UNIVERSIDADE ABERTA
UNIVERSIDADE NOVA DE LISBOA

Perceção de acento e redução vocálica na aquisição fonológica em PL2. Implicações para a prática pedagógica

Word stress and reduced vowels perception and acquisition in L2 Portuguese. Pedagogical implications

Gabriela Pereira Tavares

Tese de Doutoramento em Didática das Línguas – Multilinguismo e Educação para a Cidadania Global

PhD dissertation in Language Teaching - Multilingualism and Education for Global Citizenship

2024

UNIVERSIDADE ABERTA
UNIVERSIDADE NOVA DE LISBOA

Perceção de acento e redução vocálica na aquisição fonológica em PL2. Implicações para a prática pedagógica

Word stress and reduced vowels perception and acquisition in L2 Portuguese. Pedagogical implications

Gabriela Pereira Tavares

Tese de Doutoramento em Didática das Línguas – Multilinguismo e Educação para a Cidadania Global

PhD dissertation in Language Teaching - Multilingualism and Education for Global Citizenship

Supervision:

Professor Susana Correia (NOVA University Lisbon)
Professor Andrea Deme (Eötvös Loránd University)

Janeiro 2024

This work is licensed under a Creative Commons
NonCommercial-NoDerivatives 4.0 International licence (CC BY-NC-ND 4.0)

Acknowledgements

Writing this dissertation has been an adventure and, as with all adventures, it was filled with possibilities, discoveries, but also anxieties and moments of distress. Slowly, the feeling of a prolonged tiredness gives place to the realization of the knowledge gathered during the way, and the consciousness that ending this work is only the turning point to begin a new adventure.

My first acknowledgment goes to the people who, throughout the last three years, have made this journey alongside me. To my supervisors – Susana Correia e Andrea Deme – I owe my sincere gratitude for constantly teaching me the way of rigorous scientific thought and meticulous work. With you, I discovered my passion for phonetics, phonology, psycholinguistics, as well as for the academic world. You have nurtured my intellectual well-being, always supportive and encouraging. You became my role-model, from the very beginning, as teachers and as researchers. I sincerely hope one day I will be able to inspire students as you inspire me.

My heartfelt 'thank you' also goes to the people who, without understanding exactly the *what* (nor the *why*) of this work, understood how important it was to me, and stood by my side. To my father, for teaching me the value of intellectual honesty and work ethic. To my sister, for her unquestionable devotion to the family. To my daughter Alice and my son Vicente, for pushing me to be an example and for all the times you encouraged me, and objectively helped. To László, who arrived when my adventure had already begun, but who was by my side when it ended, tirelessly helping me, and waiting. To my friends who understand the *what* and the *why*: Aude, Brigitta, Clara, Isabel, Laura, and Móni. To you all, family and friends, my apologies for the time I should have been taking care of you and I was not.

To Antónia Fodor, Ágnes Keresztes, Daniella Szabó, Emese Rasztovich, Eszter Andorka, Gergely Sámson, Krisztina Babos, and Zsuzsanna Lakatos-Báldy, the responsible teachers for the L2 Portuguese courses in the Hungarian universities. Thank you for making this project possible.

I'm also in debt to the many volunteers – Portuguese and Hungarian – who recorded the stimuli, translated instructions and participants' responses, piloted the tests, and tested the tasks.

This dissertation would have not been possible without the grant from the Portuguese Foundation for Science and Technology (2020.05740.BD). I'm also grateful to CLUNL - Linguistics Research Centre of NOVA University Lisbon, who greatly contributed to my education, by supporting my participation in conferences and courses. I owe a special thanks to the University Corvinus of Budapest, for allowing me to take a pause from teaching, and to Mónika Bense, for taking such good care of the students in my absence.

Finally, to my students, those who were and those who will be. I'm always amazed by your motivation and commitment to learn a language that does not have the prestige of other more popular languages. After many years of teaching, you continue to make me feel fulfilled with this profession. I'm also grateful for all I learn with you at the end of each class. It was you who led to this work, and it is to you that I also dedicate this dissertation.

To my students

To Marcolina

UNIVERSIDADE ABERTA
www.uab.pt

DECLARAÇÃO DE INTEGRIDADE

STATEMENT OF INTEGRITY

Declaro ter atuado com integridade na elaboração da presente dissertação/tese. Confirmo que em todo o trabalho conducente à sua elaboração não recorri à prática de plágio ou a qualquer outra forma de falsificação de resultados.

Mais declaro que tomei conhecimento integral do Regulamento Disciplinar da Universidade Aberta, publicado no Diário da República, 2.ª série, n.º 215, de 6 de novembro de 2013.

I hereby declare having conducted my thesis with integrity. I confirm that I have not used plagiarism or any form of falsification of results in the process of the thesis elaboration.

I further declare that I have fully acknowledged Disciplinary Regulations of the Universidade Aberta (regulation published in the official journal Diário da República, 2.ª série, N.º 215, de 6 de novembro de 2013).

Universidade Aberta, 27 de junho de 2023

Nome completo/Full name: Gabriela Pereira Tavares

Assinatura/Signature:

Assinado por: **Gabriela Pereira Tavares**
Num. de Identificação: 10521369
Data: 2023.07.26 21:24:27+01'00'

manuscrita ou digital / handwritten or digital

CHAVE MÓVEL

Abstract

Word stress and reduced vowels perception and acquisition in L2 Portuguese. Pedagogical implications

L2 speech perception has regained attention from researchers in the last two decades. However, studies in L2 Portuguese are still scarce. This absence curtails the design of adequate didactic material for phonological acquisition. The present work aims at contributing to lessen this gap, by observing the acquisition of reduced vowels and word stress by Hungarian learners of L2 Portuguese. To this purpose, two experiments were conducted.

First, a forced-choice identification task was run to observe how Hungarian native speakers map the EP oral vowels into their L1 inventory. Results showed multiple cases of perceptual overlap, with higher rates for the vowels [ɐ], [ɛ] and [e]. The second experiment consisted of a high variability perceptual training designed to help learners with the difficulties observed in the identification task, as well as with word stress discrimination, in which Hungarian speakers have been reported to exhibit 'deafness'. Three groups of Hungarian learners of L2 Portuguese were recruited at the onset of learning: one group was trained on EP vowels discrimination, the second group received training on EP word stress without vowel reduction, and the third group trained EP word stress with vowel reduction. Results showed that the interventions focused on word stress were effective in stress discrimination, but no improvement was observed in EP vowel discrimination as a result of training.

The findings of the present work offer us an insight on L2 speech perception, namely that L2 vowel perceptual tunning is not a homogeneous process and that perceptual training aids learners to overcome problems with word stress discrimination.

Keywords: L2 speech perception; L2 Portuguese; cross-language influence; perceptual training; reduced vowels; word stress;

Resumo alargado [Portuguese]

Perceção de acento e redução vocálica na aquisição fonológica em PL2. Implicações para a prática pedagógica

Atualmente, a promoção oficial da língua portuguesa em universidades no estrangeiro abrange cerca de 90 países, através de cátedras académicas, leitorados e acordos de cooperação, e Centros de Língua Portuguesa (Camões – Instituto da Cooperação e da Língua, 2016a). Na Hungria, o português está presente em nove universidades, duas das quais com um professor nativo de português europeu (PE). Os materiais didáticos para o Português L2 (PL2) utilizados na Hungria incluem livros publicados em Portugal para um público mais vasto de aprendentes e livros criados especificamente para aprendentes húngaros. Os manuais publicados em Portugal incluem uma variedade de material auditivo, mas não se destinam especificamente a falantes nativos de húngaro. O material concebido para os alunos húngaros de PL2 é constituído por dois manuais e um volume de exercícios, mas nenhum deles inclui exercícios auditivos.

Uma breve consulta ao material disponível para o ensino do PL2 revela lacunas no que respeita à aquisição fonológica e, especificamente, a exercícios auditivos e de pronúncia. Embora nas últimas décadas a investigação sobre perceção de fala em L2 tenha atraído a atenção de investigadores, e a produção científica nesta área se tenha multiplicado e desenvolvido consideravelmente, os estudos sobre PL2 nas áreas da perceção da fala e aquisição fonológica ainda são escassos, o que condiciona a criação de material didático que reflita conhecimento baseado em evidência. O presente trabalho tem como objetivo contribuir para diminuir esta lacuna.

O tema desta tese foi motivado pela observação em sala de aula de problemas específicos observados em aprendentes húngaros na aquisição da fonologia do PE, nomeadamente com as vogais reduzidas [ɐ] e [ɨ], bem como com o acento de palavra. Relativamente a este último, estudos anteriores demonstraram que os falantes nativos húngaros apresentam insensibilidade a contrastes de acento de palavra, isto é 'surdez acentual' (Honbolygó et al., 2017; Peperkamp et al., 2010; Peperkamp & Dupoux, 2002).

Quanto às vogais átonas do PE [ɐ] e [ɨ], estas encontram-se ausentes no inventário vocálico húngaro e, até ao momento, não foram publicados estudos sobre a perceção destas vogais por falantes nativos húngaros. No entanto, as dificuldades e o percurso de aprendizagem dos nativos desta língua podem ser facilmente identificados nas produções escritas dos aprendentes, em sala de aula. No caso de [ɐ], os estudantes apresentam o hábito sistemático de transcrever esta vogal reduzida como <e>, que no sistema húngaro representa inequivocamente a categoria /ɛ/. Este problema é observável tanto em níveis de aprendizagem iniciais como em níveis mais avançados, embora menos frequentemente. Já no que diz respeito a [ɨ], embora se observem algumas dificuldades iniciais, os aprendentes parecem aperceber-se rapidamente de que, na fala, esta vogal é frequentemente apagada, ao contrário do que acontece com [ɐ].

Finalmente, vale a pena mencionar que os problemas descritos acima – vogais reduzidas e acento de palavra – estão fortemente interligados, uma vez que, em português, as vogais reduzidas ocorrem principalmente em sílabas átonas. Além disso, a redução das vogais é a principal pista acústica para a perceção de acento para os falantes nativos de PE (Correia et al., 2015; Delgado-Martins, 1986). A relação entre vogais reduzidas e acento de palavra é, portanto, de grande importância na aquisição do PL2.

A questão específica que orienta este trabalho é a de compreender como falantes cuja língua não inclui as vogais reduzidas [ɐ] ou [ɨ] e que não distingue contrastes de acento de palavra adquirem estas propriedades do PE. Para além desta questão, pretendemos também investigar qual o método mais eficiente para adquirir estas características. Para averiguar estas questões, foram conduzidos dois estudos.

O primeiro estudo consistiu num teste de identificação (*forced-choice identification task with goodness-of-fit rating*). Neste estudo foram analisados resultados de 78 participantes, que identificaram as nove vogais orais do PE a partir das nove vogais húngaras. Os resultados mostraram que as vogais portuguesas [ɐ] e [ɨ] são identificadas maioritariamente como as categorias húngaras /ɛ/ e /y/, respetivamente. Estes resultados sugerem os seguintes problemas: primeiro, existe um conflito na perceção de [ɐ], uma vez que tanto [ɐ] como [ɛ] são identificados pelos falantes húngaros como /ɛ/. Segundo, os ouvintes húngaros também apresentaram problemas na perceção das vogais [ɛ], [e] e, em menor grau, [i]. Com base nos resultados recolhidos nesta experiência, selecionámos então as vogais portuguesas a treinar por aprendentes húngaros: as vogais-alvo da presente tese, [ɐ] e [ɨ], e as vogais [ɛ], [e] e [i].

O segundo estudo consistiu num treino percetivo, que incluiu três intervenções, conduzidas paralelamente em três grupos de aprendentes húngaros de PL2: uma intervenção focada nas vogais orais do PE acima mencionadas (Grupo *Vowels*), uma intervenção focada em contraste de acento suprassegmentais (Grupo *Stress*), e uma intervenção focada nas vogais orais e acento do PE (Grupo *Vowels & Stress*). Os treinos foram completados por 66 aprendentes húngaros inscritos em cursos de PL2 Nível Iniciação, recrutados em nove universidades da Hungria. Cada intervenção teve a duração de seis semanas, tendo os aprendentes de completar uma sessão por semana. As sessões consistiram em tarefas de discriminação AX e AXB, com cerca de 10 ou 15 minutos cada. Antes e depois do treino, os participantes completaram um pré-teste e um pós-teste. Estes consistiram em tarefas de intruso (*oddity tasks*), e testaram a discriminação das vogais orais centrais e anteriores do PE, assim como do acento de palavra a nível suprassegmental. Todas as etapas – testes, sessões de treino e questionários – foram conduzidas online.

No que diz respeito a alterações na discriminação das vogais, a análise dos resultados mostrou que, no geral, nenhum dos três grupos melhorou significativamente na discriminação das vogais-alvo, nem mesmo os grupos cujas intervenções focaram contrastes vocálicos – *Vowels* e *Vowels & Stress*. No entanto, na análise dos resultados de cada contraste observou-se que os três grupos melhoraram na discriminação dos contrastes [ɐ]-[ɛ] e [ɐ]-[ɨ], embora neste último caso a melhoria tenha sido apenas ligeira. Relativamente à discriminação de contrastes de acento de palavra, os grupos cujo treino focou este aspeto – Grupo *Stress* e Grupo *Vowels & Stress* – apresentaram uma melhoria significativa na perceção de acento de palavra, comparativamente ao Grupo *Vowels*. A análise mostrou ainda que o Grupo *Stress* apresentou melhorias mais visíveis do que o Grupo *Vowels & Stress*.

Os resultados do estudo do treino percetivo permitiram-nos apresentar diferentes conclusões. Primeiro, a trajetória de aquisição de vogais não é homogénea, uma vez que os participantes apresentaram melhorias apenas alguns contrastes vocálicos.

Especificamente, os aprendentes húngaros de PL2 apresentaram uma melhoria mais visível no contraste que prevíamos ser o mais problemático: [ɐ]-[ɛ]. Contrariamente, nos contrastes envolvendo as vogais [ɛ], [e] e [i], os participantes revelaram mais dificuldades.

Segundo, a melhoria observada na discriminação de contrastes de acento de palavra mostra que a 'surdez acentual' pode ser ultrapassada com um treino adequado, e em pouco tempo (6 semanas).

Terceiro, um treino que combine vogais e acento de palavra não é mais eficaz do que treinos que foquem estas propriedades separadamente. No entanto, acrescentamos que a análise dos resultados também não apresenta evidência estatística de que um treino que combine vogais e acento de palavra seja menos eficaz.

Finalmente, os resultados mostram que aprendentes húngaros de PL2 no nível inicial da aprendizagem da L2 beneficiam de um treino percetivo que foca uma propriedade suprassegmental (acento de palavra), mas não de um treino percetivo que foca uma propriedade segmental (vogais). Este facto sugere que, tal como para a L1, na aquisição da perceção de fala em L2, as propriedades suprassegmentais precedem as segmentais. No entanto, esta conclusão deve de ser interpretada com precaução, uma vez que os resultados podem ser específicos do contexto do estudo realizado.

Os estudos apresentados nesta tese apresentam evidência empírica que, esperamos, possa vir a contribuir para uma melhor compreensão do processo de aquisição de fala do PL2. Além disso, com este trabalho, pretende-se também contribuir para o desenvolvimento de material didático a partir de conhecimento cientificamente adquirido.

Palavras-chave: perceção de fala em L2; PL2; influência interlinguística; treino percetivo; vogais reduzidas; acento de palavra

Table of Contents

PART II: EMPIRICAL STUDIES

LIST OF FIGURES

List of Tables

SYMBOLS AND ABBREVIATIONS

EP	European Portuguese
L1	first language
L2	second language
C	consonant
V	vowel
< >	orthographic form for character (e.g., <e>) note: orthographic form for words is noted in italic (e.g., *casa*)
'	gloss in English (e.g., 'house')
]	morphological boundary (e.g., $\text{cas]}_{\text{ROOT}}\text{a]}_{\text{CLASS MARKER}}$)
ˈ	primary stress (e.g., ['kazɐ])
ˌ	secondary stress (e.g., [ˈfoʊtəˌgræf])
*	ungrammatical form (e.g., *kasa)
f0	fundamental frequency
F	formant
ISI	inter-stimulus interval
HVPT	high variability perceptual training
LVPT	low variability perceptual training
T^F	female talker (e.g., $T1^F$)
T^M	male talker (e.g., $T4^M$)

Introduction

The official promotion of the Portuguese language beyond national boarders dates back from the 1920's, with the creation of a first lectureship in France, followed by Germany and the United Kingdom (Guedes, 1998). Today, the network supporting Portuguese at universities covers nearly 90 countries through academic chairs, lectureships and cooperation agreements, and Portuguese Language Centers (Camões – Instituto da Cooperação e da Língua, 2016a).[1]

In Hungary, Portuguese is present in nine universities, two of which employ a native Portuguese lecturer. Only one institution – Eötvös Loránd University (Budapest) – offers a university degree in Portuguese studies, specifically, the degree of *Philologist in Romanistics – Specialized in Portuguese studies*. Didactic materials for L2 Portuguese used in the courses include books published in Portugal for a wider audience of learners, and books created specifically for Hungarian learners. The coursebooks published in Portugal include a variety of auditory material, but they do not specifically target Hungarian native speakers. The material designed for Hungarian L2 Portuguese learners consists of two coursebooks and an exercise volume, but neither include auditory exercises.

A brief consult to the material available to L2 Portuguese teaching shows relevant gaps when it comes to phonological acquisition, and, specifically, auditory and pronunciation exercises. As Allegro (2014) had already pointed out, in L2 Portuguese coursebooks, most exercises still consist of list of segments or words that the learner must listen and repeat, in the hope that he/she will 'magically' internalize the phonetics of Portuguese. The popularity of this type of drill – a heritance from the Audiolingual Method – is not limited to L2 Portuguese. To the day, this approach continues to be a source for learning tasks, although resorting to new techniques (Moyer, 2013). The stagnation of L2 phonology (mostly referring to pronunciation), after the Audiolingual Method fell out of favor, was caused by a conjunction of different factors (Levis & Sonsaat, 2017; Moyer, 2013). First, cognitivist ideas viewed errors as a natural consequence of learning and, presumably, an adequate exposure to the input would be enough to surpass them. Second,

[1] *Instituto Camões* also supports a network of teachers at primary and secondary schools.

the assumption of a *critical period* for L2 learning (Lenneberg, 1967) contributed to discourage teachers to focus on pronunciation mastering. Additionally, the Communicative Language Teaching approach caused a shift from linguistic competence (understood as formal knowledge of the L2) to communicative competence (i.e., the ability to use the L2 in real communicative situations). Since pronunciation was considered a linguistic competence, the focus was then transferred from discrete-segmental accuracy to global fluency (Rato, 2013). Furthermore, still in line with the Communicative approach, as a result of the debates on the principles for L2 English teaching promoted by the TESOL International Association, the aim set for the end-state of pronunciation teaching was no longer native-like productions but rather intelligibility. Interestingly, as a consequence of these changes, suprasegmental properties were prioritized rather than segmentals.

An example of the 'peripheral' place of phonology in L2 Portuguese profiles and guidelines, as compared to lexicon and grammar, is the *Referencial Camões PLE*, a didactic document conceived as a support tool for teachers of L2 Portuguese (Camões – Instituto da Cooperação e da Língua, 2016b). It consists of an inventory of *functions* (e.g., *expressing wishes and emotions*), *notions* (e.g., *localization*), and *grammar* (e.g., *demonstrative pronouns*), for each proficiency level of the Common European Framework of Reference for Languages. The Portuguese document does not include (yet) any instructions for auditory perception or pronunciation. Such absence can be explained with the scarcity of research in L2 Portuguese phonology.

As for the Common European Framework, although the 2001 edition crucially lacked descriptors for pronunciation teaching and training (Moyer, 2013), the revised version (Council of Europe, 2020) implemented important changes with respect to *Phonological control*, namely by presenting a more detailed and realistic progression between levels.

Despite recent developments, more research and time are needed to implement new guidelines for L2 curricula and didactic material. To the moment, the lack of proper guidance for L2 phonology results in a tendency to avoid pronunciation topics in the L2 classroom. Reed & Levis (2015) remind us that although L2 teachers acknowledge the importance of teaching phonetics and phonology, few are familiar or even comfortable with this linguistic area. According to Derwing & Munro (2005), in preparation courses for L2 teachers, pronunciation is frequently left out and, consequently, teachers must decide themselves what to teach in this topic and how to teach it. The lack of preparation of L2 teachers in the area of phonology is oftentimes aggravated by a certain degree of unease with new technologies and/or programs (Allegro, 2014).

The topic of this dissertation emerged as a reflection of the problems mentioned above, also observed in the context of L2 Portuguese teaching for Hungarian learners. The search on how to surpass the difficulties in the acquisition of specific Portuguese features, namely reduced vowels and word stress, led to the observation that these features are not adequately treated – if they are treated at all – in L2 Portuguese coursebooks and classrooms. Furthermore, considering the scarce research in L2 Portuguese phonology, one may question on which basis the exercises found in the coursebooks were designed – teaching experience or adaptation from other L2? It must be pointed out that this twofold problem – scarce research in L2 phonology and lack of reflection of research findings in teaching material – is also observed in L2 English (Derwing & Munro, 2005).

In sum, L2 Portuguese is in a developing moment, but empirical evidence to support didactic materials is still scarce. This scarcity is even more evident in the context of research in pronunciation teaching. This dissertation aims at contributing to evidence-based L2 Portuguese teaching and learning of pronunciation.

Why perception?

Despite a lasting dominance of grammar and vocabulary in L2 research, a certain rehabilitation of phonology has been observed in the last decades, also coinciding with the emergence of technological advances in computers, accessible software of speech analysis and safer techniques for neurophysiological testing (Edwards & Zampini, 2008). Also, speech perception gained importance as a relevant factor in L2 pronunciation difficulties. The role of perception in pronunciation had already been underlined by Trubetzkoy (1958/1969), who suggested that properties of the L1 act as a *phonological filter* in non-native perception, leading to inadequate production. Although some studies do not find a straightforward perception-production link (Chan, 2014; Sheldon & Strange, 1982), an extensive body of evidence gathered in the last two decades seems to supports the fact that problems in production can be related to poor discrimination abilities (Chao et al., 2019; Flege et al., 1999; Perkell et al., 2004; Rauber et al., 2010, 2005). Studies using magnetic resonance imaging techniques provide further evidence for the perception-production link by observing an overlap in the cortical areas activated in listening and in speech production (Reiterer et al., 2013; Wilson et al., 2004). Based on the positive correlation between perception and production, the Speech Learning Model (Flege, 1999) posited that accuracy in L2 production is dependent of accuracy in L2 perception.[2] Later on, this position was revised, considering some methodological issues in the analyses of correlations between perception and production results, and instead the existence of a bidirectional coevolution of perception and production in L2 acquisition was proposed (Flege & Bohn, 2021).

Although researchers were not yet able to categorically demonstrate that perception precedes production in L2 acquisition (as overwhelmingly shown for L1; Gervain, 2014), results in L2 speech research do show a tendency for the precedence of the former over the latter (Akahane-Yamada et al., 1996; Bradlow et al., 1997, 1999; Lengeris & Hazan, 2010; Rato, 2013; Wang et al., 2003). Additionally, studies have shown that non-native speakers are able to rate foreign-accent productions similarly to native speakers, although they may present themselves a strong accent in that L2 (Flege, 1988; Mackay et al., 2006).

Considering that perception may modulate L2 phonological acquisition and be a cause for difficulties in production, this dissertation will focus on the perception of L2 Portuguese speech perception and acquisition.

The case of Hungarian learners

As previously mentioned, the topic of this dissertation is motivated by the observation in the classroom of specific problems shown by Hungarian learners with

[2] However, Flege (1999) does not claim that all production problems have roots exclusively in perceptual factors.

Portuguese phonology, namely with the reduced vowels [ɐ] and [ɨ], as well as with word stress.

With respect to [ɐ] and [ɨ], the difficulties and learning path can be easily identified in written productions. The frequent responses given by students at the Beginner level, in an auditory task in which they are asked to identify, for instance, the proper feminine name *Ana* ['ɐnɐ], are the following:

(1a) **Anna*

(1b) **Ene*

(1c) **Ena*[3]

Each of these responses suggests a different strategy or learning stage. In (1a), students identify the proper name, and simply write it in Hungarian. In (1b), the students do not identify the proper name and eventually write the word as they perceive it. This suggests that they may be categorizing the Portuguese [ɐ] as the Hungarian /ɛ/. In Hungarian, the grapheme-phoneme correspondence is transparent, and the Hungarian vowel /ɛ/ has the unequivocal orthographic form <e>. Finally, in (1c), it may be the case that students have already learnt that the word-final <e> in Portuguese is frequently deleted. In this situation, learners may be conscious that **Ene* would sound [ɛn] and, accordingly, they self-correct the response to **Ena*. The problems with the [ɐ] – written with <e> – seem to be persistent even in more advanced learning stages, as demonstrated by the transcription of novel words (e.g., **calemidade* instead of *calamidade* [kɐlɐmi'dadɨ] 'calamity').[4] As for [ɨ], although some early difficulties are observed, learners seem to rapidly realize that, in speech, the vowel [ɨ] is frequently deleted, contrary to [ɐ].

Problems with word stress are more difficult to assess based on students' productions, since EP stress mostly involves both suprasegmental and segmental cues. Previous studies have shown that Hungarian listeners, whose language has fixed-stress in

[3] This is an anecdotal example observed by the author of the dissertation, in the last years of practice. In the auditory task, a female Portuguese speaker introduces herself as "Ana da Silva".

[4] The first [ɐ] is correctly transcribed by students presumably due to a self-monitoring strategy, since more advanced learners have better control EP orthographic rules and know that by writing **celemidade*, the initial syllable would correspond to [sɨ], since in EP, <c> followed by <e> is realized as [s] and not [k].

the first syllable, exhibit persistent ‘deafness’ in the perception of stress contrasts (Honbolygó et al., 2017; Peperkamp et al., 2010; Peperkamp & Dupoux, 2002). In fact, students show persistent ‘deafness’ to stress in Portuguese even when explicitly instructed about it in the classroom. For instance, students struggle with perceiving the stress contrast in the words *partiram* [pɐɾˈtiɾɐ̃w̃] 'they left' vs. *partirão* [pɐɾtiˈɾɐ̃w̃] 'they will leave'. Indeed, observations in the classroom suggest that the learning path for EP stress acquisition is not straightforward. For example, there are some anecdotal examples of persistent incorrect pronunciation of *sábado* [ˈsabɐdu] 'Saturday'. Many students tend to pronounce it stressed in the second syllable, which does not obey neither the Hungarian stress rule (stress in the first syllable of the word) nor to the word they listen to in Portuguese or that they recognize from English (*Saturday* [ˈsæt̬ɚdeɪ]), but it is consistent with EP regular stress (stress in the right edge of the root, frequently in the penultimate syllable – here, *sábado* is an exception).

Finally, it is worthwhile mentioning that the problems described above – reduced vowels and word stress – are in fact interconnected, since in Portuguese, reduced vowels mainly occur in unstressed syllables. Furthermore, vowel reduction is the main cue to EP native stress perception (Correia et al., 2015; Delgado-Martins, 1986). The link between reduced vowels and stress is, therefore, of great importance in L2 Portuguese acquisition.

The present dissertation is guided by the general question of how native speakers of a fixed-stress language acquire variable word stress and reduced vowels. This implies, it is presumed, mastering both segmental and suprasegmental properties that are absent in their native language, therefore providing insight on the learning mechanisms and interaction between different phonological domains in L2 speech. In this line, the following questions arise:

(i) How do L2 learners map unfamiliar segmental and suprasegmental features at the onset of learning?

(ii) Is there a hierarchy in L2 acquisition between segmental features and suprasegmental features?

(iii) Should auditory tasks focus on segmental and suprasegmental separately or combine both features?

To investigate these issues, we recruited learners enrolled in Beginner courses of L2 Portuguese at Hungarian universities. Two aspects must be mentioned regarding the recruited learners. First, these participants reflect an important context of L2 Portuguese teaching worldwide: adult L2 acquisition in a formal learning environment (the classroom). Second, studies on L2 perceptual training[5] rarely involve learners at the early stages of the acquisition process, rather observing students with some knowledge of the L2 that are able to read and write in the target language (Aliaga-García, 2017; Aliaga-García & Mora, 2009; Brawerman-Albini, 2012; Bundgaard-Nielsen et al., 2011; Dupoux et al., 2008; Grenon et al., 2019; Kissling, 2015; Michaux, 2016; Ou, 2011; Rato, 2013; Ylinen et al., 2010). The choice of having early beginners as research subjects is grounded in the fact that advanced learners have already undergone a process of perceptual recategorization of L2 phonological properties, that is, the initial categorization of L2 sounds has changed due to auditory tunning, lexical acquisition, and other linguistic processes. However, in order to investigate the early changes in L2 phonological knowledge, one must observe the initial perceptual mapping of the L2 features, which can only be achieved at the onset of acquisition.

It is important to point out that strictly speaking, European Portuguese (EP) configures an L3 to the Hungarian learners recruited for this study, since all of them reported a good knowledge of at least one L2, at the time they started learning Portuguese. The term *L3* is used to identify any language in the process of acquisition after the first L2, that is, any L2+n. Research on L3 acquisition is motivated by the assumption that the acquisition of an L3 is facilitated by the fact that the learner has already experienced the acquisition of another language, other than the L1 (Hammarberg, 2007). However, given

[5] In L2 phonology research, three terms are used: *perceptual training*, *auditory training* and *phonetic training.* We opted for *perceptual training* for the following reasons. First, *phonetic training* does not specifically entail training of perception, since it is also use for training of production. Second, although *auditory training* is also found in the literature, the key issue in our work is speech perception (differences are described in Chapter 1). Accordingly, *perceptual training* is a more adequate term for the present dissertation.

the fact that the term *L3* is used in a specific research context that is not the aim of the present work, the term *L2* is going to be used instead, since it is more in line with the research approach that guide this dissertation.

Structure of the dissertation

This work is divided into two Parts: **(I) Theoretical Background** and **(II) Empirical Studies**. In Part I, a literature review on L2 speech perception will be presented (**Chapter 1**). In this chapter, a description of factors on L2 speech perception will be reported, followed by an account of studies in L2 vowel and stress perception and acquisition. In **Chapter 2**, we will describe the phonetics and phonology of EP and Hungarian. In the last chapter of Part I, **Chapter 3**, the Research Questions and Hypotheses will be put forth.

Part II includes two studies. The first study consists of a perceptual identification experiment conducted with naïve Hungarian listeners and aimed at observing how these listeners map the EP oral vowels into their L1 system. This study is described in **Chapter 4**.

Chapter 5 reports the main study of this dissertation, a perceptual training study focused on vowels and word stress, conducted with Hungarian learners of L2 Portuguese at the onset of learning.

The two studies reported in Chapter 4 and Chapter 5 will be reviewed in the **General discussion**, and their results argued in light of the hypotheses and theoretical background that guided this dissertation.

In the **Conclusion**, our findings will be summarized, and the relevance of our work for the design of didactic materials for L2 Portuguese will be examined. We will also present some considerations on future research.

PART 1

THEORETICAL BACKGROUND

CHAPTER 1

L2 SPEECH PERCEPTION

1.1. Speech perception

Sound perception involves a chain of neurophysiological events, processed by the peripheral auditory system (Figure 1.1) and the central auditory system (Figure 1.2). As put by Rost "[t]he anatomy of hearing is elegant in its efficiency" (2013, p. 12). *Hearing* engages a chain of events that starts with the reception of sound waves from the exterior, captured and amplified by the outer ear, travelling through the middle ear, until they arrive to the inner ear. In the inner year, the fluid of the cochlea moves as a reaction to the sound waves, which are transformed into electrical pulses. In turn, these electrical pulses travel along the auditory nerve to the brain. Sound is *perceived* if the electrical pulses reach and are processed by a cortical area of the brain. Speech perception holds the physiological and neurological mechanisms mentioned above, as well as computational and cognitive operations, necessary to transform the acoustic input into meaning. To this purpose, several areas of the brain are or may be activated (Figure 1.2, next page).

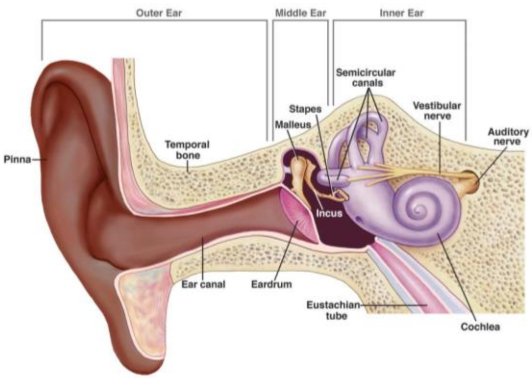


Figure 1.1. The peripheral auditory system (National Institute of Deafness and Other Communication Disorders, 2015)

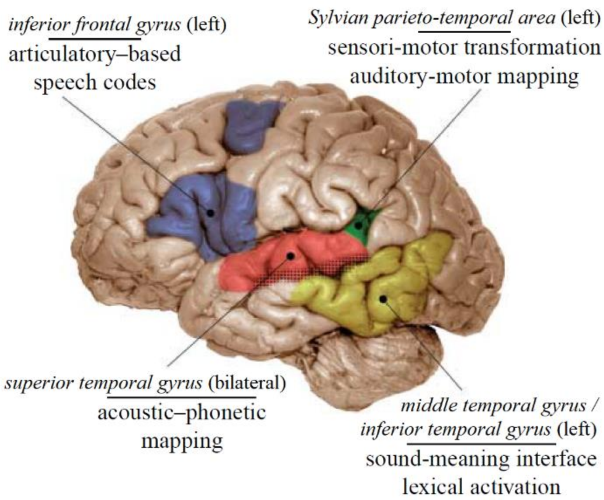


Figure 1.2. The central auditory system (Poeppel et al., 2008, p. 1075)

Although listeners with no hearing impairments present remarkable auditory discrimination abilities, and are able to discriminate even a very slight difference in the auditory input, as Lotto & Holt (2016) explain, speech perception is more than just discriminating sounds, it rather consists of recognizing patterns in a continuum. The authors add that "[t]hese patterns must be robustly encoded and perceptually discriminable for efficient speech communication" (p. 185).

Several theorical approaches have been put forward to describe the process of speech perception. The *motor theory*, proposed by Liberman and colleagues (Liberman et al., 1967; Liberman & Mattingly, 1985), was one of the most popular theory in the last century. In this theory, it is argued the articulatory nature of speech perception (rather than acoustic or auditory). Specifically, the authors posit that speech perception is the process of recognizing the neuromotor commands underlying vocal tract gestures. Moreover, the theory claims that speech perception is innate and human-specific. In line with this articulatory-based model, Fowler (1986) proposed the *direct realist* theory, according to which, speech perception is not mediated by representations (concepts or ideas), hence the term 'direct'. Contrary to the motor theory, the direct realism approach rejected the idea that speech perception is special, that is, unique to humans.

In the mid-seventies, new empirical findings showed that contrary to the previous assumption, some similarities can be found between speech and nonspeech sounds' perception (Pisoni, 1977). Additionally, other experiments revealed that perception of certain acoustic features that discriminate speech sounds is not limited to humans, but can be observed in some mammals or even birds (Kuhl & Miller, 1978). This led to a new approach, which proposed that speech perception is part of *general auditory* mechanisms, that is, access to spoken messages is operated from the acoustic signal, and not from articulatory gestures.

Another important finding by Liberman and colleagues (Liberman et al., 1957) is the notion of *categorical perception*. While conducting identification experiments, the researchers observed that listeners labelled sounds that varied along a continuum in abruptly distinct categories. Furthermore, participants also showed better performance in discriminating sounds belonging to different categories than sounds from a single category. A clear explanation of categorical perception is presented below.

Categorical perception occurs when stimuli distributed along a continuum are perceived categorically rather than continuously. An example is the contrast /t/-/d/, as in /taɪ/-/daɪ/ (*tie* vs. *die*). Although the voice onset time changes in a continuum, listeners will identify either /t/ or /d/ (Psy, 2002).

Categorical perception – as in general speech perception – is inherently *language-specific*, that is, varies between languages. The most obvious proof of this fact is the observed 'deafness' of some adult speakers when it comes to discriminating non-native contrasts. Studies have shown that from very early development stages, infants' speech perception is adapted to their environmental language. For example, Polka & Werker (1994) compared the discrimination ability of non-native contrasts between 6-8 months infants and 10-12 months infants. The authors confirmed the existence of a shift from language-general to language-specific patterns, and that this shift occurs during the 1st year of life. Moreover, based on their results and the results of previous studies, the authors suggested that the mentioned shift is more likely to occur first for vowels than for consonants. This and other studies led to the *perceptual magnet effect* theory (Kuhl, 2000; Kuhl & Iverson, 1995). This framework posits that infants map statistical features of their ambient language, and that this language filter begins to form in their 1st year, even before children start to speak. The experience with the ambient language influences infants' discrimination abilities, consequently altering perception in favor of that language.

Having introduced the speech perception features central to this dissertation – its *auditory*, *categorical* and *language-specific* nature –, we will focus next on how those features operate in second language (L2) perception.

1.2. L2 speech perception

Perceptual acquisition in the first language (L1) starts even before lexical acquisition. The attunement of sounds in linguistic categories demands "a long-term exposure to statistically defined distributions of sound tokens" (Flege & Bohn, 2021). This distributional learning – predicted to exist in L2 acquisition too – is ruled by an optimal weighting of different auditory cues. As Flege and Bohn (2021) explain, distributional learning and optimalization takes several years in L1 acquisition, and it is reasonable to predict that in L2 acquisition this process can last as long or even longer than in L1.

When comparing L1 to L2 acquisition, there is a growing consensus on the idea that adult L2 learners preserve the learning devices used for L1 acquisition. However, it is also clear that adults struggle to acquire a language and infants do not. Several factors have been pointed out to explain this fact.

In the sixties, the idea that adults would not be able to acquire an L2 with the same proficiency as their L1 or a native speaker of the L2 – the ultimate attainment – was explained on biological and neurological grounds. The *critical period hypothesis* (Lenneberg, 1967) argued that the completion of the hemispheric lateralization, occurring around puberty, would be the cause for the loss of learning abilities. Accordingly, after puberty, ultimate attainment in L2 would be virtually impossible. However, in more recent years, the importance of such maturational constraints in L2 acquisition was questioned, as well as the existence of a specific age limit for biological predisposition for language acquisition. Instead, some proposals defend a more continuous process of maturation, and consequently, a continuous decrease of the abilities to acquire a native-like L2 (Hyltenstam & Abrahamsson, 2003).

With the increase of research in L2 acquisition, other factors – social, psychological or linguistic – were given equal or more importance than maturational constraints to explain learning difficulties. As a result of this shift, multiple topics have been investigated, among others, the *onset of learning* (Abrahamsson & Hyltenstam, 2009), *age of arrival* (Flege et al., 1999), *language experience* (Bohn & Flege, 1990), *attention* (Guion & Pederson, 2007), *markedness* (Eckman, 2008), *L1 transfer* (Major, 2008), *L1-L2 cross-linguistic effects* (Best & Strange, 1992), *input* (Flege, 2008, 2018), *perceptual overlap* (Faris et al., 2018), *cue-weighting* (Cebrian, 2006) and *individual variability* (Dörnyei, 2014). All these aspects may contribute to our understanding why some linguistic features are more difficult than others to perceive and/or acquire by non-native speakers, or why some learners are better than others in the acquisition of the same L2. However, considering that the last four factors listed above – *L1-L2 cross-linguistic effects*, *input*, *perceptual overlap*, *cue-weighting*, and *individual variability* – are specifically relevant for our work, we will look closer into these. Cross-language perceptual models will be described in detail in section 1.3, and the remaining factors are introduced below.

Input has been one of the most studied issues, given its fundamental importance for L2 (and L1) acquisition. Considering that the L2 input to which adults are exposed is different than the input a native speaker is exposed to as an infant, the end states of L1 and L2 speech perception necessarily will present differences. This is especially true if we acknowledge the fact that L2 adult learners are often exposed to different varieties or dialects, or even foreign-accented productions of the L2, which renders the L1 input less uniform (Bohn & Bundgaard-Nielsen, 2008). As pointed out by Colantoni et al. (2015), input is probably one of the main factors for individual variability in L2.

Both quantity and quality must be considered, when it comes to the input necessary for the L2 learner to reach optimal L2 speech perception (Boersma & Escudero 2004). In line with this idea, high variability perceptual training (HVPT) has been established as the most adequate to L2 perceptual learning (Lively et al., 1993, 1994; Logan et al., 1991; Pisoni & Lively, 1995). On the other hand, some studies present some conflicting results, since no evidence of significant benefits of HVPT over low variability perceptual training (LVPT) were found (Brekelmans et al., 2022), or the benefits were not observed across conditions

(Wade et al., 2007). Apart from mixed results, it is safe to assume that HVPT may not always lead to significantly better results in L2 perception learning than LVPT, but it will lead to better or similar results as a LVPT study. For example, Wong (2012) found evidence that learners trained with LVPT can significantly improve their L2 perception, whereas participants training with HVPT improved significantly better than the ones who completed the LVPT.

Besides variability, other aspects must be considered when designing experimental stimuli. A question that arises immediately concerns the ecological nature of the stimuli: Which one is more efficient, natural or synthetic stimuli? Nobre-Oliveira (2007) conducted a study with Brazilian learners of L2 English, who received training on L2 vowels, some with natural stimuli, and others with synthetic stimuli. Results showed no significant difference between training with natural stimuli and training with synthetic stimuli, although the author refers that participants trained with the synthetic stimuli achieved better results. Rosa (2022) also compared the efficiency of the two types of stimuli with Brazilian learners of L2 English, aiming at the acquisition of /æ/ and /ɛ/[6], and, as Nobre-Oliveira (2007), did not find significant differences between the two types of training. However, the author reports that participants trained with natural stimuli improved more than participants who were trained with synthetic stimuli.

The question of stimuli manipulated from natural tokens has also been mentioned in the literature. Iverson et al. (2005) conducted a study comparing natural stimuli and stimuli manipulated with specific enhanced acoustic cues. After ten training sessions, all the groups showed improvement, and no significant differences were found between the results of each group. Thus, the authors concluded that there are no benefits of training with high variability natural stimuli over acoustically manipulated stimuli, or vice-versa.

Perceptual overlap, that is, overlap in the perception of more than one L2 sound as a single L1 category, has also been signaled as an important issue to understand difficulties in L2 speech perception (Elvin et al., 2021; Faris et al., 2018; Flege & Mackay, 2004; Tyler

[6] In this work, we will maintain the notation conventions (including phonetic, phonological, and orthographic notations) used by the authors in their publications to indicate the target L2 sounds.

et al., 2014). Overlap is usually calculated based on results of identification tasks[7], by looking at situations when two or more L2 vowels are partially or completely identified as the same L1 category. In L2 perceptual experiments, overlap values are usually calculated "as the smaller percentage of responses when two members of a pair of non-native (or L2) speech sounds are assimilated to the same native category" (Levy, 2009). Figure 1.3 provides an example, with results from a forced-choice identification task conducted by Elvin et al. (2021).

ES Vowels	BP Vowels						
	i	e	ɛ	a	o	ɔ	u
i	**1.0**	**0.56**	0.01				
e		0.44	**0.94**				
a			0.05	**1.00**		0.02	
o					**0.69**	**0.97**	0.01
u					0.31	0.01	**0.99**

Figure 1.3. Identification of Brazilian Portuguese vowels by European Spanish native speakers (Elvin et al., 2021, p. 14)

In this task, Spanish native speakers had to identify in their L1 vowel system seven Brazilian Portuguese vowels. According to the results, the contrast /i/-/a/ does not present any probability of perceptual overlap, since the two sounds are identified as different L1 categories. The contrast /o/-/ɔ/ presents the highest level of perceived overlap, 69%, since /o/ is identified in 69% of the cases as /o/ and /ɔ/, in 97% of the cases. Therefore, the smallest percentage is 69%. /o/-/u/ presents a lower level of overlap, 31%.

