success of an intelligent tutoring system for lexical practice and reading comprehension. In *Proceedings of the 9th International Conference on Spoken Language Processing, ICSLP*, Pittsburgh, Pennsylvania.
- Henrichsen, L. (2006). Sandhi-variation: A filter of input for learners of ESL. *Language Learning*, 34(3):103 – 123.
- Johnsrude, I. S. and Buchsbaum, B. R. (2016). Representation of speech. In Gaskell, G. and Mirković, J., editors, *Speech Perception and Spoken Word Recognition*, pages 5–22. Psychology Press, London, England.
- Lass, R. (1984). *Phonology: An introduction to basic concepts*. Cambridge University Press.
- Lin, Y.-H. (2007). *The Sounds of Chinese*. Cambridge University Press, Cambridge.
- Maastricht, L. (2018). *Second Language Prosody: Intonation and Rhythm in Production and Perception*. PhD thesis, Tilburg University, Tilburg, Netherlands.
- Magnuson, J. (2016). Mapping spoken words to meaning. In Gaskell, G. and Mirkovic, J., editors, *Speech Perception and Spoken Word Recognition*, pages 76–96. Psychology Press, London, England.
- Marujo, L. (2009). REAP. PT. Master’s thesis, Instituto Superior Técnico, Universidade de Lisboa, Lisboa.
- Marujo, L., Lopes, J., Mamede, N., Trancoso, I., Pino, J., Eskenazi, M., Baptista, J., and Viana, C. (2009). Porting REAP to european portuguese. In *Proceedings of International Workshop*

- on *Speech and Language Technology in Education, SLaTE*, pages 69–72, Wroxall Abbey Estate, Warwickshire, England.
- Mata, A. I. and Moniz, H. (2016). Prosódia, variação e processamento automático. In Martins, A. M. and Carrilho, E., editors, *Manual de linguística portuguesa*, pages 116–155. De Gruyter, Berlin, Boston.
- Mateus, M. H. M. and Andrade, E. (2000). *The Phonology of Portuguese*. Oxford University Press, Oxford.
- Mateus, M. H. M., Falé, I., and Freitas, M. J. (2005). *Fonética e Fonologia do Português*. Universidade Aberta, Lisboa.
- Mattys, S. L. and Bortfeld, H. (2016). Speech segmentation. In Gaskell, G. and Mirkovic, J., editors, *Speech Perception and Spoken Word Recognition*, pages 55–75. Psychology Press, London, England.
- Mattys, S. L., Melhorn, J. F., and White, L. (2007). Effects of syntactic expectations on speech segmentation. *Journal of Experimental Psychology: Human Perception and Performance*, 33(4):960–967.
- Mitterer, H. and Tuinman, A. (2012). The role of native-language knowledge in the perception of casual speech in a second language. *Frontiers in Psychology*, 3:Article 249.
- Moniz, H. (2006). Contributo para a caracterização dos mecanismos de (dis) fluência no português europeu. Master’s thesis, Faculdade de Letras, Universidade de Lisboa, Lisboa.
- Moniz, H., Batista, F., Mata, A. I., and Trancoso, I. (2014). Speaking style effects in the production of disfluencies. *Speech Communication*, 65:20–35.
- Nenadic, F., Kelley, M. C., Podlubny, R. G., and Tucker, B. V. (2019). Speech perception and the role of semantic richness in processing. *Canadian Acoustics*, 47(3):96–97.
- Nunes, A. (2015). European portuguese phonetics: Difficulties for chinese speakers – considerations. *Procedia - Social and Behavioral Sciences*, 192:92–100.