In a study conducted with American English speakers, Tyler et al. (2014) found evidence of the effect of perceptual overlap in discrimination of non-native contrasts, when comparing performances between a forced-choice identification task and an AXB discrimination task.[8] The results showed that contrasts with no overlap or with a very low degree of overlap were significantly better discriminated than highly overlapping contrasts.

[7] In identification tasks, listeners are asked to identify a sound among different possibilities, in the L1 or L2, usually presented in orthographic form.

[8] In an AXB tasks, listeners hear a sequence of three tokens and decide if the second (X) matches the first (A) or the third token (B).

Flege & Mackay (2004) also observed effects of overlap in discrimination performance. In this study, Italian native speakers were tested in English vowels, by means of a forced-choice identification task and an oddity discrimination task[9]. A comparison of the results showed that participants presented poorer discrimination performances for contrasts with high degrees of overlap, and, contrarily, more accurate discrimination for the contrasts with low values of overlap. However, the authors also found some conflicting results. The authors suggest that individual variability, or acoustic features (proximity in formant characteristics) may explain the deviant results. Faris et al. (2018) observed discrimination of non-native vowels by Australian English speakers, presented with Danish monophthongs and diphthongs. In their study, the authors concluded that perceived phonological overlap is a more reliable predictor for discrimination difficulties than the classification of assimilation types according to the Perceptual Assimilation Model (Best, 1995)[10]: discrimination was very good to excellent for contrasts with absence of overlap and poor to moderated for contrasts where a partial overlap was observed.

The issue of *cue-weighting* is another relevant topic to this dissertation, especially that in Hungarian (the L1 of participants in the studies presented in the dissertation) vowel length is lexically contrastive, which is not observed in Portuguese (L2 of participants). Besides investigating the ability of change from relying more in an acoustic cue to another, this topic addresses the more general question of accessing features in L2 that are absent in the L1 of the speaker. Bohn & Flege (1990) observed the perception of the English contrasts /i/-/ɪ/ and /ɛ/-/æ/ by German speakers with different L2 experience levels, and conclusions were different for each contrast.[11] In the case of /i/-/ɪ/, which has a similar counterpart in German, no effect of L2 experience was found. Contrarily, in the contrast /ɛ/-/æ/ (not present in German), participants with more L2 experience relied mostly on spectral cues, similar to the native speakers, whereas the inexperienced German participants relied mostly on temporal cues. The authors suggest that the reliance on durational cues is not due to an L1 strategy but rather to a more general strategy in speech perception that emerges if spectral information is not sufficient (Bohn & Flege, 1990,

[9] In oddity tasks, listeners hear a sequence of sounds and must identify the odd (different) one.
[10] This model's assimilation types will be described in section 1.3.1.
[11] *L2 experience* refers to length of residence/immersion in the L2.

p. 326). This view was confirmed by Bohn (1995) in a series of experiments with English, Germain, Spanish, and Mandarin native speakers, in perception of the English contrast /æ/-/ɛ/ or /i/-/ɪ/. Results showed that all language groups relied on duration, which was expected for the German, but not for the Spanish and Mandarin listeners, since duration is not contrastive in these languages. The author proposed the *desensitization hypothesis*, which states that "whenever spectral differences are insufficient to differentiate vowel contrasts because previous linguistic experience did not sensitize listeners to these spectral differences, duration differences will be used to differentiate the non-native vowel contrast" (Bohn, 1995, pp. 294–295). Cebrian (2006) also investigated the relation between L2 experience and cue-weighting, with the English lax-tense contrast /i/-/ɪ/, in an experiment with native speakers of Catalan, a language that do not have temporal contrasts in its vowel system. Results showed that both experienced and inexperienced Catalan participants relied significantly on duration, confirming theories that posit the availability of duration as a cue in perception, even for speakers of languages which do not use duration contrastively.

The studies of Grenon et al. (2019) and Ylinen et al. (2009) aimed at a shift from a reliance on durational cues to a reliance in spectral cues. The participants were, respectively, Japanese and Finnish native speakers, languages with contrastive vowel length. In both cases the L2 was English. After a perceptual training program, the intended shift was achieved, suggesting that changes in cue-weighting are possible.[12]

The last issue we address in this section is *individual variability.* Individual differences include a wide range of factors, related to personality (e.g., learning motivation, learning styles, anxiety), cognitive abilities (e.g., phonological awareness, auditory acuity, musical ability, working auditory memory) and factors related with L2 experience (e.g., length of residence, language experience, knowledge of other L2). Although in this dissertation we do not aim at investigating how individual differences contribute to differences in the results, numerous authors have point out the importance of looking at individual results instead of assuming participants as a homogeneous group. For example, in a perceptual training study with Greek native speakers, Lengeris & Hazan (2010) found a

[12] These two studies will be reported with more detail in section 1.3.2.

clear effect of L1 across participants. However, subjects who presented higher levels of frequency acuity in the discrimination of synthetic vowels in L1 and L2 and nonspeech stimuli also presented higher scores on the identification of natural L2 vowels, not only after training but also before. Mayr & Escudero (2010) point out to an important aspect: individual variability can also entail different learning paths for each learner. Accordingly, investigating perception at individual level entails some methodological aspects. For example, Tyler et al. (2014) found significant differences between individual performances in a perceptual identification task. Consequently, the authors decided to establish assimilation patterns for each participant and test discrimination predictions based on those individual patterns rather than on group averages. Elvin et al. (2021) also account for individual differences in the L1 when predicting behavior in L2 speech perception. Specifically, the authors collected production data for each participant, and established individual L1-L2 acoustic comparisons. Categorization and discrimination predictions were also tested based on those individual comparisons. Interestingly, the authors did not found evidence for the benefit of a methodology accounting for individual differences in favor of group averages.[13]

1.3. L2 vowel perception and acquisition

We start this section by presenting models on L1-L2 cross-linguistic perception. The reason to introduce here these theorical frameworks is two-fold. First, these models were designed and tested mainly for perception and acquisition of speech segments (vowels and consonants), and are rarely mentioned in the context of perception and acquisition of suprasegmental features. Second, along with the characterization of the models, we will present some studies on L2 vowel perception, that support the models' theorical assumptions, while providing us with practical information on methodologies and analysis of the results. These studies, however, mainly observe speakers on a specific moment of the learning process (e.g., early or advanced proficiency levels) and even when they approach the acquisition process, the authors compare speakers at different points in the

[13] This study will be reported in detail in the next section.

course of learning. Therefore, in section 1.3.2, we will report studies that investigate the effect of interventions in the acquisition of L2 vowels.

1.3.1. Cross-language models on L2 perception

Although the first theories on L2 acquisition date back to the fifties (VanPatten & Williams, 2015), it is only after the mid-nineties that we observe a special interest in L2 perception, with the emergence of different theorical frameworks, namely, the *Phonological Interference Model* (Brown, 2000) the *Ontogeny Phylogeny Model* (Major, 2001), the *Speech Learning Model* (Flege, 1995), the *Perceptual Assimilation Model* (Best, 1995), and the *Second Language Perception Model* (Escudero, 2005; Escudero & Boersma, 2004). From these approaches, we will not include the first two in the description below, for the following reasons: The *Phonological Interference Model* approaches L2 perception from the phonological theory of feature geometry and focus on perception of contrasts in which one of the L2 sounds is a phoneme in the speaker's L1 (e.g., the perception of the English contrast /ɹ/-/l/ by Chinese speakers). Consequently, the acquisition of contrasts in which neither vowel is present in the L1 is not approached by the model.[14] The *Ontogeny Phylogeny Model* is focused on the more general question of changes in language (L1 included), as a result of language contact, and it does not provide specific predictions for L2 perception, therefore, it is not adequate to our study. Considering these facts, we will focus our attention on three models: the *Speech Learning Model*, the *Perceptual Assimilation Model*, and the *Second Language Perception* model.

The *Speech Learning Model* (Flege, 1995; henceforth: SLM) attempts to account for L2 perception and production in the context of naturalistic learning by sequential bilingual speakers. SLM is primarily concerned with explaining difficulties in the ultimate attainment in L2 pronunciation, and the relation between those difficulties and perception. One of the key points of SLM is that it presented an alternative to the *critical period hypothesis* theory. Based on comparisons between early and late learners, and/or with different ages of arrival, Flege (1995) proposes that even late learners have access to the mechanisms and processes for perceptual learning available for L1 acquisition, rejecting the idea that

[14] This is the case of the contrast [ɐ]-[ɨ] for Hungarian learners of L2 Portuguese.

neurocognitive aspects constrain L2 learning after puberty. Instead, SLM emphasizes the importance of phonetic input in L2 acquisition. However, although assuming that mechanisms and processes for perceptual learning are available in adulthood as well, Flege suggests that limitations arise, and the ability of L2 speakers to distinguish L1-L2 sounds may decrease as age of first exposure increases. This is due to the creation of higher-order invariants, that stabilize with the increase of age. According to SLM, features that are absent or have little importance in the L1 may be disregarded in L2 perception.

From SLM's perspective, L2 speech perception is shaped by a process of *equivalence classification* (Flege, 1987), also referred to as *interlingual identification* (Flege & Bohn, 2021), that is, L2 sounds are equated into existing L1 categories. Although this cross-linguistic assumption is also present in other models, the originality of SLM is the linguistic level on which identification takes place. According to Flege (1995), L1 and L2 categories are matched at an allophonic level. The idea that "learners perceptually relate positional allophones in the L2 to the closest positionally defined allophone in the L1" (Flege, 1995, p. 238) is grounded on the observation of variability in allophonic distribution among languages, and variability of acoustic-articulatory properties within a language. Empirical evidence also suggests that L2 learners show different levels of difficulty with the same phone when it is located in different positions in the word (initial, middle or final), reinforcing the position-sensitive allophonic assumption.

According to SLM, *equivalent classification* may block formation of new categories for L2, specifically when the L2 category is *similar* to the matched L1 category. In the event of an L2 phonetic category being similar to an L1 phonetic category, this may result in a *merged category*, or a *diaphone*. Formation of *new* categories is therefore more likely to occur to L2 sounds that are not similar to L1 categories. However, dissimilar sounds can also present difficulties, as shown in Flege et al. (1994). In this study, the authors tested Spanish native speakers' perception of English vowels. Half of the participants were experienced learners, with averages of seven or ten years of length of residence in the United States, while the other half were inexperienced learners. All the subjects had to complete a dissimilarity rating task and an oddity discrimination task.[15] As predicted by SLM, the results showed that acoustic distance between vowels is related to the perception of dissimilarity, that is, the bigger the distance between vowels, the more dissimilar they were judged by the participants. However, comparisons between performances of the inexperienced participants with that of the experienced learners showed no significant differences, which suggests that mere exposure to L2 input is not enough to promote perceptual learning.

SLM assumptions are grounded on the premise that both L1 and L2 phonetic systems coexist in a common phonetic space[16], which results in bidirectional cross-linguistic influence. From the model's perspective, in the common phonetic space, L1 and L2 phonetic categories interact in two ways: *phonetic category assimilation* and *phonetic category dissimilation*. *Assimilation* is due to the development of a merged or composite category, a consequence of similarity between L1 and L2 categories. In this case, production of the L2 sound approaches L1-like quality while the matched L1 sound may became L2-like. In the case of *phonetic category dissimilation*, which occurs when a new category is created near an existing L1 category in the acoustic vowel space, neither the L1 nor the L2 categories will be produced as monolinguals produce them. For example, in Flege (1987), French-English bilingual speakers exhibited a voice onset time in /t/ that did not correspond neither to their L1 nor the L2.

The *Revised Speech Learning Model* (Flege & Bohn, 2021; henceforth: SLM-r) aims to adapt SLM having in consideration empirical evidence gathered since the original model was proposed. While continuing to focus on the context of naturalistic learning by speakers who are acquiring an L2 after the establishment of the L1, the SLM-r defines a new direction for future research. One point of review is the previous perspective of perception as preceding production, and that "perception places an upper limit on the accuracy with which L2 sounds are produced" (Flege & Bohn, 2021, p. 28). Instead of these claims, SLM-r

[15] In the dissimilarity rating task, participants were presented with two L1 vowels, or two L2 vowels, or one L1 vowel and one L2 vowel, and had to rate dissimilarity between the vowels in each pair, in a scale from 1 to 9 (1 = very similar, 9 = very dissimilar).

[16] "Bilinguals strive to maintain contrast between L1 and L2 phonetic categories, which exist in a common phonological space" (Flege, 1995, p. 239). While SLM used the expression *common phonological space*, Flege & Bohn (2021) classify it as a misnomer, adopting instead the term *common phonetic space*.

proposes a bidirectional perception-production dynamic. This idea is based on evidence from more recent studies, which show, for example, that the brain areas involved in speech perception and in speech production overlap and are both activated during perception or production.

In addition, in SLM-r, the focus of SLM's target population – highly experience learners who struggle to achieve the ultimate attainment in L2 production – is also broadened, given the evidence that "it is virtually impossible for L2 learners to produce and perceive an L2 sound exactly like mature monolingual native speakers of the target L2" (Flege & Bohn, 2021, p. 25). Comparisons between early and late learners are therefore judged to be of no interest, since studies show that early learners can also differ from native speakers.

SLM-r also introduces the *category precision hypothesis*. The authors define *category precision* "as the variability of acoustic dimensions measured in multiple productions of a phonetic category" (Flege & Bohn, 2021, p. 36). Category precision is related to intercategory distances, but also to individuals' endogenous factors, such as auditory acuity. According to the *category precision hypothesis*, the level of precision of the L1 categories at the time of first exposure to the L2 have an effect on the formation of *new* categories. Specifically, the better defined the L1 categories are, the easier it is to discern L1 and L2 sounds and create *new* categories for the L2 sounds. In line with this, the authors emphasize the importance of investigating intersubject variability in the L1 phonetic categories to understand and predict how subjects will map L2 sounds.

Finally, SLM-r rejects the original model's assumption that features not exploited in the L1 are also disregarded in L2. This is based on empirical evidence of changes in cue weighting, and use of features in L2 which are redundant in the L1, even immediately after perceptual training. Considering that cue weighting changes are available for both early and late learners, SLM-r adopts the *full access hypothesis*, which states that adult learners have access to the same learning mechanism available for L1 (Flege 2005; Escudero & Boersma, 2004).

The *Perceptual Assimilation Model* (Best, 1995; henceforth: PAM) approaches cross-linguistic perception from the epistemological perspective of direct-realism, that is, from the principle that humans perceive directly their environment, and not through mental constructs. According to PAM, during L1 acquisition, individual articulatory gestures are attuned into higher-order invariants, or gestural constellations. Perceptual L1 learning is "to discover the optimal gestural invariants that constitute native language structures" (Best, 1995, p. 185). This has, however, implications for L2 perceptual learning, since listeners may have difficulty in perceiving non-native sounds that do not fit their L1 gestural invariants.

PAM is based on three assumptions. First, human gestural possibilities are limited, therefore we can assume that there will be overlapping situations between different languages, that is, a certain pair of L1 and L2 usually share a certain number of phonetic categories. Second, perception of non-native sounds is constrained by similarities and dissimilarities between the L1 and the L2. Third, these similarities/dissimilarities consist of similarities/differences in articulatory gestures and constellations.

Best (1995) posits that one of three circumstances occurs when perceiving a non-native sound: the sound can be perceived as a speech sound and, in this case, is said to be *assimilated*, or the sound is not perceived as a speech sound and, in this case, is said to be *not assimilated*. In the case of assimilation, two further situations are possible. If the non-native sound is assimilated to a specific native category, it is designated as *categorized*. If, however, the non-native sound is not matched with a specific L1 category it is deemed to be *uncategorized*.

Whereas SLM is focused on the perception of non-native segments, PAM observes perception of contrasts in the L2. If a non-native contrast is perceived as a contrast (i.e., two different categories), the listener does not present perceptual problems. However, if there is no perception of contrast, no phonological distinction is made, and discrimination can pose difficulties to the listener. Best (1995) outlines five *L2 assimilation types*, resulting from the combination of *categorized* and *uncategorized* sounds. For each of these types, a prediction for discrimination is also proposed:

(i) When each sound of the non-native contrast is assimilated to different L1 categories, the type of assimilation is termed as *two-category assimilation* (*TC*). In this case, discrimination is expected to be excellent.

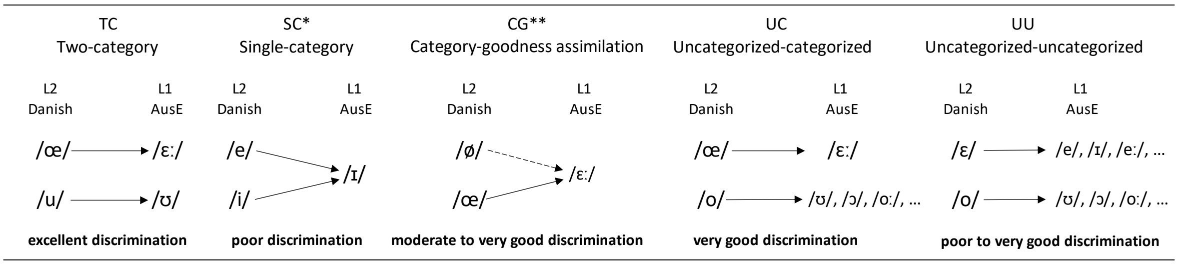
(ii) If both sounds of the L2 contrast are assimilated to one single L1 category and both are perceived as an equally deviant, good, or very good fit to the L1 category, the type of assimilation is termed as *single-category assimilation* (*SC*). In this case, discrimination is predicted to be problematic.

(iii) If both sounds of the non-native contrast are assimilated to a unique L1 category, but one of them is perceived as a better exemplar to the L1 category, the type of assimilation is termed as *category-goodness assimilation* (*CG*). In this case, discrimination can range from moderate to very good.

(iv) If one of the two sounds in the L2 contrast is assimilated into an L1 category and the other is uncategorized, the type of assimilation is termed as *uncategorized-categorized assimilation* (*UC*). in this case, discrimination is expected to be very good.

(v) When both sounds of the L2 contrast are uncategorized, the type of assimilation is termed as *uncategorized-uncategorized assimilation* (*UU*). In this case, discrimination can range from poor to very good, depending on the similarity/dissimilarity between the L2 categories, and between these and the L1 categories at hand.

Figure 1.4 (next page) illustrates the five assimilation types. The examples are retrieved from the study of Faris et al. (2018), which tested native Australian English speakers on perception of Danish vowels and diphthongs.



*In the case of the contrast /e;i/, a t-test revealed both segments equally fitted the AusE /ɪ/.

**Although /ø/-/œ/ are both categorised as the AusE /ɛː/, a t-test showed that /œ/ was perceived as a better exemplar than /ø/ for the L1 /ɛː/.

Figure 1.4. Examples of assimilation types, according to PAM, based on Faris et al. (2018)

Tyler et al. (2014) conducted a study with university students who were native speakers of American English, and the results were consistent with the assimilation types and predictions outlined in PAM. In this study, participants were tested in vowel contrasts from three L2, produced by female speakers of each language: Norwegian (/i/-/y/ and /i/-/ʉ/), Thai (/ɯ/-/ɤ/) and French (/o/-/õ/, /ø/-/œ/, and /y/-/ø/). The experiment compared the results from two tasks: identification and discrimination. In the identification task, the American listeners were asked to match the non-native sounds to American English real words, and to rate similarity on a scale from 1 to 5 (1 = *unlike*, 5 = *identical*). Based on the results of the identification task, and considering a categorization threshold of 70%, the target non-native sounds were classified as *categorized* or *uncategorized*. The authors then determined the assimilation types, following PAM. Most of the cases fell into the category of TC, UC or UU assimilation types, some CG were also identified, and only one contrast was classified a SC assimilation. Confirmation of PAM predictions for discrimination performances were found when the authors crossed the assimilation types with discrimination, tested by means of an AXB task. Discrimination was excellent for TC and UC contrasts, good to very good in CG contrasts, and poor for SC contrasts. UU assimilations were excluded from the analysis due to the uncertainty of predictions for this scenario. Final results were thus (from better to poorly discriminated): TC/UC > CG > SC.

Considering the number of uncategorized vowels in the above study, Faris et al. (2016) proposed a subclassification for uncategorized cases: *focalized*, *clustered* and *dispersed.* In Figure 1.5 (next page), examples of these situations are presented, using results from the study of Faris et al. (2016) with native speakers of Egyptian Arabic and

Australian English as L2. A *focalized* situation occurs when the non-native sound is assimilated into mainly one L1 category, but below the categorization threshold. In the *clustered* situation, the listeners choose a small set of L1 categories in the L2 sound assimilation, all being below the categorization threshold but above chance level. If listeners identify the L2 sound as diverse L1 categories below chance level, the assimilation is said to be *dispersed*.

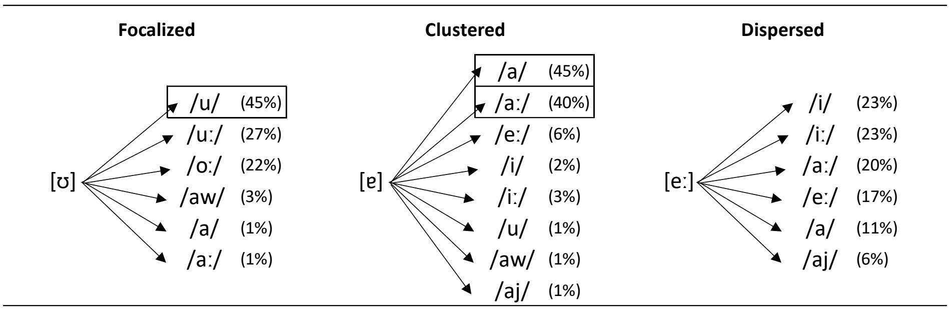


Figure 1.5. Examples of *focalized*, *clustered* and *dispersed assimilations,* based on Faris et al. (2016). Rectangles mark assimilations below the 50% categorization threshold established by the authors, but significantly above chance level

In a series of experiments, Faris et al. (2018) observed uncategorized assimilation types (UC and UU), comparing discrimination of these cases with discrimination of categorized assimilation types (TC, CG, SC). Recall PAM's discrimination prediction: TC assimilations will be the better discriminated, followed by UC and CG assimilations. SC assimilations are expected to be the most problematic in discrimination. As for UU, no specific prediction is established.

Considering the probability of perceptual overlap cases in uncategorized phones assimilated as focalized or cluster, the authors considered that parameter – perceptual overlap – to predict discrimination for these cases. In their study, the authors tested perception of Danish monophthongs and diphthongs by naïve Australian English native speakers.[17] Participants completed forced-choice identification tasks and AXB discrimination tasks. Results revealed that UU contrasts with no overlap were the most

[17] *Naïve listeners* are speakers who have no contact with the L2.

accurately discriminated, followed by TC and UC contrasts with no overlap observed, and then by UU contrasts with partial overlap. Although the results were overall in accordance with PAM's predictions, UC contrasts with partial overlap and SC contrasts were equally poorly discriminated, contrary to the predictions. The authors suggest that this result was due to the lenient 50% categorization criterion. If a 70% threshold had been used, for example, UC contrasts may had been classified also as a SC. With this study, the authors concluded that perceived phonological overlap is a reliable predictor for discrimination difficulties: discrimination was very good to excellent for contrasts with absence of overlap and poor to moderated for contrasts where a partial overlap was observed.[18] Additionally, the authors raise the question of the arbitrariness of thresholds for assimilation classification purposes. Although aware of the theoretical consequences that disregarding categorization thresholds would lead to, the authors question the adequacy of using such methods to establish if the L2 sound is assimilated or not to the L1 system.

PAM was originally developed to describe perception of non-native sounds by naïve listeners. However, Best & Tyler (2007) proposed an adaptation of the model to L2 context (PAM-L2), extending PAM's discrimination predictions to L2 acquisition, while discussing SLM's assumptions. Although Best & Tyler agree with Flege's postulate of the preservation of L1 acquisition's mechanisms during adulthood, their theory diverge from Flege's on the nature of speech sound representations. Whereas SLM approach is that learners create categories from acoustic-phonetic cues, PAM-L2 is in line with PAM on the assumption that learning involves identifying higher-order invariants in the L2 speech, and that listeners draw their cues by focusing, separately or combined, on the phonetic, phonological, and gestural levels.

To illustrated how PAM can be extended to the L2 context, Best & Tyler describe four possible situations that will occur during L2 learning:

(i) If only one L2 category is assimilated with an L1 category (*two-category* or *categorized-uncategorized* assimilations), L2 listeners do not show discrimination problems. Being this the case, shifts in the L2 phonetic realization

[18] No contrasts with complete overlap were found in this study.

are unlikely to occur, unless one of the phones is considered a deviant realization of the L1 category, in which case, phonetic dissimilation is observable.

(ii) In the case of a *category-goodness* assimilation, PAM-L2 predicts the possibility of the learner creating a new phonetic and phonological category for the more deviant L2 phone. Best & Tyler proposed that the creation of this new category occurs with exposure to the L2, which allows the learner to discern contrasts that are lexically functional.

(iii) When both L2 phones equally fit a single L2 category (*single-category* assimilation) –, learners struggle with discrimination. PAM-L2 predicts that, in this case, formation of new categories depends on the frequency and/or phonological neighborhood density.[19] As the authors explain, if the L2 phones occur in several words, and their discrimination is relevant for communication, this will motivate learning of the perceptual differences. If the contrary is observed, L2 contrasting words remain homophonous, without prejudice for the learner.

(iv) Finally, in the case of *uncategorized-uncategorized* assimilations (when the listener identifies both L2 contrasting phones as several L1 categories), learning depends on overlapping probabilities. If there is a distance between the L2 phones in the phonological space, then creation of new categories should not pose problems. However, if that distance is small and the L2 sounds are assimilated to the same L1 categories, then learners will struggle with perceiving lexical-functional differences. In this case L2 learning can also depend on frequency and/or phonological neighborhood density, but it is also possible that difficulties remain.

[19] "*[P]honological neighborhood* density refers to the number of words that differ from a target word by a single phoneme." (Munson & Solomon, 2004, p. 1049)

Considering that PAM and PAM-L2 focus on discrimination of minimally contrasting lexical items, L2 access to vocabulary will be paramount to perceptual learning and tuning. Both SLM and PAM-L2 refer to input as an essential factor to L2 perceptual learning, but while Flege's model is focused on the phonetic tuning, PAM-L2 claims that interaction between the phonetic and the phonological levels is the key for perceptual learning, as "establishing lexical items in the target language is likely to exert forceful linguistic pressure for the L2 learner to "re-phonologize" perception of the target contrasts" (Best & Tyler, 2007, p. 33).

A longitudinal study with Japanese learners of Australian English by Bundgaard-Nielsen et al. (2011) showed evidence of positive and negative relation between vocabulary size and discrimination accuracy. The authors propose the *Vocabulary-Tuning Model of L2 Rephonologization*, according to which an increase in vocabulary size may benefit learning or tuning for some contrasts (e.g., UU or UC) but may inhibit for other cases (e.g., SC). In the study, students were divided into two groups, according to their vocabulary size in L2, and were tested in two moments. In the first testing moment, they completed an identification task with goodness rating, which was followed by an AXB discrimination task. In the second testing moment, about six months after the first, students repeated the same tasks. The goal was to observe if perceptual learning/tunning occurred after some time of exposure to L2. Results showed that the group with the greater vocabulary performed better in both tasks, and in both testing moments, supporting the assumption that vocabulary is a relevant factor in the reattunement of the perception of L2 sounds. However, cross-analysis between the two testing moments also showed that some effect of L1, observed in the first testing moment, persisted. This suggests that while some segments may be easier to attune, perception of others is unlikely to change with the acquisition of more vocabulary, even in context of immersion in the L2. According to the authors, this implies that if learners reach a level of perceptual discrimination that allows them to function in their environment, changes in the L2 perception will slow down, even if they are immersed in the L2. Additionally, no improvement was detected for the TC contrast's discrimination, supporting PAM-L2 prediction that changes in discrimination of this type of contrast are unlikely to occur. SC contrast remained at chance level (about 50%), confirming the greater discrimination difficulty in this type of contrast.

The *Second Language Linguistic Perception* model (Escudero, 2005; Escudero & Boersma, 2004, henceforth: L2LP model) is an alternative theorical framework to SLM and PAM, that considers both cross-linguistic as well developmental factors. The L2LP model goes beyond the question of speech perception/sounds categorization, focusing also on the difficulties and learning tasks speakers face when acquiring an L2. Therefore, contrary to SLM and PAM, this model was originally designed for L2 acquisition.

The L2LP model is grounded on several assumptions. First, native listeners categorize sounds based on the maximum-likelihood that those sounds will match the categories intended by the speaker – the *optimal perception hypothesis*. L1 speakers are optimal perceivers, and the aim is that L2 learners also become optimal perceivers. Second, the LPL2 separates *representations* from *grammars*, both for L1 as for L2 speech perception. It proposes the existence of two levels for each: *phonological* and *lexical representations* and *perception* and *recognition grammars*. (Figure 1.6).

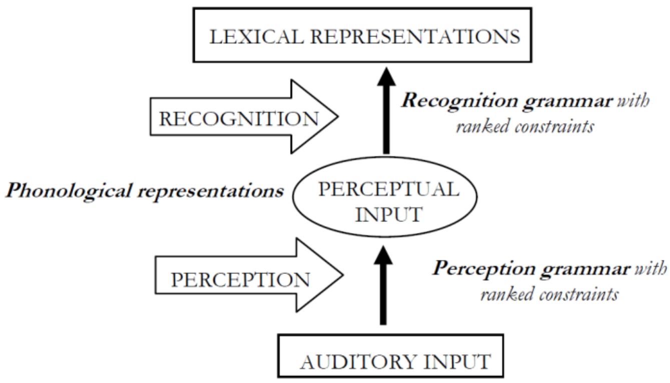


Figure 1.6. Levels in speech perception (Escudero, 2005, p. 43)

In the L2LP model it is proposed that perceptual categories are created through *auditory-driven perceptual learning*. The perception grammar contains constraint rankings, created based on production distributions, which lead to optimal categorization of the auditory input. Perceptual categories will become optimal with *lexicon-driven perceptual* *learning* and optimal lexical storage will be accomplished through *message-driven recognition learning*.

At the onset of L2 acquisition, learners *copy* their L1 perceptual grammar, as well as the L1 perceptual categories: this constitutes the *full copying* assumption of the model. As for the acquisition process, similarly to SLM and PAM, the L2LP model stipulates that L2 speakers have access to the same mechanisms used for L1 acquisition: the *full access* assumption. According to the L2LP model, the learning processes such as category formation, category boundary shifting, and recognition learning are driven by the *gradual learning algorithm*, "a general learning device, innate and blind, that acts upon different kinds of input" (Escudero, 2005, p. 68). This learning instrument is, therefore, responsible for the creation and change of the constraint rankings in the perceptual grammar. Finally, the L2LP model posits the possibility of *full proficiency* in L2 along with full proficiency in L1.

This last premise is related to one of the central assumptions of the L2LP model, which is the separation of the L1 and L2 systems. This view is different from SLM's 4th postulate, that states the coexistence of L1 and L2 phonetic categories in a common phonological space (Flege, 1995). Flege's assumption entails one important consequence, predicted in his model, which is that L1 is subject to changes during L2 acquisition. Escudero (2005) points out that in the case of a common phonological space, the possibility of maintaining optimal proficiency in both languages will be low. Contrarily, the L2LP model proposes that the L1 and L2 have separated perceptual grammars. However, the author combines this dual system with the proposal of *language modes* by Grosjean (2001).[20] Depending on the context or on certain parameters (e.g., language used in the instructions of the experiment), a parallel activation of both perceptual grammars may occur. In such cases, perception can result in intermediate responses, as a consequence of temporary merging of the perception grammars at the moment of the categorization. Consequently, Escudero (2005) predicts that, as long as both languages are regularly used, the learner can

[20] Grosjean defines *language mode* as "the state of activation of the bilingual's languages and language processing mechanisms at a given point in time" (2001, p.3). Language mode occurs in a continuum, from L1 monolingual mode to L2 monolingual mode, through L1-L2 bilingual mode. Activation of the modes is situational and depends on the L2 experience.

reach optimal L2 perception *(full proficiency)* while maintaining the stability of L1. Evidence of the effect of language modes in the perception of L2 contrasts is presented by Yazawa et al. (2020). The authors observed the relative weighting of vowel duration and spectral cues in the perception of the English contrast /iː/-/ɪ/, by Japanese learners. Language modes were manipulated by conducting two sessions with the same stimuli: one with instructions in Japanese and the other with instructions in English. In the first session, participants were informed that they would hear sounds from Japanese, and in the second session, sounds from English. Stimuli, the same in both sessions, consisted of synthetic productions of target L2 vowels, with different duration and spectral values. Results of the experiment were in line with the L2LP model language modes for L2 perception: while participants relied more on duration in the Japanese session, in the English session they relied significantly less on duration and more on spectral cues. This is an indicator of activation of different language modes, as in English spectral cues are more heavily weighted, while Japanese uses mainly duration cues.[21]

Considering that L2 learners initiate development in the L2 phonology by copying their L1 categories and grammars, a detailed description of the L1 perceptual system is necessary to understand and predict L2 learning problems and respective learning tasks. A thorough acoustic comparison of L1 and L2 sounds at issue is, therefore, in order, and studies have shown that this method can successfully predict L2 categorization and discrimination. Elvin et al. (2014) tested the perception of six Brazilian Portuguese vowel contrasts by naïve Australian English and Iberian Spanish listeners. The aim was to test if acoustic similarity is indeed the most reliable factor to explain non-native perception. In their study, Elvin et al. (2014) compare acoustic similarity with vowel inventory size Considering that the Spanish vowel system includes only five oral vowels, and Australian English 12, if a bigger inventory size is more beneficial for perception, then Australian speakers have better performance than that of Iberian Spanish speakers. If, however, acoustic similarity is more favorable to L2 perception, then Spanish speakers outperform Australian speakers, since their vowel categories are closer to the Brazilian categories (Figure 1.7, next page).

[21] However, the authors point out the considerable individual variability in the results.

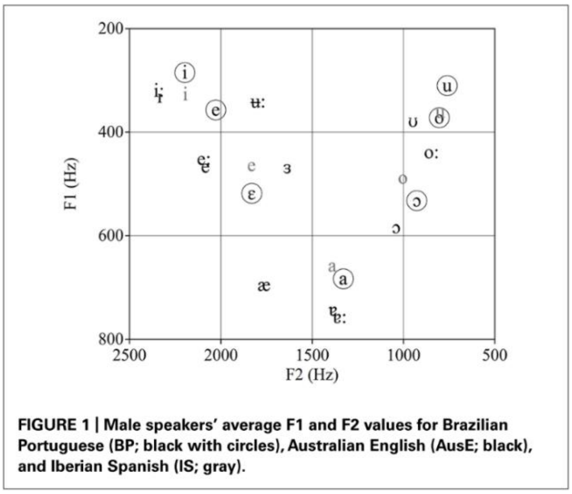


Figure 1.7. Vowel space map for Brazilian Portuguese, Australian English, and Iberian Spanish (Elvin et al., 2014, p. 3)

Another key-aspect when it comes to predicting the learning tasks L2 learners face is the distinction between representations (*perceptual* and *lexical*) and grammars (*perceptual* and *grammar*), previously mentioned. This separation allows us to distinguish *perceptual learning tasks* from *representational learning tasks*. In certain cases, it is possible that the learner has only a perceptual task, whereas in other cases, acquisition can be a more complex process, involving both perceptual and representational tasks. With this in consideration, the L2LP model proposes three *learning scenarios*: *new*, *similar*, and *subset*.

In the *new scenario*, learners map L2 categories into fewer L1 categories. This is the case of the English contrast /i/-/ɪ/, perceived by speakers of Spanish as /i/ (Escudero, 2005). In this case, it is possible that the learner will use the L1 phoneme /i/ for the lexical representations of *ship* and *sheep*, for example. This means that in this scenario, the L2 learner has perceptual and representational tasks. First, he/she must create a new category (or split an existing category), after which an adjustment of the category boundaries takes place, in order to reach optimal perception. This is achieved with the supervision of the lexicon. According to the L2LP model, this scenario is the most difficult, since it involves tasks of different nature: auditory-driven category formation and lexicon-driven category boundary shifting. According to Elvin et al. (2021), the higher level of difficulty for *new scenarios* is also posited by PAM, since in this model the *new scenario* corresponds to a *single-category assimilation*, to which PAM predicts more discrimination problems will occur. However, this is not the point of view of SLM, which states that *new scenarios* are easier.

In the *similar scenario*, learners start with an identical number of perceived categories, as in the case of the Canadian French contrast /æ/-/ɛ/, perceived as /æ/-/ɛ/ by Canadian English speakers (Escudero & Vasiliev, 2011). In this case, learners only have to adjust their perceptual grammar, by shifting the categories' boundaries, so that these increasingly match the optimal location. Elvin et al. (2021) explain that the L2LP model's *similar scenario* corresponds to PAM's *two-category assimilation*, and in both models, it is predicted to be less problematic than a *new scenario*/*single-category assimilation*. This contradicts SLM's proposal, in which the *similar scenario* is unlikely to undergo changes.

Finally, in the *subset scenario*, learners initially match more L1 categories to fewer L2 sounds. This is the example of Dutch learners when using three Dutch vowels, /i/, /ɪ/ and /ɛ/, to categorize the Spanish contrast /i/-/e/ (Escudero & Boersma, 2002). In this case, the learner must reduce the number of perceptual categories and of lexical categories. Therefore, the *subset scenario* also involves perceptual and representational learning tasks. However, differently than for the *new scenario*, this "occurs through the parallel adjustment of the recognition and perception grammars. Learning starts when recognition has to change due to a semantic-driven error, and this in turn triggers the recognition-perception mismatch that is needed to get lexicon-driven perceptual learning started" (Escudero, 2005, p. 312). In PAM, the *subset scenario* can be matched to *focalized*, *clustered* or *dispersed uncategorized assimilations*. As we previously explained, difficulties in *uncategorized assimilations* can be better predicted if we take into account the level of perceptual overlap. Considering the variety of *subset* cases that may emerge, Elvin et al. (2021) proposed two *subset scenarios*: *subset easy* and *subset difficult.* For the latter, the authors suggest that it should be regarded as presenting almost the same difficulty and the same tasks as the *new scenario*.

Figure 1.8 (next page) displays examples for each learning scenario, based on the results of Elvin et al. (2021).

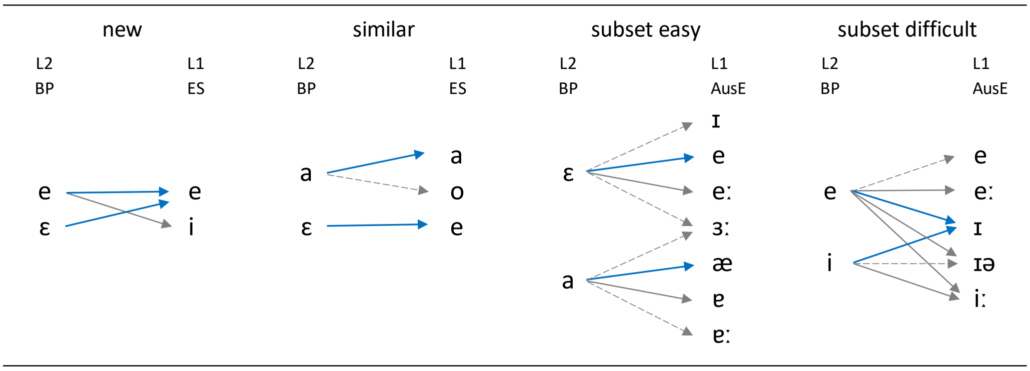


Figure 1.8. L2LP learning scenarios, based on Elvin et al. (2021). Blue arrows indicate top-ranked choice categorizations, grey continuous arrows indicate categorizations above chance-level, and grey dashed arrows indicate categorizations below chance-level

Elvin et al. (2021) conducted a perceptual study with Australian English native speakers and European Spanish native speakers, presented with Brazilian Portuguese oral vowels as L2 input. The study pursued three goals. First, confirming the results of the study previously described, by Elvin et al. (2014). Second, the authors aimed at investigating the better method to predict the learning scenarios. Specifically, they compared two methods of calculating overlap: *acoustic overlap*, calculated from acoustic L1-L2 comparisons, and *perceptual overlap*, calculated from the categorization results of the force-choice identification task the participants completed. Third, they compared individual results with group average results, in order to test the importance of individual learning paths, posited by the L2LP model. To observe individual variability, and following the L2LP principles, the authors established individual L1-L2 acoustic comparisons. To this purpose, production data were collected for each participant.

Once production data were collected, the authors established individual L1-L2 acoustic comparisons and, based on these, they also established predictions for categorizations. To test these predictions, the authors conducted a forced-choice identification task, in which participants had to match the seven Brazilian Portuguese vowels, inserted in /fVfe/ nonwords, with L1 vowels. Australian participants were given 13 English real words, while Spanish listeners, who have a transparent orthography, were simply presented with a list of their five vowels. Results collected for the Spanish listeners strongly matched with the predictions based on the cross-linguistic acoustic comparisons.

As for Australian listeners, however, authors found disparities between predictions and results, in some of the participants. The authors offer different explanations for these results, such as the possibility that Australian listeners could have been influenced by orthography or by the higher level of response possibilities, since their L1 has considerably more vowel categories than Spanish. Moreover, factors such as cue weighing could have also affected the results of the Australian participants.

With the data collected from the L1-L2 acoustic comparisons, the identification task and the discrimination task, the authors were then able to investigate the most effective method to predict discrimination difficulties – the *acoustic overlap* parameter or the *perceptual overlap* parameter. For both overlap parameters, contrasts with higher overlap values were predicted to be more difficult to discriminate (e.g., /o/-/u/ for both Australian and Spanish listeners) and, conversely, contrasts with lower overlap values were predicted to be easier (e.g., /a/-/ɛ/, also for all participants). However, in some cases the overlap values were very different between the two groups, and in this case, specific discrimination difficulties were predicted for each L1 (e.g., Australian listeners should have difficulties with the contrast /a/-/ɔ/ but not Spanish participants). Analysis of the results confirmed that both methods – acoustic overlap and perceptual overlap – were reliable in predicting discrimination accuracy/difficulties. As for establishing the best parameter, results were conflicting: while acoustic overlap scores were a better predictor in the case of Australian listeners, Spanish participants' results were more in line with the perceptual overlap values. The authors suggest that due to the Spanish smaller vowel inventory, perceptual overlap could be a better predictor in this case, and that for Australian speakers, acoustic predictions may be more appropriate, since, as previously mentioned, there were some deviant results in the categorization task for these participants.

Additionally, the authors aimed at investigating if individual overlap scores were better at predicting discrimination accuracy than group averages. Interestingly, the results showed that group averages were better predictors, especially in the case of the Spanish group. Such results go against the L2LP model assumption. As explanation, the authors suggest that this is due to statistical analysis, since in group averages' outlier results have a weaker effect than at individual level analysis.

Finally, the study confirmed the results of Elvin et al. (2014), on acoustic similarity versus vowel inventory size, since the richer vowel inventory of Australian listeners (13 vowels), when comparing with the Spanish listeners who only have five vowels available, did not constitute an advantage. As the authors point out, the discrimination results were more in line with the different scenarios (predicted by a comprehensive acoustic L1-L2 comparison) than with the vowel inventory size.

1.3.2. Perceptual training in L2 vowel acquisition

Although L2 phonetic training is a relatively recent area of research, it has been the focus of a large number of studies in the past few years. The review presented in this section will include a detailed analysis of methods that aim mostly at L2 vowel acquisition, and will focus mainly on methods and training effectiveness for improvement in vowel perception. Furthermore, except for Ylinen et al. (2009), the studies reported here concern late L2 learners, often university students, which is in line with the context of the experiments conducted and described in the present dissertation. In addition, the selected studies adopt the HVPT paradigm and include natural stimuli, also similarly to what was used in this dissertation. Finally, and again except for the study of Ylinen et al. (2009), the studies described in this section provide explicit phonetic or phonological instructions to the participants, although to different extents: some include articulatory-visual explanations, while others only include written instructions on the importance of the trained L2 sounds. All studies included feedback in the training trials. It must be mentioned that some of the studies aim not only at improving perception, but also production, and others observe vowels and consonants. However, since neither production nor consonant perception are within the scope of this study, these aspects will be mentioned only briefly. The last two studies reported in this section – Oliveira (2006) and Castelo & Santos (2017) – do not follow the selection criteria listed above, since they do not involve perceptual training. However, since they are the only studies on acquisition of EP oral vowels by L2 learners, we judged important to included them also in this literature review.

The effect of a perceptual and production training on English as L2 was tested by Aliaga-García & Mora (2009), with 29 advanced bilingual Catalan-Spanish learners of

English Philology at the University of Barcelona. Of these speakers, eighteen received training, the remaining eleven were used as controls. The target L2 sounds were two English vowel contrasts (/iː/-/ɪ/ and /æ/-/ʌ/) and two English consonant contrasts (/p/-/b/ and /t/-/d/). Training consisted of six two-hour training sessions (one hour for vowels and one hour for consonants), and lasted six days. Each session included a theoretical component, with explicit instructions (articulatory-visual explanations, L1-L2 comparisons) and exposure to native productions, followed by different perception and production tasks. Perceptual training included discrimination and identification tasks, as well as phonetic transcriptions exercises. Sessions were conducted in groups and were complemented with short individual sessions. L2 vowel perception/acquisition was tested by means of an AX discrimination task, which included the target vowels inserted in monosyllabic words in different contexts (CVC, CVCC, CCVC and CCVCC).[22] Comparisons of pretest and post-test results showed a significant positive effect of the phonetic training, since the trained group presented significant improvement, while the control group did not. The improvement was similar for both vowel contrasts. However, comparisons with the baseline native group showed that performance of the Catalan-Spanish learners remained nonnative-like.

Rato (2013) conducted a five-week training program with 34 Portuguese learners of L2 English, to investigate the effects of perceptual training in L2 perception and production. The target contrasts consisted of the pairs /i/-/ɪ/, /ɛ/-/æ/, and /ʊ/-/u/. The vowels were inserted in CVC context and produced by different native speakers of American English, to introduce variability. Participants were undergraduate students attending an English Phonetics and Phonology course, in University of Minho, and were divided into two groups: the experimental group (n = 22) trained the target vowels, while the control group (n = 12) trained consonants. Improvement was analyzed by comparing results of four testing moments: pretest, post-test, generalization test, and delayed post-test.[23] Perception was tested by means of a seven-alternative forced-choice identification task with goodness-rating. Perceptual training combined articulatory-visual instructions, followed by discrimination AX and ABX tasks and after that, identification

[22] In AX tasks two tokens are presented and participants must decide if they are different or the same.
[23] Generalization tests are conducted usually in the same session as post-tests and delayed tests (also known as retention tests) are conducted sometime after. Section 5.1 provides more details on these types of tests.

tasks. Tasks were presented with increasing difficulty from the 1st to the 5th session. Each session lasted about 45 minutes and was conducted in group, in a computer lab. Results confirmed the positive effect of training in perception, as well as production. Although the author found positive evidence of long-term and delayed effects, generalization was only partial, since it was not observed for the back vowels. According to the author, this was due to difficulties with the vowel /ʊ/. Analysis of results at individual level showed a considerable variability in the ability of correct identification of the L2 sounds. While some participants achieved native-like or near native-like perceptual mapping, others presented more difficulties and less improvement. Overall, the author concluded that high variability perceptual training has a positive effect on acquisition.

The question of which training tasks are more effective – identification tasks or categorical discrimination tasks – was the aim of the phonetic training study of Carlet (2017). To this purpose, the author observed 63 Catalan-Spanish bilingual graduate students, enrolled in an English Phonetics and Phonology course, at Universitat Autònoma de Barcelona. The L2 target sounds were the English vowels /i/, /ɪ/, /æ/, /ʌ/ and /ɜː/, and English initial and final stop consonants. Participants were divided into five groups. Two groups were trained with identification tasks (one training aimed at vowels and the other at consonants), two groups trained with discrimination tasks (here also, one training targeted vowels and the other consonants), and one group, the control group, performed phonetic transcription exercises, so that they would be exposed to similar instructions as the experimental groups, but without receiving perceptual training. Other than investigating the best method for L2 segmental acquisition (identification or discrimination tasks), the author aimed at addressing the issue of *training attention*. To this purpose, stimuli were identical for the groups training consonants and the groups training vowels, but contrasts were distributed differently. For example, the contrasts *vap-vab* and *vup-vub*, presented to the groups who trained consonants, explicitly aimed to the contrast /b/-/p/, while exposing the learners to the vowels /æ/ and /ʌ/. On the other hand, the contrasts *vab-vub* and *vap-vup*, used in vowel training, directed the attention to the contrast /æ/-/ʌ/, while implicitly exposing learners to /b/ and /p/.[24] From the two groups who trained with

[24] The author named this method *cross-training*.

categorical discrimination method, one explicitly trained the L2 vowels and the other explicitly trained the L2 consonants. The same principle was applied to the two groups training with identification tasks. Perception was tested with forced-choice identification tasks. Other than pretest and post-test, training effect was also evaluated in generalization trials and in a retention test. As for results, both training methods had a positive effect, since the four experimental groups' performances were significantly better than that of the control group. With respect to the most effective method, results differed for consonants and vowels. While for consonants, both discrimination tasks and identification tasks were equally efficient, for vowels, the group trained with forced-choice identification tasks outperformed the group trained with discrimination tasks. Moreover, the groups trained with identification tasks achieved better results in generalization and retention, for the L2 vowels. This suggests that different training methods may be adequate to vowel or consonant acquisition in the L2. Regarding training attention effect, only the groups who trained with discrimination tasks presented significant improvement in the perception of the unattended L2 sounds, which in turn, implies that discrimination tasks are preferable for acquisition of explicit and implicit target L2 features. A final remark, regarding the results: in the evaluation of the training program, students favored identification tasks over the discrimination tasks.

Grenon et al. (2019) investigated changes in cue-weighting. The authors observed the effect of an adaptive training, aiming the English contrast /i/-/ɪ/, with 23 Japanese learners of L2 English, from Seijo University in Japan. The training goal was to cause a shift in cue-weighting – from durational differences to spectral differences – and thus the creation of a new category. According to the authors, such train is necessary since Japanese learners rely more on durational differences than on formant frequency differences, due to the fact that in Japanese vowel length is contrastive, and is expressed primarily by duration. Training consisted of two-alternative force-choice identification tasks, using increasingly difficult stimuli. These stimuli included words with different complexity of structure, presented to participants in citation form and in carrier sentences, and produced at different speech rates. Training included eight levels of difficulty, and tasks were divided into ten sessions of 30 minutes each. Participants received training and testing in a quiet room in the university, and they could do more than one level per session, if they wished to. In general, they completed one or two sessions per week. The pretest and the post-test consisted also of two-alternative force-choice identification tasks, with tokens retrieved from training, as well as new words and sentences produced by new talkers.[25] Additionally, in the same day as the pretest and the post-test sessions, a complementary two-alternative forced-choice identification task was conducted using tokens of the real L2 contrast *beat-bit*, in which duration and formants of the vowel section were manipulated to enhance either duration or spectral features. With this task, the authors aimed at having a better insight on the effect of individual differences in cue-weighting shift ability. The analysis of the results showed that training contributed to improvement, both in tasks with trained stimuli and generalization tasks. However, as for cue-weighting, only half of the Japanese students displayed a shift from durational cues to spectral cues after training. To account for this mixed result, the authors looked into aspects such as L2 experience and age, but did not find that these factors significantly affected the scores. Another factor analyzed was length of exposure, since some participants took less time to complete the training than others. However, the results were contrary to the expectations, since participants completing the training in less time showed better results.[26] The authors concluded that the different outcomes may be related with individual aspects, such as acoustic sensitivity or musical experience.

Ylinen et al. (2009) also investigated the effect of phonetic training on changes in cue weighting, for the same English contrast, /i/-/ɪ/. In their study, the authors observed ten Finish native speakers with intermediated or advanced level of English knowledge. Similar to Japanese, in Finnish, vowel length is contrastive, and listeners of this language rely more on durational cues for the L2 contrast. The study involved both behavioral and electrophysiological (EEG) tasks, two-alternative forced-choice identification tasks and passive oddball tasks, respectively. By including this last method, the authors aimed at observing possible plastic changes in the cortex, that is, the creation of long-term memory representations for the L2 sounds. Stimuli consisted of minimal pairs of real words with the

[25] In L2 research, the term 'talker' refers to a speaker who produces tokens for the experiments' auditory stimuli.
[26] Grenon et al. (2019) suggested that maybe intensity was a more relevant factor than length of training.

target L2 contrast, recorded by male and female native speakers of English. Tokens were presented in both their original form and a manipulated form with neutralized durational differences. The modified stimuli were included to observe identification when vowel duration was ambiguous. A subset of the tokens was used for testing and training, and another subset only for training, so tests comprised tokens that were used in the training sessions and untrained tokens. Training included two-alternative forced-choice identification tasks exclusively, and consisted of ten sessions, each lasting between 20 and 25 minutes, conducted within three weeks. Results showed improvement in the identification task, both for trained and untrained words, since differences between the scores of the Finnish group and the native group in post-test were not significant, as opposed to what was observed in the pretest. Mismatch negativity (MMN) measures were in line with the identification task results, since an enhancement was observed from the pretest to the post-test, over the left hemisphere and midline sites of the brain, for speech sounds. Overall, the results demonstrated that phonetic training can effectively cause a shift in cue-weighting, from durational to spectral cues, and promote the creation of new categories in long-term memory.

The study of Oliveira (2006) aimed at determining the order of the acquisition of neutralization processes resulting from vowel reduction: changes in tongue movements in height or in place of articulation/backness/fronting. To this purpose, the author recorded productions of learners enrolled in courses of L2 Portuguese, at the University of Lisbon, with different L1. In the recordings, participants had to name 39 pairs of EP real words, presented in pictures, that included the target vowels in stressed and correspondent unstressed production (e.g., <pedra> [ˈpɛdrɐ] 'stone', <pedrinha> [pɨˈdriɲɐ] 'small stone'). Data collection was conducted in two groups (B1 and C2 levels), each with five students, and in two moments: in the beginning and immediately before the end of the semester. Based on the analysis of the data, Oliveira reported that stressed oral vowels are acquired and stabilized before the unstressed system.[27] Second, the author observed more difficulties in the acquisition of [ɨ] (as a result of neutralization of /ɛ/ or /e/) then in the acquisition of [ɐ] (resulting of the reduction of /a/), which lead to the conclusion that

[27] [ˈɔ] constituted an exception, since it was classified as ‘not acquired’ for both groups.

reduction occurring along the place of articulation/backness/fronting dimension is acquired later than reduction occurring along the height dimension.

The hierarchy in the acquisition of vowel reduction reported by Oliveira (2006) was not confirmed by Castelo & Santos (2017). In their study, the authors analyzed a written and oral corpus produced by Chinese learners of L2 Portuguese. Specifically, the authors accounted for 67% of errors involving vowel height (e.g., [e] instead of [ɛ]) against only 26% of errors involving place of articulation/backness/fronting (e.g., [ɐ] instead of [e]), and 8% in vowel heigh and place of articulation/backness/fronting (e.g., [ɨ] instead of [e]). Furthermore, 68% of errors were related to unstressed vowels, which suggests that other than problems with articulation, students also have difficulties in the mastering the rules of EP vowel reduction.[28]

1.4. L2 stress perception and acquisition

Other than a system of speech segments, languages are also distinguishable by suprasegmental or prosodic features, such as stress, tone, and intonation, that contribute to meaning. Traditionally, two main groups of languages are identified: *stress languages* and *tone languages*. With rare exceptions, stress and tone are mutually exclusive in a language (van Heuven, 2019). According to the World Atlas of Linguistic Structures (Maddieson, 2013), of the 527 languages listed in this database, 307 (around 58%) are stress languages and 220 (around 42%) are tone languages. Stress languages are found mainly in Europe, America and Oceania, while tone languages are mostly present in Africa and Asia.

At word level, stress can be defined as “nothing more than the fact that in a succession of spoken syllables or words some will be perceived as more salient or prominent than others” (Couper-Kuhlen, 1986, p. 19). Physically, word stress may be signaled by a combination of acoustic cues, such as duration, pitch and/or intensity. The greater effort invested in the pronunciation of stressed syllables using the mentioned cues, leads to the prominence of the stressed syllable in comparison with unstressed syllables.

[28] It should be mentioned that the analysis was conducted on productions of only seven learners, with different levels of L2 Portuguese knowledge.

Some languages also resort to vowel reduction in unstressed syllables, which contributes to emphasize the stressed syllable (van Heuven, 2019).

While stress features vary between languages, some characteristics can be observed across stress languages, constituting *stress universals* (Hayes, 2009; Pereira, 2020).[29] Stress is *obligatory*: a word always carries a more prominent syllable. It is also a *hierarchical* phenomenon: primary or main stress is more prominent than secondary stress, and secondary stress is more prominent than tertiary stress. Only one syllable can be the prosodic head of a word. For this reason, stress is said to be *culminative*: a word can have multiple stresses, but they all culminate in one primary stress. Additionally, a stressed syllable cannot induce stress in the neighboring syllables, meaning that stress is *not assimilative*. The word *photograph* /ˈfoʊtəˌgræf/ exemplifies these stress universals. Only the 1st syllable receives the main stress. The last syllable receives a secondary stress and [tə] is unstressed (with a schwa). Thus, this word displays a hierarchy of prominence. Similar to English, EP also admits more than one level of stress (d'Andrade & Viana, 1989). However, since the present dissertation focuses on main stress perception and acquisition, we will refer only to main word stress, using the term 'stress' for this purpose.

Languages differ in their preferred position for stress, and within stress languages, stress can occur in different positions. However, stress is always located close to the border of the stress domain, that is, stress is *delimitative*. For example, in Hungarian, stress falls on the left edge of the word (Siptár & Törkenczy, 2000) and in EP it falls on right edge of the root (Pereira, 2020), as exemplified in (1).

(1) *gabona* (Hungarian) [ˈgɒbonɒ] 'cereal'
sapato (EP) [sɐˈpatu]: sapat]$_{\text{ROOT}}$ + o]$_{\text{CLASS MARKER}}$ 'shoe'[30]

Acoustically speaking, stress is always a *relative* dimension: a syllable is perceived as stressed only by comparison with the unstressed syllable(s). For example, in the EP word *cucu* [ˈkuku] 'cuckoo', the first syllable is perceived as stressed because it is longer than the

[29] We refer to languages in which stress domain is the word. French is not included in these languages, since stress domain is the phrase. In this dissertation, we use 'stress' to refer word stress.
[30] Whenever the example word is not from English, we will specify the language between round brackets.

second. Stress is *opportunistic*, in the sense that it resorts to acoustic dimensions that are used for other phonological purposes. For instance, duration can be used at suprasegmental level, to signal stress, but also, at segmental level in languages with contrastive vowel length, as, for example, Hungarian. Additionally, the combination of features that signal the stressed syllable, duration, pitch, and intensity are also used to signal prosodic functions, such as differentiating interrogative from declarative sentences.

Stress rules and acoustic correlates are language-specific and may differ among languages of the same family. For example, EP, English and Russian, which are Romance, Germanic and Slavic languages respectively, have contrastive stress, contrarily to French and Polish, which are Romance and Slavic languages (Hayes, 2009). Consequently, a typology of stress languages cannot follow family criteria. Instead, classification and description of stress languages usually follows parameters such as functionality, predictability, metricallity, and rhythmicity. Predictability is a key-difference in the target L2 and L1 of this dissertation. While Hungarian is a *predictable* (or *fixed stress*) language, EP is classified as a *non-predictable, lexical,* or *variable stress* language, as we will report in Chapter 2.

Other than stress typologies, stress characterization must include a description of the acoustic correlates used to signal prominence. In every language, different acoustic cues behave interactively, as a stronger use of one of them compensates for a weaker use of the other. However, this relative weighting of acoustic features is language-specific, even among languages within the same stress typology. For example, English and Dutch, both lexical stress languages, have vowel reduction. However, this seems to be a more important cue for stress perception for English listeners than for Dutch, probably due to a higher frequency of reduced vowels in English comparing to Dutch (Braun et al., 2008; Cooper et al., 2002). According to Hayes (1995), the language-specific weighting of stress cues is also related with the acoustic cues used at the segmental level. The author posits that there is a tendency in languages to avoid relying in the same acoustic features at segmental and suprasegmental level. The author gives the example of Czech, a language in which vowel length is contrastive, and consequently, duration is avoided as a cue for word stress, with speakers resorting instead to pitch and intensity. The author explains that "[t]his makes sense, since using duration to mark stress in these languages would obscure the phonemic vowel length contrast" (Hayes, 1995, p. 7). However, as pointed out by Chrabaszcz et al. (2014), while acoustic cues involved in stress are usually well documented in languages, the importance of each and the interaction between them are less frequently studied, particularly when it comes to L2 perception.

Although it is generally accepted that speakers of fixed stress languages display weaker abilities in stress perception than speakers of variable stress languages, the later can also exhibit difficulties in certain conditions, for example, when specific acoustic cues are absent. In the next section, we will address this question, and factors that affect L2 stress perception, by describing several perceptual studies on these issues.

1.4.1. Studies on L2 stress perception

Contrarily to L2 vowel/consonant perception, to the moment no theoretical model has been proposed for L2 stress perception. One reason for this absence is the interdependency and relativity of acoustic cues that may express stress. Moreover, not only stress patterns vary from one language to the other, but they also vary between speakers of the same language as well as within a single speaker, depending on the context. Although this variability is also observed in vowels and consonants, stress cues – intensity, duration, fundamental frequency (f0), vowel reduction, are not specific to this feature, and are used at segmental level or in the domain of the phrase as well. Such complexity and variability would demand a strong body of empirical evidence, collected from an ample number of studies. Although experiments on L2 stress perception, and more specifically on L1 effects on L2 stress perception, have been increasing in the last decade, they vary in their focus, investigating aspects such as acoustic cue-weighting, effect of rhythmic properties, or stress 'deafness'. Additionally, among studies with the same approaches, we find conflicting results. Below we will present a selection of studies, representative of these questions.

In a cross-linguistic study comprising seven languages, Salsignac (1998) attempted to determine if non-native speakers rely more on acoustic cues ('bottom-up processing') or transfer from their L1 ('top-down processing') in detecting stress. The author predicted that in case of a more prominent stress, listeners rely first on acoustic cues. If, however, stressed syllables are pronounced with less prominence, listeners resort to their L1 stress patterns. To test these hypotheses, the author recruited speakers of Turkish, Polish, Hungarian, Czech, Russian, and Spanish, all with good knowledge of French. Table 1.1 summarizes the L1 features and the tested L2.

Table 1.1. L1 and L2 tested in Salsignac (1998, p. 88)[31]

			Tested L2	
Participants' L1	Stress typology / acoustic cues	Level of prominence	L2 with a stress system similar to the L1	L2 with a stress system different from the L1
French	fixed final / duration and eventually f0	median	Turkish	Polish, Hungarian, Czech, Spanish, and Russian
Turkish	fixed final / duration and eventually f0	week	French	x
Polish	fixed paroxyton / duration and/or intensity	week	x	French
Hungarian	fixed inicial / duration and/or intensity (vowel lentgh contrastive)	strong	Czech	French
Czech	fixed inicial / duration and/or intensity (vowel lentgh contrastive)	median	Hungarian	French
Russian	free-stress / multiple acoustic cues (intensity dominant)	strong	Spanish	French
Spanish	free-stress / multiple acoustic cues (intensity dominant)	strong	Russian	French

As shown in this table, each language group was tested in French and, in most of the cases, in another language with some similarity to the L1 (e.g., Hungarian speakers were also tested in Czech, since both languages have fixed stress, and rely on duration and f0 to signal it). The study also included two French participants, who were tested in all the L2 included in the study. Participants were presented with two sets of auditory stimuli, one consisting of false friends and the other with indigenous words. After hearing each token, as many times as needed, participants had three tasks: transcription, syllabification, and identification of the stressed syllable. The effect of stress prominence was analyzed according to the level of prominence assigned to each language. As predicted, results showed a general tendency of a bottom-up strategy (attention to acoustic signals) when the L2 had a more prominent stress. As for the effect of L1, it varied across languages and

[31] We maintained the author's nomenclature of "free-stress" for variable stress.

between participants of the same language but, in most of the subjects, the author observed an effect of their “stress habits”. For example, Hungarian and Czech listeners (both with word-initial stress) performed better with stimuli not stressed on the final syllable. Additionally, in the case of the non-French participants, results showed that not only their L1 affected stress perception, but also, at some level, their knowledge of French (their adopted language). Based on the results, Salsignac suggests that L1 affects L2 stress perception more than acoustic cues. Generalization of Salsignac’s conclusions, however, need caution, since the study presents some methodological deficiencies, for example, the fact that only two speakers of each language were tested.[32] More importantly, one can also argue that the more or less attention given to acoustic cues that signaled stress may in itself be an effect of the L1, as we will see in the studies we review below.

When it comes to stress perception studies, one of the central concepts is stress ‘deafness’. In L2 phonology, the term ‘deafness’ refers to the difficulties listeners show when discriminating pairs of non-native segmental or suprasegmental contrasts (Peperkamp & Dupoux, 2002).[33]

The studies conducted by Dupoux and colleagues (Dupoux et al., 1997, 2001) with speakers of French (a predictable stress language) were important to establish the idea that stress ‘deafness’ is a categorization problem and not an auditory one. Dupoux et al. (1997) first observed French native speakers in a series of perceptual discrimination experiments, using CVCVCV nonwords, recorded by female and male speakers of Dutch. Stimuli contrasted in stress location (e.g., *bópelo*/*bopélo* or *bopélo*/*bopeló*), phoneme (e.g., *fídape*/*lídape*) or stress and phoneme (e.g., *fídape*/*lidápe*). Spanish speakers (a non-predictable stress language) were also recruited, as a control group. For each experiment, the authors compared differences between mean error rates for each condition (*stress*, *phoneme*, *stress and phoneme contrasts*), and between the experimental and control group. The first three experiments consisted of ABX tasks. Results revealed a strong stress ‘deafness’ in the French participants, comparing to phoneme contrast trials,

[32] Furthermore, some important details on the experiment are not described in the article, such as number of trials, use of speaker variability, and physical conditions of the experiment.

[33] Following these authors, we will use the term ‘deafness’ in quotes for two reasons. First, it is only a partial phenomenon, since speakers with normal hearing abilities can perceive, at least to a small degree, stress contrasts. Second, as we will report, ‘deafness’ refers to a problem with representation.

and to the Spanish speakers' performance. In the fourth experiment, the authors tested only French speakers, this time using an AX task. In this experiment, besides phoneme contrast and stress contrast, one other condition was tested: the inter-stimulus interval (ISI). While in some trials this parameter was set to 200 ms, in others it was set to 2200 ms. Results showed that French speakers can acoustically discriminate stress contrasts as well as phoneme contrasts (3,2% and 2,2%, respectively). Although a higher error rate was found in trials with the 2200ms ISI, the difference was not significant. The authors explain that this result was due to the task being less demanding than the previous, in two aspects. First, the AX task requires less memory load than the AXB, since it involves only two stimuli. Second, the authors used tokens from only one of the female Dutch speakers, thus greatly reducing phonetic variability. The authors concluded that French speakers are able to acoustically discriminate stress contrasts, conserving a working memory for the acoustic correlates for more than 2 seconds.

The results of this study were later confirmed in Dupoux et al. (2001), using sequence-recall tasks,[34] which, according to the authors, configure a more robust methodology to assess individual results. The authors tested stress 'deafness' also in French speakers (with Spanish as control group), but other than a new method, they also introduced more manipulation of phonetic variability. The different conditions were created either varying speaker (using different tokens from different speakers, different tokens from the same speaker, or one token from the same speaker) or manipulating pitch by using speech resynthesis. To prevent participants from resorting to recoding strategies, a brief ISI (80 ms) was used. Other than phonetic variability, different lengths of sequences were tested, to assess the effect of memory load. Results were analyzed by comparing each participants' performance as expressed in mean error rate in phoneme contrast trials (e.g., [kúpi]-[kúti]) with mean error rate in stress contrasts trials (e.g., [mípa]-[mipá]), as well as scores for each sequence length. Mean error rates of each language group, for each condition, were also analyzed (Figure 1.9, next page).

[34] In these tasks, participants had to memorize two tokens from a minimal pair, associating each with a key in a computer keyboard. They then listened to randomized sequences of the tokens and had to reproduce them, using the respective keys.

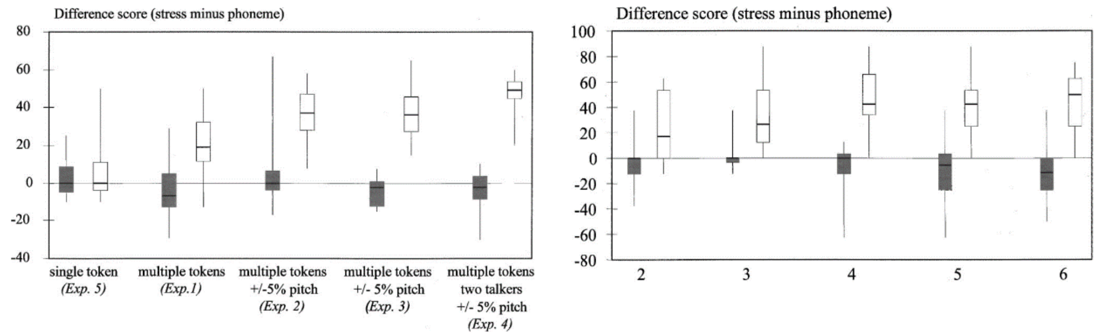


Figure 1.9. Results for French participants (in white) and Spanish participants (in grey), as a function of phonetic variability (on the left), and of sequence length (on the right), (Dupoux et al., 2001, p. 1615)

Based on the results, the authors draw three conclusions. First, the size of the 'deafness' effect in French speakers increases with the amount of phonetic variability, but only up to a certain point. Performance in trials with a single token significantly differed from performance in trials with multiple tokens. However, no significant differences were found when pitch variability and pitch+speaker variability were introduced. Second, although error rates for shorter sequences were lower than for longer sequences, differences between stress contrast scores and phoneme contrast scores were not significant between sequences of 4, 5, and 6 tokens. Considering this, the authors concluded that "it is the presence of memory load that matters, not the amount of it" (Dupoux et al., 2001, p. 1615). Third, results suggest that phonetic variability alone was not responsible for the stress 'deafness', the same applying to memory load. This suggestion is based on the fact that in the tasks with zero phonetic variability no problems emerged in the stress contrast, not even in longer sequences trials. On the other hand, in the training sessions (with one-token sequence), also no problems were observed, even in the trials with more phonetic variability. Instead, the authors propose that it is the presence of both factors – phonetic variability and memory load – that affects perception.

The findings of Dupoux et al. (1997, 2001), summarized in Figure 1.10 (next page), suggest, as mentioned above, that stress perception difficulties in French speakers does not have an acoustic root, but they are rather a problem of encoding the acoustic information at an abstract level.

	French participants		Spanish participants	
	Stress	Phoneme	Stress	Phoneme
AX, low phonetic variability[a]	3.8%	4.0%	—	—
ABX, high phonetic variability[b]	*19%*	5.8%	4.0%	8.8%
Sequence recall, low phonetic variability[c]	27%	23%	19%	17%
Sequence recall, high phonetic variability[d]	*53%*	15%	20%	24%

Figure 1.10. Summary of error rates in different experiments in Dupoux et al. (1997) and Dupoux et al. (2001) (Dupoux et al., 2008, p. 686)

Peperkamp & Dupoux (2002) extended the previous study to other predictable stress languages (Finnish, Hungarian, Fijian and Polish), with the intention of establishing a typology for stress 'deafness' based on stress rules of these type of languages. The authors' assumption was that stress 'deafness' depends on the moment when the s*tress parameter* is set. The *stress parameter* is a binary parameter, proposed by the authors, as to whether stress is coded as contrastive or non-contrastive. According to the authors, during the first two years of life, infants must decide, based on the limited language knowledge they have and on acoustic cues on stress at utterance boundaries, if stress is contrastive or not in their L1. If by the time the *stress parameter* is set – hypothetically around the same age in every language[35] – infants are not yet able to infer stress regularity in their L1, then they will not encode stress as non-contrastive and, consequently, maintain a good ability at discriminating stress contrasts. If, on the contrary, when the *stress parameter* is set, infants have already encoded stress as non-contrastive, then stress 'deafness' will be observed. Based on universal L1 acquisition steps (segmental inventory → function words → content words), the authors predict five situations:

(i) if the *stress parameter* is set at a late point in L1 acquisition, after the lexicon is formed (or a sizeable part of it), speakers of non-predictable stress (Spanish) and of predictable stress (French, Finnish, Fijian, Hungarian and Polish) will behave in a similar way, neither showing stress 'deafness'. This will occur since neither

[35] According to the authors, "[t]his follows from a more general assumption that languages are equally learnable. Along the same lines, we assume that the tuning of the phonological representation is finished at a certain age, regardless of the language under consideration." (Peperkamp & Dupoux, 2002, p. 5)

of these speakers extracted stress regularities at that time. The authors named this hypothesis the *Lexical Parameter Setting hypothesis*.

Alternatively, if this situation is not observed, then one of the *Non-lexical Parameter Setting hypotheses*, described below, will be observed (Table 1.2, next page).

(ii) if the *stress parameter* is set in an earlier stage of L1 acquisition, when only phonetic information is perceived by infants, speakers of French and Finish show stress 'deafness'. Since in French stress is phrase-final and in Finish is word-initial, with no exceptions, infants can infer this non-contrastive nature of stress without accessing any language-specific features, and will ignore stress contrasts from very early on. On the contrary, speakers of the other languages (Fijian, Hungarian and Polish) maintain the ability to discriminate stress contrasts, since at this moment they were not yet able to extract stress regularities.

(iii) if the *stress parameter* is set when phonological rules were already extracted but no acquisition of function nor content words occurred, speakers of languages such as Fijian also present stress 'deafness'. In Fijian, stress depends on syllable weight, and once infants acquire the distinction between light and heavy syllables, they encode stress regularity.

(iv) if the *stress parameter* is set after the acquisition of function words (but before content words), Hungarian speakers will show stress 'deafness', since in this language, utterance-initial functions words are unstressed. Consequently, while acquiring function words, Hungarian infants will also be able to extract stress regularities, from which moment they will ignore stress contrasts.

(v) if the *stress parameter* is set after the acquisition of the content words lexicon, Polish speakers will also show stress 'deafness', since in this language stress regularities can be inferred only with the perception of content word boundaries.

Table 1.2. Overview of the *Non-lexical Parameter Setting hypotheses*

	L1 features aquired by the infant, by the time the stress parameter is set to 'non-contrastive'			
	phonetics	phonetics phonology	phonetics phonology function words	phonetics phonology function words content words
Languages in which stress 'deafness' will be present or absent in L2 perception, according to the L1 knowledge of the infant red = present; blue = absence	French Finish Fijian Hungarian Polish	French Finish Fijian Hungarian Polish	French Finish Fijian Hungarian Polish	French Finish Fijian Hungarian Polish
	(ii)	(iii)	(iv)	(v)

To test these hypotheses, the researchers recruited Finnish, Hungarian and Polish speakers, who completed different sequence-recall tasks, following Dupoux et al. (2001).[36] The perception of stress contrasts (e.g. [mípa]-[mipá]) was compared with the perception of phonemic contrasts (e.g. [kúpi]-[kúti]) across different levels of phonetic variability and memory load. Results from this study were compared with those from Dupoux et al. (2001), Table 1.3 displays results for both studies.

Table 1.3. Overview of results (adapted from Peperkamp & Dupoux, 2002, p. 17)

	'stress deafness' index*	**stress rule**	**vowel length**	**stimuli**
French	38.10%	phrase-final	non-contrastive	set 1
Finish	24%	word-initial	contrastive	set 2
Hungarian	23.70%	word-initial, except if the phrase starts with a function word	contrastive	set 2
Polish	11.60%	content word-final	non-contrastive	set 1
Spanish	-4.40%	non-predictable	non-contrastive	set 1

*ER stress contrast - ER phoneme contrast

The first conclusion is that stress 'deafness' is not equally observed in all the non-contrastive stress languages. Accordingly, the *Lexical Parameter Setting hypothesis* is falsified. However, scores significantly differed across the tested non-predictable

[36] The authors were not able to test speakers of Fijian. Additionally, in the trials with Finish and Hungarians, stimuli were further manipulated, with durational differences between stressed/unstressed syllables being mitigated, to avoid conflict with vowel duration differences, which in those two languages is contrastive.

languages. *T*-tests revealed that, with exception of Finish speakers, all the other languages' results were significantly different from French speakers' results. Additionally, when compared to Spanish speakers' performance, only Polish scores did not differ significantly. This means that, in terms of stress 'deafness', French and Finish speakers formed a group, with strong stress 'deafness', and Polish and Spanish speakers another group, with weak stress 'deafness'.

The intermediate behavior of the Hungarian speakers posed a problem, since it cannot be explained by neither one of the *Non-lexical Parameter Setting hypotheses*. If the stress parameter is set before access to function words, Hungarian speakers should not show stress 'deafness', pairing with the Polish and Spanish speakers. If, on the contrary, coding of stress as non-contrastive is set after access to function words, in this case Hungarian participants should present a similar level of stress 'deafness' than French and Finish participants. The authors propose two explanations for the ambiguity of the Hungarian speakers' results. First, the stress parameter may not be binary, as they propose. Second, even if the stress parameter is binary, individual variation may be a significant factor.

The results of this study were partially confirmed in Peperkamp et al. (2010). Although the methodological approach in this study is in line with the previous one, the researchers address the question of stress 'deafness' in predictable languages from the perspective of four possible influential factors: domain of stress, lexical use of stress cues, variability in the stress position, and presence of lexical exceptions (Figure 1.11).

Language	Domain of stress	Lexical use of stress cues	Variability in stress position	Lexical exceptions (%)
Standard French	Phrase	None	No[a]	0
Southeastern French	Phrase	None	Moderate[b]	0
Finnish	Word	Duration	No[c]	0
Hungarian	Word	Duration	No[c]	0
Polish	Word	None	Moderate[d]	0.1
Spanish	Word	Duration, F0 and intensity	High[e]	17

[a] Final.
[b] Last non-schwa syllable.
[c] Initial.
[d] Penultimate in polysyllables, and on the only syllable of monosyllables.
[e] One of the last three syllables.

Figure 1.11. Stress features of the tested language (Peperkamp et al., 2010, p. 424)

Based on differences between the languages in what respects the influential factors, the authors established four hypotheses.

(i) if stress 'deafness' is influenced mainly by the *domain of stress*, then speakers of French (of the two varieties) present more problems, since in French stress is assigned to the phrase final syllable, thus is expected that French listeners ignore stress at word level.

(ii) if stress 'deafness' is influenced mainly by the *lexical use of stress cues*, French and Polish present a stronger stress 'deafness'. As for Finish and Hungarian speakers, since duration (vowel length) is lexically contrastive in these languages, participants display an intermediate case of stress 'deafness': weaker than French and Polish, but stronger than Spanish. Additionally, the authors predict that if durational contrasts are removed from the tokens, performance of Finish, Hungarian and Spanish participants decrease more than for French and Polish. As the authors explain, "[i]ndeed, whereas all populations should have more difficulties when one of the phonetic correlates is absent, this disadvantage should be larger for those who make lexical use of the missing correlate in their native language" (Peperkamp et al., 2010, p. 423).

(iii) if stress 'deafness' is influenced mainly by *variability in stress location*, Standard French, Finish and Hungarian participants present a high degree of stress 'deafness', due to the fixed location of stress in their L1 (phrase-final in Standard French, and word-initial in Finish and Hungarian). Southeastern French and Polish speakers, on the other way around, will have less difficulties. Regarding Southeastern French, unlike the Standard variety, stress can be located in the penultimate syllable of the phrase, if the final vowel is a schwa. As for Polish, stress is not in a fixed position, although it is more regularly assigned to words' penultimate syllable.

(iv) if stress 'deafness' is influenced mainly by the *number of lexical exceptions*, speakers of Standard French, Southeastern French, Finnish and Hungarian, all languages with no lexical exceptions to stress rule, reveal a stronger stress 'deafness'. Polish is an intermediate group, since some lexical exceptions are present in this language. Spanish presents a higher percentage of lexical exceptions; therefore, native speakers of this language have less problems perceiving stress contrasts.

To test these hypotheses, twelve participants of each language group completed two-token and five-token sequence-recall tasks, in three conditions: *phonemic* contrast (e.g., /ˈmuku/-/ˈmunu/), *stress-with-duration* contrast (e.g., /ˈnumi/-/nuˈmi/), with stress marked by duration, f0 and intensity, and *stress-without-duration* contrast, with stress marked by f0 and intensity. As in Peperkamp & Dupoux (2002), tokens were separated by 80 ms silence, to prevent participants from resorting to encoding strategies. Results (Figure 1.12) showed a robust 'deafness' effect in all non-contrastive languages.

	Phoneme		Stress-with-duration		Stress-without-duration	
	% error	SE	% error	SE	% error	SE
Standard French	34.2	3.9	77.5	4.4	84.6	2.2
SE French	45.8	4.3	85.4	2.0	86.7	2.3
Finnish	49.6	5.3	90.0	2.3	87.1	2.2
Hungarian	45.4	6.0	80.4	4.3	85.4	3.9
Polish	42.5	6.4	59.2	7.9	65.4	8.0
Spanish	47.5	5.6	47.1	5.4	48.8	7.4

Figure 1.12. Results according to tested conditions (Peperkamp et al., 2010, p. 427)

Statistical analysis revealed that Standard French, Southeastern French, Finnish, and Hungarian speakers' performance was not significantly different. Polish speakers' results, however, were significantly different from the ones above, and also from Spanish; consequently, they formed an 'intermediate' group, regarding stress 'deafness'. In respect to the effect of absence of durational cues, the analysis showed that there was a main effect of this condition, since error rates were higher in trials in the *stress-without-duration* condition. However, this affected all the groups, no significant difference being found between speakers that uses duration lexically (Finnish, Hungarian, Spanish) and the others.

The findings of this study clarify the ambiguity in the results of Hungarian speakers in the study of Peperkamp & Dupoux (2002). While in Peperkamp & Dupoux (2002) a different set of stimuli was used to test Finish and Hungarian speakers than the one used to test Polish speakers, in the present study the same set was used to test all language groups. The authors concluded that Hungarian speakers displayed a strong stress 'deafness', comparable to French and Finish speakers. Considering the three groups of results – Standard French, Southeastern French, Finnish and Hungarian, in one side, Spanish in the opposite side, and Polish as intermediate –, the authors propose three typologies for stress 'deafness', according to the listeners L1: (1) speakers of predictable stress languages with no lexical exceptions, and independently of stress domain and variability (e.g., French, Finish, Hungarian), display a strong stress 'deafness', (2) speakers of predictable stress languages with few lexical exceptions (e.g., Polish) display a weak stress 'deafness', and (3) speakers of non-predictable stress with a significant amount of lexical exceptions (e.g., Spanish) do not display stress 'deafness'.

In the studies described so far, the focus was predictable stress languages (with Spanish as a control group), and stress contrasts signaled only by suprasegmental properties (duration, pitch, and intensity). However, the study of Chrabaszcz et al. (2014) presents a different case. Specifically, the authors investigated the relative weighting of different acoustic cues used in L2 stress perception – segmental and suprasegmental –, by native speakers of Russian and Mandarin, a predictable stress language and a tone language, respectively. The acoustic cues observed were vowel quality, pitch, duration, and intensity. Stimuli were produced by a female native speaker of American English, and speakers of this language were also recruited and tested, as baseline. Based on previous studies, the authors predict a hierarchy of stress cues for each of the three languages (Table 1.4, next page,"Prediction" columns). All participants completed a forced-choice auditory identification task, in which they were asked to identify the location of stress in minimal pairs of non-words, rating their confidence in the answer (Figure 1.13, next page). Stimuli included variations of the word "maba" (/mɑbɑ/, /məbə/, /mɑbə/, and /məbɑ/)[37]. To test

[37] Only tokens of /'mɑbə/ and /'məbɑ/ were recorded. The tokens /mɑbɑ/ and /məbə/ were created by combining syllables of the first two tokens.

effect of pitch, duration, and intensity, different tokens were created from the original tokens, by manipulating these features.

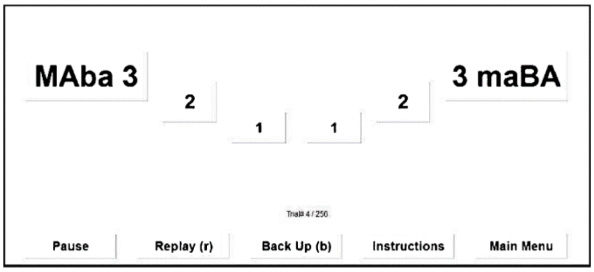


Figure 1.13. Computer screen of the trials (Chrabaszcz et al., 2014, p. 1473)

Statistical analysis showed that in all the languages each of the four acoustic cues was significant for perception, and that languages significantly differed in the use of each of the four cues. However, the predictions for the cues' importance hierarchies were refuted by the results (Table 1.4).

Table 1.4. Predictions and results for hierarchy of importance of acoustic cues (1 = most important, 4 = least important) (Chrabaszcz et al., 2014)

	A. English		Mandarin		Russian	
	prediction	result	prediction	result	prediction	result
Vowel quality	2	1	2	1	2	1
Pitch	1	2	3	2	4	4
Intensity	4	3	4	3	3	2
Duration	3	4	1	4	1	3

Contrary to the predictions, vowel quality was the strongest cue across languages. The authors suggested that this is due to the 'independence' of vowel quality, in the sense that a schwa always signals an unstressed syllable.[38] As the authors pointed out, pitch, intensity and duration are relative cues: a longer syllable, for example, is only perceived as longer if the other(s) is(are) perceived as shorter. Additionally, the strong reliance on vowel

[38] A similar case is observed with the EP [ɨ], which can only occur in unstressed position. However, the case of [ɐ] is different, since it can occur in stressed position too (before the nasal consonants [m], [n] or [ɲ]).

quality by the speakers of Mandarin was also unexpected. According to the authors, this may be due to previous knowledge of English by those speakers. Another interesting outcome in this study is the fact that, regarding the use of cues to signal stress, more similarities emerged between a stress language and a tonal language (English and Mandarin) than between two stress languages (English and Russian).

1.4.2. Perceptual training in L2 stress acquisition

If much is still to be done regarding research on non-native stress perception, studies focusing specifically on L2 stress acquisition are even scarcer. Below we report five studies that aimed at observing changes in stress discrimination ability. Differently than for the studies in acquisition of L2 vowels we presented in 1.3.2, in L2 stress acquisition we could not set strict selection criteria for our literature review, given the low number of studies published to the moment. However, the experiments we report next have methodological aspects or results of interest for this dissertation project. The first three studies we report (Dupoux et al., 2008; Michaux, 2016; Schwab & Llisterri, 2014) observe results of only one session of training. Although participants in the study of Schwab & Llisterri (2014) were not learners of the L2, we include it due to relevant findings presented by the authors. We also review the study of Honbolygó et al. (2017), in which stress perception in Hungarian learners of Germain with different L2 knowledge is observed. The last two studies we describe, Ou (2011) and Brawerman-Albini et al. (2017), conducted longer trainings with university students.