- Oliveira, P. (2016). A prosódia da iode epentética para quebrar hiatos em variedades do português. Master's thesis, Faculdade de Letras, Universidade de Lisboa, Lisboa.
- Oliveira, P., Paulino, N., Cruz, M., and Vigário, M. (2014). Onde (ainda [j]) há o fenómeno? contributo para o estudo da inserção de glide entre vogais centrais. In Moreno, A., Silva, F., Falé, I., Pereira, I., and Veloso, J., editors, *XXIX Encontro Nacional da Associação Portuguesa de Linguística*, pages 419–436, Porto. APL.
- Paulino, N. (2016). Fenómenos de sândi vocálico em variedades do português europeu. Master's thesis, Faculdade de Letras, Universidade de Lisboa, Lisboa.
- Paulino, N. and Frota, S. (2016). Variação prosódica no português europeu: análise comparada de fenómenos de sândi vocálico. *Revista da Associação Portuguesa de Linguística*, 10(1):651–674.
- Pellegrini, T., Correia, R., Trancoso, I., Baptista, J., and Mamede, N. (2011). Automatic generation of listening comprehension learning material in European Portuguese. In *Proc. Interspeech 2011*, pages 1629–1632.
- Pellegrini, T., Ling, W., Silva, A., Correia, R., Trancoso, I., Baptista, J., and Mamede, N. (2012). Overview of computer-assisted language learning for European Portuguese at L2F. In *Proceedings of the 4th International Conference on Computer Supported Education, CSEDU*, volume 2, pages 538–543.
- Reed, M. (2019). Listening skills instruction: practical tips for processing aural input. In Levis, J., Nagle, C., and Today, E., editors, *Proceedings of the 10th Pronunciation in Second Language Learning and Teaching Conference*, pages 401–412, Ames, IA. Iowa State University.
- Ribeiro, C., Jepp, P., Pereira, J., and Fradinho, M. (2010). Lessons learnt in building serious games and virtual worlds for competence development. In Taisch, M., Cassina, J., and Smeds, R., editors, *Proceedings of the 14th Workshop on Experimental Learning on Sustainable Management, Economics and Industrial Engineering*, pages 52–62, Milano, Italy.
- Richards, M. and Cotos, E. (2019). Teaching tip segmental accuracy: A recommended training sequence for moving learners from accurate perception to accurate and automatic production in

- the stream of speech. In Levis, J., Nagle, C., and Todey, E., editors, *Proceedings of the 10th Pronunciation in Second Language Learning and Teaching Conference*, pages 413–422, Ames, IA. Iowa State University.
- Santos, R., Freitas, M. J., and Veloso, J. (2014). Grupos consonânticos na escola: desenvolvimento fonológico e conhecimento ortográfico. *Diacrítica*, 28(1):407–436.
- Sawusch, J. R. (2005). Acoustic analysis and synthesis of speech. In Pisoni, D. B. and Remez, R. E., editors, *The Handbook of Speech Perception*, pages 7–27. Blackwell Publishin, 1st edition.
- Severino, C. S. (2011). Fronteiras prosódicas e desambiguação no português europeu. Master’s thesis, Faculdade de Letras, Universidade de Lisboa, Lisboa.
- Silva, A., Mamede, N., Ferreira, A., Baptista, J., and Fernandes, J. (2011). Towards a serious game for portuguese learning. In *Proceedings of the 2nd International Conference on Serious Games Development and Applications, SGDA*, volume 6944 of *Lecture Notes in Computer Science*, pages 83–94. Springer.
- Silva, A., Marques, C., Baptista, J., Jr., A. F., and Mamede, N. J. (2012). REAP.PT serious games for learning portuguese. In Caseli, H., Villavicencio, A., Teixeira, A., and Perdigão, F., editors, *Proceedings of the 10th International Conference on Computational Processing of the Portuguese Language, PROPOR*, volume 7243 of *Lecture Notes in Computer Science*, pages 248–259, Coimbra, Portugal. Springer.
- Snijders, T., Kooijman, V., Cutler, A., and Hagoort, P. (2007). Neurophysiological evidence of delayed segmentation in a foreign language. *Brain Research*, 1178(6):106–113.
- Tatham, M. and Morton, K. (2006). *Speech production and perception*. Palgrave Macmillan, New York.
- Van Engen, K., Baese-Berk, M., Baker, R., Choi, A., Kim, M., and Bradlow, A. (2010). The wildcat corpus of native- and foreign-accented english: Communicative efficiency across conversational dyads with varying language alignment profiles. *Language and speech*, 53(Pt.4):510–540.

- Vandergrift, L. and Goh, C. C. (2012). *Teaching and Learning Second Language Listening: Metacognition in Action*. Routledge, New York.
- Veloso, J. (2016). O sistema vocálico e a redução e neutralização das vogais átonas. In Martins, A. M. and Carrilho, E., editors, *Manual de linguística portuguesa*, pages 636–662. De Gruyter, Berlin, Boston.
- Vigário, M. (2011). *The prosodic word in European Portuguese*, volume 6 of *Interface Explorations [IE]*. De Gruyter Mouton.
- Weber, A. and Cutler, A. (2006). First-language phonotactics in second-language listening. *Journal of the Acoustical society of America*, 119(1):597–607.
- Zhou, C. (2017). Contributo para o estudo da aquisição das consoantes líquidas do português europeu por aprendentes chineses. Master's thesis, Faculdade de Letras, Universidade de Lisboa, Lisboa.
- Zhou, C., Cruz, M., and Frota, S. (2017). O ritmo da interlíngua na produção do português europeu por falantes chineses. *Revista da Associação Portuguesa de Linguística*, (3):423–435.
- Zhou, C., Freitas, M. J., and Castelo, A. (2019). A aquisição das consoantes líquidas do português europeu em coda por aprendentes chineses. In Alonso, C. P., Russo, V., Vecchi, R., and André, C. A., editors, *De Oriente a Ocidente: estudos da Associação Internacional de Lusitanistas*, volume V, Estudos da AIL sobre Ciências da Linguagem (Língua, Linguística, Didática), pages 87–117. Angelus Novus, Coimbra.
- Zhou, C., Freitas, M. J., and Castelo, A. (2021). On the acquisition of european portuguese liquid consonants by L1-Mandarin learners. In Pinto, J. and Alexandre, N., editors, *Multilingualism and third language acquisition: learning and teaching trends*, chapter 2, pages 25–45. Language Science Press, Berlin.