Michaux (2016) tested the ability of French-speaking Belgian learners of Dutch (n = 30) to perceive Dutch word stress. Contrary to French, Dutch has a lexically contrastive stress. The author conducted two experiences. First, the participants had to complete an AXB discrimination task, testing three stress contrasts: *1st vs. 2nd syllable*, *2nd vs. 3rd syllable*, and *1st vs. 3rd syllable*. After that, they completed an identification task, in which they were asked to identify the location of stress (1st, 2nd or 3rd syllable). In both tasks, two set of stimuli were used: one with real words and matching nonwords resulting from a change in the stress location, the *correct-incorrect* condition (e.g., *PAgina*, **paGIna*, **pagiNA*)[39], and

[39] Only the first exists in Dutch: /ˈpaːɣinaː/.

a second set of stimuli with nonwords, created with the substitution of the consonants for [s] and the vowels for [a] (e.g., **SAsasa*, **saSAsa*, **sasaSA*). Stimuli were recorded by a male and a female Dutch speaker. Both the AXB and the identification task were presented to participants in three blocks and in each block, each set of stimuli was tested separately. By dividing trials in three consecutive blocks, the author aimed at observing a learning effect. The analysis revealed both similar and different behaviors between the two language groups. For example, both groups scored worse in the trials with the 'created' nonwords (e.g., **SAsasa*) than in the trials with the real words with *correct* or *incorrect* stress placement, in both tasks. According to the authors, this may be due to the absence of phonotactic information. However, Dutch speakers were more affected by incorrect stress location in real words (e.g., **paGIna* and **pagiNA*) than French participants, who treated *correct and incorrect* stress stimuli in a similar way. The results showed that French participants performed better in the AXB task (which tested phonetic discrimination ability) than in the identification task (which tested phonological representation), a result that contradicts previous findings by Dupoux and colleagues (Dupoux et al., 1997). Regarding a learning effect in the AXB task, while French speakers improved from block 1 to block 3, for the real words with *correct* and *incorrect* stress location, this was not observed in the case of the 'created' nonwords. According to Michaux (2016), this reinforces the idea that segmental information is necessary to process stress contrasts. As for the identification task, the authors did not discover significant improvements in either group.

An additional result of interest was the fact that French and Dutch participants revealed a similar behavior when it came to the effect of stress in the median syllable. In the AXB task with real words, sequences with words stressed in the median syllable had better scores. In the identification experiment, the result was the opposite, with words stressed in the middle with the worst scores. The author presented two possible explanations. It may be that the similar behavior results from two different causes, specifically, a delayed lexical access, in the case of the Dutch participants, and the unusualness of stress in median position, in the case of French speakers. Alternatively, the author also suggests that prosodic perception is perhaps not only limited by the L1, and maybe there are other perceptual constraints common to languages which are typologically different.

In two experiments focusing on Spanish stress perception by French learners of Spanish, Dupoux et al. (2008) found no differences in the ability to discriminate stress contrasts among late learners with different degrees of L2 knowledge. The study included four groups of French speakers: French speakers with no contact with Spanish (n = 20), and French learners of Spanish in the Beginner level (n = 14), in the Intermediate level (n = 14), and in the Advanced level (n = 11). Twenty Spanish speakers were also recruited, as control group. The first experiment consisted of sequence-recall tasks and was divided into two parts, the first aimed at phoneme contrast, and the second, at stress contrast. In each part, participants were first familiarized with different auditory tokens of two nonwords (/fiku/-/fitu/, or /númi/-/numí/), associated with the keys [1] or [2]. The familiarization trials were followed by a training block, in which they were asked to identify several tokens from the familiarization task with [1] or [2]. In this block, participants had to complete as many training trials as necessary to reach seven correct answers in a row. When this goal was achieved, they were tested with 28 four-item sequence-recall trials, a replication, with some adjustments, of the high phonetic variability experiment of Dupoux et al. (2001). Results confirmed a robust stress 'deafness' effect in the French speakers, independent of the contact with Spanish or knowledge of Spanish. The authors concluded that "[s]tress 'deafness' emerges here as a robust processing limitation, which cannot be eliminated with a significant exposure to a language with contrastive stress" (Dupoux et al., 2008, p. 695). The second experiment consisted of a lexical decision task, in which participants had to listen to different pairs of words, consisting of a real-word and a matching non-word, obtained by change in the stress location (e.g., *górro*-**gorró*). Segment contrasts were included as control condition (e.g., *grupo*-**brupo*). In this experiment, no training block was included. Analysis of the results showed no significant difference between the three subgroups of French leaners (Beginner, Intermediate and Advanced), nor between these and the French participants with no contact with Spanish, leading to the conclusion that "the stress 'deafness' effect observed with the sequence recall task is not limited to the encoding of stress in short-term memory but extends to lexical access" (Dupoux et al., 2008, p. 699). Furthermore, results showed that French participants performed significantly better for real words than for non-words (24% vs. 58%), but this difference was found across levels of Spanish knowledge, meaning no improvement in stress perception was observed with increase of lexicon. When comparing their results with the ones from previous studies (Dupoux et al. 1997, 2001), the authors concluded that, although in previous studies French speakers showed some ability to discriminate stress, this ability is restricted to the phonetic level. When subject to more demanding tasks, these speakers failed at encoding stress differences at an abstract phonological level. Furthermore, this study also showed that even with metalinguistic awareness of stress – in their course, these learners receive explicit instruction on contrastive stress – stress 'deafness' persisted.

It is worth mentioning the remarks made by the authors as for the adequacy of speech perception models to suprasegmental perception, specifically SLM (Flege, 1995) and PAM (Best, 1995). According to Dupoux and colleagues, although these models suggest that non-existing categories are more easily perceived (and acquired) than similar ones, this does not seem to be the case for stress perception by French speakers. The authors proposed that this may be due to differences in the way the human brain processes non-native segmentals and suprasegmentals. According to the authors, in L1 acquisition suprasegmental features are processed by the left hemisphere, whereas in adults, non-native suprasegmentals are processed by the right hemisphere. Consequently, in the latter, no phonological representations are formed.

The persistence of 'deafness' was also attested in Honbolygó et al. (2017). The authors administered a sequence recall task to Hungarian learners of Germain, with two L2 knowledge levels: Intermediate and Advanced.[40] A control group with Hungarian speakers with no previous contact with Germain was also recruited. In the sequence recall task, stimuli consisted of stress contrasts ([LIdu-liDU]) and phoneme (consonant) contrasts ([mäge-mäse]). Unlike in Dupoux et al. (2001, 2008), no manipulation of the acoustic cues was included. Sequences of 4, 5, and 6 stimuli were presented to the participants. Results showed a significantly higher error rate for stress contrast than for phoneme contrast, and

[40] The sequence recall task was accompanied with a series of other tests (a digit span test, a stroop test, a phoneme deletion, and a linguistic test).

this difference slightly increased with the increase of sequence length. These results confirm a clear 'deafness' in perception of stress contrasts by Hungarian listeners. Furthermore, no significant differences were found between learners, and between these and the control group, suggesting the persistence of 'deafness' independently of L2 proficiency, as already suggested by Dupoux et al. (2008).

Schwab & Llisterri (2014) conducted a short perceptual training, also with French speakers, to test the ability to improve Spanish stress perception. Forty-nine French naïve speakers with no previous contact with Spanish were recruited and divided into two groups, with only one receiving training. A group with 22 bilingual Catalan-Spanish speakers was also included, in testing and training. All participants completed a pretest first, in which they heard a Spanish real word and had to identify the stressed syllable. Stimuli consisted of triplets, each including a proparoxytone word (*médico* [ˈmeð̞iko], 'doctor'), a paroxytone word (*medico* [meˈð̞iko], 'I give a medicine') and an oxytone word (*medicó* [með̞iˈko], 'he/she gave a medicine').[41] Tokens were recorded by a single Spanish native speaker. Training comprised one session only, divided in five blocks, and trials consisted of shape/nonword matching tasks (Figure 1.14), a type of task that other than a low-level acoustic process requires also lexical access.

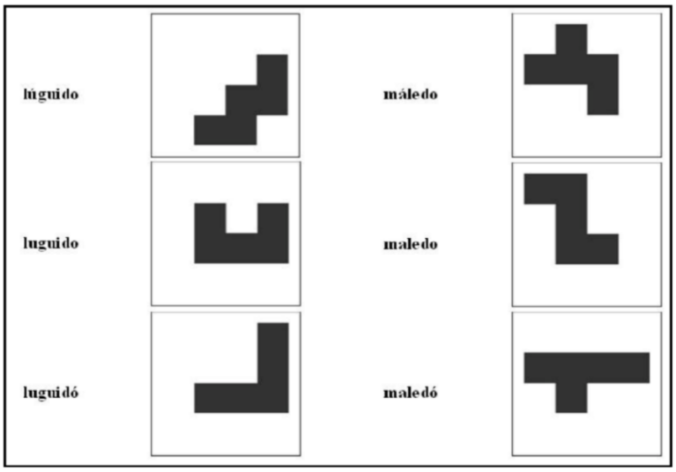


Figure 1.14. Example of a shape/nonword matching trial for the training session (Schwab & Llisterri, 2014, p. 628)

[41] Proparoxytones are words stressed in the antepenultima syllable, paroxytones are stressed in the penultima syllable, and oxytone words receive stress in the last syllable.

In the first four blocks, participants received feedback, but in the last no feedback was provided. The results of this fifth block were used to evaluate participant's progress. After training, participants repeated the pretest trials. Results showed a perceptual improvement, even in the case of the native Spanish speakers, which, according to the authors, "suggests that the mere fact of repeating the exact same task improves the performance" (Schwab & Llisterri, 2014, p. 631). Surprisingly, non-native participants who received training did not improve, while the non-native participants who did not receive training improved. To account for this unexpected result, the authors proposed that participants' learning abilities had a bigger effect than training. This explanation was supported by the analysis of participants' performance in post-test, considering the performances in the fifth block (with no feedback). Results indicated that, individually, a bigger improvement in the training phase corresponded to a bigger improvement from the pretest to the post-test. Thus, the authors concluded that improvement was not a result of training but rather depended on participants' learning ability.

Ou (2011) investigated the effect of a two-weeks perceptual training in the acquisition of English stress by 19 speakers of Mandarin or Taiwanese, with an average of about seven years of English studies. Six minimal pairs of English dissyllables differing only in stress (e.g., *pérmit* vs. *permít*) were recorded in two conditions, *falling intonation* and *rising intonation*. In the first one, stimuli were inserted in declarative-answer sentences (e.g., *Yes, I said pérmit*. vs. *Yes, I said permit*)., with the stressed syllable receiving a high tone. In this condition, participants were expected to rely mainly on tone. In the second condition, *rising intonation*, stimuli were inserted in yes/no questions (e.g., *Did you say pérmit?* vs. *Did you say permít?*) with the stressed syllable receiving a low tone. In this condition, participants were expected to rely on acoustic correlates other than f0, since the low tone does not signal prominence for these speakers. Stimuli were the same for testing and training, and were recorded by one female native speaker of American English. the pretest and the post-test were identical and included two-alternative forced-identification trials, in which participants were asked to identify an auditory stimulus (e.g., *pérmit*) among two possibilities, visually presented. The experimental group's training included six sessions, with an average of 15 minutes each, completed in six days. In each training session, two of the six minimal-pairs were used. This meant that in the first three sessions all six pairs of words were used, and the following three sessions consisted in the repetition of the first three. For each pair of words, participants had to complete a multiple-step task. First, they were presented with auditory tokens of a minimal pair and the corresponding visual form, and could hear each token as many times as needed to recognize the differences. They were then asked to listen to one of those tokens and identify it as noun or a verb, verifying their choice with a pre-recorded answer (e.g., *It's a noun*.). After, they heard the same token and pronounced it at least once. Finally, they had to answer orally to a pre-recorded question (e.g., *The noun's* ______.), and check their answer by listening to the correct token. The control group also received training, but not related to stress perception. Results showed that the control group did not show changes from the pretest to the post-test. As for the experimental group, two contradictory results emerged: although a 33% improvement was observed in the *rising intonation* condition, in the *falling intonation* condition results were worse in the post-test than in the pretest. The author suggested that this decline in stress perception in the falling contour was due to an adjustment, that is, learners were in the process of adjusting their perception of stress from pitch to other acoustic cues, but they still couldn't perform native-likely.

The last study we report here is from Brawerman-Albini (2012), who conducted a perceptual/production training with native Brazilian speakers, learners of L2 English.[42] In the study, the author investigated if training would be effective in improving these learners' ability to perceive and produce English preantepenultimate stress, uncommon in the L1 of the learners. Although both the L1 and L2 have lexical contrastive stress, the authors explain that since Brazilian speakers have rare access to preantepenultimate pattern in their L1, training is needed to compensate the lacking input. Thirty students enrolled in a course of Portuguese and English languages, at a Brazilian university, were recruited and divided into two groups: the experimental group (n = 20) and the control group, who did not received training (n = 10). The study lasted around three weeks, and included pretest,

[42] This study was conducted for a PhD dissertation, and was after published in Brawerman-Albini et al. (2017).

training, post-test, and generalization test.[43] Perceptual tasks were completed by the participants in a computer laboratory room at the university. Stress perception was tested by means of an identification task, with participants being asked to identify the stressed syllable of real words. Stimuli included 40 words with preantepenultimate stress (e.g., *fortunately* [ˈfɔrʧunətli]) and ten distractors (e.g., *prejudicial* [prɛʤəˈdɪʃəl]), all recorded by five native American English speakers (male and female). Post-test included the repetition of pretest trials and a generalization test, with new stimuli produced by new speakers. As for training, this was administrated in five sessions, divided into six blocks, and sessions lasted between 10 and 35 minutes. The first two sessions were conducted in two days in a row (one in each day), then a pause of five days was introduced, followed by the remaining three sessions, that were completed in three days (also one per day). Training tasks were similar to test tasks, but in this case, feedback was provided, after each trial and at the end of each block, and participants could hear the tokens as many times as necessary. Training stimuli included preantepenultimate stressed words as well as distractors (with antepenultimate or penultimate stress) and were produced by two of the American English speakers. Statistical analysis revealed significant differences between the pretest and the post-test, in the case of the experimental group, but not for the control group. However, when comparing the pretest's results with the generalization test's results, the differences were found not to be significant, that is, learning was not transferred for new words/new talkers. The authors suggested that the absence of generalization of learning was possibly due to the novel speakers, who, according to the participants, were more difficult to understand.

1.5. Summary

In this chapter, we reviewed fundamental aspects of L2 speech perception, namely that it is inherently an *auditory*, *categorical* and *language-specific* phenomenon. With respect to L2 acquisition, the preservation of the L1 learning mechanisms in L2 – *full access*

[43] The study also included a retention test for production, that was administrated two months after post-test. Although the study of Brawerman-Albini et al. (2017) focus both perception and production, we will only describe the information pertinent for perception.

hypothesis – is nowadays a consensual assumption. However, this does not mean that acquisition in L2 will be identical as in L1, since a multiplicity of factors have to be accounted for the L2 context, differently than for L1, specifically when we refer to late learners.

One of the main factors that contributes to differences between L1 and L2 acquisition is *input*. The input the adult learner is exposed to is considerably different than the L1 input, in terms of quality and quantity. Researchers agree that the distributional learning observed in infants for the L1 is also the mechanism that leads the L2 learner to optimal perception. Consequently, in order to subtract the L2 relative frequency distributions, the learner needs long-term exposure to L2 input (which is also a fundamental condition in L1 acquisition). Considering the key-role of input, when planning a perceptual training, stimuli must be attentively designed. Although some mixed-results emerged in studies investigating the most effective experimental stimuli, high variability perceptual training has been established as the most adequate method for perceptual training, as it is more likely to promote categorical learning over phonetic discrimination. The question of including natural, synthetic and/or manipulated stimuli has also been subject to empirical observations. However, evidence provided from studies is inconclusive, with no clear proof of one being significantly more effective than the other.

With respect to cross-linguistic effects, and more specifically, the effect of L1 in L2 perception, several models have been proposed in the last three decades. In section 1.3.1 we described three models, SLM, PAM and the L2LP model. Although the three models recognize the importance of L1 in L2 perception, they diverge in some key aspects. SLM posits that L1 and L2 categories occupy one common phonological space, which entails an L1-L2 bidirectional influence. As opposed to this, the L2LP model posits the existence of separated phonological spaces and identifies mixed perceptions and/or productions as an effect of a bilingual mode. Consequently, improvement in L2 perception does not affect L1 stability. SLM, PAM and the L2LP model also differ in their target subjects. While SLM is particularly concerned with learners in naturalistic contexts, PAM observes naïve non-native listeners. As for L2LP model, it considers learners of different proficiency levels and has been tested with subjects from formal contexts of L2 acquisition. Additionally, these models also diverge on which L2 sounds are easier/more difficult to perceive. SLM posits that it is unlikely that the L2 learner changes his/her perception of L2 sounds that are similar to L1 sounds and, conversely, new categories are more easily created for more dissimilar sounds. By contrast, the L2LP model posits that learning tasks for new categories are more demanding, since they involve not only perceptual learning but also changes in representation. As for PAM, since it was originally developed for L2 perception on naïve listeners, the model is unclear on the acquisition of L2 sounds. Comparisons between the three models must be made with caution, however, since they approach speech perception in different ways. SLM posits an allophonic identification process between L1 and L2 sounds, and establishes comparisons between L1 and L2 segments. PAM investigates perception of L2 contrasts from an articulatory perspective. As for the L2LP model, it also compares L2 to L1 contrasts, however, as SLM, comparisons are based on acoustic analyses. The divergencies between the three models summarized above are reflected in the different methodologies used to test the models. For example, while PAM uses perceptual identification tasks to establish the L2 contrasts' assimilation types, and respective predictions for discrimination, the L2LP model bases discrimination predictions on L1-L2 acoustic comparisons.

Apart from theoretical differences, there is a consensus on the importance of considering *perceptual overlap* situations when accounting for L2 speech perception. As explained in section 1.2, perceptual overlap can be determined using different methods. However, studies' results consistently demonstrated that the amount of perceptual overlap between L2 and L1 segments is a reliable predictor in L2 perceptual discrimination difficulties.

With respect to the learning process in L2 perception, we presented training studies for vowel and for stress acquisition (sections 1.3.2 and 1.4.2, respectively). The studies on L2 vowel acquisition we reported in this chapter present clear evidence of the positive effect of training. These studies differ in some aspects, such as the L1 of the learners and the target L2, the tasks and stimuli, and planned duration. However, independently of these factors, training seems to lead to immediate improvement. In most of the studies reported, learning was also generalized to new stimuli/speakers, although in Rato (2013) this was only partially observed. It should also be pointed out that even in significant improvement cases with advanced learners of L2 (Aliaga-García & Mora, 2009; Rato, 2013), students or some of the students did not reach a native-like performance. The existence of individual differences is pointed out as a possible explanation for differences in learners' improvement, which emphasizes the need to look beyond group results. For example, Grenon et al. (2019) observed that although improvement from the pretest to the post-test was overall achieved by participants, analysis of cue-weighting tasks showed that only half of the students were able to shift from durational to spectral cues.

The situation of L2 word stress acquisition is not as straightforward as for L2 vowel acquisition. Not only fewer studies have been conducted, but also results are conflicting. For example, Dupoux et al. (2008) compared French learners of Spanish of three levels (Beginner, Intermediate and Advanced) and found no significant differences in stress perception, which suggests a persistent stress 'deafness', even when students are explicitly instructed about stress on class. Moreover, results also showed that there was no effect of the size of the lexicon on stress perception. Schwab & Llisterri (2014) obtained a contrary outcome in their study, conducted also with French learners of Spanish. As result of a one-session training, the authors observed improvement, even in the native group (who also completed the training). However, the effect of training was more related with individuals' learning abilities than training itself. Brawerman-Albini et al. (2017) and Ou (2011) also found a significant effect of training in improvement of stress perception. Interestingly, in both studies the effect was not generalized to new stimuli/new talkers.

Overall, the theoretical questions approached in this chapter provide us with a useful insight on L2 speech perception. However, the issues here reported can only be employed taking into account the specific L2 and L1, which we will describe in the next chapter.

CHAPTER 2

PHONETICS AND PHONOLOGY OF EUROPEAN PORTUGUESE AND HUNGARIAN

The aim of the present chapter is to describe the phonetics and phonology of the target language (L2) in the present research, EP. However, given the potential importance of the effect of the L1 in L2 acquisition, a description of the phonetics and phonology of the L1 language relevant to the present research and its participants, Hungarian, is also included.

In the present chapter, we will focus on the linguistic features pertinent to the dissertation: the vowel systems (in the case of EP, with more detail for the reduced vowels [ɐ] and [ɨ]), and word stress.

The description of EP phonetic and phonological system we present in this chapter, as well as the examples, are based on "The Phonology of Portuguese" of the Oxford University Press (Mateus & d'Andrade, 2000), "Gramática da Língua Portuguesa" of Editorial Caminho (Mateus et al., 2003), and "Gramática do Português" of Fundação Calouste Gulbenkian (Raposo et al., 2013a, 2013b, 2020). As for the description and examples of Hungarian, they are based "Hangtan" (Markó, 2017) and on "The Phonology of Hungarian" of the Oxford University Press (Siptár & Törkenczy, 2000). Information from other sources will be specifically referenced.

2.1. Phonetics and phonology of European Portuguese

Portuguese is an Indo-European, Romance and Iberian language (Ilari, 2013). At the time the present dissertation is written is reportedly spoken by over 230 million people, from which only about ten and a half million are Portuguese living in Portugal, and approximately five million are Portuguese emigrants living abroad. The remaining speakers of Portuguese are spread throughout the former colonies, mainly in Brazil (around 210 million speakers). Other than Portuguese varieties overseas, Portuguese has also regional dialects within Portuguese borders. These include continental dialects (northern, southern, and central) and insular dialects (Azores and Madeira). Standard EP is spoken in the region of Lisbon and Coimbra, and it is considered a subdialect of the southern and central variety, referred to as the central-coast dialect. This dissertation will focus on this dialect, since it is "accepted as a reference for teaching Portuguese as a second language" and it is "the most commonly heard on radio and television" (Mateus & d'Andrade, 2000, p. 4).

2.1.1. The vocalic system of European Portuguese

EP has a large vocalic inventory when compared with other Romance languages, and has oral and nasal diphthongs, a rare feature among languages. The vocalic system of EP consists of 14 vowels and 14 diphthongs:

(i) Oral vowels: [a], [ɐ], [ɛ], [e], [ɨ], [i], [ɔ], [o], and [u]

(ii) Nasal vowels: [ɐ̃], [ẽ], [ĩ], [õ], and [ũ]

(iii) Oral diphthongs: [aj], [ɐj], [ɛj], [oj], [ɔj], [uj], [aw], [ew], [ɛw], and [iw]

(iv) Nasal diphthongs: [ɐ̃j̃], [ɐ̃w̃], [õj̃], and [ũj̃]

Traditionally, EP oral vowels are classified according to tongue height and backness (Figure 2.1), while the features length (long/short), tenseness (tense/lax), and lip rounding (rounded/unrounded) are not used for their description, as these features play no role in any phonological vowel contrasts in the language.

<table>
<tr><td></td><td></td><td colspan="3">tongue position</td></tr>
<tr><td></td><td></td><td>front</td><td>central</td><td>back & round</td></tr>
<tr><td rowspan="4">openness</td><td>close</td><td>[i]</td><td>[ɨ]</td><td>[u]</td></tr>
<tr><td>mid-close</td><td>[e]</td><td rowspan="2">[ɐ]</td><td>[o]</td></tr>
<tr><td>mid-open</td><td>[ɛ]</td><td>[ɔ]</td></tr>
<tr><td>open</td><td></td><td>[a]</td><td></td></tr>
</table>

Figure 2.1. Phonetic oral vowel system of EP, adapted from Andrade (2020, p. 3255)

Although a lot of attention has been paid to describing the EP vowel system, acoustic measurements are scarce. Figure 2.2 (next page) displays the EP oral vowel space with the values calculated from Andrade (2020, p. 3251; see Appendix IV).

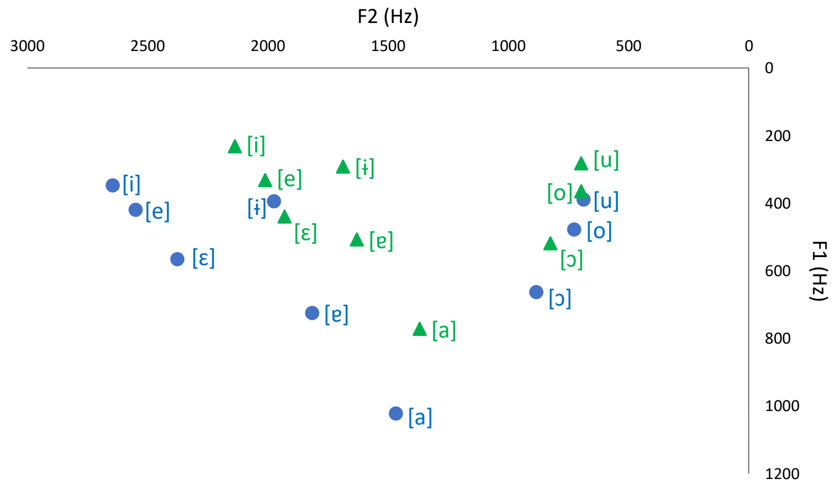


Figure 2.2. EP production vowel space. Blue dots correspond to productions by female speakers, green triangles to productions by male speakers

With respect to differences between female and male productions, Figure 2.2 confirms the observation of Escudero et al. (2009): female productions occupy a wider area in the vowel space, considering both the first formant (F1) and the second formant (F2).[44] For both female and male productions, [ɐ] seems to be positioned in a central place, equidistant to [a], [ɛ] and [ɨ].

The high, central vowel [ɨ] has the same level of tongue height as [i] and [u], but is more fronted, closer to [i]. This contradicts other authors' proposals. Delgado-Martins (1986) assumes [ɨ] as a schwa (<ə>). However, Cruz-Ferreira (1999) observes that [ɨ] does not have the quality of a schwa. The author locates it near [u], proposing to classify it instead as [ɯ]. Finally, in an MRI study, Martins et al. (2008) found that [ɨ] has the highest tongue dorsum point among the central vowels, but it cannot be considered a high central vowel. However, the results of this study are limited, since all acoustic data were collected from the production of one single female speaker.

The controversy regarding [ɨ] is a consequence of the instability of this vowel, since its production can vary a lot, with frequent deletion in colloquial speech, as the term "mute *e*", traditionally used to name this vowel suggests. For instance, the word *pequeno*

[44] These differences are probably observed in all languages as this is not the result of articulatory characteristics but the result of short vocal tracts in females, compared to males (Whiteside, 2001).

[pɨˈkenu] 'small' is produced in spontaneous speech as [pˈkenu]). Furthermore, [ɨ] may function as an epenthetic vowel, for example in a very formal register in word-initial position, if this starts with a fricative followed by consonant (e.g., *espécie* [ʃˈpɛsi] produced as [ɨʃˈpɛsi] 'species'), or at the end of an intonational, interrogative sentence for emphasis purposes (e.g., *hipotecar?* [iputɨˈkaɾ] 'to mortgage?' produced as [iputˈkaɾɨ̠]).

Contrary to [ɨ], the vowel [ɐ] can never be deleted, and it cannot function as an epenthetic vowel either. However, we also observe a lot of variation in its production. As explained by Andrade (2020):

> [t]he use of [ɐ] (and [ɐ̃]) takes into consideration a whole literature tradition about Portuguese. It must be understood as a symbol corresponding to an entire area of considerable variation, in the acoustic space defined by F1 and F2, and that it corresponds to sounds in which F1 can present a higher frequency than [ɛ] (p. 3255).[45]

The variability of [ɐ] and [ɨ] reported in the literature is consistent with the variability in the stimuli recorded for the two studies included in the present dissertation (see F1×F2 vowel spaces presented in Appendix VI).

With respect with phonological representations, the nine oral phones [a], [ɐ], [ɛ], [e], [ɨ], [i], [ɔ], [o], and [u] correspond to seven phonemes: /a/, /ɛ/, /e/, /i/, /ɔ/, /o/, and /u/.[46] Figure 2.3 (next page) presents the correspondence between the mentioned phones and phonemes, ruled by the process of unstressed vowel reduction. In this dissertation, the term *vowel reduction* refers solely to the phonological process in which specific vowels are subject to quality changes (centralization or/and raising) in unstressed position versus stressed position and are assumed to be produced with less effort. Besides vowel reduction induced by absence of stress, there are also cases of phonetic reduction, that is, changes in the articulation of some vowels, or even deletion, that depend on the speed or informality

[45] In the original: "A utilização de [ɐ] (e [ɐ̃]) toma em consideração toda uma tradição de bibliografia sobre o português. Deve ser entendido como um símbolo correspondente a toda uma área de variação considerável no espaço acústico definido por F1 e F2, que compreende sons cujo F1 pode ter uma frequência mais alta do que de [ɛ]." (Andrade, 2020, p. 3255)

[46] Although a line of research proposes the existent of phonological /ɐ/ and /ɨ/ (Veloso, 2012), it is not within the scope of this dissertation to discuss this approach in further detail.

of speech. However, these cases are not within the scope of this dissertation.[47] As Figure 2.3 shows, vowel reduction results in a simplification of the vowel system at the surface level: the seven categories are realized as only four phones, when unstressed.

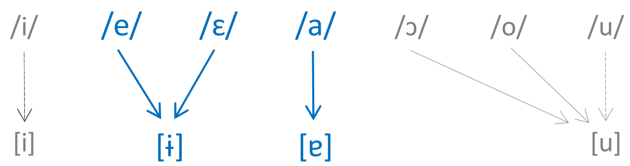


Figure 2.3. Vowel reduction in EP

It must be mentioned that the process of vowel reduction is not homogeneous (Figure 2.4). In EP, while /a/ undergoes raising into [ɐ], the production of /ɛ/ and /e/ as [ɨ] entails first a change in tongue height, follow by centralization.

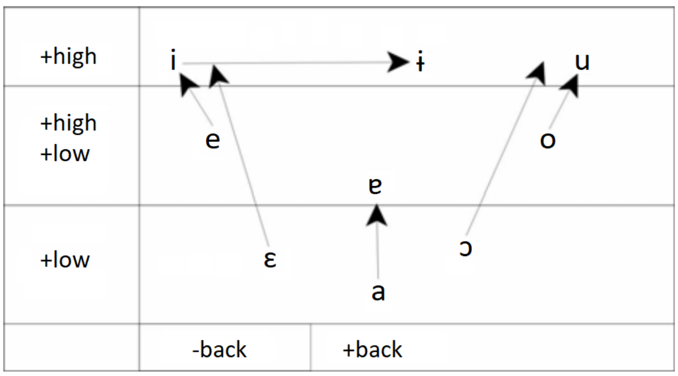


Figure 2.4. Vowel reduction processes in EP (Mateus, 2003)

Additionally, it's important to note that while /i/ and /u/ do not undergo vowel reduction in unstressed position, this does not mean that there are no quality changes when stress is absent. However, these changes do not entail significant acoustic and articulatory variation. The words in (1) exemplify the reduction of /a/, /ɛ/, and /e/, both in

[47] The importance of this distinction between *vowel reduction* and *phonetic reduction* (in Portuguese: *redução vocálica* and *redução fonética*, respectively) is attested by the mention of these phenomena, separately, in the Portuguese version of the Common European Framework (Conselho da Europa, 2001), in the section about *Phonologic competence*.

pre- and post-stressed position. We will signal the stressed syllable in bold and underline the vowel we wish to point out.[48]

(1a) /a/ → [ɐ] ***ca***_sa_ [ˈkazɐ] 'house'
*casa**men**to* [kɐzɐˈmẽtu] 'marriage'

(1b) /e/ → [ɨ] ***se***_de_ [ˈsedɨ] 'thirst'
*se**den**to* [sɨˈdẽtu] 'thirsty'

(1c) /ɛ/ → [ɨ] ***cé***_lebre_ [ˈsɛlɨbrɨ] 'famous'
*celebri**da**de* [sɨlɨbriˈdadɨ] 'celebrity'

We must add two further observations, regarding [ɐ]. First, this vowel can also result from an underlying /e/. This is due to a process of dissimilation before the coronal vowel [i] or the coronal consonants [ʒ], [ʃ], ɲ], and [ʎ] (e.g., *a**be**lha* [ɐˈbɐʎɐ] 'bee'). This restriction is mainly observed in Standard EP (the central coast dialect). Second, [ɐ] may occur in stressed position when /a/ is placed before the nasal consonants [m], [n] or [ɲ] (e.g., ***ca***_ma_ [ˈkɐmɐ] 'bed').[49]

Before describing EP stress, a word on frequency and orthography. Figure 2.5 (next page) shows the frequency of each EP phonetic vowel in stressed and unstressed positions (Frota et al., 2010). As we can see, [ɐ] is the most frequent vowel in EP, and it occurs mostly in unstressed position. As for [ɨ], this vowel can only occur in unstressed position.

As for orthography, previous studies place Portuguese in the center of the orthographic depth spectrum, close to French and German (Justi et al., 2023; Seymour et al., 2003). Concerning the target vowels of this dissertation, while [ɨ] is represented as <e> and, more rarely, as <i>, [ɐ] can correspond to the graphemes <a>, <â> or <e>. It is worth mentioning that, in EP, among all vocalic graphemes, <e> has the bigger number of possible phonetic realizations (Figure 2.6, next page).

[48] Henceforward in this chapter, we will use this convention.

[49] This restriction is not observed in the 3rd person plural of simple past indicative, due to the need of lexical contrast between this form and the 3rd person plural of the indicative present tense (e.g., *fa**la**mos* [faˈlɐmuʃ] 'we speak' vs. *fa**lá**mos* [faˈlamuʃ] 'we spoke').

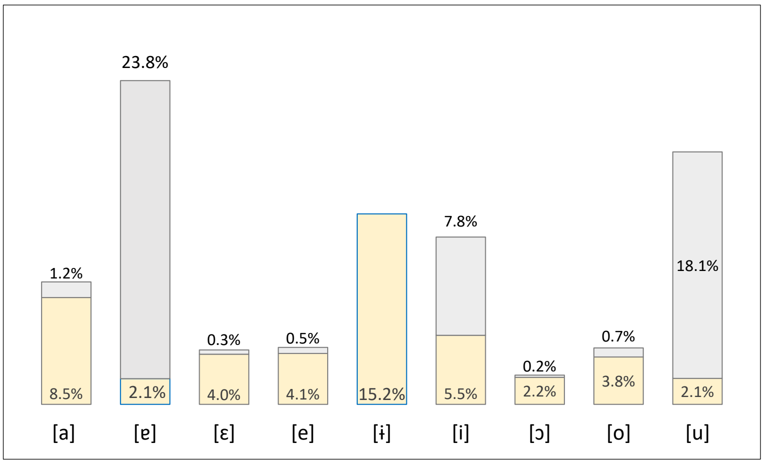


Figure 2.5. Frequency of the oral vowels in EP speech, in stressed (in yellow) and unstressed (in grey) position. Percentages were calculated from the total occurrences of oral vowels, based on Frota et al. (2010)[50]

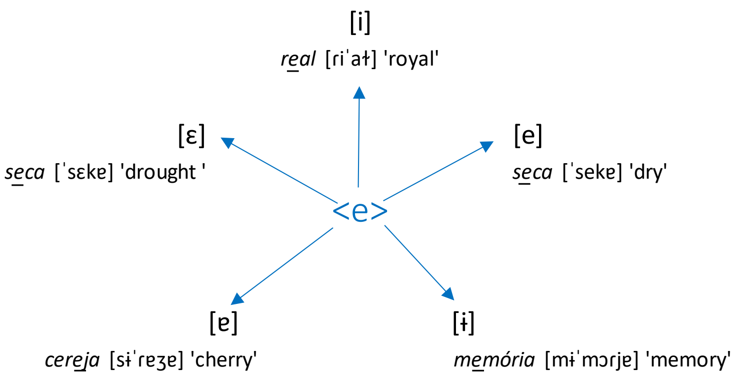


Figure 2.6. EP phonetic realizations for the grapheme <e>

To complicate this picture, <e> is frequent in homographic words with the pair/ɛ/-/e/, as for example, *sede*: [ˈsɛdɨ] 'headquarters' *vs*. [ˈsedɨ] 'thirst'; *este*: [ˈɛʃtɨ] 'east' *vs*. [ˈeʃtɨ] 'this'; colher: [kuˈʎɛr] 'spoon' *vs*. [kuˈʎer] 'to collect'; *acerto*: [ɐˈsɛrtu] 'I adjust' *vs*. [ɐˈsertu] 'adjustment'.

[50] The values displayed in the *corpus* FrePOP (Frota et al., 2010) were obtain from oral and written texts, formal and informal productions.

2.1.2. Word stress in European Portuguese

Latin was a weight-sensitive and fully predictable stress language, with stress falling at the right edge of the words, without, however, being placed on the last syllable: Latin polysyllabic words were paroxytonic or proparoxytonic, with only monosyllables being oxytonic. In its evolution to Romance, several changes occurred. Among them was the loss of vowel length contrasts and declensions, which implied a simplification of the word structure and led to shorter words.[51] Along these changes, new features of word stress appeared, as for example, the emergence of oxytonic polysyllables. Similar to Spanish and Italian, prosodic words in Portuguese can bear stress on three possible locations: last syllable (oxytonic or acute words), penultimate syllable (paroxytonic or grave words) or antepenultimate syllable (proparoxytonic words), as exemplified in (2a), (2b), and (2c), respectively. This is known as the *three-syllable window rule* and is considered universal.

(2a) Oxytonics words:
*ca*_**fé**_ [kɐˈfɛ] 'coffee'; *capa*_**taz**_ [kɐpɐˈtaʃ] 'foreman'; *par*_**ti**_ [pɐɾˈti] 'I left'

(2b) Paroxytonic words:
_**sa**_*co* [ˈsaku] 'bag'; *ra*_**pa**_*zes* [ʀɐˈpazɨʃ] 'boys'; *du*_**vi**_*da* [duˈvidɐ] 'he/she doubts'

(2c) Proparoxytonic words:
_**sá**_*bado* [ˈsabɐdu] 'Saturday'; _**ál**_*cool* [ˈalkuɔɫ] 'alcohol'; _**dú**_*vida*[ˈduvidɐ] 'doubt'

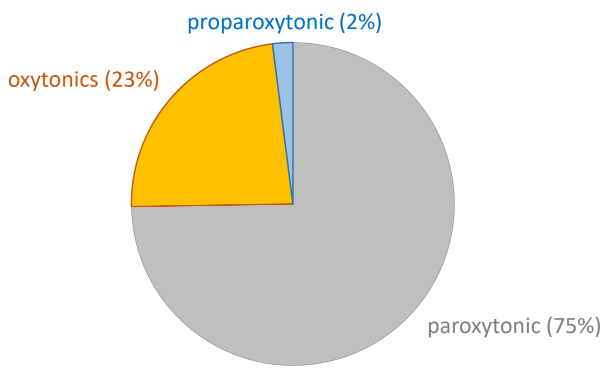


Figure 2.7. Distribution of EP word stress (Frota et al., 2010)

[51] While Classical Latin was a synthetic language, Vulgar Latin tended to an analytic structure.

Most Portuguese words are paroxytonic (stressed on the penultimate syllable) – about 74%. Oxytonic words represent around 23% of the lexicon, and only 2% are proparoxytonic (Frota et al., 2010; Figure 2.7, in the previous page).

EP stress is contrastive, that is, the position of stress determines or contributes to meaning contrasts, as shown in (3)

(3a) ***tú**nel* [ˈtunɛɫ] 'tunnel' *vs*. *to**nel*** [tuˈnɛɫ] 'barrel'

(3b) ***bam**bo* [ˈbɐ̃bu] 'lax' *vs*. *bam**bu*** [bɐ̃ˈbu] 'bamboo'

EP primary stress is determined mainly by morphological and lexical aspects (Castelo, 2005; Pereira, 2020; Vigário, 2003), and non-verbs and verbs must be described separately: in the later, stress domain is the lexical word whereas in non-verbs words the domain is the root.[52] Morphologically, nouns and adjectives exhibit a structure of [root]+[class marker]+[number].[53] In regular cases, stress falls on the last vowel of the root. This is considered the 'regular stress rule' for EP, since the majority (about 80%) of non-verbs concurs with it. The application of the regular stress rule may result in paroxytonics or oxytonics words:

(i) non-verbs ending in <o>, <a> or <e> (phonetically produced as [u], [ɐ] or [ɨ], respectively) are paroxytonics (e.g., ***ga**to* [ˈgatu] 'cat': gat]$_{\text{ROOT}}$ + o]$_{\text{CLASS MARKER}}$);

(ii) non-verbs ending in consonant, except for the plural marker /s/, are usually oxytonic. In these cases, an underlying class marker is assumed to exist, although not phonetically realized in the singular form (e.g., *ani**mal*** [ɐniˈmaɫ] 'animal': animal]$_{\text{ROOT}}$ + _]$_{\text{CLASS MARKER}}$);

[52] In studies on Portuguese stress, two approaches are debated: the morphology-based perspective and the weight-sensitive algorithm. Both approaches present supporting evidence, and both have been debated recently. However, for the purposes of the present dissertation, the morphological approach is adopted. The reason for this choice is the fact that when instructions are provided in the classroom context, the rules for Portuguese stress are usually supported by morphological explanations.

[53] In Portuguese, /a/, /o/, and /e/ are the class markers. These are always unstressed, and surface as [ɐ], [u], and [ɨ], respectively, or in Ø (as zero). In general, [ɐ] identifies feminine words, while [u], [ɨ] and Ø denote masculine words. /s/ is the marker of plural, and it's produced as [ʃ], [ʒ] or [s].

(iii) non-verbs ending in oral or nasal diphthongs, or in full vowels, are usually oxytonics. In these cases, it is assumed that no class marker exists (e.g., *chami**nés*** [ʃɐmiˈnɛʃ] 'chimneys': chaminé]$_{ROOT}$ + _]$_{CLASS\ MARKER}$ + s]$_{NUMBER\ MARKER}$).

In addition, loanwords are almost always adapted following the regular rule, a fact that strengthens the idea of the regular stress pattern (e.g., *estore* [ˈʃtɔɾ] 'roller blind', adapted from the English word *store*; the Portuguese word follows the rule displayed in (ii)).

There are also less frequent cases of non-verbs which do not follow the regular stress pattern. However, most of these words present an underlying rule: when the last vowel of the root is marked as unstressed, then stress falls on the penultimate syllable of the root (e.g., ***plás**tico* [ˈplaʃtiku] 'plastic': plástic]$_{ROOT}$ + o]$_{CLASS\ MARKER}$). Non-verbs which do not comply with this rule, nor with the regular EP stress, are rare, and native EP speakers tend to attenuate their irregularity phonetically, producing them as paroxytonic (e.g., ***ál**cool* [ˈaɫkuɔɫ] 'alcohol' produced as [ˈaɫkɔɫ]: álcool]$_{ROOT}$).

With respect to verbs, stress is assigned depending on tense. Verbs are morphologically structured as [root] + [theme vowel] + [tense/mode/aspect] + [person/number].[54] In past tenses, stress falls on the syllable of the theme vowel, irrespective of its position (e.g., *can**tas**te* [kɐ̃ˈtaʃtɨ] 'you sang': cant]$_{R}$ + a]$_{TV}$ + _]$_{TMA}$ + ste$_{PN}$).[55] The infinitive, gerund, and participle also comply with this rule. In future and conditional forms, stress falls on the tense marker (e.g., *canta**rás*** [kɐ̃tɐˈɾaʃ] 'you will sing': cant]$_{R}$ + a]$_{TV}$ + ra]$_{TMA}$ +s$_{PN}$). As for present tenses, it appears that in these cases stress is not linked to any of the morphologic constituents, but rather falls on the penultimate syllable (e.g., ***can**tas* [ˈkɐ̃tɐʃ] 'you sing': cant]$_{R}$ + a]$_{TV}$ + _]$_{TMA}$ +s$_{PN}$).[56]

Whereas EP content words (nouns, verbs, adjectives, and some adverbs) are always stressed, the majority of function words are unstressed (e.g., ***dá**-me* [ˈdamɨ] 'give-me').

[54] Portuguese verbs are divided in three conjugational groups, according to the theme vowel: /a/ (*cantar* 'to sing'), /e/ (*bater* 'to hit') and /i/ (*partir* 'to leave').
[55] R = root, TV = theme vowel, TMA = time/mode/aspect marker, and PN = person/number marker.
[56] Since EP imperative conjugation derives from present indicative forms, it presents the same stress pattern.

When function words are stressed, most of them comply with the regular stress rule (e.g., *como* [ˈkomu] 'as').

As for the acoustic properties of stress, and as described in Chapter 1, in stress languages, different cues may be used to signal differences between prominent and non-prominent syllables: duration, intensity, pitch, and vowel quality. Despite the importance of word stress in EP, very few authors have focused on the acoustic cues for stress in this language.

The acoustic parameters of EP stress were empirically observed by Delgado-Martins (1986), through an analysis of Portuguese real words in production. To determine EP stress acoustic features, the author conducted acoustic analyses of EP triplets (examples in (4)), each of them recorded by native Portuguese speakers in three conditions: isolation, inserted in a fixed carrier sentence and inserted in a free sentence. For each syllable, values of f0, duration, intensity, and energy were collected (where energy is the intensity over time).

(4a) *ex**plí**cito* [ʃˈplisitu] 'explicit'
*expli**ci**to* [ʃpliˈsitu] 'I make explicit'
*explici**tou*** [ʃplisiˈto] 'he/she made explicit'

(4b) ***exér**cito* [iˈzɛrsitu] 'army'
*exer**ci**to* [izɨrˈsitu] 'I exercise'
*exerci**tou*** [izɨrsiˈto] 'he/she exercised'

The author concluded that, when vowel reduction is absent, duration and energy are the main acoustic cues for stress, but only when it falls on the antepenultimate or the last syllable of the word (i.e., in case of irregular stress, the marked pattern). If stress falls on the penultimate syllable (regular stress, unmarked pattern), there are no systematically appearing acoustic cues. This indicates that the basic rule in EP is manifested in the acoustic dimension, that is, stress is cued if the word does not comply with the regular stress pattern (the marked case). Results of Delgado-Martins (1986) showed that vowel reduction is the main cue for stress, with no need for other cues. Additionally, the author found that f0 and intensity are not cues to word stress, but rather operating at the phrase and sentence level.

These EP stress parameters were later confirmed in perceptual experiments by the same author.

Castelo (2005) investigated the role of lexical and morphological knowledge, as well as the role of phonological cues (vowel reduction) in EP native speakers' perception of stress. In the study, the author presented an auditory corpus consisting of 100 real words. To observe differences between non-verbs and verbs in the activation of stress cues, the author included nouns and adjectives, as well as verbs. The stimuli were recorded by two female speakers, one from the EP variety, and the other from the Brazilian Portuguese variety.[57] Ten native Portuguese speakers were asked to identify stress. The results showed that Portuguese native speakers could correctly identify stress location in most of the cases, although they presented some difficulties with the Brazilian corpus. Regarding morphological and lexical cues, the author concluded that "word primary stress is more easily identified in words with a marked stress pattern or a suffix lexically marked for stress (i.e. when the lexical exceptions override the morphological structure)" (Castelo, 2005, p. 163). Results also emphasized the importance of vowel reduction in stress perception by EP listeners: reduced vowels were mostly identified with the absence of stress, and, on the contrary, full vowels were identified as being stressed.

The importance of vowel reduction for stress perception by EP native speakers was also attested in a perceptual study by Correia et al. (2015). Three experiments were conducted by the researchers: an AXB task and two sequence-recall tasks, with and without vowel reduction. The authors used disyllabic and trisyllabic nonwords, recorded in citation form and within a carrier sentence. Following previous cross-linguistic studies, phoneme contrasts were used as control condition. Stimuli are exemplified in (5).

(5a) stress contrast without vowel reduction: [ˈmipu]/[miˈpu]; [ˈdɐmitu]/[dɐˈmitu], [dɐˈmitu]/[dɐmiˈtu]

(5b) stress contrast with vowel reduction: [ˈnemi]/[nɨˈmi]

(5c) phonemic contrast: [ˈsiɾu]/[ˈseɾu]; [ˈmupɐ]/[ˈmunɐ]

[57] The author used tokens produced by the two varieties, since they differ in the way vowel reduction operates. Additionally, from the 83 non-verbs forms used, 17 were compounds, which were included to observe the effect of secondary stress.

Results showed that, in the absence of vowel reduction, Portuguese native speakers had difficulties discriminating stress contrasts, performing similarly to French speakers (Dupoux et al., 1997, 2001; Peperkamp et al., 2010). By contrast, when vowel reduction was present, Portuguese participants made significantly less errors. Overall, results suggest that, in the absence of vowel reduction, Portuguese listeners display stress 'deafness'.

2.2. Phonetics and phonology of Hungarian

Hungarian ("magyar") is a Finno-Ugric language, spoken by over fourteen million people: ten million within the Hungarian borders, three million living in the surrounding countries as Hungarian speaking minorities, and more than a million in other countries all over the world.

Hungarian vowel inventory comprises fourteen phones [ɒ], [aː], [ɛ], [eː], [i], [iː], [o], [oː], [ø], [øː], [u], [uː], [y], and [yː] (Figure 2.8). These phones correspond, respectively, to the phonemes /ɒ/, /aː/, /ɛ/, /eː/, /i/, /iː/, /o/, /oː/, /ø/, /øː/, /u/, /uː/, /y/, and /yː/.

	tongue position							
	front				central		back	
	unrounded		rounded		unrounded		rounded	
openness	short	long	short	long	short	long	short	long
close	[i]	[iː]	[y]	[yː]			[u]	[uː]
close-mid		[eː]	[ø]	[øː]	[ə]		[o]	[oː]
open-mid	[ɛ]							
open						[aː]	[ɒ]	

Figure 2.8. Phonetic vowel system of Hungarian, adapted from Markó (2017, p. 156)[58]

In addition to the features that are traditionally used for classification/description of vowels (tongue height, backness and lip position), in Hungarian, phonological length is also distinctive (e.g., örül [ˈøryl] 'rejoiced' vs. *őrül* [ˈøːryl] 'getting crazy'). On this dimension, we can readily observe five short-long pairs contrasting only in length (examples in (6)), and

[58] The schwa ([ə]) does not correspond to a phoneme and has no representation in the Hungarian orthography, as the word *gnu* [gənu] 'gnu' exemplifies (Markó, 2017).

two pairs contrasting in length but also in quality (examples in (7)) (Markó, 2017; Siptár & Törkenczy, 2000; Szende, 1990).

(6a) [i]-[iː]: *hidra* [ˈhidrɒ] 'beckon' vs. *hídra* [ˈhiːdrɒ] 'to (the) bridge'

(6b) [o]-[oː]: *koma* [ˈkomɒ] 'friend' vs. *kóma* [ˈkoːmɒ] 'coma'

(6c) [ø]-[øː]: *növel* [ˈnøvel] 'increases' vs. *nővel* [ˈnøːvel] 'with (a) women'

(6d) [u]-[uː]: *szurok* [ˈsurok] 'tar' vs. *szúrok* [ˈsuːrok] 'I stab'

(6e) [y]-[yː]: *büntettek* [ˈbyntɛtːɛk] 'they punished' vs. *bűntettek* [ˈbyːntɛtːɛk] 'crimes'

(7a) [eː]-[ɛ]: *kéz* [ˈkeːz] 'hand' vs. *kezek* [ˈkɛzɛk] 'hands'

(7b) [aː]-[ɒ]: *nyár* [ˈɲaːr] 'summer' vs. *nyarak* [ˈɲɒːrɒk] 'summers'

Hungarian orthography is highly transparent (Borgwaldt et al., 2004), and even the short-long vowel system is unambiguously represented in the orthographic system. Short vowels may only be marked by overdot, umlaut or have no diacritics, while long vowels are marked by acute or double acute diacritics (Table 2.1).

Table 2.1. Phone-grapheme correspondence in Hungarian, according to vowel length

short vowels		long vowels	
[ɒ]	<a>	[aː]	<á>
[ɛ]	<e>	[eː]	<é>
[i]	<i>	[iː]	<í>
[o]	<o>	[oː]	<ó>
[u]	<u>	[uː]	<ú>
[ø]	<ö>	[øː]	<ő>
[y]	<ü>	[yː]	<ű>

As for the frequency of occurrence (Figure 2.9, next page), [ɛ] and [ɒ] are the most frequent vowels (above 20%), followed by [o], [i], [aː] and [eː] (between 7.3% and 11.9%). The remaining vowels have a frequency lower than 3.3%. It is interesting to note that short vowels are considerably more frequent than long vowels: 78.9% against 21.1%.

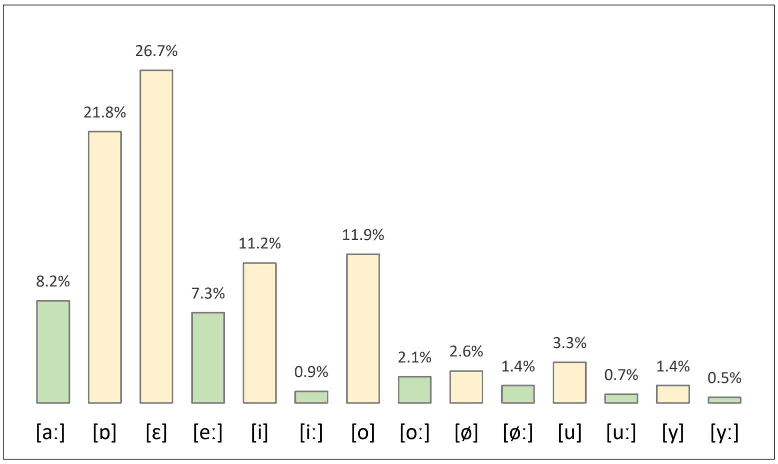


Figure 2.9. Frequency of short vowels (in yellow) and long vowels (in green) in Hungarian speech. Percentages calculated from the total occurrences of vowels, based on Gósy (2004, p. 89)[59]

Although not directly related to the topic of this dissertation, it should be mentioned that Hungarian shows vowel harmony, since it is its most broadly debated feature. In few words, in Hungarian, harmony refers to vowels in the stem agreeing in backness. This harmonic behavior expands to the suffixes. Suffixes in Hungarian can present two or three forms (e.g., *-ban*/*-ben* [bɒn]/[bɛn] 'in'; -hoz/-hez/-höz [hoz]/[hɛz]/[høz] 'to'), and the selection of the form is governed by the last vowel of the stem. In the case of three-form suffixes, roundness harmony is also at work. In most cases, stems include only back or front vowels, and are suffixed accordingly, as the examples in (8a) show. Stems that include both front and back vowels are suffixed according to the last vowel in the stem (8b). Stems that include [i], [iː], [ɛ], and [eː] in this last position may induce back suffix selection, and for this reason they are considered to be neutral/transparent: they do not themselves govern suffix selection, but they lend this role to the vowel that precedes them (and if it is also neutral, to the vowel that precedes that vowel and so on), as in example (8c).

(8a) *tükör*]$_{\text{STEM}}$*től* [ˈtykørtøːl] 'from the mirror'
város]$_{\text{STEM}}$*tól* [ˈvaːroʃtoːl] 'from the city'

[59] These values were calculated from a corpus consisting of monologues and dialogues (spontaneous speech), produced by 12 Hungarian speakers.

(8b) *sofőr*]STEM*től* [ˈʃoføːrtøːl] 'from the chauffeur'

(8c) *kastély*]STEM*tól* [ˈkaːʃteːjːtoːl] 'from the castle'

In citation form, and with very few exceptions (such as interjections), Hungarian words are stressed in the first syllable, independently of being simple (9a), derived (9b) or compound words (9c).

(9a) *ember* [ˈɛmbɛr] 'man'

(9b) *emberek* [ˈɛmbɛrɛk] 'men'

(9c) *fűnyíró* [ˈfyːɲiːroː] 'lawnmower'

Word stress is not contrastive in Hungarian, and studies have shown that Hungarian listeners present a high level of stress ‘deafness’ when it comes to perception of stress contrasts (Honbolygó et al., 2017; Peperkamp et al., 2010; Peperkamp & Dupoux, 2002).

Traditionally, f0 and intensity are regarded as acoustic cues of stress in Hungarian in production (Fónagy, 1958). In earlier studies these cues were also assumed to be the most important in perception, but more recent studies point to a bigger role of f0 than intensity. It must be noted that in Hungarian, sentence level accent and words stress always align (Szalontai et al., 2016), a fact that curtails the study of word stress on itself.

Finally, traditionally there is also a consensus that vowel quality does not undergo variation as a result of present/absence of stress. However, Markó et al. (2019) alludes to the absence of reliable studies on this matter.

2.3. Comparison between the Portuguese and Hungarian systems

In sum, both Hungarian and EP have a rich vowel system, but they present a paramount difference. In EP, the quality of the vowel cues their stressed or unstressed status: [a], [ɛ], [ɔ], and [o] are mostly perceived as stressed, [ɐ] and [ɨ] are perceived as unstressed, and [i] and [u] are perceived as stressed or unstressed. By contrast, in Hungarian, vowel quality do not relate to stress and vowels are contrastively divided into short ([ɒ], [ɛ], [i], [o], [ø], [u], [y]) and long vowels ([aː], [eː], [iː], [oː], [øː], [uː], [yː]). Apart from a different pattern between the two vowel systems (*stressed-unstressed* in EP vs. *short-long* in Hungarian), in the comparison between the two inventories we observe that the EP vowels [ɐ] and [ɨ] are absent from the Hungarian inventory, and the Hungarian vowels [ø] and [y] are not available in EP.[60]

As for word stress, it is not a contrastive feature in Hungarian, and consequently, native speakers exhibit 'deafness' when it comes to the perception of stress contrasts. In EP, word stress is not only lexically contrastive, but also impacts vowel quality (e.g., *exército* [iˈzɛrsitu] 'army' vs. *exercito* [izɨrˈsitu] 'I exercise'). Furthermore, acquiring EP stress entails morphological and lexical knowledge, which only comes with the expansion of vocabulary, and it is also influenced by metalinguistic knowledge, as suggested by Castelo (2005).

We must mention the issue of duration. As we described, this acoustic cue is used to signal EP stress only in marked cases (i.e., when stress does not follow the regular cases) or/and if the word does not include reduced vowels (e.g., *túnel* [ˈtunɛɫ] 'tunnel'). Contrarily, in Hungarian vowel length is contrastive (e.g., *örül* [ˈøryl] 'rejoiced' vs. *őrül* [ˈøːryl] 'getting crazy'). In order words, in EP, duration signals relevant contrasts at suprasegmental level whereas in Hungarian, duration operates lexically at segmental level.

Finally, it is worth mentioning that as opposed to Hungarian, EP orthography is not transparent. A vowel (but also many consonants) can be represented by different graphemes (e.g., *cer<u>e</u>j<u>a</u>* [sɨˈr<u>ɐ</u>ʒ<u>ɐ</u>] 'cherry') and one grapheme may have different phonetic realizations (e.g., *s<u>e</u>d<u>e</u>* [ˈsedɨ] 'thirst' or [ˈsɛdɨ] 'headquarters'). The rules behind the phone-grapheme correspondence are phonological and dependent of the presence/absence of stress, or, in some cases, are determined by etymological aspects.

Having described the phonetic and phonological relevant aspects of the L1 and the L2 systems, and supported by the literature review presented in Chapter 1, we can now establish the hypotheses for our work.

[60] Recall that we only refer to standard EP, the norm to which learners of L2 Portuguese are more exposed. [y] is present in the dialect of the island of São Miguel (Azores), for example.

CHAPTER 3

RESEARCH QUESTIONS AND HYPOTHESES

The review of the studies in L2 perception shows that some segmental and suprasegmental features are easier to perceive than others. Findings strongly suggest that acoustic (dis)similarity in vowels determines difficulty or ease in perception. The *stress parameter hypothesis* (Peperkamp & Dupoux, 2002) predicts that speakers of languages that encode stress as non-contrastive in L1 acquisition (e.g., Hungarians) will exhibit stress 'deafness' in L2 perception. While in the case of L2 vowels, research points to the positive effects of training, findings for stress acquisition are less clear. Additionally, studies on L2 targeting languages which display variable stress and vowel reduction crucially exclude the issue of vowel reduction.

As previously pointed out, in EP, word stress and vowel quality are connected, since the latter (full vs. reduced vowels) is determined by the presence/absence of stress. To the best of our knowledge, the question on *how do speakers who have a different vowel inventory, predictable stress (thus, expected to be stress 'deafness'), and with no stress induced vowel reduction acquire the phonology of an L2 which includes these features* has not yet been addressed in existing research.

In this dissertation, we aim at answering the above question by observing Hungarian native speakers learning Portuguese. As described in Chapter 2, in the Hungarian vowel inventory, the EP reduced vowels [ɐ] and [ɨ] are absent. As for word stress, while in EP it is variable, in Hungarian is fixed (on the first syllable). Furthermore, EP word stress is accompanied by stress-induced vowel reduction, whereas in Hungarian, it is not. In the sections below, the research questions guiding this dissertation will be presented and motivated.

3.1. How do Hungarian learners categorize the EP reduced vowels [ɐ] and [ɨ], at the onset of learning?

In Chapter 1 we reported cross-language models which posit that L2 perception is modulated by the L1 phonological system, and that L2 sounds are perceived in relation to the closest L1 categories. However, these models differ in their approach. While PAM predicts L2 perception based on articulatory-phonetic similarity, SLM and the L2LP model construct their frameworks on acoustic similarities/dissimilarities. Elvin et al. (2021) point out the close relationship between acoustics and articulation, suggesting that both methods are equally pertinent. L1-L2 acoustic comparisons have been widely used to predict and analyze how listeners map non-native sounds (Chládková & Podlipský, 2011; Elvin et al., 2014, 2021; Escudero & Vasiliev, 2011; Flege et al., 1994; Mayr & Escudero, 2010). Considering the above-mentioned reasons, we establish our predictions on comparisons of the F1×F2 vowel space between the L1 (Hungarian) and the L2 (EP).

To conduct the acoustic comparison between the Hungarian and the EP vowel spaces, we placed the EP target vowels, [ɐ] and [ɨ], in the Hungarian perceptual vowel map (Figure 3.1). The values for the EP vowels were calculated from Andrade (2020, p. 3251), and are reported in Appendix IV. The values for Hungarian data were retrieved from Kiss (1985, p. 166) and modified according to Markó (2017).[61]

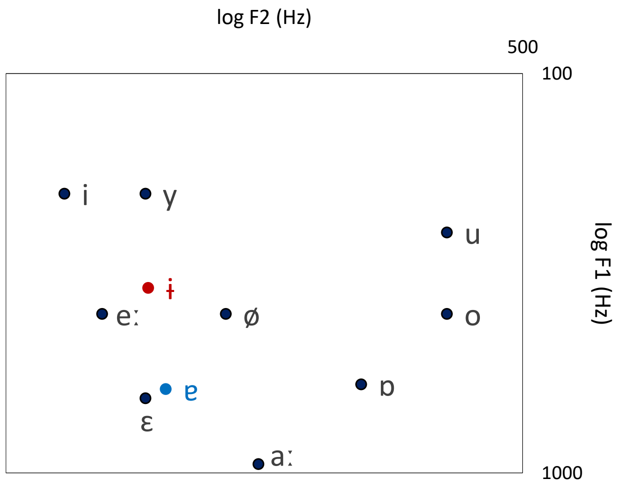


Figure 3.1. Location of EP [ɐ] and [ɨ] in the Hungarian perceptual vowel map

As shown in Figure 3.1., /ɛ/ is the closest Hungarian category to the EP vowel [ɐ], which can explain Hungarian learners' tendency to write the reduced vowel [ɐ] as <e>, a grapheme that unequivocally represents /ɛ/ in Hungarian, as previously explained. The next

[61] Henceforward, when we mention the Hungarian perceptual map, we will refer to the vowel space with values for Hungarian perception data retrieved from Kiss (1985, p. 166) and modified according to Markó (2017). Additionally, references to standard EP vowels are always based on the values retrieved from Andrade (2020, p. 3251), as reported in Appendix IV.

closest categories are /eː/, /ø/ and /aː/. Considering that the later differs from [ɐ] in the height of the tongue relative to the palate – [aː] is low, while /eː/and /ø/ are mid vowels, as [ɐ] –, we expect Hungarian /eː/ and /ø/ to be the alternate choices in the perception of [ɐ].

As for [ɨ], this vowel is located in the space between /y/, /eː/, /ɛ/ and /ø/, therefore we expect these categories to be used by Hungarian listeners.

Hypothesis 1

Hungarian learners of L2 Portuguese, at the onset of learning, categorize the EP [ɐ] into /ɛ/, /eː/ or /ø/, and the EP [ɨ] into /y/, /eː/, /ɛ/ or /ø/.

3.2. Does a perceptual training improve discrimination abilities with the reduced vowels [ɐ] and [ɨ]?

The studies reviewed in section 1.3.2 demonstrate that L2 vowel perception improves immediately after training. Therefore, we expect a similar outcome when training Hungarian learners of EP, that is, we expect that Hungarian learners show improvement in the perception of the EP vowels [ɐ] and [ɨ], after a training program. Moreover, according to PAM-L2, the onset of learning is the ideal moment for phonetic attunement, since vocabulary acquisition and morphosyntactic properties may curtail perceptual attunement (Best & Tyler, 2007).

It is worth mentioning that phonetic attunement is not equivalent to phonological acquisition, and that L2 segmental acquisition entails the perception of relevant lexical L2 contrasts (Best & Tyler, 2007). This assumption is shared by the L2LP model, in which it is proposed that acquisition involves not only auditory-driven learning, but also lexicon-driven learning. Thus, the advantage that learners may have in phonetic attunement at the onset of learning is counterbalanced with the disadvantage of the none or negligible amount of L2 word knowledge. The lack of a solid basic L2 lexicon in learners affects the possibilities of the training itself. As mentioned in Chapter 1, studies have shown that the combined use of identification and discrimination tasks is ideal to promote learning. Moreover, the study of Carlet (2017) demonstrates that students who are trained with identification tasks achieve better scores in the generalization test for L2 vowel perception than the students who train with discrimination tasks, which suggests that identification tasks are more efficient in promoting long-term learning. However, in our training, identification tasks cannot be included, since for these, a stable knowledge of a minimum L2 lexicon is needed. This limitation in the training methodology may thus restrict the level of improvement in the perception of the target EP vowels as well.

In Chapter 1 (section 1.3.1) we detailed three cross-language models which establish predictions for the difficulty/ease for L2 vowel perception and acquisition, based on the identification of the L2 segments into the L1 system. Both L2LP and PAM-L2 posit that when learners map a pair of sounds that contrast in L2 into just one single L1 category (the *new scenario*, in the L2LP model, or a *single-category* assimilation, in PAM), learners present more difficulties in acquisition. According to the L2LP model, in this scenario, learners have a perceptual task, but also must adjust the categories in terms of lexical representations. A different learning path emerges when learners map the contrasting L2 sounds in more than two L1 categories. In the L2LP model, this corresponds to a *subset scenario*, while in PAM, it is considered to be an *uncategorized (focalized, clustered or dispersed) assimilation*. In this case, the degree of difficulty in acquisition depends on the degree of perceptual overlap found for the sounds in question. The wider the overlap, the more difficult it is for learners to adjust categorization of the L2 sounds. However, this scenario should not be more difficult than the *new* scenario. Finally, both the L2LP model and PAM propose that when the L2 contrast is perceived as an L1 contrast – a *similar scenario* (L2LP) or *two-categories assimilation* (PAM) –, discrimination is not problematic, and learning consist only of a perceptual adjustment task. The SLM presents different predictions. Recall that the model approaches L2 perception from (dis)similarities between L1 and L2 sounds, and not contrasts. According to SLM, *new* categories are more likely to be created than *similar* categories. In sum, the three models propose different learning paths while agreeing on the importance of the (dis)similarity between L1 and L2 segments or contrasts in L2 speech perception. Another aspect to have in consideration is the target population they were designed to account. Both SLM and the L2LP models propose to explain L2 perception by L2 learners, while PAM focuses on non-native perception by naïve listeners. However, the focus of SLM's is the naturalistic learning process, whereas the L2LP model is more concerned with formal learning contexts.

Considering that (1) studies provide strong evidence on the positive effect of perceptual training in L2 vowel acquisition, (2) the low level of L2 knowledge limits the training tasks that can be used, which can also limit the improvement in training, (3) models on L2 perception differ in their approaches and in their predictions, and to date, no conclusive findings were published in which model is more adequate to explain L2 speech acquisition, and (4) observations in the classroom suggest a more persistent difficulty for [ɐ] than for [ɨ] (described in Introduction), we propose the following hypothesis:

> *Hypothesis 2*
>
> A perceptual training focusing on vowel discrimination promotes the acquisition of the EP vowels [ɐ] and [ɨ] by Hungarian learners, although to different extent for each of the vowels.

3.3. Does a perceptual training improve discrimination abilities with word stress?

With respect to L2 word stress acquisition, empirical evidence is scarce. In Chapter 2, we reported that in EP, stress is distinctive, whereas in Hungarian, it does not mark lexical contrasts. The consequence of this fact is that Hungarian speakers disregard stress contrasts, thus presenting stress 'deafness'. As reviewed in Chapter 1 (sections 1.4.1 and 1.4.2), results of perceptual experiments are conflicting. In studies with French listeners (who, similar to Hungarian speakers, present stress 'deafness'), researchers obtained different outcomes. Dupoux et al. (2008) tested French learners of L2 Spanish and did not find significant differences between learners with different levels of L2 knowledge. A similar result was obtained in Honbolygó et al. (2017). In this study, no significant differences in stress perception discrimination abilities were found between Hungarian Beginner and Advanced learners of Germain. Michaux (2016) observed French learners of Dutch L2 and found some evidence of learning. A comparison between results of the first block and the last block of the experiment showed that the French participants improved in trials that included contrasts of pseudowords in which stress location was manipulated (e.g., *PAgina vs.* **paGIna*).[62] Finally, Schwab & Llisterri (2014) also tested changes in the perception of Spanish by French speakers, but in this case participants were naïve listeners. A single training session was conducted, and the results showed improvement. However, a more extensive analysis suggested that progress was more dependent on individual learning abilities than an effect of training. Finally, two studies conducted a longer training program: Brawerman-Albini (2012) and Ou (2011). In Brawerman-Albini (2012) both learners' L1 (Brazilian Portuguese) and L2 (English) were variable stress languages. The results show that although participants improved from the pretest to the post-test, no generalization of learning was achieved. In the study of Ou (2011), with Chinese learners of L2 English, improvement was observed only in one of the conditions (when rising intonation cued word prominence). It must be pointed out that the participants in the studies of Brawerman-Albini (2012) and Ou (2011) had an intermediate or advanced knowledge of the L2. In sum, although empirical evidence is conflicting about the persistence of 'deafness' in perception of stress contrasts, the studies who included perceptual training (Brawerman-Albini, 2012; Ou, 2011) presented positive results, that is, learners were able to improve in stress perception. Considering this, we establish the following hypothesis:

> *Hypothesis 3*
>
> Hungarian learners benefit from perceptual training focusing on stress discrimination to overcome stress 'deafness'.

3.4. Is a perceptual training that combines word stress and vowels more effective in improving discrimination abilities of reduced vowels and word stress than a training that focuses on one of those features only?

[62] The other tested condition included nonwords, created from real words, with low segmental variability (e.g., **SAsasa* vs. **saSAsa*). According to the authors, the lack of phonotactic information might have contributed to the poor results in this condition.

As explained in Chapter 2, in EP, vowel reduction is a paramount feature, and stress is intrinsically connected to it. In the absence of vowel reduction, Portuguese listeners present 'deafness' in the perception of stress contrasts. Therefore, the link between stress and vowel quality is significantly reflected in the EP vocabulary. Although Hungarian L2 Portuguese learners are exposed to a limited amount of input in the classroom (when comparing with learners in a naturalistic context, for example), that input is, nevertheless, representative of the Portuguese phonological system and, consequently, learners are exposed to frequency and distributional patterns of EP stressed and unstressed vowels. We assume that a training with stimuli designed to combine stress and vowel discrimination will draw learners' attention to these properties of the L2. Moreover, given the above-mentioned link, reduced EP vowels cannot be acquired without mastering word stress. Having these facts in consideration, we postulate the hypothesis below:

Hypothesis 4

Hungarian learners of L2 Portuguese benefit more from a perceptual training that combines stress and vowel reduction than from a training focused on stress or vowels only.

3.5. Is there a hierarchy in L2 acquisition between segmental features and suprasegmental features?

The *full access hypothesis* posits that the mechanisms used for L1 acquisition are available and adopted for L2 acquisition as well, even in L2 late learners (Flege, 2005; Escudero & Boersma, 2004). Therefore, it seems legit to assume that a similar order of acquisition can be observed for L1 and L2 at the onset of learning. If so, according to the *phonological bootstrapping hypothesis* (Teixidó et al., 2018; van Heugten et al., 2014), learners rely on prosodic information, such as rhythm, intonation, and stress, for L2 speech segmentation, which will allow them to acquire L2 grammar and vocabulary. Although we have to account for the influence of formal learning context in which our study is conducted, the input provided in classroom reflects the properties of the L2, as already mentioned. In the case of EP, stress is located in the right-edge of the prosodic word. To be able to extract words from the speech continuum, learners will have to identify word boundaries, and word stress is a critical cue to it.

With this is consideration, we establish our 5th hypothesis:

Hypothesis 5

Hungarian learners of L2 Portuguese at the onset of learning focus first on the L2 suprasegmental features. Accordingly, stress precedes vowels in L2 acquisition.

In the present chapter we presented the hypotheses, grounded on the literature review reported in Chapter 1 and Chapter 2, for the research questions that guide our dissertation. Next, we will present Part II – Empirical studies, in which we described two experiments designed to answer the research questions, and accordingly, corroborate or refute the established hypotheses.

PART 2

EMPIRICAL STUDIES

CHAPTER 4

PERCEPTION OF EUROPEAN PORTUGUESE ORAL VOWELS BY HUNGARIAN NATIVE SPEAKERS

In Chapter 1, we described a selection of studies aimed at L2 vowel acquisition. These studies present perceptual trainings designed to help learners overcome difficulties previously observed. For example, in Aliaga-García (2017) and Rato (2013), the authors starting point were difficulties already attested in the perception of English vowel contrasts (e.g., /i/-/ɪ/). In the case of Hungarian learners of L2 Portuguese, to the best of our knowledge, no perceptual study was yet published; therefore, we don't have information on the difficulties those learners display with the EP vowels. Consequently, we must conduct a study to be able to collect the necessary data to design a perceptual training. The experiment we report here was design to that purpose, and took place between April and May 2021.[63]

Based on a comparison of the F1×F2 vowel spaces of the L1 and the L2, we established the following hypothesis for the perception of reduced vowels [ɐ] or [ɨ] by Hungarian listeners (see Chapter 3):

Hungarian learners of L2 Portuguese, at the onset of learning, categorize the EP [ɐ] into /ɛ/, /eː/ or /ø/, and the EP [ɨ] into /y/, /eː/, /ɛ/ or /ø/.

Although our aim is to observe the perception of the EP reduced vowels [ɐ] and [ɨ], we decided to test all EP oral vowels, and thus have a complete picture of how Hungarian listeners map the EP oral vowel system. Accordingly, we included the nine EP oral vowels: [a], [ɐ], [ɛ], [e], [ɨ], [i], [ɔ], [o], [u]. As for the Hungarian categories, as reported in Chapter 2, Hungarian system contains fourteen categories, seven long vowels and seven short vowels. However, there are only nine different vocalic qualities. Since we aimed at investigating perception on the basis of vowel quality, we only included the following Hungarian vowels: /ɒ/, /aː/, /ɛ/, /eː/, /i/, /o/, /ø/, /u/, and /y/.

Before reporting the study, a clarification is in order, regarding the terms *perception*, *identification*, *categorization*, and *assimilation*. As described in the beginning of Chapter 1, speech perception is inherently *categorical*, that is, speech sounds are

[63] The preliminary results of this study were published in Tavares et al. (2022).

perceived as linguistic categories. However, in this chapter, c*ategorization* will be used strictly as proposed by PAM (Best, 1995). According to this model, a sound is classified as *assimilated* (as speech sound) if it is identified above chance level, and it is classified as *categorized* if it is identified above a specific categorization threshold. As for *identification*, we will use this term to refer to results of the forced-choice identification task. In the context of the present study, *perception* is, thus, a broader concept that includes *assimilation*, *identification* and *categorization*.

4.1. Study design

To observe how Hungarian listeners map the Portuguese oral vowels in their L1 system, we designed a multiple forced-choice identification test with goodness-of-fit rating, a task frequently used in perceptual categorization studies in L2 Phonology (Bundgaard-Nielsen et al., 2011; Faris et al., 2018; Flege & Mackay, 2004; Mayr & Escudero, 2010). In these tasks, participants are first asked to match a non-native auditory stimulus among different L1 or L2 alternatives, usually presented in orthographic form. Considering that the target listeners of this study are naïve speakers (i.e., with no knowledge of the L2), it was only possible to present the answers to the participants in their L1. As for the goodness-of-fit rating scale, although studies usually employ odd-scales (e.g., Bundgaard-Nielsen et al., 2011; Cebrian, 2010; Faris et al., 2016, 2018), in our experiment we decided to use a four-point Likert scale (1 = *very bad example*, 2 = *bad example*, 3 = *good example*, 4 = *very good example*). This methodological decision was made to force participants to express a clearly positive (3 and 4) or negative (1 and 2) judgment and avoid the *face-saving don't know* effect (Sturgis et al., 2014), that is, the expression of *absence* of opinion rather than the expression of a *neutral* opinion. We are aware, however, that by excluding a midpoint, we may be biasing results towards a positive or negative assessment despite the fact that participants may have wanted to express neutral opinions.

Two groups of participants were recruited: a group of Hungarian native speakers having no knowledge of the L2, the experimental group (Group EXP), and a group of Portuguese native speakers, to control for stimuli nativeness (Baseline Group).

The experiment was divided into four parts: (1) informed consent, (2) background questionnaire and volume adjustment, (3) familiarization tasks, and (4) main trials.

The background information of the participants was collected at the beginning of the experiment, and included questions about age, gender, and academic studies, as well a section on L2 knowledge. Given the preliminary nature of this study, this section was concise and presented in the form of a multichoice questionnaire. In the case of Group EXP, we selected the languages most taught in Hungarian universities, in courses unrelated to Languages Studies, which were English, Germain, Spanish, French, Italian and Russian. To exclude participants with formal contact with Portuguese, we added Portuguese among the choices. We also established the possibility of four levels: *beginner*, *intermediate*, *advanced*, and *native level*. Participants could choose up to four L2, and when choosing an L2, the indication of the proficiency level was mandatory. Appendix II displays the questionnaires.

The second part consisted of the familiarization trials. The purpose of this part was to provide the participants the opportunity to get acquainted with the type of task they would have to complete in the main trials. In the present experiment, two familiarization tasks were created. The first task included forced-choice identification trials with feedback, and the second included the same trials followed by goodness-of-fit rating and without feedback. Each familiarization task included nine trials, randomly presented, with tokens produced in the participants' native language.

The third part of the experiment – the section containing the main trials – was similar to the second familiarization task: forced-choice identification trials followed by goodness-of-fit rating, without feedback. Auditory stimuli consisted of the nine EP vowels, produced by three native female talkers[64], in [gV] context, and the visual stimuli consisted of nine real words of the L1, also starting with [gV]. The main test included, thus, 27 trials (9 vowels × 3 talkers), which were randomized across participants.

[64] In this dissertation, we will use the term 'talker' to refer to a speaker who produces tokens for the experiments. We use 'speaker' in a broad sense (as talkers and listeners) and always referring to native speakers. We also use it as a synonym for listeners (e.g., Hungarian speakers = Hungarian native speaking listeners).

We created two versions of the experiment, one for Group EXP (exp-HU) and one for the Baseline Group (exp-PT). The two versions were structurally and graphically identical, but instructions as well as the auditory stimuli for the familiarization tasks were presented in the L1 of the target group (Hungarian for Group EXP and Portuguese for the Baseline Group). As for the auditory stimuli of the main trials, they consisted of tokens in EP (the target L2), in both groups. Regarding orthographic stimuli, they were always presented in the respective L1. A summary of the two versions is displayed in Table 4.1.

Table 4.1. Overview of the experiment

<table>
<tr><th>experiment version</th><th>section</th><th>task</th><th>language of the auditory stimuli</th><th>language of the orthographic stimuli</th><th>number of trials</th></tr>
<tr><td rowspan="6">exp-HU (EXP Group)</td><td colspan="5">consent
background questionnaire
volume adjustment</td></tr>
<tr><td>familiarization task 1</td><td>forced-choice identification (with feedback)</td><td>Hungarian</td><td>Hungarian</td><td>9</td></tr>
<tr><td>familiarization task 2</td><td>forced-choice identification and goodness-of-fit rating (with no feedback)</td><td>Hungarian</td><td>Hungarian</td><td>9</td></tr>
<tr><td>main trials</td><td>forced-choice identification and goodness-of-fit rating (with no feedback)</td><td>EP</td><td>Hungarian</td><td>27</td></tr>
<tr><td rowspan="4">exp-PT (Baseline Group)</td><td colspan="5">consent
background questionnaire
volume adjustment</td></tr>
<tr><td>familiarization task 1</td><td>forced-choice identification (with feedback)</td><td>EP</td><td>EP</td><td>9</td></tr>
<tr><td>familiarization task 2</td><td>forced-choice identification and goodness-of-fit rating (with no feedback)</td><td>EP</td><td>EP</td><td>9</td></tr>
<tr><td>main trials</td><td>forced-choice identification and goodness-of-fit rating (with no feedback)</td><td>EP</td><td>EP</td><td>27</td></tr>
</table>

Both versions were built in PsychoPy, version 2020.2.3 (Peirce et al., 2019) and hosted by *pavlovia.org*. Results were downloaded from *pavlovia.org* in csv format and raw data was prepared and analyzed in R (R Core Team, 2014).

Wilcoxon signed-rank tests were used to attest if identification means were significantly different than chance level (i.e., the probability of each L2 vowel to be identified as a specific L1 category) or significantly different than established categorization's thresholds. Additionally, we used Wilcoxon rank-sum tests to analyze goodness-of-fit ratings.[65] To assess talker effects on the results, we used χ^2 tests. We opted to exclude reaction times from the analysis, since this variable can be difficult to control on online tasks such as in this experiment, consequently presenting significant errors (Holden et al., 2019).

4.2. Participants

All participants were recruited online: a recruiting email with a link to the experiment was shared with teachers and students at Hungarian universities, who were asked to share it, in turn, with possible candidates. Participation criteria were: (1) to be aged between 18 and 45, (2) to be a native speaker of Hungarian (Group EXP) or EP (Baseline Group), and (3) to have access to a computer, internet, and headphones or earbuds. In the case of Portuguese participants, they should be from the dialectal area of Lisbon.

A total of 102 Hungarian native speakers completed the test. However, from these, 21 referred having previous or present contact with Portuguese through family, friends or attending a course. These participants were excluded from the analysis. Three other participants were not included due to incomplete or unclear background data (missing or duplicated information regarding L2). As for the Baseline Group, initially, 31 Portuguese speakers completed the task, but one reported an A2 proficiency level in Hungarian and therefore was excluded.

The final data considered for this experiment comprehended results from 108 participants: 78 from Group EXP and 30 from the Baseline Group. Table 4.2 (next page) lists the general information collected from the included participants.

[65] Although previous studies used *t-tests*, individual Shapiro-Wilk tests showed that our data were not normally distributed; therefore, we used Wilcoxon non-parametric tests.

Table 4.2. Participants' general information

	Age group					Gender			Academic degree			
	18 to 25	26 to 30	31 to 35	36 to 40	41 to 45	male	female	other	secon-dary studies	BA	MA	PhD
Group EXP (n=78)	27	14	20	9	8	27	51		22	12	25	19
Baseline Group (n=30)	17	7	2	1	3	8	21	1	8	12	9	1

Data on L2 knowledge from Group EXP is presented in Table 4.3. All participants reported English as L2, and the majority reported the *advanced level*. The second most spoken L2 was German, but contrary to the case of English, most of the participants reported only a *beginner level*.

Table 4.3. Reported spoken L2 in Group EXP

		n. of responses
English (n=78)	*beginner level*	4
	intermediate level	29
	advanced level	43
	native level	2
Germain (n=51)	*beginner level*	26
	intermediate level	15
	advanced level	10
French (n=16)	*beginner level*	5
	intermediate level	8
	advanced level	3
Italian (n=16)	*beginner level*	10
	intermediate level	4
	advanced level	2
Spanish (n=15)	*beginner level*	6
	intermediate level	5
	advanced level	3
	native level	1
Russian (n=5)	*beginner level*	2
	intermediate level	1
	advanced level	2

4.3. Materials

Forced-identification tasks require two sets of stimuli: one for the 'question' and one for the 'answer'. In the present experiment, we used auditory stimuli for the 'question' and visual (orthographic) stimuli for the 'answer'.

In the case of the auditory stimuli, three sets were created: one for the main trials for both groups, one for the practice trials of Group EXP, and one for the practice trials of the Baseline Group.

As for the visual stimuli, it consisted of sets of nine real words, presented orthographically. Four sets were created: one for the main trials of Group EXP, one for the main trials of the Baseline Group, one for the practice trials of Group EXP, and one for the practice trials of the Baseline Group. An overview of the stimuli is presented in Table 4.10.

4.3.1. Auditory stimuli

Following previous studies (Alispahic et al., 2017; Flege & Mackay, 2004), we used monosyllabic nonwords. A CV structure was chosen, based on two reasons. First, vowels are better perceived in consonantal context than in isolation (Deme, 2014; Rakerd, 1984). Second, cross-linguistically, the CV structure is the most natural, that is the most predictable, unmarked syllable (Rice, 2007; Maddieson 2013).

Considering that the auditory stimuli had to comply with phonotactic rules of both the L1 and L2, a sequential procedure was followed to select the consonant for the CV stimuli for the main trials. We started by looking for word-initial consonants common to Hungarian and Portuguese. Table 4.4 displays the word-initial consonants for EP, retrieved from Mateus & d'Andrade (2000), and for Hungarian, retrieved from Siptár & Törkenczy (2000). The common onset consonants, signaled by yellow shading, are /p/, /b/, /t/, /d/, /k/, /g/, /f/, /v/, /s/, /z/, /ʃ/, /ʒ/, /m/, /n/, /l/.

Table 4.4. Word-initial consonants for EP and Hungarian

EP	/p/	/b/	/t/	/d/		/k/	/g/	/f/	/v/	/s/	/z/	/ʃ/		/ʒ/				/m/	/n/		/l/		/ʀ/	
Hungarian	/p/	/b/	/t/	/d/	/ɟ/	/k/	/g/	/f/	/v/	/s/	/z/	/ʃ/	/h/	/ʒ/	/ʦ/	/ʧ/	/ʤ/	/m/	/n/	/ɲ/	/l/	/r/		/j/

Second, we selected the CV combinations that would entail the fewest meanings. We first looked into Portuguese CV words (Table 4.5), and selected those which entailed less meaningful possibilities: /bV/, /gV/, /fV/, /ʃV/ and /nV/.