Chapter 8

Appendix

1. Single word identification tasks (answer sheet)

Faculdade de Letras da Universidade de Lisboa

1. Preencha os espaços em branco com as palavras que vai ouvir.

1. _____

2. _____

3. _____

4. _____

5. _____

6. _____

7. _____

8. _____

9. _____

10. _____

11. _____

1. Preencha os espaços em branco com as palavras que vai ouvir.

1. _____

2. _____

3. _____

4. _____

5. _____

6. _____

7. _____

8. _____

9. _____

10. _____

11. _____

Nome do aluno: _____

Catarina Maria Realinho Rafael

2. Cross word identification tasks (answer sheet)

1. Vai assistir duas vezes a uma notícia sobre a emigração dos jovens portugueses. Complete o texto com as palavras em falta, de acordo com a mesma.

Os jovens portugueses com qualificações académicas superiores e com idades entre os trinta e os trinta e _____ continuam a emigrar. O estudo da fundação empresarial de Portugal decorreu entre dezembro de dois mil e dezasseis e abril deste ano. Conclusões? Os _____ cada vez mais em busca de _____ de trabalho, mas o desejo de regressar ao _____ nasceram continua elevado. Mas vamos a números.. sessenta e cinco por cento dos jovens pensa voltar, sendo que destes trinta e quatro por _____ que provavelmente vai regressar, apenas oito por cento tem certezas quanto ao regresso e que este _____ planeado.

Mas se a vontade de regressar é um facto evidente, isto pode demorar tempo e essa é também outra variável deste estudo. Cerca de quarenta e um por cento diz que pensa fazê-lo daqui a mais de cinco anos, vinte e sete por cento daqui a mais de três e apenas uma pequena minoria quer regressar já em _____ e dezoito. Um dos fatores que mais influencia a decisão é a saudade. A esmagadora maioria diz que quer voltar para estar mais próximo da família, logo _____ as oportunidades de emprego, o ordenado em Portugal, a oferta de trabalho e por fim, a oportunidade de desenvolver o próprio negócio. Nos últimos cinco meses, a tendência de emigração foi mais acentuada nas mulheres. Quanto ao destino, o Reino Unido continua à frente e na Europa os _____ para emigrar são a Alemanha e a Holanda.

2. Vai assistir duas vezes a uma parte de uma reportagem sobre um casal de emigrantes portugueses em Inglaterra. Complete o texto com as palavras em falta, de acordo com a mesma.

Chegamos a segunda feira. O casal precisa _____ um emprego. Contas feitas quase _____ já foram gastos entre renda da casa, transportes, comida, enxoval. Diana, _____ de trabalho de João numa operadora de telemóveis, pode ser a carta de apresentação de que precisam. Está em Londres há um ano, é chefe de cozinha numa cadeia hoteleira. Traziam na bagagem este contacto que pode ser precioso. Diana _____ grupo está a recrutar gente. Tem vinte e dois hotéis.

João vai à primeira entrevista, aguardamos pelo _____ com a diretora de recursos humanos. Quarenta e cinco minutos depois.. *“ Correu bastante bem, ou seja, vou neste momento dirigir-me a Kensington a um hotel que estão a precisar de um rececionista. _____ gerente inclusive com quem ela falou é portuguesa, portanto vou _____, ou seja, uma segunda entrevista. ”*

João tem esperança. Rececionista de um hotel seria excelente para começar. *“À partida começo _____ . Portanto, correu bastante bem e em menos de uma semana penso que tenho emprego fulltime que era o que procurava com alguma _____ portanto acho que vai correr bem”*

Seis euros e _____, trinta e cinco horas semanais de trabalho mas quanto mais horas trabalhar, mais receberá. A amiga Diana chegou a trabalhar oitenta por semana quando aqui chegou, queria juntar dinheiro. Amanhã João terá que se apresentar de cabelo e barba cortados e impecavelmente vestido. Mais barato não encontrava. Cristiana vai tentar a sorte na mesma _____.

Nove da manhã; terça-feira; uma semana depois de ter chegado, João está a trabalhar. Cristiana também foi contratada para rececionista.

Nome do aluno: _____

Catarina Maria Realinho Rafael