Table 4.5. CV structured syllables with meanings as words in EP

EP oral vowels	Possible common word-initially consonants: p	b	t	d	k	g	f	v	s	z	ʃ	ʒ	m	n	l
[a]	[pa]		[ta]	[da]	[ka]		[fa]	[va]	[sa]		[ʃa]	[ʒa]	[ma]		[la]
[ɐ]				[dɐ]									[mɐ]	[nɐ]	[lɐ]
[ɛ]	[pɛ]						[fɛ]		[sɛ]	[zɛ]			[mɛ]		
[e]	[pe]	[be]	[te]	[de]	[ke]	[ge]		[ve]	[se]	[ze]	[ʃe]	[ʒe]			[le]
[ɨ]	[pɨ]	[bɨ]	[tɨ]	[dɨ]	[kɨ]	[gɨ]	[fɨ]	[vɨ]	[sɨ]	[zɨ]	[ʃɨ]	[ʒɨ]	[mɨ]	[nɨ]	[lɨ]
[i]	[pi]					[gi]		[vi]	[si]			[ʒi]	[mi]		[li]
[ɔ]	[pɔ]		[tɔ]	[dɔ]				[vɔ]	[sɔ]			[ʒɔ]	[mɔ]		
[o]			[to]	[do]				[vo]	[so]			[ʒo]			
[u]		[bu]	[tu]	[du]	[ku]				[su]	[zu]		[ʒu]	[mu]	[nu]	

The selected Portuguese CV combinations were then cross-checked with Hungarian CV structured words and examined in terms of frequency (Table 4.6). To obtain this information, we consulted the Hungarian corpus "Szószablya" (Halácsy et al., 2003).

Table 4.6. CV structured syllables with meanings in Hungarian

bV	freq.	**gV**	freq.	**fV**	freq.	**ʃV**	freq.	**nV**	freq.
[ba]	894	[ga]		[fa]		[ʃa]		[na]	
[bɛ]	754961	[gɛ]		[fɛ]		[ʃɛ]	151473	[nɛ]	473047
[be]	640	[ge]		[fe]		[ʃe]	68	[ne]	367
[bi]		[gi]		[fi]		[ʃi]	584	[ni]	1631
[bɔ]		[gɔ]		[fɔ]	29399	[ʃɔ]		[nɔ]	75179
[bo]		[go]	66	[fo]		[ʃo]	7141	[no]	39126
[bu]	370	[gu]		[fu]	292	[ʃu]		[nu]	

As we can see in Table 4.5, the syllables [bV], [gV], [fV], [ʃV] and [nV] have each three meanings in Portuguese. As for these syllables in Hungarian (Table 4.6), [gV] is the least frequent, and only [go] appears in the Hungarian *corpus* as a word, referring either to a board game or a programing language, in both cases, a loanword. As for [gV] meanings in EP, [gi] is an abbreviation of *Guilherme* 'proper name' and [ge] the name of the letter <g>. Although both [gi] and [ge] are used in EP, they are nonwords in Hungarian. This fact combined with the rare occurrence of a [gV] word in Hungarian determined that this should be the CV used in the experiment. The final stimuli were therefore [ga], [gɐ], [gɛ], [ge], [gɨ], [gi], [gɔ], [go], [gu].

For practice trials, a [tV] structure was selected for the auditory stimuli. The choice of this consonant was due to its high frequency in both languages. For the experiment with the Hungarian listeners, the stimuli were [ta], [tɛ], [te], [ti], [tɔ], [to], [tø], [tu], and [ty], and for the Portuguese group, [ta], [tɐ], [tɛ], [te], [tɨ], [ti], [tɔ], [to], [tu].

Recordings took place in two sessions, one in Budapest and one in Lisbon, between January 2021 and March 2021. The lockdown restrictions imposed at that moment did not allow us to record the auditory stimuli in sound attenuated rooms, therefore, all the recordings had to be done individually, at each talker's home.[66]

We asked each talker to pay special attention to the recording conditions, specifically, to record early in the morning or late in the evening to prevent background noise, and preferably in a location where reverberation was the lowest possible.

Recordings in Lisbon involved five native Portuguese female talkers, all from the standard dialectal area.[67] The equipment, a TASCAM DR-05 V2 digital recorder with a Beyerdynamic MCE 85 BA condenser microphone, was provided by LabCC–Communication Sciences Laboratory of the NOVA University Lisbon. The equipment was taken to each talker's home.

Recorder device settings were regulated and verified before each recording: file format was set to wav, sampling frequency to 44100 Hz, mono, and 32-bit depth.

To provide an example to the talkers for the recordings, the first tokens were recorded by a trained phonetician, the EP speaking supervisor of this dissertation. These

[66] A preliminary session of recordings was held at the Department of Applied Linguistics and Phonetics of the Eötvös Loránd University, in July 2020, with two Portuguese female talkers, using the equipment of the MTA—ELTE Momentum Lingual Articulation Research Group. The purpose of this session was to test the adequacy of the stimuli and to determine and test instructions for future recordings.

[67] All the talkers reported living in Lisbon area at the moment of recruitment and for at least the previous 10 years.

tokens were then edited and sent to the other talkers. Talkers were asked to listen carefully to the guideline recording, and then, make their own recording, speaking at a normal conversational rate.

The [gV] or [tV] stimuli were inserted in a carrier sentence, with an example of a real word containing a target syllable, underlined, as exemplified below in (1). The purpose was to control the quality of the vowels, and the intonation. Talkers were specifically instructed to not delete [ɨ], as usually occurs in EP speech.

(1a) <u>GA</u>VETA **[gɐ]**. Diga **[gɐ]**, por favor: **[gɐ]**. 'DRAWER [gɐ]. Say [gɐ], please: [gɐ].'

(2a) <u>TA</u>REFA **[tɐ]**. Diga **[tɐ]**, por favor: **[tɐ]**. 'TASK [tɐ]. Say [tɐ], please: [tɐ].'

Only tokens in boldface were retrieved for posterior selection. Appendix V presents the complete list of carrier sentences.

Talkers recorded one token of each phrase. After each recording, the equipment was collected, and the tokens transferred from the recorder's SD card to a personal computer. Tokens were verified and assessed by the author of the dissertation. If quality criteria were not met, the talker was asked to repeat the recordings.

The recordings of the [tV] stimuli for Group EXP's familiarization task were made in Budapest. Three female native speakers of Hungarian (aged 28 to 45) were recruited for this purpose. Recordings were also made individually at each talker's home, with an external sound card on a PC or laptop, and an omnidirectional condenser head-mounted microphone. Since the three Hungarian talkers were trained phoneticians, and these stimuli was created only for familiarization tasks, they were recorded in simple citation form and in isolation.

Once all the recordings were concluded, a multiple-step selection and editing procedure took place.

First, from the five Portuguese talkers, three were chosen, on the basis of the following. We chose those that had the best (and similar) sound quality, which were assessed to be homogenous with respect to intonation, and that showed homogeneity of speech tempo across tokens. The three female talkers – $T1^F$, $T2^F$ and $T3^F$ –, were 46, 42 and 46 years old, respectively.

After selecting the talkers, the background noise was eliminated. Since recording conditions differed from talker to talker, this procedure had to be done manually for each token sequence, with Audacity (Audacity Team, 2020).

The full sound files were then segmented manually and saved as individual wav files using a script in Praat (Boersma & Weenink, 2020). After this, new TextGrids were created for each segment, where the author labelled the vocalic sections of the [gV] and [tV] structures: vowel onset was identified at the first voicing bar of a complete formant structure. We then proceeded with the acoustic measurements of the tokens, on the vocalic sections, also with a script in Praat. We measured f0, frequency of F1 and F2, vowel duration and maximum intensity. The values for f0, F1 and F2 were calculated from a median of the values measured in a 10% window in the midsection of the vowel. We set time step to 0.01 ms, number of formants below 5.5 kHz to 5, and window length to 25 ms. For f0 measurement, pitch floor was set to 100 Hz and ceiling to 500 Hz, as recommended for female voices (Boersma & Weenink, 2020).

From these measurements and also from perceptual judgments of sound quality, one token from the three stimuli retrieved from each carrier sentence was chosen for each of the six talkers. EP and Hungarian tokens were verified by the author and the EP speaking supervisor of the dissertation, and the Hungarian supervisor of the dissertation, respectively. All selected tokens had a falling intonational contour. The final set included a total of 81 tokens: 54 produced by Portuguese female talkers ((9 [gV] + 9 [tV]) × 3 talkers) and 27 produced by Hungarian female talkers (9 [tV] × 3 talkers).

Once the tokens were selected, the final step was to normalize the duration and intensity. Normalization of the duration was paramount, considering that Hungarian listeners are sensitive to vowel length, and it was important to reduce the likelihood of Hungarian participants base their decisions on this feature. To do so, we manipulated vowel length manually when necessary, using Praat.

We equalized mean intensity across tokens to 70 dB, using a script written in Praat. Finally, we inserted 200 ms of silence before and after each token, also with a Praat script.

Table 4.7 presents a summary of the acoustic data for the stimuli used in the main trials ([gV]). Appendix VI reports the detailed acoustic measurements.

Table 4.7. Mean values of acoustic measures for the [gV] tokens' vocalic segment

stimuli (syllable)	**vowel**				
	duration (ms)	intensity (db)	f0 (Hz)	F1 (Hz)	F2 (Hz)
[ga]	261	85	174	850	1543
[gɐ]	260	88	180	636	1967
[gɛ]	265	88	178	528	2450
[ge]	285	89	181	427	2749
[gɨ]	244	89	197	381	2236
[gi]	240	91	206	384	2841
[gɔ]	275	89	186	590	1024
[go]	247	90	195	391	760
[gu]	261	90	203	350	718

4.3.2. Orthographic stimuli

In this experiment, speakers had to listen to an auditory stimulus from the L2, and match it with a real word of their L1, presented visually in orthographic form. Contrary to the auditory stimuli selection, low frequency was not a criterion for the orthographic stimuli, but rather homogeneity of structure and frequency. To provide the best set of existing Hungarian words, a multistep search was conducted. First, the adequate syntactic category and syllabic structure was selected. Based on the collected data, a disyllabic structure, preferably with nouns, was chosen. A new search revealed that the structure CV.CVC was the most frequent.[68]

One last query was run to select a set of real Hungarian words with the structure [gV.CVC]. To ensure a similar degree of familiarity among words in Hungarian, we conducted a familiarity questionnaire with 25 Hungarian native speakers (aged 18-45).

[68] The Hungarian corpus Szószablya (Halácsy et al., 2003) revealed the following CV-structure frequencies: CVVCCC=37, CVVCC=609, CVVC=2 486, CVCCV=54 673, CVCV=113 783, CVCCVC=156 633, CVCVC=173 717. According to the Maximal Onset Principle, CV.CVC should be more frequent than CVC.VC, and thus preferable to that.

Participants were asked to rate the familiarity of different [gV.CVC] words on a scale from 1 (not at all familiar) to 4 (very familiar). The nine words we selected for the main test showed similar familiarity ratings (median = 4). Regarding the orthographic stimuli for the practice trials, a set of real words with the structure [tV.CVC] was selected. Table 4.8 displays the orthographic stimuli selected for the exp-HU version.

Table 4.8. Orthographic stimuli presented to Group EXP

vowel	real word for main trials	real word for familiarization trials
[ɒ]	GARÁZS, [ˈgɒraːʒ], 'garage'	TANÁR, [ˈtɒnaːr], 'teacher'
[aː]	GÁBOR, [ˈgaːbor], 'proper name'	TÁBOR, [ˈtaːbor], 'camp'
[ɛ]	GERELY, [ˈgɛrɛj], 'javelin'	TEREM, [ˈtɛrɛm], 'room'
[eː]	GÉPÉSZ, [ˈgeːpeːs], 'machinist'	TÉTEL, [ˈteːtɛl], 'item'
[i]	GITÁR, [ˈgitaːr], 'guitar'	TITOK, [ˈtitok], 'secret'
[o]	GONOSZ, [ˈgonos], 'wicked'	TOJÁS, [ˈtojaːʃ], 'egg'
[ø]	GÖRÉNY, [ˈgøreːɲ], 'ferret'	TÖMEG, [ˈtømɛg], 'mass'
[u]	GULYÁS, [ˈgujaːʃ], 'goulash'	TUDÁS, [ˈtudaːʃ], 'knowledge'
[y]	GÜGYÖG, [ˈgyɟøg], 'to babble'	TÜKÖR, [ˈtykør], 'mirror'

Table 4.9. Orthographic stimuli presented to the Baseline Group

vowel	real word for main trials	real word for familiarization trials
[a]	GATO, [ˈgatu], 'cat'	TAPA, [ˈtapɐ], 'cover'
[ɐ]	GAVETA, [gɐ'vetɐ], 'drawer'	TAREFA, [tɐ'rɛfɐ], 'task'
[ɛ]	GUERRA, [ˈgɛʀɐ], 'war'	TELA, [ˈtɛlɐ], 'canvas'
[e]	GUÊ, [ˈge], 'gee'	TEMA, [ˈtemɐ], 'theme'
[ɨ]	GUERREAR, [gɨʀi'ar], 'to fight'	TERROR, [tɨ'ʀor], 'terror'
[i]	GUITO, [ˈgitu], 'money'	TINA, [ˈtinɐ], 'bowl'
[ɔ]	GOLA, [ˈgɔlɐ], 'collar'	TOMA, [ˈtɔmɐ], 'take'
[o]	GOTA, [ˈgotɐ], 'drop'	TODA, [ˈtodɐ], 'every'
[u]	GULA, [ˈgulɐ], 'gluttony'	TUDO, [ˈtudu], 'everything'

With respect to the real words for the Baseline Group version, a less exhaustive search was run, since the purpose of the tasks with this group was only to control for stimuli nativeness. The guidelines for the selection were words with [gV] or [tV] as initial syllable, easily identifiable by Portuguese speakers. In some cases, a trisyllabic word was chosen (Table 4.9, above).

Each of the four sets of words was placed in a 3×3 grid, written in capital letters and with the first syllable ([gV] or [tV]) highlighted (Figure 4.1, in the next section).

Table 4.10 displays a summary of the auditory and orthographic stimuli.

Table 4.10. Summary of the stimuli. In green: stimuli identical in the two groups

experiment version/group	task	auditory stimuli	orthographic stimuli
exp-HU (Group EXP)	familiarization task	[tV], recorded by Hungarian native talkers	Hungarian real words starting with [tV]
	main test	[gV], recorded by Portuguese native talkers	Hungarian [gV.CVC] real words
exp-PT (Baseline group)	familiarization task	[tV], recorded by Portuguese native talkers	Portuguese real words starting with [tV]
	main test	[gV], recorded by Portuguese native talkers	Portuguese real words starting with [gV]

4.4. Procedure

When accessing the experiment's link, participants were asked to choose a username, and after entering it, they had to read the consent form with general information about the experiment and the inclusion criteria. After consenting on the participation, they could start the experiment. If they did not agree with the terms, they only needed to close the window of the experiment.

Following the informed consent, they were given some general instructions such as the structure of the experiment. They were asked to use headphones or earbuds, and they were also informed of the importance of running the experiment in a quiet environment without disruptions, and with notifications turned off. They were then presented with the background questionnaire.

Once the questionnaire was completed, participants were informed that the experiment includes a short familiarization section with two tasks, followed by the main test. They were instructed that they would have the opportunity to regulate the sound volume, but once they've done it, they should not change it for the entire duration of the experiment. Participants were also asked to answer as quickly as possible, and never interrupt the experiment.

No time limit was preset for the display of instructions, that is, participants could familiarize themselves with them as long as they felt necessary. To move forward, participants just had to enter the space bar. We also did not set a time limit for trial completion. Upon completion of one trial, the experiment automatically moved forward to the next trial.

For sound volume adjustment, participants were presented with several repetitions of the token [te], produced by one of the L1 talkers. Participants were again warned that volume must not be changed during the rest of the experiment.

Once the volume was set, participants received instructions for the first familiarization task. It included a brief explanation of the task, with an image of the grid. Participants were asked to pay attention to the image, since they would use it in the following task. The image of the grid was always shown in the instructions, so that participants could be familiarized with the words for the trials. Hungarian listeners were particularly instructed to ignore vowel length and to focus on vowel quality exclusively. In each trial, participants were presented with the grid containing the nine words selected for the task. After hearing an auditory stimulus, participants had to match the syllable they heard with the Hungarian syllables presented visually, by clicking on the chosen word that included the chosen syllable. After each response, participants received feedback (in their L1) – "Correct" or "Incorrect… try again!". If the answer was correct, the following trial began automatically. In case of an incorrect answer, participants were presented with the same trial again immediately after the first presentation, as many times as needed until they arrived at the correct answer. After successfully completing the first familiarization part, participants received new instructions, this time for the second part of the familiarization section. As before, instructions included a brief explanation of the task, with the image of the grid (the same as presented before), and a reminder to ignore vowel length, in the case of Hungarian listeners. The task consisted of the repetition of the trials of the first part, but without feedback. Since trials were randomized, in the second part of the familiarization task stimuli were presented in a different order. After each forced-choice identification trial, participants heard the same auditory token again and were asked to evaluate how good the example of the previously heard syllable was of the syllable in the chosen word of the grid, on a scale from 1 to 4 (1 = *very bad example*, 2 = *bad example*, 3 = *good example*, 4 = *very good example*).

The third section of the experiment – the section of the main trials – started automatically after participants had completed the familiarization section. They were informed that they would be presented with a task similar to the previous task (the second familiarization task), but they would receive a new grid with a different set of words, and different sounds too. However, they were not informed about the language of the sounds. As before, Hungarian participants were asked to ignore vowel length. Figure 4.1 displays a computer screen of the grid, as presented to Group Exp in the main trials.[69]

Figure 4.1. Computer screen of the main trials presented to Group EXP

Before activating the experiment, eight pilot sessions were conducted. Version exp-HU was first piloted by the author and the Hungarian speaking supervisor of the dissertation, and then by three other volunteers, Hungarian native speakers (two of them working in phonetics, phonology, or psycholinguistics). Version exp-PT was piloted by the

[69] This computer screen is from the second part of each trial, the goodness-of-fit rating: below the grid, we can see the rating options.

author and the EP speaking supervisor of the dissertation, and by two other Portuguese native speakers working in linguistics.

Piloting the Hungarian version took longer and involved more sessions, since it was the first version created. The exp-PT version was created after the piloting of exp-HU was completed, and it was adapted from the Hungarian final version. Piloting and adjustments were done sequentially: after the first volunteer completed the experiment, corrections and adjustments were made, and only after that the second volunteer was asked to pilot the experiment, and so on.

4.5. Results

In total, 2916 answers were collected from the participants' results in the main trials: 2106 answers from Group EXP (78 participants × 9 vowels × 3 repetitions) and 810 answers from the Baseline group (30 participants × 9 vowels × 3 repetitions).

Tables 4.11 and 4.12 (next page) present the mean percent of identification of EP stimuli into each specific L1 categories (rounded to the nearest decimal number) and the median of goodness-of-fit ratings (in brackets). Orthographic stimuli presented in the grids are marked by capital letters; above these, using IPA symbols, phonemic transcription of the first syllable is shown. Table 4.11 presents results for the Baseline Group, with rectangles with solid line edges marking the expected answer. Table 4.12 shows Group EXP's results, with rectangles with dashed line edges marking the predicted choices for Hungarian listeners. Following previous studies (Bundgaard-Nielsen et al., 2011; Faris et al., 2016, 2018), we will focus only on identifications above change level (11,1%), marked in bold in both tables .[70]

[70] Chance level was calculated by dividing one response by nine possible categories: 11.1%. The studies mentioned assume that responses below chance level suggest random responding.

Table 4.11. Results for the Baseline Group

auditory stimuli	/ga/ GATO	/gɐ/ GAVETA	/gɛ/ GUERRA	/ge/ GUÊ	/gɨ/ GUERREAR	/gi/ GUITO	/gɔ/ GOLA	/go/ GOTA	/gu/ GULA
	Orthographic stimuli								
[ga]	**98.9% (4)**	1.1% (2)							
[gɐ]	2.2% (3)	**95.6% (4)**			2.2% (2.5)				
[gɛ]			**90.0% (3)**	8.9% (3)	1.1% (2)				
[ge]			4.4% (3)	**87.8% (3)**	5.6% (2)	2.2% (2)			
[gɨ]			2.2% (3.5)	5.6% (2)	**92.2% (3)**				
[gi]				8.9% (2)	1.1% (2)	**90.0% (3)**			
[gɔ]							**94.4% (4)**	5.6% (3)	
[go]					1.1% (4)		2.2% (3)	**50.0% (4)**	**46.7% (3)**
[gu]	1.1% (1)			1.1% (3)				4.4% (3)	**93.3% (3)**

In the Baseline Group, except for [go], which fell between /go/ and /gu/, all the EP vowels were identified as the expected EP category. χ^2 tests revealed that the misidentification of [go] was a consequence of a talker effect. A χ^2 test run in both groups pooled, showed that perception of [go] was affected by the identity of the talker ($\chi^2 = 166.57$, $p < .001$), and χ^2 tests run on the two samples separately (Hungarian and Portuguese listeners) revealed that this effect was present in both groups of participants (Group EXP: $\chi^2 = 122.7$, $p < .001$; Baseline Group: $\chi^2 = 45.07$, $p < .001$).

Table 4.12. Results for Group EXP

auditory stimuli	/gaː/ GÁBOR	/gɒ/ GARÁZS	/gɛ/ GERELY	/geː/ GÉPÉSZ	/gi/ GITÁR	/go/ GONOSZ	/gø/ GÖRÉNY	/gu/ GULYÁS	/gy/ GÜGYÖG
	Orthographic stimuli								
[ga]	**97.4% (4)**	2.6% (2.5)							
[gɐ]	0.9% (1.5)		**68.4% (3)**				**29.9% (3)**	0.4% (2)	0.4% (4)
[gɛ]			**57.7% (2)**	**41.0% (2)**			1.3% (1)		
[ge]		0.4% (3)	1.3% (3)	**82.1% (3)**	**15.8% (2)**		0.4% (2)		
[gɨ]				1.7% (1.5)			**25.2% (2)**		**73.1% (3)**
[gi]				1.7% (1)	**98.3% (3)**				
[gɔ]		**78.2% (3)**				**21.8% (2)**			
[go]						**57.7% (3)**	0.4% (3)	**42.9% (3)**	
[gu]						11.5% (3)		**88.5% (3)**	

Table 4.12 shows that dispersion of responses was higher in Group EXP than in the Baseline group. While Portuguese listeners matched the nine EP vowels among 10 categories, Hungarian participants used 16 categories.

In Group Exp, EP vowels [a], [e], [i] and [u] were systematically identified as /aː/, /eː/, /i/, and /u/, respectively. Similar to what was observed in the Baseline Group's results, [o] was identified as /o/ (57.4%) or /u/ (42.2%). As for [ɔ], identification of this stimulus fell between /ɒ/ (78.2%) and /o/ (21.8%). The vowel [ɐ] was identified as /ɛ/ (68.4%) or /ø/ (30.0%), but never as /eː/. Regarding [ɨ], the identification of this vowel was split between /y/ (73.4%) and /ø/ (24.9%). Identification as /eː/ was below chance level, and /ɛ/ received no responses. It's important to notice that Hungarian participants showed indecision when identifying the Portuguese [ɛ]: a Wilcoxon rank-sum test showed that the identification of [ɛ] as /ɛ/ (57.7%) was not significantly different from the identification of [ɛ] as /eː/ (41.0%).

Regarding the goodness-of-fit ratings, as previously mentioned, we displayed a four-point Likert scale (1 = *very bad example*, 2 = *bad example*, 3 = *good example*, 4 = *very good example*). All EP tokens were perceived by Portuguese participants as being a very good example (4) or a good example (3) of the given category. As for Group EXP, most of the stimulus vowels had a median value of 3, reflecting a positive fit of the category (considering only identifications above chance levels). In the case of [e], [ɨ] and [ɔ], the more systematic choice (respectively, /eː/, /y/, and /ɔ/) was evaluated as a "good example" (median = 3), while less systematic options (/i/, /ø/, and /o/) had a negative rating (median = 2). The cases of [a] and [ɛ] were an exception: while the Portuguese [a] received a very good rating (median = 4) by the majority of participants, [ɛ] was perceived as "not a good example" (median = 2) in both categories chosen (/ɛ/, and /eː/).

In the following section, we will analyze specifically the perception of [ɐ] and [ɨ] by the Hungarian participants. In line with the method that guided our predictions for the EP vowel identifications and as advised by Colantoni et al. (2015), we will look into acoustic information to provide an explanation for the results.

4.5.1. Perception of [ɐ]

In *Hypothesis 1*, we predicted that EP vowel [ɐ] is perceived by Hungarian listeners as the closest L1 categories: /ɛ/, /eː/, or /ø/. Results showed that /ɛ/ was the preferred choice (68,4%), while /ø/ was the second choice, with a considerably lower identification rate (29.9%). However, no responses were registered for /eː/.

To account for these results, we plotted mean F1 and F2 formant frequency values estimated in the [ɐ] tokens for the three Portuguese talkers used in the experiment ($T1^F$, $T2^F$, and $T3^F$), on the Hungarian perceptual vowel map used for the predictions (Figure 4.2).

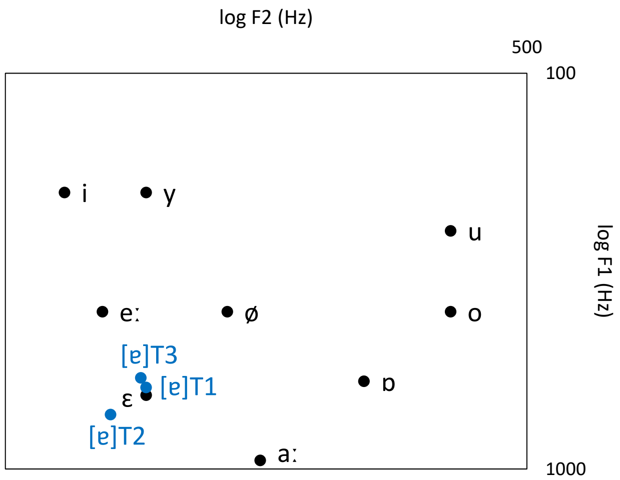


Figure 4.2. Hungarian perceptual vowel space with production data of the three [ɐ] tokens used in the experiment (in blue)[71]

The proximity of the EP [ɐ] tokens to the Hungarian /ɛ/ explains the overall preference for this category. However, it does not explain the 29.9% of the answers for /ø/, nor the lack of /eː/ responses.

We considered two possible explanations. Firstly, the characteristics of the EP tokens used in the experiment could be deviant from the standard Portuguese [ɐ] vowel. To examine this possibility, we placed the standard Portuguese [ɐ] produced by female speakers (noted as ɐ(f)), in the Hungarian perceptual map, along with the tokens we used in the present experiment ([ɐ]$T1^F$, [ɐ]$T2^F$, and[ɐ]$T3^F$; Figure 4.3, next page). As we can see

[71] For a clearer visualization, in the vowel space figures we omit ‘F’ in the talkers’ notation.

in Figure 4.3, the tokens used in the present experiment are close to the standard female [ɐ]. Therefore, this explanation does not hold for our results.

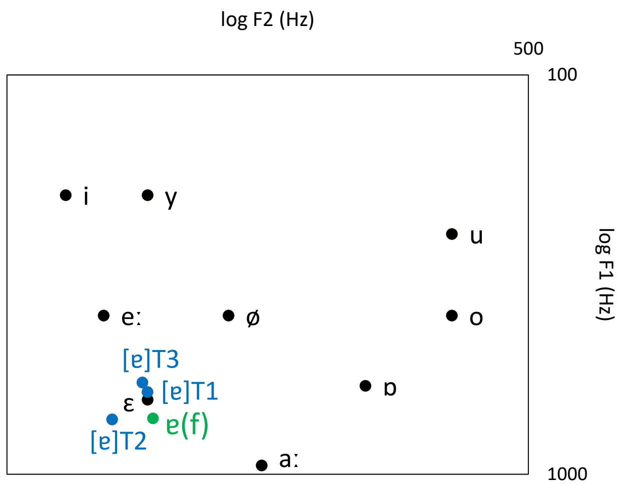


Figure 4.3. Hungarian perceptual vowel space with production data of the three [ɐ] tokens used in the experiment (in blue) and the standard EP [ɐ] measured for Portuguese female speakers (in green)

The second explanation lies in the Hungarian perceptual vowel map we used to establish our predictions.

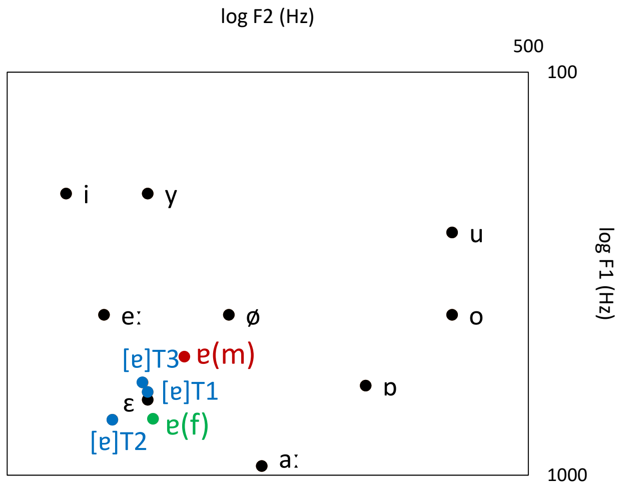


Figure 4.4. Hungarian perceptual vowel space with production data of the three [ɐ] tokens used in the experiment (in blue), the standard EP [ɐ] measured for Portuguese female speakers (in green), and the standard EP [ɐ] measured for Portuguese male speakers (in red)

If we place the standard EP [ɐ] produced by Portuguese male speakers (noted as ɐ(m)) in the Hungarian vowel space, we are able to see that, while [ɐ]T1^{F}, [ɐ]T2^{F}, and [ɐ]T3^{F}, and ɐ(f) are placed in the bottom left side of the vowel map, near /ɛ/, ɐ(m) is located between /ɛ/ and /ø/, and distant from /eː/ (Figure 4.4, previous page).

This suggest that the F1×F2 map from Kiss (1985) may be based on values for male Hungarian speakers, with a larger vocal tract that results in lower formant frequencies. To investigate this possibility, we calculated the formant dispersion metric, as an indicator of body size (Fitch, 1997).[72] Our results supported this explanation. Figure 4.5 displays the mean formant dispersion calculated for the tokens used in the present experiment as well for Kiss's perceptual study. Portuguese tokens' formant dispersion (between 1138 Hz and 1282 Hz) was higher than that calculated for the Hungarian tokens of the perceptual study of Kiss (1985) (1017 Hz).

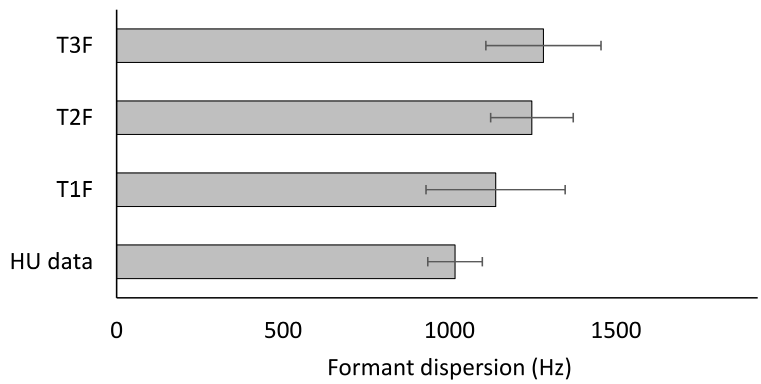


Figure 4.5. Formant dispersion for the Portuguese female talkers, estimated from the tokens used in the present experiment, for each talker individually (T1^{F}, T2^{F} and T3^{F}), and for Hungarian data, calculated from the perception data of the study of Kiss (1985, p. 166)

Three additional observations are important. As for the preference of /ɛ/, this could have been a result of the frequency of [ɛ] in Hungarian speech. According to Gósy (2004, p. 89), [ɛ] has a frequency of 11.4%, while [ø], [øː], and [eː] have frequencies of only 1.1%, 0.6%, and 3.1%, respectively.[73]

[72] The formant dispersion metric consists of the "average distance between each adjacent pair of formants" (Fitch, 1997, p. 1216).

[73] If we consider frequency only among vowels, [ɛ] has a frequency of approximately 27.0% (Gósy, 2004, p. 87; see Figure 2.9 in Chapter 2 of this dissertation).

The second additional observation is related to vowel length. As mentioned before, vowel length is a contrastive feature in Hungarian, hence Hungarian listeners are sensitive to the duration of the vowels. The EP tokens presented to the participants had they duration equalized, to avoid responses based on vowel length differences. The manipulation of token duration was made so that they would be perceived neither as short nor long vowels. However, [e:] exists only as a long vowel and [ɛ] as a short vowel. According to Gósy (1997, p. 110), in stressed position the duration of [ɛ] can vary between 78 to 220 ms, while [e:] ranges from 115 to 323 ms. Consequently, it's possible that the Portuguese [ɛ] did not meet the duration criterion to be identified as a long [e:] by Hungarian participants.[74]

Given the higher distance between the token [ɐ]T2^F from the two other tokens, we decided to investigate the possibility of a talker effect in the results. Surprisingly, the results show that the token [ɐ] recorded by talker T3^F was different: while this token was clearly identified as /ø/ (in 79.5% of all cases), the tokens [ɐ] produced by Portuguese talkers T1^F and T2^F were identified in a systematic way as /ɛ/ (89.7% and 97.4% respectively) (Table 4.13).

Table 4.13. Identification scores and median of ratings for [ɐ], according to talker

	T1^F	T2^F	T3^F
[ɛ]	89.7% (3)	97.4% (3)	17.9% (2)
[e:]	-	-	-
[ø]	10.3% (2)	-	79.5% (3)
[a]		2.6% (1.5)	
[y]			1.3% (4)
[u]			1.3% (2)

A χ^2 test confirmed a talker effect in the responses (χ^2 = 152.19; p < .001). Individual χ^2 tests between talkers T1^F and T2^F, T2^F and T3^F, and T1^F and T3^F showed a significant talker

[74] The mean duration of the tokens [ɛ] recorded for the experiment was 265 ms, which is closer to duration of the Hungarian [e:]. However, values calculated by Gósy (2004) were collected from spontaneous speech, whereas vowels used in our experiment were produced at a controlled speech rate.

effect only when T3^F was involved (χ^2 = 79.85; p < .001 for T1^F vs. T3^F, and χ^2 = 99.62; p < .001 for T2^F vs. T3^F).

A possible explanation for the fact that participants identified the [ɐ] produced by talker T3 as [ø] can be that there was an effect of f0. According to Traunmüller (1981), tonotopic distance between the first formant and f0 influences vowel openness's perception. To investigate this possibility, we calculated the tonotopic distances for the three [ɐ] tokens (Table 4.14).

Table 4.14. Tonotopic distances between F1 and f0 for the tokens [ɐ] produced by the three Portuguese female talkers[75]

	[ɐ]T1^F		[ɐ]T2^F		[ɐ]T3^F	
	Hertz	bark	Hertz	bark	Hertz	bark
f0	162.0	1.7	176.0	1.8	199.0	2.1
F1	622.0	5.9	729.0	6.7	589.0	5.7
F1−f0 (bark):		**4.2**		**4.9**		**3.6**

As we can see from Table 4.14, the tonotopic distance between f0 and F1 in [ɐ] is smaller for talker T3^F than for talkers T1^F and T2^F, which might have led participants to perceive the token [ɐ]T3^F as less open than in the case of the other [ɐ] tokens, and, thus, closer to the Hungarian /ø/.

4.5.2. Perception of [ɨ]

With respect to [ɨ], we predicted that this EP vowel would be identified as /y/, /eː/, /ɛ/, or /ø/, which are the closest vocalic categories in the Hungarian vowel system. Identification and goodness-of-fit ratings showed that /y/ was a more systematic choice (73.4%, rating median = 3) than /ø/ (24.9%, rating median = 2). /eː/ was identified below chance level and /ɛ/ was never chosen by the Hungarian participants.

We placed [ɨ]T1^F, [ɨ]T2^F, [ɨ]T3^F (the tokens used in the experiment), and ɨ(f) and ɨ(m) (EP standard female and male values, respectively) in the Hungarian perceptual map

[75] Values in bark were calculated using Praat's formula: bark = 7.0 * log (hertz/650 + sqrt (1 + (hertz/650)^2).

(Figure 4.6). As Figure 4.6 shows, stimuli used in the experiment were not deviant from the EP standard female [ɨ]. Figure 4.6 also shows that the standard male [ɨ] is located between the Hungarian /y/ and /ø/. This fact and the results obtained for [ɐ], thus confirm that the Hungarian vowel map used for our predictions presents values that are most probably representative of a male Hungarian speaker's acoustic space for vowels.

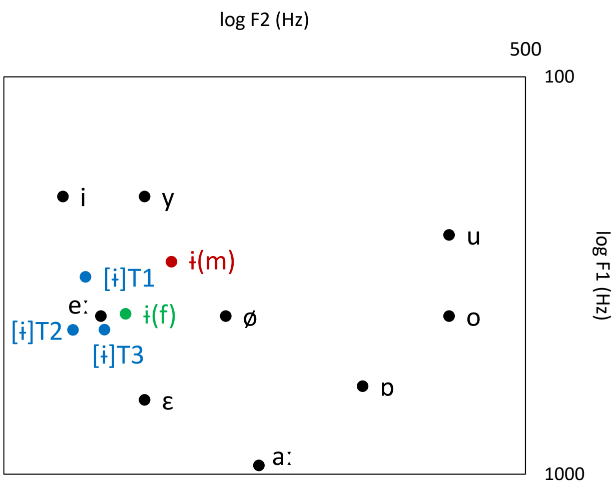


Figure 4.6. Hungarian perceptual vowel with production data of the three [ɨ] tokens used in the experiment (in blue), the standard EP [ɨ] measured for Portuguese female speakers (in green), and the standard EP [ɨ] measured for Portuguese male speakers (in red)

Similar to the analysis of [ɐ], we further investigated possible talker effects in the perception of [ɨ]. Table 4.15 shows that while the tokens produced by talkers $T1^F$ and $T3^F$ were systematically identified as the Hungarian /y/ (100% and 98.7%, respectively), the token produced by talker $T2^F$ was mainly identified as /ø/ (74.4%), although not in such a systematic way.

Table 4.15. Identification scores and median of ratings for [ɨ], according to talker

	$T1^F$	$T2^F$	$T3^F$
[ø]		74.4% (2.5)	1.3% (2)
[y]	100% (4)	20.5% (2)	98.7% (3)
[e:]		5.1% (1.5)	

A pooled χ^2 test confirmed a general effect of talker in the responses of Hungarian participants (χ^2 = 164.34; p < .001). Individual χ^2 tests between talkers showed a significant talker effect between $T1^F$ and $T2^F$ (χ^2 = 99.60, p < .001) and between $T2^F$ and $T3^F$ (χ^2 = 95.04, p < .001), but not between $T1^F$ and $T3^F$.

However, data in Figure 4.6 does not explain why the token [ɨ] produced by the talker $T2^F$ was identified as /ø/.

We calculated the tonotopic distance between the first formant and f0 to look for an effect of perceived openness (Table 4.16). The results show a higher value for talker T2, which may explain the preference for the perception of [ɨ]$T2^F$ as a more open category than /y/, in this case, /ø/.

Table 4.16. Tonotopic distances between F1 and f0 for the tokens [ɨ] produced by the three Portuguese female talkers

	[ɨ]$T1^F$		[ɨ]$T2^F$		[ɨ]$T3^F$	
	Hertz	bark	Hertz	bark	Hertz	bark
f0	192.0	2.0	176.0	1.9	225.0	2.4
F1	319.0	3.3	434.0	4.4	433.0	4.4
F1−f0 (bark):		**1.3**		**2.5**		**2.0**

The aim of this study was to observe the perception of EP vowels by Hungarian listeners, in order to have an idea of the difficulties Hungarian learners of L2 Portuguese will display at the onset of learning. Accordingly, the results of the identification experiment allow us to put forward predictions for discrimination difficulties for those learners, which we present and discuss next.

4.6. Predictions for discrimination

In Chapter 1, we described three cross-language models that provide theorical frameworks for discrimination predictions for L2 sounds. For the present study, we will outline our predictions based on the Perceptual Assimilation Model (Best, 1995). Although other models equally contribute to understanding L2 perception, PAM has been primarily tested with naïve listeners, which form the target subjects of our experiment.

In line with Best's model, determining how L2 sounds are mapped into the L1 is a fundamental step to define contrasts' assimilation patterns, and in turn, discrimination predictions. According to the model, non-native sounds can be *categorized*, *uncategorized*, or *not assimilated*. To establish this classification, two thresholds are used: chance level and categorization boundary. Chance level is defined as the probability of a specific non-native sound being matched with one L1 category. In our study, as shown in Tables 4.11 and 4.12, this was calculated as 11,1%, since we had nine candidates (nine different vocalic qualities in the Hungarian vowel system) for each EP vowel. According to PAM, when a sound is identified below chance level, it is classified as *not-assimilated*, that is, not perceived as speech a sound in the given language. If the sound is identified above chance level, then it is said to be *assimilated*. As for the categorization threshold, and contrary to chance level, there is no value unanimously defined. Previous perceptual assimilation studies vary in the categorization criteria, from 50% (Faris et al., 2018) to 70% (Tyler et al., 2014), or even 90% (Harnsberger, 2001). Table 4.17 displays the assimilation patterns according to thresholds of 50%, 70%, and 90%.

Table 4.17. Assimilation patterns' results according to different threshold criteria. Between round brackets: top-ranked choices

L2 stimulus	threshold: 50%	threshold: 70%	threshold: 90%
[a]	Categorized (into /a/)	Categorized (into /a/)	Categorized (into /a/)
[ɐ]	Categorized (into /ɛ/)	Uncategorized	Uncategorized
[ɛ]	Categorized (into /ɛ/)	Uncategorized	Uncategorized
[e]	Categorized (into /eː/)	Categorized (into /eː/)	Uncategorized
[ɨ]	Categorized (into /y/)	Categorized (into /y/)	Uncategorized
[i]	Categorized (into /i/)	Categorized (into /i/)	Categorized (into /i/)
[ɔ]	Categorized (into /ɒ/)	Categorized (into /ɒ/)	Uncategorized
[o]	Categorized (into /o/)	Uncategorized	Uncategorized
[u]	Categorized (into /u/)	Categorized (into /u/)	Categorized (into /u/)

A one-sample Wilcoxon signed-rank test conducted to compared identification of [ɛ] as /ɛ/ (57.7%) against the 50% threshold revealed no significant difference; therefore, we considered [ɛ] as categorized, for that threshold. Regarding categorization of [ɐ] into /ɛ/ (68.4%) when a 70% threshold is established, the Wilcoxon test showed significant difference (V = 889; p < .001), consequently, we consider the EP vowel as uncategorized at a 70% threshold. Regarding the categorization of [u] (88.5%) at the 90% threshold, the Wilcoxon test revealed no significant difference, that is, the EP [u] can be classified as categorized into the Hungarian /u/ at the mentioned threshold.

The definition of the threshold usually depends on the aim of the study, and the establishment of different categorization criterion can determine a different assimilation pattern, and consequently, different discrimination predictions according to PAM. Looking, for example, at the contrast [ɐ]-[ɛ] when both vowels are identified as /ɛ/, if we assume a 50% threshold, and both EP vowels equally fit the Hungarian /ɛ/, this contrast will be classified as a *single-category assimilation type* (SC), with discrimination predicted to be poor. However, if we consider the same threshold, but one of the EP vowels fits better the Hungarian /ɛ/ than the other, this contrast will be classified as a *category-goodness assimilation type* (CG), with discrimination predicted to range from moderated to good. Finally, if we consider the 70% or 90% thresholds, an *uncategorized-uncategorized assimilation type* (UU) occurs. In such cases, PAM is less clear in the predictions, with discrimination ranging from poor to very good, depending on the distance between the two non-native contrasts, and between these and the L1 categories at hand.

In a study aiming at UC and UU assimilations, Faris et al. (2018) found that perception of overlapping sounds can be a more accurate way of predicting discrimination performances than the arbitrary establishment of a categorization threshold. To account for overlap scores, three methods are proposed. Elvin et al. (2021) propose the *acoustic overlap* method, in which overlap values are calculated from L1-L2 acoustic comparisons. *The classification overlap score*, proposed by Flege & Mackay (2004), is calculated based on scores of an identification task, and it includes all identification scores, even those below chance level. Faris et al. (2018) use a similar method, however the authors suggest that scores below chance level "constitute anything more systematic than random responding"

(p. 4) and thus should not be considered. The authors proposed the *phonological overlap method*, which focuses on the analysis of perceived overlap just for assimilated sounds. We will consider this last method.

The results obtained in the present study show four EP contrasts that can be perceived as overlapping (considering only identifications above chance-level): [ɐ]-[ɛ], when the two vowels are identified as /ɛ/ (57.7% of overlap), [ɛ]-[e], when the two vowels are identified as /eː/ (41.0% of overlap), [e]-[i], when the two vowels are identified as /i/ (15.8% of overlap), and [ɐ]-[ɨ], when the two vowels are identified as the Hungarian /ø/ (25.2% of overlap) (Table 4.18).[76]

Table 4.18. Identification scores for the categories which present overlap

L2 stimuli	L1 categories: /ɛ/	/eː/	/i/	/ø/
[ɐ]	68,4%			29,9%
[ɛ]	57,7%	41,0%		
[e]		82,1%	15,8%	
[ɨ]				25,2%
[i]			98,3%	

Other than calculating the perceived overlapping scores, it is also important to analyze the goodness-of-fit between sounds involved in those overlapping situations, that is, to determine if both L2 sounds are either an equally good fit to the L1, an equally a poor fit to the L1, or if one is a better fit than the other. To assess this, we run two-sample Wilcoxon rank-sum tests on the goodness-of-fit ratings in the four contrasts mentioned above.

In the case of the EP contrast [ɐ]-[ɛ], in 57.7% of the cases both vowels were identified by the Hungarian participants as /ɛ/. However, the Wilcoxon rank-sum test showed that the Portuguese [ɐ] (median = 3) was a significantly better fit to the Hungarian /ɛ/ than the Portuguese [ɛ] (median = 2) (W= 15572, p < .001; Figure 4.7, next page).

[76] We exclude overlaps in the identification of [u] and [o], since they are not involved in the perception of the target vowels of the study, [ɐ] and [ɛ]. Moreover, no cases of complete overlap were found.

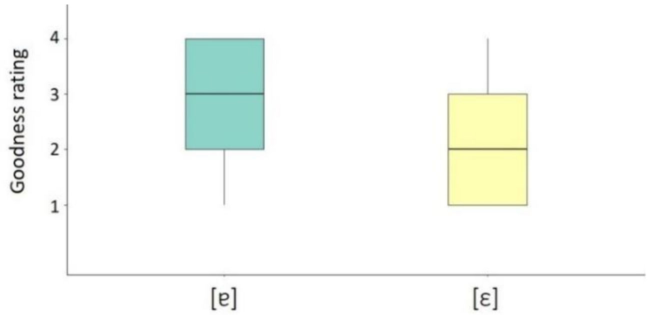


Figure 4.7. Goodness-of-fit ratings for identification of EP [ɐ] and [ɛ] as the Hungarian /ɛ/

In 41.0% of the cases, both vowels of the EP contrast [ɛ]-[e] were identified as /eː/, but according to the Wilcoxon test, [e] was a better example of the Hungarian vowel (median = 3), than [ɛ] (median = 2) (W = 15368, p < .001; Figure 4.8).

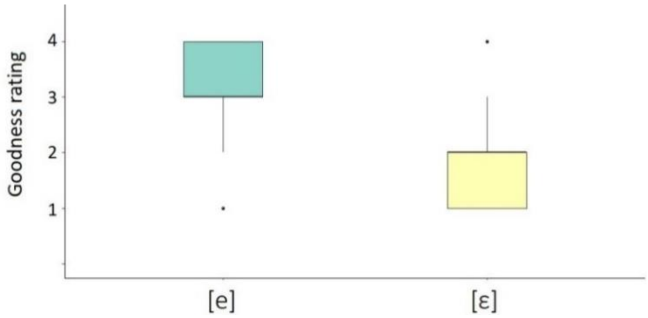


Figure 4.8. Goodness-of-fit ratings for identification of EP [e] and [ɛ] as the Hungarian /eː/

Regarding the contrast [e]-[i], in 15.8% of the times, both vowels were identified as /i/. However, the test showed that [i] (median = 3) was a significant better fit for /i/ than [e] (median = 2) (W = 1963, p < .001; Figure 4.9).

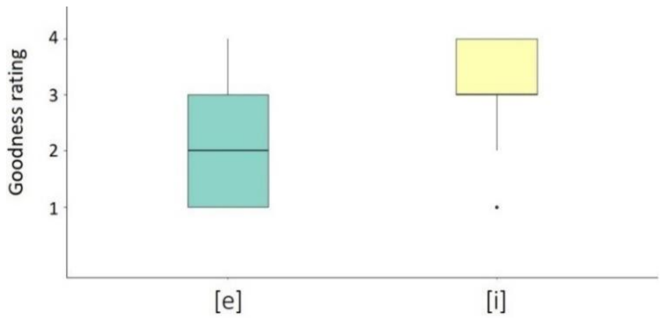


Figure 4.9. Goodness-of-fit ratings for identification of EP [e] and [i] as the Hungarian /i/

Finally, in 25.2% of the trials, the vowels in the contrast [ɐ]-[ɨ] were both identified as the Hungarian /ø/. In this case we also found a significant difference in the goodness-of-fit, with [ɐ] (median = 3) being a better example of /ø/ than [ɨ] (median = 2), although to a lower extent than the previous contrasts (W = 2578, p = .011; Figure 4.10).

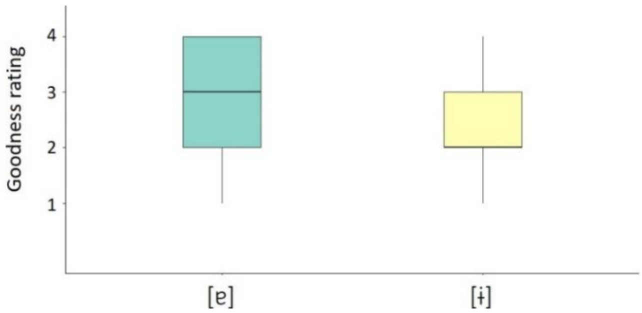


Figure 4.10. Goodness-of-fit ratings for identification of EP [ɐ] and [ɨ] as the Hungarian /ø/

As mentioned earlier, to determine assimilation patterns of contrasts, we must first define a categorization threshold. We opted for the 70% value, for two reasons. First, Table 4.14 shows that if we consider the 50% threshold, all EP vowels would be categorized by the Hungarian listeners. Conversely, at a 90% threshold, only the vowels in the limit of the vowel space triangle (/a/, /i/, and /u/) are categorized. A 70% categorization boundary is probably a more suitable criterion for the uncategorized situations. Second, the 70% threshold seems a proportional measure considering that in our study only nine L1 categories were available to identify the EP vowels.

After setting a 70% categorization threshold and determining the goodness-of-fits, it is possible to establish assimilation types for the contrasts: [e]-[i] configures a CG type, [ɛ]-[e] and [ɐ]-[ɨ] form UC types, and the contrasts [ɐ]-[ɛ], a UU assimilation type. Both the UC assimilations present partial overlap (Table 4.19, next page).[77]

[77] The acronyms mentioned here are explained in section 1.3.1, in the description of PAM.

Table 4.19. Assimilation types for the target contrasts ('P' indicates partial overlapping)

L2 contrast	identified L2 category	contrast assimilation pattern
[ɐ]-[ɛ]	/ɛ/	UU-P
[ɛ]-[e]	/e/	UC-P
[ɐ]-[ɨ]	/i/	UC-P
[e]-[i]	/ø/	CG

According to Faris et al. (2018), the following hierarchy for discrimination performance should occur (from excellent to poor, 'N' indicates non-overlapping): "{TC = UC-N = UU-N} > {CG = UC-P = UU-P} > {SC = UC-C = UU-C}" (p. 4). Applying this proposal to our study, the four contrasts should present a similar difficulty in discrimination: CG = UC-P = UU-P. However, this contradicts empirical evidence from the classroom, since Hungarian learners struggle more with [ɐ]-[ɛ] than with other EP contrasts.

As mentioned in Chapter 1, some authors suggest that the perceived overlap score might be a better predictor for discrimination performance than the arbitrary classification of assimilation types (Faris et al., 2018; Tyler et al., 2014). Following this suggestion, we would have (from better to poorer discrimination) /e/-/i/ > /ɐ/-/ɨ/ > /ɛ/-/e/ > /ɐ/-/ɛ/.

To assess these predictions, we analyzed data collected at the beginning of the perceptual training experiment (reported in the next chapter). Since the perceptual training will be described in the next chapter, we will not include here methodological details of the task. We used data collected in the pretest, from 50 Hungarian participants with no previous or present contact with the Portuguese language.[78] The participants mean age was 21.5 years, 31 were female, and the majority (n = 41) was attending the first year of a bachelor's degree. The task consisted of an oddity discrimination task with catch trials, testing the following contrasts: [gɐ]-[gɛ], [gɛ]-[ge], [ge]-[gi], [gɐ]-[gɨ], [gɐ]-[ga], [gɛ]-[ge], and [ge]-[gɨ]. The mean percentage of error rate across contrasts was 27.3%. As we can see in Table 4.20, the contrasts with an error rate above the mean (in grey) were those which presented perceptual overlap in the identification experiment.

[78] Some of the participants included in this analysis completed the pretest but not the training.

Table 4.20. Mean error rates in percentage for vowel discrimination, retrieved from the pretest of the training study

contrast	ER	sd	se	ci
[ɐ]-[ɛ]	49.4%	23.8%	3.4%	6.8%
[ɛ]-[e]	47.7%	25.5%	3.6%	7.2%
[ɐ]-[ɨ]	32.3%	17.0%	2.4%	4.8%
[e]-[i]	29.1%	19.6%	2.8%	5.6%
[ɛ]-[ɨ]	13.7%	12.3%	1.7%	3.5%
[a]-[ɐ]	10.3%	11.7%	1.7%	3.3%
[e]-[ɨ]	8.6%	10.1%	1.4%	2.9%

Table 4.21 summarizes the results of the identification scores and phonological overlap values obtain from the identification task (reported in the present chapter) and the preliminary scores for the discrimination oddity task completed in the pretest of the training (displayed in Table 4.20).

Table 4.21. Summary of the results from the identification task (present chapter) and the preliminary results of the discrimination task (next chapter)

L2 contrast	**identification task** contrast's assimilation patern	**identification task** overlap score	**ER in the discrimination task**
[ɐ]-[ɛ]	UU-P	57.7%	49.4%
[ɛ]-[e]	UC-P	41%	47.7%
[ɐ]-[ɨ]	UC-P	25.2%	32.3%
[e]-[i]	CG	15.8%	29.1%

As we can see from the values in the last two columns, the results confirm that the phonological overlap score is an effective predictor for discrimination in our experiment.

4.7. Conclusion

This preliminary study was designed to answer to the question of how Hungarian listeners map the EP oral vowels, and more specifically, the EP reduced [ɐ] and [ɨ], into their L1 system. According to the assumption that listeners identify non-native sounds as the closest category or categories of their L1, we expected the Hungarian listeners to identify the EP [ɐ] as /ɛ/, /eː/ or /ø/, and the EP [ɨ], as /y/, /ø/, /ɛ/ or /eː/.

The results partly corroborated these predictions: identification of [ɐ] fell into /ɛ/ or /ø/, and [ɨ] was identified into /y/ or /ø/. An explanation for these results may lay on the Hungarian perceptual map we used to establish our predictions, which is likely to show perceptual categories for male productions, while the tokens used in the experiment were all produced by female talkers. This explanation suggests that acoustic comparisons may be suitable to predict how speakers map L2 sounds in their L1, however, the L1 and L2 vowel spaces must be attentively selected considering the specific purpose of the experiment. In other words, in our study, predictions should have been made considering that we were using stimuli produced by female Portuguese speakers, and accordingly, crossing these with F1 and F2 values for Hungarian female speakers.

Additionally, it must be mentioned that other factors may have also contributed to the results, such as talker effects, vowel relative frequency, and vowel length. For example, it would be interesting to compare the results of our experiment with a similar experiment, but including the full Hungarian vowel inventory (i.e., short and long vowels).

Regarding perceptual overlap, we observed several situations, from which, the most problematic case is [ɐ]-[ɛ]. Based in the results of our experiment, this EP contrast presents 57.7% chances of being perceived by Hungarian listeners as a single category, /ɛ/, with [ɐ] being a better fit to it than [ɛ]. Accordingly, this contrast will present the most difficulties at discrimination level. This prediction was confirmed with preliminary results of the discrimination task conducted in the pretest of the perceptual training, since in this task, the contrast [ɐ]-[ɛ] obtain the highest mean error rate.

In respect to [ɨ], results from the identification tasks suggest a less problematic situation for perception of this vowel, since no conflict with Portuguese vowels emerged.

That is, [ɨ] was consistently identified as a vowel that does not exist in the EP system (/y/ and, to less extend, /ø/).

Overall, the results suggest a higher level of difficulty in the perception of [ɐ] than of [ɨ], which is in line with the observations in the classroom, described in Introduction. In the next chapter, we will describe the perceptual training we designed to aid Hungarian learners surpass, among others, the difficulties we observed in the present study.

CHAPTER 5

ACQUISITION OF EP STRESS AND THE REDUCED VOWELS [ɐ] AND [ɨ] BY HUNGARIAN LEARNERS OF EP: A PERCEPTUAL TRAINING STUDY

In the previous chapter, we described a perceptual experiment which provided us with a picture on how Hungarian naïve listeners map the EP oral vowels into their L1 vowel system. As predicted, Hungarian participants identified the L2 vowels into the closest L1 categories.

The results showed a complex situation for the perception of [ɐ], since both EP [ɐ] and [ɛ] were identified by Hungarian listeners as /ɛ/ (in 68.4% and 57.7% of the cases, respectively). Additionally, EP [ɛ] also presented some difficulties to Hungarian participants, since in 41.0% of the trials it was perceived by Hungarian listeners as /eː/. In turn, EP [e] also displayed some problems in perception, although to a lesser extent: in 15.8% of the trials, a value that is slightly above chance level, this vowel was identified by Hungarian participants as /i/. These results reveal a picture of the problems Hungarian learners of EP face when acquiring [ɐ]. With exception for the vowels in the limit of the vowel space – [a] and [i] –, a perceptual conflict seems to arise with the mid-central and mid-front EP vowels, [ɐ], [ɛ], and [e].[79]

The perception of the EP vowel [ɨ], however, seems to be less problematic to the Hungarian listeners. Results of the identification experiment presented in Chapter 4 showed that, in 73.1% of the cases, this vowel was perceived as /y/ by the Hungarian listeners, and in 25.2% of the cases as /ø/. In both cases, no conflict with other EP vowels arises, since neither [y] nor [ø] are present in standard EP.

The acquisition of [ɐ] and [ɨ] will depend on changes in perception of specific EP contrasts, by the learners. To acquire the vowel [ɐ], Hungarian learners must create a new category or split one existing category (/ɛ/) into two. Moreover, EP vowels [ɐ] and [ɛ] must be matched with separate perceptual categories and representations in the lexicon. This learning process will not be independent of the perceptual adjustment of the contrasts [ɛ]-[e], [e]-[i], and [ɐ]-[ɨ].

Two aspects may hinder the learning tasks predicted above – the influence of orthography, and the stress 'deafness' displayed by Hungarian speakers which is, in the case of EP, not independent from the question of segment acquisition.

[79] As previously mentioned, the back vowels are not within the scope of this dissertation, since no perceptual conflict emerges between these vowels and the target vowels [ɐ] and [ɨ].

Regarding the effect of orthography, we must point out the fact that our target learners have contact with the L2 through formal learning; therefore, their attention is explicitly drawn to written words and texts. As described in Chapter 2, grapheme-phoneme correspondence in EP is not straightforward, a fact that may interfere with phonological acquisition when learners are exposed to input from written sources. For example, the grapheme <e> can phonetically be realized as [e], [ɛ], [ɐ], [ɨ] or [i]. Phonetic realization is determined by factors like presence/absence of stress, etymology, or phonological constraints.

With respect to the influence of stress, in EP, with the exception of /i/ and /u/, vowel quality depends mostly on the presence/absence of stress. Consequently, 'deafness' to stress contrasts displayed by Hungarian speakers adds a level of difficulty to the acquisition of the EP vowels. Hungarian learners must match both the unstressed [ɐ], and the stressed [a] vowel phones with the phonemic category /a/. Similarly, learners will have to match the unstressed [ɨ], the stressed [ɛ], and the stressed [e] all to one of two phonemic categories, /ɛ/ or /e/. Moreover, along with the above interrelations, Hungarian learners are also exposed to phonological constraints and lexical exceptions, that involve circumstances in which, for example, [ɐ] occurs in stressed position, or belongs to the phonemic category of /e/, underlyingly (see 2.1.1).

In regard to EP stress, when vowel reduction is absent, duration and intensity are the acoustic cues on which Portuguese listeners rely among all suprasegmental features. However, on the production side, the modulation of duration, for example, is not straightforward, that is, the durational ratio between the stressed syllable(s) and the unstressed syllable(s) depends on the word. Specifically, the ratio is lower in unmarked cases (when stress complies with the general rule) and higher in marked cases (when stress does not follow the general rule). In sum, when planning the training of the suprasegmental aspect of EP stress we had one purpose: to draw learners' attention to the fact that, contrary to their L1, stress is variable, that is, it may fall in different positions within a word.

Having the issues above stated in consideration, we designed a perceptual training program that provides specific tasks to help learners improve their discrimination abilities at both the segmental and suprasegmental levels of speech. The study included three training programs, each targeted to a different topic: (1) EP vowel perception, (2) EP stress suprasegmental cues, and (3) EP stress and vowel reduction, combined. These interventions were administered to three groups of Hungarian learners of L2 Portuguese, respectively: Group *Vowels*, Group *Stress*, and Group *Vowels & Stress*.

A note is in order regarding the label 'Group *Stress*'. Considering that EP stress is signaled by suprasegmental features as well as by vowel reduction, 'Group *Stress: suprasegmentals*' would be a more correct name for this group. However, for simplicity, we opted to name it Group *Stress*. This is also in line with the use of the word 'stress' throughout this chapter, which refers only to prosodic features at word level (and only to primary word stress).[80] Accordingly, in this chapter, we use the expression 'stress discrimination' to refer to discrimination of stress contrasts signaled by suprasegmental cues, which, in the case of EP, are duration and energy, as shown by previous perceptual studies (Delgado-Martins, 1986).

To the best of our knowledge, no perceptual training study has been published on L2 stress acquisition by Hungarian learners. L2 vowel acquisition by Hungarian learners was addressed by Peltola et al. (2014), in a study with Finish as L2. However, this study aimed at pronunciation. Moreover, the few perceptual training studies on L2 stress described in Chapter 1 are mostly focused on suprasegmental features and do not relate stress with the acquisition of vowel reduction. Consequently, the design of our training task for stress perception is built upon scarce methodological background.

In the following sections, we describe the design, structure, and timeline of the study (tests and training) (section 5.1), the participants (section 5.2), and the materials (questionnaires and stimuli) we used in the different tasks (section 5.3). This section is followed by the description of the procedures for each phase of the auditory training program (section 5.4). Regarding the results, we start by presenting the exploratory data (section 5.5), and conclude this chapter with a summary and discussion of the results (section 5.6).

[80] When we specifically refer to EP stress, we are mentioning both segmental and suprasegmental features.

5.1 Study design

The experiment we report on this chapter consisted of an intervention study following previous studies (Aliaga-García, 2017; Brawerman-Albini et al., 2017; Grenon et al., 2019; Iverson & Evans, 2009; Kissling, 2015; Nobre-Oliveira, 2007; Ou, 2011; Rato, 2013).

Our study presented two originalities compared with the studies above-mentioned. First, it recruited learners at the onset of the L2 acquisition process. As explained in Introduction, we aimed at observing changes in perception in the first moments of contact with the L2, which, according to Best & Tyler (2007), is the moment when phonetic attunement is less restricted by other learning processes, such as morphosyntactic or vocabulary acquisition. Second, the study was conducted online. The decision to do so was imposed by the lockdown restrictions at that beginning of the present dissertation.

In line with previous perceptual trainings, our study includes three phases: pretest, perceptual training, and post-test, consisting of a repetition of the pretest. To better assess effectiveness of training, studies frequently include generalization and retention tests. While the first ones are usually conducted along with the post-test or soon after, retention tests are presented sometime after the post-test. Considering the duration of our training program – ten weeks –, and the high rate of students who usually do not continue their studies after one semester, we decided not to include a retention test, only generalization trials. The aim of generalization trials is to evaluate the ability to transfer the acquired knowledge to new contexts. To this purpose, different possibilities are available, such as the use of new talkers, the inclusion of new stimuli, or even the addition of tasks of different nature. Although some authors separate the repetition of pretest trials from the generalization test (Brawerman-Albini et al., 2017; Rato, 2013), in the present study, repeated trials and generalization trials were included in one post-test session for two reasons. First, participants who completed the repeated pretest trials would forcefully have to complete the generalization trials. Second, this study used different sets of stimuli, but the same talkers, in the pretest and in the training sessions. The novel trials included new stimuli produced by new talkers. Table 5.1 (next page) presents the distribution of stimuli and talkers in testing and training.

Table 5.1. Distribution of stimuli and talkers in the testing moments and in training

pretest	training	post-test repeated trials (post-test^R)	 novel trials (post-test^N)
stimulus set A	stimulus set B	stimulus set A	stimulus set C
talker set 1	talker set 1	talker set 1	talker set 2

As we can see in Table 5.1, the repeated trials of the post-test included the same talkers used in the training but different stimuli, which means that we introduced a degree of generalization from training to the repeated trials (untrained stimuli). The novel trials, however, introduced two levels of generalization: untrained stimuli and novel talkers, different from both the pretest and the training. Therefore, we considered that both the repeated trials as well as the novel trials of the post-test represented a generalization, and thus, we opted to present both in one single session. Henceforward, we will use *post-test*^{R} to refer to the post-test trials which are repeated from the pretest, and *post-test*^{N} to refer to the novel trials.

The present study aimed at observing four issues, detailed in Chapter 3:

(i) Does a perceptual training improve discrimination abilities with the reduced vowels [ɐ] and [ɨ]?

(ii) Does a perceptual training improve discrimination abilities with word stress?

(iii) Is a perceptual training that combines word stress and vowels more effective in improving discrimination abilities of reduced vowels and word stress than a training that focuses on those features separately?

(iv) Is there a hierarchy between segmental and suprasegmental acquisition in L2?

To this purpose, we designed three training regimens, one aimed at *vowel discrimination*, the second aimed at *stress discrimination*, and the third aimed at *vowel and stress discrimination*. The effectiveness of these regimens was assessed by comparing results of the pretest and the post-tests, and with comparisons between groups' performances. Although we aimed at observing the effectiveness of training in the three regimens above-referred, in the pretest and the post-test we only included vowel discrimination and stress discrimination. Testing stress and vowel discrimination combined would not allow us to determine which factor contributed to the improvement or decline in the discrimination ability: changes in vowel discrimination, changes in stress discrimination, or in both. In sum, *vowel discrimination* and *stress discrimination* were tested and trained, while *vowel and stress discrimination* was only a training condition.

Vowel discrimination included contrasts in which, on the basis of the perception study's results we reported in Chapter 4, *perceptual overlap* may occur when Hungarian listeners perceive EP vowels: [ɐ]-[ɛ], [ɛ]-[e], [e]-[i], and [ɐ]-[ɨ]. Considering the importance of vowel reduction in the L2, we judged necessary to include also *stressed-unstressed* minimal pairs. To this purpose, three other contrasts were added, corresponding to /a/ and /ɛ/ in *stressed-unstressed* condition: [a]-[ɐ], and [ɛ]-[ɨ] and [e]-[ɨ], respectively.

Stress discrimination included pseudowords contrasting only in stress position. To this purpose, and following Correia et al. (2015) and Dupoux et al. (1997), two conditions were tested: *1^{st} syllable - 2^{nd} syllable* contrast (e.g., [ˈdutiku]-[duˈtiku]), and *2^{nd} syllable - 3^{rd} syllable* contrast (e.g., [duˈtiku]-[dutiˈku]).

Vowel and stress discrimination included pseudowords contrasting in stress location and in vowels. To this purpose, we used disyllabic words contrasting in *1^{st} syllable - 2^{nd} syllable* stress location and that included the same vowels as for *vowel discrimination* ([a], [ɐ], [ɛ], [e], [ɨ], and [i]). The choice of disyllabic pseudowords instead of trisyllabic was based on the assumption that a trisyllabic contrast with vowel reduction would be too difficult for the students. The target vowel contrasts were included in the 1^{st} syllable, the 2^{nd} syllable, or in both syllables. Consequently, three cases of vowel reduction were included: (1) contrasts in which vowel reduction only occurs on the first syllable (e.g., [ˈza̲zi]-[zɐ̲ˈzi]), (2) contrasts in which vowel reduction only occurs on the second syllable (e.g., [ziˈza̲]-[ˈzizɐ̲]), and (3) contrasts in which vowel reduction occurs on both syllables (e.g., [ˈza̲zɐ̲]-[zɐ̲ˈza̲]).

A summary of the conditions observed for *vowel discrimination*, *stress discrimination*, and *vowel and stress discrimination* is presented in Table 5.2 (next page).

Table 5.2. Conditions observed for vowel discrimination, stress discrimination, and vowel and stress discrimination

vowel discrimination	stress discrimination	vowel and stress discrimination
perceptual overlap contrasts	*1st syllable - 2nd syllable contrast*	*1st syllable - 2nd syllable contrast with*
stressed-unstressed contrasts	*2nd syllable - 3rd syllable contrast*	*vowel reduction on the 1st syllable*
		vowel reduction on the 2nd syllable
		vowel reduction on both syllables

Regarding the selection of tasks, in perceptual training programs, two types of tasks are commonly used for auditory testing and training: identification tasks and discrimination tasks. These tasks complement each other, since the first aim at acquiring "distinctiveness" and the second, "equivalence" (Pisoni & Lively, 1995, p. 445). However, identification tasks can only be used when participants already have some knowledge of the L2, which is not the case in the present study, since our goal is to observe the effect of training in the early stages of L2 acquisition. Considering this, in the present study only discrimination tasks were possible, even though they may be lead to a phonetic training rather than promote phonological acquisition, as Pisoni & Lively (1995) point out.

To prevent participants from paying attention solely to phonetic differences, we resorted to a longer ISI, since longer ISI promote phonemic processing. Based on previous studies (Escudero & Wanrooij, 2010; Flege & Mackay, 2004; Mayr & Escudero, 2010), we set the ISI in all trials – pretest, post-test, and training trials – to 1200 ms. Additionally, the stimuli used in our study were recorded by different talkers, promoting phonetic variability, which in turn also contributed to categorical processing.

Discrimination tasks differ in difficulty, increasing from AX to AXB, to oddity tasks and to sequence recall tasks. In AX discrimination tasks, the listener must decide if the second stimulus (X) is either of the same category of the first (A) or different. In AXB discrimination tasks, participants are asked whether the second stimulus (X) is of the same category as the first (A) or the third (B). Therefore, listeners must retain the three stimuli in memory, before making their decision. In the three-stimuli oddity discrimination tasks, participants hear a sequence of three stimulus and must decide which one is different (the first, the second, or the third). Typically, oddity tasks include catch trials as well, in which all stimuli belong to the same category. Inclusion of catch trials induces more difficulty for the listeners, since there is an additional task for them to do, which is to decide whether the three stimuli are from the same category or not. In sequence recall tasks, participants must recall a sequence of stimuli auditorily presented to them. Since the participants had no knowledge or almost no knowledge of the L2 at hand, sequence recall tasks were judged too difficult. Consequently, AX and AXB tasks were used for training, and three-stimuli oddity tasks were used for testing. Table 5.3 gives an example of sequences for each discrimination task type we included in this study.

Table 5.3. Tokens' sequences for the contrast [ɐ]-[ɛ], as a function of type of task

task	trial type	example	answer
AX	AA	[ɐ]-[ɐ]	both tokens are of the same category
	AB	[ɐ]-[ɛ]	the two tokens are of a different category
	BA	[ɛ]-[ɐ]	the two tokens are of a different category
	BB	[ɛ]-[ɛ]	both tokens are of the same category
AXB	AAB	[ɐ]-[ɐ]-[ɛ]	the 1st token is of the same category as X
	ABB	[ɐ]-[ɛ]-[ɛ]	the 3rd token is of the same category as X
	BBA	[ɛ]-[ɛ]-[ɐ]	the 1st token is of the same category as X
	BAA	[ɛ]-[ɐ]-[ɐ]	the 3rd token is of the same category as X
Oddity with catch trials	ABB	[ɐ]-[ɛ]-[ɛ]	the 1st token is the odd
	AAB	[ɐ]-[ɐ]-[ɛ]	the 3rd token is the odd
	ABA	[ɐ]-[ɛ]-[ɐ]	the 2nd token is the odd
	BBA	[ɛ]-[ɛ]-[ɐ]	the 3rd token is the odd
	BAB	[ɛ]-[ɐ]-[ɛ]	the 2nd token is the odd
	BAA	[ɛ]-[ɐ]-[ɐ]	the 1st token is the odd
	AAA	[ɐ]-[ɐ]-[ɐ]	the three tokens are the of the same category
	BBB	[ɛ]-[ɛ]-[ɛ]	the three tokens are the of the same category

Two principles guided the design of the stimuli. The first was phonetic variability, previously reasoned. Phonetic variability may be achieved by different methods. It may be included by using stimuli recorded by different L1 talkers. Second, it may be promoted with the inclusion of stimuli different in nature: natural versus synthetic. Third, it may also refer to the use of stimuli in which specific features were manipulated (e.g., duration). Since studies demonstrated the positive effect of high variability in phonetic training with different talkers (Lively et al., 1993, 1994; Wong, 2012), we opted to apply this method in the present study. A total of seven EP talkers were used to record our stimuli: $T1^F$, $T2^F$, $T3^F$, $T4^M$, $T5^F$, $T6^F$, $T7^M$.[81]

The second principle was to use pseudowords over real words. This decision was based on the fact that participants were beginners, that is, at the early onset of the L2 acquisition process. Although beginners are expected to learn a significant amount of L2 vocabulary during one semester, they would probably not have acquired enough L2 vocabulary words, at least in the first training sessions. Furthermore, considering that the recruited learners were majoring in very different fields (Linguistics, Medicine, Economics, among others), there was also the probability that the purpose of the courses would influence the vocabulary acquired in class.

All tasks (questionnaires, instructions, trials) were built and conducted with Gorilla Experiment Builder (Anwyl-Irvine et al., 2018).[82]

The training program lasted twelve weeks, with recruitment included. Perceptual trainings differ in intensity and length. While some researchers condense the training length to a short period of time (Alves & Luchini, 2017; Bettoni-Techio, 2008; Schwab & Llisterri, 2014), others spread the training for a longer period of time, with shorter training sessions (Aliaga-García, 2017; Kissling, 2015). We opted for the latter alternative, for two reasons. First, most of the recruited participants were having Portuguese classes as an optional subject and may had no interest in completing an intensive training on a topic that is not central to their studies, particularly in the beginning of the academic semester.

[81] The subscript 'F' and 'M' indicate the gender of the talker, female or male, respectively.

[82] Access to this program was granted by LiFE group — Formal and Experimental Linguistics, of the Linguistics Research Centre of NOVA University Lisbon (CLUNL).

Second, it was expected that a combination of longer training and contact with EP in classes would be advantageous for acquisition.

Regarding the number of training sessions, it was determined that at least six sessions would be necessary to train all target contrasts. The six training sessions were presented over six weeks, one session per week. Other than these six weeks, additional time was planned for recruitment, pretest, and post-test. The timeline of the training program, with the tasks carried out in each phase, is displayed below, in Table 5.4.

Table 5.4. Timeline of the training program

WEEK 1 WEEK 2		recruitment
WEEK 3	PHASE 1	pretest session 1: background questionnaire + vowel discrimination test pretest session 2: stress discrimination test
WEEK 4		break
WEEK 5	PHASE 2	training session 1
WEEK 6		training session 2
WEEK 7		training session 3
WEEK 8		training session 4
WEEK 9		training session 5
WEEK 10		training session 6
WEEK 11		break
WEEK 12	PHASE 3	post-test session 1: follow-up questionnaire + vowel discrimination test post-test session 2: stress discrimination test

Recruitment was conducted in the first two weeks. Once contact was made with the students who volunteered to participate, they were given a week to complete the pretest – Phase 1 – divided into two sessions. In the first session we administrated the background questionnaire and the vowel discrimination test, and in the second session, participants completed the stress discrimination test. After Phase 1, we collected information on the students and assigned them to a training group, and students were given a weeklong break. From week five to ten, the training sessions were conducted – Phase 2. Participants were given a week to complete each session, and they could only access a new session after completing the previous one. Before Phase 3, another weeklong break was made. Phase 3, – the post-test –, included three parts, which were presented in two sessions. First, participants completed the follow-up questionnaire, and, in the same session, the vowel discrimination test. Second, in a different session, participants were presented with the stress discrimination test.

Students were recruited from different universities. To recruit a higher number of students, the training program was repeated in three moments, at each time, including new participants, all starting Beginner L2 Portuguese courses: in the first semester of the 2021/22 academic year, in the second semester of the same year, and in the first semester of the 2022/23 academic year.

In the next sections we describe in more detail the design for test and training sessions. The description is focused on the tasks presented to the intervention groups, that is, the Hungarian participants. Although the *Baseline* group, which included Portuguese participants, completed the same trials, the tasks we presented to this group were, to some extent, different from the tasks presented to the intervention groups. We describe the structure of the tasks completed by the *Baseline* group in 5.4.4.

5.1.1 Vowel discrimination tests

Vowel discrimination was assessed by means of oddity tasks. As mentioned above, seven vocalic contrasts were tested in two conditions. In the first condition, vowels showed *perceptual overlap* ([ɐ]-[ɛ], [ɛ]-[e], [e]-[i], and [ɐ]-[ɨ]), while in the second condition, vowels were contrasted in *stressed vs. unstressed* pairs ([a]-[ɐ], [ɛ]-[e], and [e]-[ɨ]). This resulted in a total of 42 change oddity trials: 7 contrasts × 3 order possibilities (AAB/ABA/AAB) × 2 odd targets (A or B). The tests also included 6 catch trials, with the vowels [a], [ɐ], [ɛ], [e], [i], and [ɨ]. Thus, each participant completed 48 trials altogether.

Pretest stimuli consisted of the [gV] tokens we used for the categorization experiment reported in Chapter 4. The post-test included the repetition of the pretest trials (post-test^R) as well as novel trials (post-test^N). For the novel trial we chose a syllable type that is frequent in EP speech: [sV]. As a result, vowel discrimination post-test consisted of 96 trials: 84 change trials (7 contrasts × 2 contexts ([gV] or [sV]) × 3 order possibilities (AAB/ABA/BBA) × 2 odd targets (A or B)) and 12 catch trials (6 vowels ([a], [ɐ], [ɛ], [e], [i] and [ɨ]) × 2 contexts ([gV] or [sV])). For familiarization tasks in the pretest and the post-test, we used the [tV] tokens recorded for the identification experiment (see Chapter 4).

Although in the pretest vowel discrimination was assessed with female talkers only, in the novel trials a male voice was also included to add phonetic variability. The order of talkers was the same in all trials: T1^F-T2^F-T3^F, in the pretest and the post-test^R, and T5^F-T6^F-T7^M, in the post-test^N.

Both the pretest and the post-test were designed with the same structure: (1) informed consent, (2) introduction, (3) familiarization task, and (4) main task. The introduction presented general instructions about the test, as well as the opportunity to adjust the sound volume. For volume adjustment, three tokens, produced by three Hungarian female talkers, were selected from the familiarization stimuli – [ta], [tɛ], and [ti] –, which the participant could listen to as many times as needed.

Familiarization consisted of instructions for the task followed by eight trials with immediate feedback. This number of trials was decided taking into consideration that it was the minimum for understanding the task without causing fatigue or boredom. Familiarization included two trials for each possible answer (1st is the odd, 2nd is the odd, 3rd is the odd, or the three are identical). The familiarization trials included easier sequences to discriminate (e.g., [to]-[ti]-[to]) along with sequences acoustically more similar (e.g., [tɛ]-[te]-[tɛ]). Feedback was provided in written form ("Correct!" or "Incorrect… try again"), in Hungarian, and no limit for try outs were set.

The main task of the pretest included instructions followed by 48 trials, presented in a single block. As for the post-test, the 96 trials were equally divided into three blocks and separated by 30-second pauses. In the main trials, no feedback was provided to the participants.

In Appendix VII, we list the trials created for vowel discrimination tests.

5.1.2. Stress discrimination tests

Stress discrimination was also evaluated with oddity tasks, using pseudowords contrasting only in stress position. To this purpose, we use a set of tokens created from nine trisyllabic pseudowords, retrieved from Correia et al. (2015), and contrasting in adjacent stress position (*1^st^-2^nd^ syllable* and *2^nd^-3^rd^ syllable*). The pseudowords included the vowels [i] and [u], exclusively, which are the only EP vowels not prone to quality change due to the absence of stress. Furthermore, the results of the identification experiment (Chapter 4) showed that Hungarian listeners have no problems in the perception of these two vowels. Table 5.5 displays the trisyllabic contrasts used for stress discrimination.

Table 5.5. Tokens used for stress discrimination assessment

	contrasts		
pseudoword	**1st-2nd syllable**	**2nd-3rd syllable**	**experimental task**
/tudiki/	[ˈtudiki]-[tuˈdiki]	[tuˈdiki]-[tudiˈki]	pretest
/kiruʃi/	[ˈkiruʃi]-[kiˈruʃi]	[kiˈruʃi]-[kiruˈʃi]	pretest and post-test^R^
/dutiku/	[ˈdutiku]-[duˈtiku]	[duˈtiku]-[dutiˈku]	pretest and post-test^R^
/tuduri/	[ˈtuduri]-[tuˈduri]	[tuˈduri]-[tuduˈri]	pretest and post-test^R^
/tikuru/	[ˈtikuru]-[tiˈkuru]	[tiˈkuru]-[tikuˈru]	pretest and post-test^R^
/kitumi/	[ˈkitumi]-[kiˈtumi]	[kiˈtumi]-[kituˈmi]	post-test^N^
/lutinu/	[ˈlutinu]-[luˈtinu]	[luˈtinu]-[lutiˈnu]	post-test^N^
/bikulu/	[ˈbikulu]-[biˈkulu]	[biˈkulu]-[bikuˈlu]	post-test^N^
/mudini/	[ˈmudini]-[muˈdini]	[muˈdini]-[mudiˈni]	post-test^N^

In the pretest, five pseudowords were used and participants had to complete 75 trials: 60 change trials (5 pseudowords × 2 stress contrasts (*1^st^-2^nd^ syllable* and *2^nd^-3^rd^ syllable*) × 3 order possibilities (AAB/ABA/BBA) × 2 odd targets (A or B)) and 15 catch trials (5 pseudowords × 3 stress positions (*1^st^ syllable*, *2^nd^ syllable*, and *3^rd^ syllable*)). In the post-test, we selected 4 pseudowords used in the pretest, and added 4 novel words, resulting in a total of 120 trials: 96 change trials (8 pseudowords × 2 stress contrasts (1^st^-2^nd^ syllable and 2^nd^-3^rd^ syllable) × 3 order possibilities (AAB/ABA/AAB) × 2 odd targets (A or B)), and 24 catch trials (8 pseudowords × 3 stress positions (*1^st^ syllable*, *2^nd^ syllable*, and *3^rd^ syllable*)).

As in Correia et al. (2015), two female talkers and one male talker recorded the stimuli, and a unique order of talkers was included: $T1^{F}$-$T2^{F}$-$T4^{M}$, in the pretest and the post-$test^{R}$, and $T5^{F}$-$T6^{F}$-$T7^{M}$, in the post-$test^{N}$.

Following Correia et al. (2015), Dupoux et al. (1997, 2001, 2008), Honbolygó et al. (2017), and Peperkamp et al. (2010), perception of consonant contrasts was also tested, as a control condition. The aim of including phoneme contrasts is to observe if listeners are able to perceive speech differences at segmental level, while exhibiting difficulties in discriminating contrasts at suprasegmental level. Stimuli consisted of CVCVCV pseudowords, stressed on the second syllable (e.g., [buˈpili]/[buˈʃili]; Table 5.6). Stimuli stressed on the second syllable was judged to be the most suitable, since this is the unmarked case in EP, therefore, acoustic differences between the syllables within a stimulus are less prominent. We also intended to make participants aware of the fact that they were listening to non-native words, since in Hungarian stress is fixed on the 1st syllable. The pairs of pseudowords used included the contrast [ʃ]-[p], [f]-[m], [b]-[p], [b]-[ʃ], [b]-[v], and [m]-[n]. In total, 48 trials were created for consonant contrast: 36 change trials (6 pairs of pseudowords × 3 order possibilities (AAB/ABA/BBA) × 2 odd targets (A or B)) and 12 catch trials (12 pseudowords).

Table 5.6. Contrasts and respective tokens used for consonant discrimination assessment

consonant contrast	tokens
[ʃ]-[p]	[buˈʃili]-[buˈpili]
[f]-[m]	[kuˈfitu]-[kuˈmitu]
[b]-[p]	[diˈbuki]-[diˈpuki]
[b]-[ʃ]	[tiˈburi]-[tiˈʃuri]
[b]-[v]	[tiˈfubi]-[tiˈvubi]
[m]-[n]	[zuˈmitu]-[zuˈnitu]

Stress discrimination tests were designed with a structure similar to that of the vowel discrimination tests: (1) informed consent, (2) introduction, (3) familiarization task, (4) main task. In the introduction we presented general instructions followed by sound volume adjustment. For this, three tokens were selected from the familiarization stimuli: *alma* [ˈɒlmɒ] 'apple', *cékla* ['ʦeːklɒ] 'beetroot', and *hintó* [ˈhintoː] 'carriage'. Each token was produced by a different Hungarian talker, the first two tokens were produced by two female talkers, while the third token was produced by a male talker, so that the familiarization task was consistent with the main task. Familiarization task had the same structure as the familiarization task for vowel discrimination tests: task explanation followed by eight trials (2 for each possible answer), with immediate feedback. As it will be detailed in section 5.3.2, we used sequences of real words, with different degrees of similarity.

Main trials of the pretest consisted of two blocks, separated by a 30-second pause. The first block included the 75 trials for stress discrimination, and the second block the 48 trials for consonant discrimination. Regarding the post-test main task, it consisted of 120 trials equally distributed for four blocks, with 30-second pauses in between. Consonant contrasts were not included in the post-test phase, since pretest results showed that Hungarian speakers easily perceive these contrasts (10.7% of error rate, 3.3% above the *Baseline* group).[83]

In Appendix VII all the trials created for stress and consonant discrimination tests are listed.

5.1.3. Training sessions

The six training sessions created for each group were designed so that they increase in difficulty from the first to the last session. In the first four sessions, participants completed AX discrimination tasks, and each contrast to be trained was presented in a separated block, that is, if four contrasts were trained in one session, four blocks of trials were created, one for each contrast. The 5th and 6th sessions aimed at practicing all the contrasts trained before, and in AXB discrimination tasks, which represented an increase in difficulty from the AX tasks. In the 5th session, each contrast was trained in a separate block, and in the last session, trials were randomized across contrasts. This randomization was

[83] Detailed results are displayed in Appendix X.

performed once before uploading the tasks to the experiment platform, so all participants who completed the task received the same order of trials. Every session included 48 trials, to achieve balance between sessions, as well as in the time devoted to the training by each group.

Talkers in the training sessions were the same as in the pretest. However, the distribution of talkers in the training tasks varied. In the AXB tasks (5^{th} and 6^{th} sessions), we follow the same method as for the oddity tasks in testing, that is, the order of the talkers in the trials remained the same throughout the trials (e.g., $T1^F$-$T2^F$-$T3^F$). AX tasks (1^{st} to 4^{th} sessions) presented different sequences of talkers, in order to achieve 48 trials for each session. A special attention was given to the equal distribution of the talkers. For example, for the contrast ['bifuʃi]-[bi'fuʃi] presented in the 1^{st} session to Group *Stress*, we use the sequences $T1^F$-$T2^F$, $T2^F$-$T4^M$, $T4^M$-$T1^F$). With a balanced inclusion of talkers, we aimed at avoiding participants being more exposed to a particular talker, thus getting more familiarized with one voice. Table 5.7 displays the distribution of talkers by the type of training task, and intervention group.

Table 5.7. Distribution of talkers by training tasks, for each intervention

intervention groups	talkers in AX sessions (1^{st} to 4^{th})	talkers in AXB sessions (5^{th} and 6^{th})
Group *Vowels*	$T1^F$, $T2^F$, $T3^F$, in alternate order	order: $T1^F$ - $T2^F$ - $T3^F$
Group *Stress*	$T1^F$, $T2^F$, $T4^M$, in alternate order	order: $T1^F$ - $T2^F$ - $T4^M$
Group *Vowels & Stress*	$T1^F$, $T2^F$, $T3^F$, in alternate order	order: $T1^F$ - $T2^F$ - $T3^F$

Participants were given a week to complete each session. To reduce participant attrition, in the end of each training session participants could complete a ludic task on Portuguese idiomatic expressions or popular sayings. This is referred to as the *bonus task* (see Appendix XII).

All sessions were structured in the same way: (1) informed consent, (2) instructions, (3) training trials, and (4) bonus task. Informed consent was the same as for the tests. Instructions included interactive examples, in which participants could try out the task, while reading the instructions. In AX sessions, we included two examples, and in AXB sessions, five. The main task always included 48 trials, with immediate feedback, equally divided into two parts, with a 30-second pause between them. In the end of the main trials, participants could exit the session, or continue to the bonus task.

After the first four sessions, we looked into the results, for the three groups, in order to decide if it was necessary to equally include in the last two sessions all contrasts previously trained, or if we should focus in some contrasts more than others.[84]

5.1.3.1. Training regimen for Group *Vowels*

Training for Group *Vowels* addressed discrimination of *perceptual overlap* and *stressed-unstressed* vowel contrasts ([ɐ]-[ɛ], [ɛ]-[e], [e]-[i], [ɐ]-[ɨ], and [a]-[ɐ], [ɛ]-[ɨ], [e]-[ɨ], respectively). To this purpose, target vowels were inserted in a [zV] context. This context was selected to match partially with the stimuli for Group *Vowels & Stress*, which consisted of [zVzV] pseudowords. Furthermore, [zV] is a structure with low occurrence in Hungarian (Halácsy et al., 2003), which avoids prominence of certain stimuli during training. All tokens were produced by talkers $T1^F$, $T2^F$, and $T3^F$.

The seven minimal pairs were trained in increasingly difficult order. In the first two sessions, we provided training for the pairs [a]-[ɐ], [ɛ]-[ɨ] and [e]-[ɨ], corresponding to the *stressed-unstressed* condition. We considered these contrasts less difficult for the learners, since no perceptual overlap was observed for them in the identification experiment (Chapter 4). To achieve the 48 trials, a filler contrast was included, [ɔ]/[o], in the second session. The 3rd and 4th sessions included the contrast of the vowels that showed *perceptual overlap*: [ɐ]-[ɛ] and [ɐ]-[ɨ], in the 3rd session, and [ɛ]-[e] and [e]-[i] in the in the 4th session. As explained before, the 5th and 6th sessions were used to practice all contrasts (excluding fillers). However, before creating the tasks for these sessions, we looked at the preliminary results of the training obtained in the first four sessions (Table 5.8, next page).

[84] These results were preliminary and aimed only at giving us a general idea of the results of the training. They were collected in November 2021.

Table 5.8. Preliminary results for the AX tasks (1st to 4th sessions), for Group *Vowels*

change trials		catch trials	
vowel contrast	**ER**	**vowel**	**ER**
[e]-[i]	50.0%	[ɐ]	82.0%
[ɛ]-[e]	50.0%	[e]	58.0%
[ɐ]-[ɨ]	38.0%	[ɨ]	52.0%
[ɐ]-[ɛ]	20.0%	[i]	40.0%
[a]-[ɐ]	12.0%	[ɛ]	28.0%
[ɛ]-[ɨ]	12.0%	[a]	4.0%
[e]-[ɨ]	6.0%		

As we can see in Table 5.8, [a]-[ɐ], [ɛ]-[ɨ], and [e]-[ɨ] display low error rates. However, in the catch trials, the vowel [ɐ] presented the higher error rate. Therefore, in the 5th and 6th sessions, the contrast [a]-[ɐ] was included in an equal number of trials as the more problematic contrasts. On the other way around, [ɛ]-[ɨ] and [e]-[ɨ] were trained in fewer trials.

Table 5.9 summarizes the training regimen for Group *Vowels*, and in Appendix VII, we list the trials used in each session of this training.

Table 5.9. Training regimen for Group *Vowels*

Session 1 AX task	*stressed-unstressed* condition [zɛ]-[zɨ], [ze]-[zɨ] 24 trials for each contrast, presented in separated blocks
Session 2 AX task	*stressed-unstressed* condition [za]-[zɐ], [zɔ]-[zo] (filler) 24 trials for each contrast, presented in separated blocks
Session 3 AX task	*perceptual overlap* condition [zɐ]-[zɨ], [zɐ]-[zɛ] 24 trials for each contrast, presented in separated blocks
Session 4 AX task	*perceptual overlap* condition [zɛ]-[ze], [ze]-[zi] 24 trials for each contrast, presented in separated blocks

Session 5 AXB task	contrasts [a]-[ɐ], [ɐ] [ɛ], [ɛ] [e], [e] [i], and [ɐ] [ɨ]: 8 trials each contrasts [ɛ]-[ɨ] and [e]-[ɨ]: 4 trials each each contrast presented in a separated block
Session 6 AXB task	contrasts [a]-[ɐ], [ɐ] [ɛ], [ɛ] [e], [e] [i], and [ɐ] [ɨ]: 8 trials each contrasts [ɛ]-[ɨ] and [e]-[ɨ]: 4 trials each all trials presented in one block, randomized

5.1.3.2. Training regimen for Group *Stress*

With respect to the training for Group *Stress*, the goal for this group was to improve discrimination of word stress, at suprasegmental level. Accordingly, stimuli for this task could only include the vowels [i] and [u], since in EP these are the only vowels that do not undergo unstressed vowel reduction. To train stress discrimination, a similar set of pseudowords used in testing was used in the training. The CVCVCV pseudowords for the training were: /vumipi/, /filuvi/, /bifuʃi/, /mupiʃu/, /ʒurigu/, and /zituli/ (Table 5.10).

Table 5.10. Tokens used for stress discrimination training

	contrasts	
pseudoword	**1st-2nd syllable**	**2nd-3rd syllable**
/vumipi/	[ˈvumipi]-[vuˈmipi]	[vuˈmipi]-[vumiˈpi]
/filuvi/	[ˈfiluvi]-[fiˈluvi]	[fiˈluvi]-[filuˈvi]
/bifuʃi/	[ˈbifuʃi]-[biˈfuʃi]	[biˈfuʃi]-[bifuˈʃi]
/mupiʃu/	[ˈmipiʃu]-[muˈpiʃu]	[miˈpiʃu]-[mupiˈʃu]
/ʒurigu/	[ˈʒurigu]-[ʒuˈrigu]	[ʒuˈrigu]-[ʒuriˈgu]
/zituli/	[ˈzituli]-[ziˈtuli]	[ziˈtuli]-[zituˈli]

In the first session, we trained the stress contrast that we expected to be easier to Hungarian speakers, *1st-2nd syllable* contrast, since this is a familiar vs. unfamiliar stressed location contrast.[85] In the second session, we trained the *2nd-3rd syllable* contrasts, which corresponded to a contrast between two pseudowords with unfamiliar stress location. The

[85] Since word stress in Hungarian is fixed on the 1st syllable, we considered this as a familiar case, and words stressed on the 2nd syllable or on the 3rd syllable as unfamiliar cases.

3rd and 4th sessions aimed at training both conditions, and we divided our stimuli set into the two sessions. Similar to the other training groups, the 5th and 6th sessions were aimed at training all contrasts and, before creating the tasks, we inspected the preliminary results from the training obtained in the first four sessions (Table 5.11).

Table 5.11. Preliminary results for the AX tasks (1st to 4th sessions), for Group *Stress*

trial type	stress contrast; stress location	mean error rate
change	1st-2nd syllables	84.3%
	2nd-3rd syllables	70.0%
catch	1st syllable	33.2%
	2nd syllable	66.9%
	3rd syllable	40.4%

Catch trials with stress located in the 1st syllable achieved the highest accuracy, as expected, but stress contrasts between the 1st and 2nd syllables were problematic and had the highest error rate, contrary to what we predicted. Therefore, it was decided that both contrasts (*1st - 2nd syllable* and *2nd - 3rd syllable*) should be equally trained in the 5th and 6th sessions. Table 5.12 summarizes the training regimen for Group *Stress,* and, in Appendix VII, we list the trials prepared for these sessions.

Table 5.12. Training regimen for Group *Stress*

Session 1 AX task	*1st - 2nd syllable* condition [ˈbifuʃi]-[biˈfuʃi], [ˈvumipi]-[vuˈmipi], [ˈfiluvi]-[fiˈluvi], [ˈmupiʃu]-[muˈpiʃu] 12 trials for each contrast, presented in separated blocks
Session 2 AX task	*2nd - 3rd syllable* condition [biˈfuʃi]-[bifuˈʃi], [vuˈmipi]-[vumiˈpi], [fiˈluvi]-[filuˈvi], [muˈpiʃu]-[mupiˈʃu] 12 trials for each contrast, presented in separated blocks
Session 3 AX task	*1st - 2nd syllable* condition and *2nd -3rd syllable* condition [ˈvumipi]-[vuˈmipi], [vuˈmipi]-[vumiˈpi], [ˈmupiʃu]-[muˈpiʃu], [muˈpiʃu]-[mupiˈʃu] 12 trials for each contrast, presented in separated blocks
Session 4 AX task	*1st - 2nd syllable* condition and *2nd -3rd syllable* condition [ˈbifuʃi]-[biˈfuʃi], [biˈfuʃi]-[bifuˈʃi], [ˈfiluvi]-[fiˈluvi], [fiˈluvi]-[filuˈvi] 12 trials for each contrast, presented in separated blocks

Session 5 AXB task	all contrasts previoulsy trained + [ˈʒurigu]-[ʒuˈrigu], [ʒuˈrigu]-[ʒuriˈgu], [ˈzituli]-[ziˈtuli], [ziˈtuli]-[zituˈli] 4 trials for each contrast, presented in separated blocks
Session 6 AXB task	all contrasts previoulsy trained + [ˈʒurigu]-[ʒuˈrigu], [ʒuˈrigu]-[ʒuriˈgu], [ˈzituli]-[ziˈtuli], [ziˈtuli]-[zituˈli] 4 trials for each contrast, all trials presented in one block, randomized

5.1.3.3. Training regimen for Group *Vowels & Stress*

Finally, Group *Vowels & Stress* was targeted at training both vowel quality contrasts and stress contrasts. To this purpose, disyllabic pseudowords that included /a/, /ɛ/ and /e/ were used. Recall that in EP, these vowels are produced as [a], [ɛ] or [e] in stressed position, and [ɐ] and [ɨ] in unstressed position. A [zVzV] structure was selected based on two reasons. First, it contains only one consonant, [z], which contributes to control consonantal effects. Second, words resulting from this structure had no meaning, thus configuring pseudowords.[86] From the [zVzV] structure we created a series of pseudowords complying with the three conditions: *vowel reduction on the first syllable*, *vowel reduction on the second syllable*, and *vowel reduction on both syllables*. Table 5.13 presents the set of pseudowords and respective contrasts, used in this training.

Table 5.13. Pseudowords and contrasts created for training of Group *Vowels & Stress*

vowel reduction location	pseudoword	contrast
1st syllable	/zazi/	[ˈzazi]-[zɐˈzi]
	/zɛzi/	[ˈzɛzi]-[zɨˈzi]
	/zezi/	[ˈzezi]-[zɨˈzi]
	/zɔzi/	[ˈzɔzi]-[zuˈzi] (filler)
2nd syllable	/ziza/	[ˈzizɐ]-[ziˈza]
	/zizɛ/	[ˈzizɨ]-[ziˈzɛ]
	/zize/	[ˈzizɨ]-[ziˈze]
	/zizɔ/	[ˈzizu]-[ziˈzɔ] (filler)

[86] The word [ˈzɒzi] (phonetically close to the EP [ˈzɔzi]) appears in Hungarian as a proper female first name. However, it is a loanword, with rare presence in Hungarian speech (Oravecz et al., 2014). Furthermore, in this study, this word is only used as a filler.

both syllables	/zaza/	[ˈzazɐ]-[zɐˈza]
	/zazɛ/	[ˈzazɨ]-[zɐˈzɛ]
	/zaze/	[ˈzazɨ]-[zɐˈze]
	/zɛza/	[ˈzɛzɐ]-[zɨˈza]
	/zɛzɛ/	[ˈzɛzɨ]-[zɨˈzɛ]
	/zɛze/	[ˈzɛzɨ]-[zɨˈze]
	/zazɔ/	[ˈzazu]-[zɐˈzɔ] (filler)
	/zɛzɔ/	[ˈzɛzu]-[zɨˈzɔ] (filler)

To comply with the method of increasing difficulty adopted for training, 1^{st} and 2^{nd} sessions comprised contrasts in which only one syllable included vowel reduction, and stimuli used the 3^{rd} and 4^{th} sessions consisted of the pseudowords in which both syllables in the pseudowords of the contrasts included EP vowels that undergo reduction. In order to achieve 48 trials in each session, we created fillers that complied with the conditions trained in each session. For example, in the sessions aimed at *vowel reduction on the 2^{nd} syllable* we included the filler /zizɔ/ ['zizu]-[zi'zɔ]. Fillers were not included, however, in the 5^{th} and 6^{th} sessions. Before creating the tasks for these two last sessions, we collected and analyzed the preliminary results (Table 5.14).

Table 5.14. Preliminary results for the AX tasks (1^{st} to 4^{th} sessions), for Group *Vowels & Stress*

trial type	vowel reduction location	mean error rate
change	vowel reduction on the 1^{st} syllable	43.1%
	vowel reduction on the 2^{nd} syllable	58.8%
	vowel reduction on both syllables	14.7%
catch	stress on the 1^{st} syllable	29.1%
	stress on the 2^{nd} syllable	39.6%

The error rate in conditions with vowel reduction on only one of the syllables was considerably higher than in the condition *vowel reduction on both syllables.* While in the latter, two vowels are contrasting (e.g., [ˈzazɨ]-[zɐˈzɛ]), vowel reduction on only one of the syllables included pseudowords that contrast in one vowel only (e.g., [ˈzazi]-[zɐˈzi]). Even if stress contrast is not perceived, the presence of two different vowels makes the words more dissimilar. As for contrasts with only one contrasting vowel, it's possible that student's attention was not directed solely or mainly to the vowels, but also to durational differences (e.g., in [ˈzazi]-[zɐˈzi], [i] changes mainly in duration). Taking this into consideration, we decided that the three conditions should be equally trained in the last two sessions. Table 5.15 presents the training regimen for Group *Vowels & Stress*, and Appendix VII displays the trials used in each session.

Table 5.15. Training regimen for Group *Vowels & Stress*

Session 1 AX task	*vowel reduction on the 2 nd syllable* conditon [ˈzizɐ]-[ziˈza], [ˈzizɨ]-[ziˈzɛ], [ˈzizɨ]-[ziˈze], and [ˈzizu]-[ziˈzɔ] (filler) 12 trials for each contrast, presented in separated blocks
Session 2 AX task	*vowel reduction on the 1 st syllable* conditon [ˈzazi]-[zɐˈzi], [ˈzɛzi]-[zɨˈzi], [ˈzezi]-[zɨˈzi], and [ˈzɔzi]-[zuˈzi] (filler) 12 trials for each contrast, presented in separated blocks
Session 3 AX task	*vowel reduction on both syllables* conditon [ˈzazɐ]-[zɐˈza], [ˈzazɨ]-[zɐˈzɛ], [ˈzazɨ]-[zɐˈze], and [ˈzazu]-[zɐˈzɔ] (filler) 12 trials for each contrast, presented in separated blocks
Session 4 AX task	*vowel reduction on both syllables* conditon [ˈzɛzɐ]-[zɨˈza], [ˈzɛzɨ]-[zɨˈzɛ], [ˈzɛzɨ]-[zɨˈze], and [ˈzɛzu]-[zɨˈzɔ] (filler) 12 trials for each contrast, presented in separated blocks
Session 5 AXB task	all contrasts previoulsy trained (except fillers) 4 trials for each contrast, presented in separated blocks
Session 6 AXB task	all contrasts previoulsy trained (except fillers) 4 trials for each contrast, all trials presented in one block, randomized

5.2. Participants

To conduct the study, we recruited Hungarian speakers enrolled in EP courses for Beginners, in Hungary. As previously mentioned, participants were divided into three groups: Group *Vowels*, Group *Stress*, and Group *Vowels & Stress*. A *Baseline* group with EP native speakers was used to validate the stimuli of both the tests and the training sessions. Additionally, we also recruited Portuguese native speakers and Hungarian native speakers to record the stimuli. The profile of these speakers (*talkers*) will be briefly described in the section "Stimuli design and preparation" (5.3.2).

By means of questionnaires, we collected information on participants' age, gender, academic qualifications, L1, and knowledge and frequency of use of L2. In the case of the Hungarian participants, we also requested information about previous and present contact with Portuguese, as well as about the reasons for attending a Portuguese language course. Additionally, Hungarian students had to complete a follow-up questionnaire, focused on the perceptual training (motivation in the program, concentration during the tasks, self-evaluation of the progress, among others).

5.2.1. Hungarian participants

The Hungarian learners were recruited from L2 Portuguese Beginner courses offered in nine Hungarian universities: Budapest Business School, Budapest Metropolitan University, Corvinus University of Budapest, Eötvös Loránd University (Budapest), Károli Gáspár University of the Reformed Church in Hungary (Budapest), University of Debrecen, University of Pécs, University of Public Service Ludovika (Budapest), and University of Szeged. From these institutions, Eötvös Loránd University offers two Portuguese courses: Portuguese as *major* and Portuguese as *minor*. Considering the differences in nature and class load between these two courses, in the present study, they were considered as two separate affiliations (ELTEBA and ELTEMI, respectively). The Portuguese courses in the other universities are optional subjects, in some cases hosting external students. Other than having enrolled in a L2 Portuguese Beginner course, the inclusion criteria included being a native speaker of Hungarian, and being aged between 18 and 45.

Recruitment was done with the consent and assistance of the teacher responsible for each course. In some cases, the teachers requested a presentation of the study to the university in written form. A personal visit by the author of the dissertation to each class was organized with the help of the teachers, in the first week of the course. When the visit could not be done in person, an online appearance in the class was set. Before the visit, the responsible teachers for each course were given more detailed information, and they were given the possibility to try out a task, but they were asked not to give any information to students, but rather wait for the author of the dissertation to do it. After the visit, teachers were only occasionally contacted, to collect some extra information on the course or about students. Teachers were encouraged to give extra marks, if possible, to the students who completed the training, to avoid attrition. The importance of these extra marks, however, was decided by each responsible teacher.

During the above-mentioned visit, the general purpose of the study was explained to the students. Information about the phases and estimated duration of tasks and training were described. The students were informed that participation required good internet connection and a computer or laptop, as well as headphones or earbuds. No detailed description of the aim of the experiment was given to the students, but there was a brief mention to the focus on auditory abilities in EP. Also, the students were informed that the data collected during the training would not be shared with any third parties and would be used solely for the purpose of this study. It was clearly mentioned that no financial reward would be provided. However, they were further informed that, upon completion of the training, they would receive a certificate of participation from CLUNL – Linguistics Research Centre of NOVA University Lisbon, and the Department of Applied Linguistics and Phonetics of Eötvös Loránd University, the host institutions for the project (see Appendix III). Names and email addresses of the students who volunteered were provided later by the teacher of each course.[87]

As mentioned earlier, the training program was repeated in three semesters: the first and the second semesters of 2021/22, and the first semester of 2022/23. In the three semesters, the program was conducted with Hungarian students enrolled in EP Beginner courses, that is, with participants with no previous formal learning of EP. Between these three semesters, no changes occurred regarding the teachers responsible for the courses, the aim of the courses, or the main textbooks chosen by the teachers. Thus, the courses conditions between the three moments did not differ substantially.

In total, 129 students were recruited and started the training, of which only 79 completed it. From the 79 participants, 13 were excluded for the following reasons: four

[87] The collecting of the volunteers' contacts was done in the absence of the author of the dissertation, so that students would feel more comfortable to decide if they wanted to participate or not.

quit the Portuguese course in the beginning or middle of the training (but completed the training), six were aged above 45, two were Germain-Hungarian early bilinguals, and one participant was born in Portugal. For analysis of the results, and hence forward, only 66 Hungarian participants will be considered: 20 allocated to Group *Vowels*, 22 to Group *Stress*, and 24 to Group *Vowels & Stress*. Table 5.16 lists the number of participants by affiliation and by intervention group.

Table 5.16. Distribution of participants by affiliation and intervention group

affiliation (institution, course load, n. of participants)	n. of participants by group
Budapest Business School (180'/week, 3 participants)	*Vowels* (n = 1) *Stress* (n = 2)
Corvinus University of Budapest (180'/week, 17 participants)	*Vowels* (n = 4) *Stress* (n = 6) *Vowels & Stress* (n = 7)
Eötvös Loránd University: PLE as *major* (270'/week, 12 participants)	*Vowels* (n = 3) *Stress* (n = 4) *Vowels & Stress* (n = 5)
Eötvös Loránd University: PLE as *minor* (180'/week, 13 participants)	*Vowels* (n = 4) *Stress* (n = 4) *Vowels & Stress* (n = 5)
Károly Gáspár University of the Reformed Church in Hungary (180'/week, 3 participants)	*Vowels* (n = 1) *Stress* (n = 1) *Vowels & Stress* (n = 1)
Budapest Metropolitan University (180'/week, 4 participants)	*Vowels* (n = 2) *Stress* (n = 1) *Vowels & Stress* (n = 1)
University of Public Service Ludovika (180'/week, 4 participants)	*Vowels* (n = 1) *Stress* (n = 1) *Vowels & Stress* (n = 2)
University of Debrecen (90'/week, 2 participants)	*Vowels* (n = 1) *Vowels & Stress* (n = 1)
University of Pécs (90'/week, 3 participants)	*Vowels* (n = 2) *Stress* (n = 1)
University of Szeged (90'/week, 5 participants)	*Vowels* (n = 1) *Stress* (n = 3) *Vowels & Stress* (n = 1)

Participants mean age was 22 (range: 18 to 44), 51 were female speakers and 15 male speakers. The majority (n = 52) were attending bachelor's degree programs. Forty-two participants reported having had previous contact with the Portuguese language, mainly during vacations or through friends. However, nine of them reported a stronger contact. These participants either had lived 2 to 5 months in Portugal (with an Erasmus program, for example), or previously studied Portuguese in the secondary school or university (started the course but quitting), or had a close relationship with a Portuguese native speaker. As described in the exploratory data (5.5.9), however, previous contact with Portuguese had no significant effect on the results in the pretest.

All the participants were asked about possible auditory and/or respiratory problems, temporary or chronic. One Hungarian participant reported to have a mild cold, during the first part of the pretest. After a personal contact with this participant and considering that the results of her pretest did not deviate from the general results, it was decided to include this participant in the study.

Finally, information about other L2 knowledge (level and frequency of use) was also collected in the background questionnaire. All students reported knowledge of English, on levels B1 (n = 3), B2 (n = 23), C1 (n = 37), and C2 (n = 3). After English, Spanish was the most reported L2 (n = 20), followed by Italian (n = 12), and German (n = 12). Other L2 were French (n = 4), Romanian (n = 3), Serbian, Slovak, Catalan, and Dutch (each reported by one participant).[88]

5.2.2. Portuguese participants

The Portuguese native speakers who tested the stimuli – the *Baseline* group – were recruited via internet, through a google form. The request for participation was mainly shared in courses of the School of Social Sciences and Humanities of NOVA University Lisbon, and with the assistance of CLUNL – Linguistics Research Centre of NOVA University Lisbon. The inclusion criteria were to be aged between 18 and 45, and to be a monolingual native speaker of EP. For some of the tasks, the dialectal area of the participant was also

[88] Languages reported bellow B1 level or 'rarely' or 'never' spoken were not included in this description.

taken into consideration. The 45 participants included in the *Baseline* group were aged between 18 and 45 (mean = 28), 31 were female speakers, 13 male speakers, and one reported “other” for gender.

Baseline participants performed one of two tasks: test stimuli validation (n = 15) or training stimuli validation (n = 30).

The group who completed the tests stimuli validation was presented with three tests: vowel discrimination, stress discrimination and consonant discrimination. As will be described in section “Stimuli validation”, the three tests were completed in a sequential manner (vowels → stress → consonants), and of the 15 participants who completed the first part, only 12 completed the second part, and 11 the last part. Additionally, two participants were not from the standard dialectal area, for which reason their results were excluded from vowel discrimination analysis. As a result, the number of participants considered for test stimuli validation was 13 for vowel discrimination, 12 for stress discrimination, and 11 for consonant discrimination.

Regarding Portuguese speakers who completed the training tasks, given the extension of the training trials, the 30 participants were equally divided into five groups, each one completed a part of the trials. Therefore, the number of participants testing each training trial was six.

5.3. Materials

In this section, we will provide detailed information about the material prepared and used for the tests and training sessions. It starts with a brief description of the questionnaires administered to the participants and follows with a section on recording, selection, and preparation of the stimuli.

5.3.1. Questionnaires

Two types of questionnaires were included in the study: a language background questionnaire, administered to all participants, and a follow-up questionnaire, administered to the Hungarian participants. These questionnaires were conducted to assess the adequacy of the participants to the study, and to collect information for data analysis. Whenever possible, questions were presented in multiple-choice form, to standardize the measurements.

In the case of the Hungarian students, the language background questionnaire was administrated in the session of the pretest, in the L1 of the participants (Hungarian), and it contained three blocks. In the first block, name (first and family names), age, gender, academic qualifications, and affiliation (university/course) were asked. The names of the participants were necessary for communication throughout the training program. Once data collection was complete, names were deleted from the database. In the end of the first block, participants were asked about possible auditory and/or respiratory problems, temporary or chronic. The second block of the questionnaire covered language knowledge. First, participants were asked about L1 knowledge (3 options were given), and after that, about L2 knowledge: language, proficiency level, and frequency of use. For language proficiency, six levels were presented ('breakthrough - A1', 'waystage - A2', 'threshold - B1', 'vantage - B2', 'advanced - C1', and 'mastery - C2'), corresponding to the levels of the Common European Framework of Reference for Languages. Participants were advised to read information about these levels, if necessary, for which purpose, we provided them a link to a relevant site.[89] For frequency of use, the following options were given: 'daily', 'weekly', 'monthly', 'sometimes in a year', 'rarely', and 'never'. The third block of the questionnaire requested information on contact with Portuguese (past and present) and on the motivation to have enrolled in a Portuguese course.

The language background questionnaire presented to the Portuguese participants was similar to the one presented to the Hungarian students, with the difference that the questions regarding name of the participants, affiliation, previous contact with Portuguese and motivation for learning Portuguese were not presented. However, we request the address (city) in the last 10 years, to control for dialectal differences.

[89] They were given the following option to examine proficiency levels in the Common European Framework of Reference for Languages:
https://hu.wikipedia.org/wiki/K%C3%B6z%C3%B6s_Eur%C3%B3pai_Referenciakeret
In the questionnaire presented to the Portuguese participants, the link was the following:
https://pt.wikipedia.org/wiki/Quadro_Europeu_Comum_de_Refer%C3%AAncia_para_as_L%C3%ADnguas

The follow-up questionnaire was administered to the Hungarian participants who completed all the training sessions, as it was included in the post-test. As in the background questionnaire, participants were asked about possible auditory and/or respiratory problems. Next, they were asked to provide information about the number of hours attended in the Portuguese course during our training program, and about changes in the contact with the Portuguese language, outside the classes. The second block of the questionnaire included six close-ended questions to assess the following aspects: motivation to participate in the program, attention during the tasks and fatigue at the end of each task, as well as comments on the adequacy of the duration of the tasks, sufficiency to achieve progress, and whether the tasks were interesting. The last block included two open questions about how participants judged the effect of the training on their auditory abilities and pronunciation abilities. The questionnaire concluded with an open question for other comments.

Appendix II displays the three questionnaires described above, as well as translations into English of open-questions' answers.

5.3.2. Stimuli design and preparation

In the section "Study design" we referred to different sets of stimuli created for testing and training: monosyllabic ([gV], [sV], [zV]), disyllabic ([zVzV]), and trisyllabic tokens (CVCVCV, with [i] or [u] as vowels). Table 5.17 summarizes the stimuli (for main trials) and corresponding tasks and talkers.

Table 5.17. Summary of the stimuli used in the study

testing/training parameters	experimental task	stimuli structure	example	talkers
	pretest and post-test^{R}	CV (gV)	[gɐ]-[gɛ]	T1^{F}, T2^{F}, T3^{F}
vowel discrimination	post-test^{N}	CV (sV)	[sɐ]-[sɛ]	T5^{F}, T6^{F}, T7^{M}
	training	CV (zV)	[zɐ]-[zɛ]	T1^{F}, T2^{F}, T3^{F}
vowel and stress discrimination	training	CVCV (zVzV)	[ˈzazi]-[zɐˈzi]	T1^{F}, T2^{F}, T3^{F}

stress discrimination	pretest and post-test^{R}	CVCVCV (V=[i] or [u])	[ˈdutiku]-[duˈtiku]	T1^{F}, T2^{F}, T4^{M}
	post-test^{N}		[ˈlutinu]-[luˈtinu]	T5^{F}, T6^{F}, T7^{M}
	training		[ˈvumipi]-[vuˈmipi]	T1^{F}, T2^{F}, T4^{M}
consonant contrast	pretest		[zuˈmitu]-[zuˈnitu]	T1^{F}, T2^{F}, T4^{M}

Other than the stimuli mentioned above, we also created stimuli for the familiarization tasks. For *vowel discrimination*, we use the tokens created for the perceptual study reported in Chapter 4 ([tV] syllables). For *stress discrimination* tasks, and *vowels and stress discrimination* tasks, we opted for real words of the L1. Considering that Hungarian speakers are supposedly stress 'deaf', it would be impossible to create a familiarization task with tokens contrasting in stress location. We opted to use real words of the L1 for the following reasons. First, using real words would prepare participants for a 'speech mode', that is, it would predispose them to perceive the sounds of the main trials as speech and, second, the real words selected were polysyllabic words, which would prepare participants to be tested in more complex words than the monosyllables they were presented to in the vowel discrimination tests.

Each familiarization stimuli set included tokens that varied in acoustic similarity: some were more different (e.g., [to]-[ti], [ˈhintɒ]-[ˈʦeːklɒ], [ˈgalu]-[ˈmasɐ]) and others more similar (e.g., [tɛ]-[te], [ˈhintɒ]-[ˈhintoː], [ˈgalu]-[ˈgatu]). Tokens were produced by female and male native speakers (Hungarian and Portuguese). We selected talkers for the familiarization tasks on the basis of the main trials. If in the main trials we presented productions by three female talkers (as for *vowel discrimination* and *vowel and stress discrimination*), then familiarization tasks also included productions by female talkers. If the main trials presented productions of two female and one male talkers (as for *stress discrimination*), then the familiarization trials also included productions of two female and a male talker. Table 5.18 (next page) lists the stimuli used in the familiarization tasks.

Table 5.18. Stimuli used in the familiarization trials (testing and training tasks)

	stimuli for familiarization tasks conducted with the Hungarian participants
[tV]	[ta]; [tɔ]; [te]; [tɛ]; [ti]; [to]; [tø]; [tu]
real polysyllabic words	*alma* ['ɒlmɒ] 'apple'; *cékla* ['ʦeːklɒ] 'beetroot'; hinta ['hintɒ] 'swing'; hintó ['hintoː] 'carriage'; *kutya* ['kucɒ] 'dog'; *macska* ['mɒʧkɒ] 'cat'; paripa ['pɒripɒ] 'steed'; póniló ['poːniloː] 'pony'; *répa* ['reːpɒ] 'carrot'; *Sándor* ['ʃaːndor] 'proper noun'; síkság ['ʃiːkʃaːg] 'plain'
	stimuli for familiarization tasks conducted with the Portuguese participants
[tV]	[ta]; [te]; [tɔ]; [to]; [tɐ]; [tɨ]; [tɛ]; [ti]; [tu]
real polysyllabic words	*amanha* [ɐ'meɲɐ] 'he/she prepares'; *homem* ['ɔmẽj] 'men'; caro ['kaɾu] 'expensive'; carro ['kaʀu] 'car'; galo ['galu] 'rooster'; gato [gatu] 'cat'; maçã [mɐ'sɐ̃] 'apple'; massa ['masɐ] 'dough'; *oiço* ['ojsu] 'I listen'; *ouço* ['owsu] 'I listen'; sede ['sɛdɨ] 'headquarters'; sede ['sedɨ] 'thirst'

Regarding recording of the testing and training stimuli, due to the pandemic restrictions, recordings were conducted individually at each talkers' home, and in three different moments. In the beginning of 2021, recordings of stimuli for the pretest were made. Later, in August 2021, stimuli for the training were recorded, and in November, the last recording session took place for the stimuli of the novel trials of the post-test. In the first and second sessions, a TASCAM DR-05 V2 digital recorder with a Beyerdynamic MCE 85 BA shotgun condenser microphone was used. In the last session, recordings were made with a TASCAM DR-22WL digital recorder with a BEYERDYNAMIC MCE 86 shotgun condenser microphone. All the equipment was provided by the LabCC–Communication Sciences Laboratory of NOVA University Lisbon.

Recording settings were fixed to wav file format, 44.100 Hz sampling frequency, mono, and 32-bit depth. For each set of stimuli, the first recording was made by the EP speaking supervisor of this dissertation. These tokens were then selected and edited, to provide an example to the other talkers[90], who were asked to hear them carefully. Along with the examples, they were instructed to make the recordings at an early or late hour, when background noise is reduced, and on a location judged to have low reverberation.

[90] These tokens were also used as stimuli for the study. The author of this dissertation was excluded from the recordings due to the fact that some participants had classes with her, thus they would be more familiarized with her voice than the other participants.

They were instructed to listen to the examples, and then record their own stimuli, at a normal conversational rate.

The stimuli were inserted in a carrier sentence, to control for intonation. In the cases of the monosyllables, an example of a real word containing a target syllable, underlined, was presented to the talkers, to control for vowel quality (1). In the case of the disyllabic and trisyllabic tokens ((2) and (3)), the orthographic forms indicated the stressed syllable, as well as vowel quality. To avoid stress misplacement, the stressed syllables were also underlined in the target words. Talkers were asked to record one token of each sentence. In (1), (2), and (3), examples of the carrier sentences for the monosyllables, disyllables, and trisyllables are presented. From the carrier sentences, we extracted tokens in boldface for further selection. The full list of carrier sentences is presented in Appendix V.

(1) (GAVETA) **[gɐ]**. Diga **[gɐ]**, por favor: **[gɐ]**.
'(DRAWER) [gɐ]. Say [gɐ], please: [gɐ].'

(2a) **ZAZA**. Diga **ZAZA**, por favor: **ZAZA**.
'ZAZA. Say ZAZA, please: ZAZA.'

(2b) **ZAZÁ**. Diga **ZAZÁ**, por favor: **ZAZÁ**.
'ZAZÁ. Say ZAZÁ, please: ZAZÁ.'

(3a) **DÚTICO**. Diga **DÚTICO**, por favor: **DÚTICO**.
'DÚTICO. Say DÚTICO, please: DÚTICO.'

(3b) **DUTICO**. Diga **DUTICO**, por favor: **DUTICO**.
'DUTICO. Say DUTICO, please: DUTICO.'

(3C) **DUTICU**. Diga **DUTICU**, por favor: **DUTICU**.
'DUTICU. Say DUTICU, please: DUTICU.'[91]

[91] According to EP orthographic rules, when stress falls on the penultimate syllable, with few exceptions, no diacritic is needed to signal the stressed syllable, since this is not the marked case (e.g., <dutico> [duˈtiku]). When stress falls on the antepenultimate syllable, a diacritic is always needed to signal stress (e.g., <dútico> [ˈdutiku]*)*. When stress falls on the last syllable and the word ends with <u>, no diacritic is needed to signal stress location (e.g., <duticu> [dutiˈku]). These rules are implicitly known by EP native speakers. It was interesting to notice that the trisyllabic stimuli – in which no vowel reduction is present – displayed more difficulties in the recordings than the disyllabic tokens – that included vowel reduction –, which is in line with the importance vowel reduction in EP stress (Correia et al., 2015; Delgado-Martins, 1986).

Portuguese tokens were recorded by a total of twelve Portuguese native talkers, eight female and four male. These speakers were recruited through CLUNL – Linguistics Research Centre of NOVA University Lisbon, or directly by the author of the dissertation. However, only the recordings of five of the female talkers and of two of the male talkers were evaluated as having good sound quality, similar speech rate, and appropriate intonation across tokens. This evaluation was made by two trained phoneticians, the author and the EP speaking supervisor of this dissertation. All selected talkers were from the region of Lisbon and were aged between 41 and 48 (Table 5.19).

Table 5.19. Portuguese native speakers selected for the study

talker	gender	age	adress in the last 10 years
$T1^F$	female	46	Lisboa
$T2^F$	female	42	Lisboa
$T3^F$	female	46	Lisboa
$T4^M$	male	42	Lisboa
$T5^F$	female	48	Loures
$T6^F$	female	41	Vila Franca de Xira, Lisboa
$T7^M$	male	45	Lisboa

Stimuli for the familiarization tasks presented to the Hungarian participants (Hungarian real words), were produced by five native-speakers of Hungarian (aged 28 to 45), four female and one male. Since all five Hungarian speakers were trained phoneticians, tokens were recorded in simple citation form and in isolation. The recordings were made using an external sound card attached to a PC or laptop, and an omnidirectional condenser head-mounted microphone. The external sound card and microphone were provided by MTA—ELTE "Lendület" Lingual Articulation Research Group, Hungary.

After each recording – both in the case of stimuli for main trials as for familiarization trials –, the tokens were transferred as wav sound files from the digital recorder's SD card to a laptop computer. All tokens were then checked by the author of the dissertation and the EP speaking supervisor, and a repetition of the recording was asked, whenever judged necessary.

After the recordings were made, preparation of the stimuli was conducted in sequential steps. First, we eliminated background noise with Audacity software (Audacity Team, 2020). This procedure was manually done, since recording conditions varied across talkers. After this, all sound sequences were segmented manually and the target tokens (three repetitions of each stimulus, in bold in the carrier sentences) were extracted as individual wav files, using TextGrids and a Praat script.

From each talker we selected one token of each item. The selection of the final tokens produced by the Portuguese talkers was different for monosyllables and polysyllables. Selection of the [gV] and [tV] tokens was as reported in Chapter 4: based on measurements of the vocalic sections' duration and perceptual judgments of intonation and sound quality. As for the polysyllabic stimuli, these were selected on the basis of perceptual judgments made by the author and the EP speaking supervisor of this dissertation, and, when needed, based on acoustic properties of the syllable or the vowels. Acoustic measures were obtained with two types of labeling for each token, conducted in Praat. In one tier we segmented syllables, while in another, vowels within these syllables. The vocalic sections' boundaries were identified as the first and last voice bar of the section where the complete formant structure could be observed. If the vowel was preceded or followed by a nasal consonant (e.g., /kitumi/), and the vowel was completely nasalized, that is, there was no clear boundary between the vowel and the consonant, the boundary was set in the middle of the syllable, and acoustic measures for the vocalic section were conducted only on the mid-section of the first or second half of the vowel. With a script in Praat, we extract syllables' duration, and, on the vocalic sections, we measured vowel duration, f0, F1, F2, and maximum intensity. f0, F1 and F2 values were calculated from a median of the values measured in a 10% window in the midsection of the vowel. Time step was set to 0.01 ms, number of formants below 5.5 kHz to 5, and window length to 25 ms. Pitch floor was set to 100 Hz or 75 Hz and ceiling to 500 Hz or 300 Hz, for female and male talkers, respectively, as recommended by Praat (Boersma & Weenink, 2020). Acoustic measurements carried out on the stimuli are reported in Appendix VI.[92]

The selection of the stimuli produced by the Hungarian talkers for the familiarization trials with Hungarian participants was made having in consideration quality judgements by the Hungarian speaking co-supervisor and the author of this dissertation.

Regarding monosyllabic stimuli, as reported in Chapter 4, vowel length of some of the tokens was manually manipulated, with Praat, to guarantee a similar vowel duration across tokens.

Mean intensity was equalized across all tokens to 70 dB, using a Praat script, and 1200 ms of silence was inserted before each token (corresponding to the ISI), also with a Praat script.

5.4. Procedure

Once the structure of the tests and the training regiments was decided, and stimuli were recorded and prepared, we proceeded to the creation of the tasks in Gorilla. A similar design was kept throughout the different test and training tasks, so participants get familiarized early on with the appearance of the tasks, and that the task design did not draw their attention from the task itself. Written instructions for the tests, training sessions and questionnaires were presented in the L1 of the participants (Hungarian or EP). Familiarization stimuli were also in the L1 of participants, given that the aim of the familiarization trials was to give the participants the opportunity to practice the experimental tasks, so that potential difficulties in the test trials would be a result of difficulties in discrimination, and not of problems in understanding the task.

Randomization was set to be performed between participants by Gorilla, but only for the test tasks (pretest, post-test^{R} and post-test^{N}). As for familiarization trials and training sessions trials, they were manually randomized before uploaded to Gorilla, so that the order presented to participants would be the same.

[92] Although we performed measurements for all the tokens used in this study, these measurements were only aimed for consultation in case of doubts in the selection of the tokens. We did not conduct an analysis on these measurements.

After creating the tasks in Gorilla, the following step was to pilot them. Testing tasks were all tested first by the author of the dissertation, and then piloted by three Hungarian native speakers: the co-supervisor of this dissertation, a Hungarian teacher of EP, and a Hungarian speaker with no studies in Linguistics. Piloting was sequential: after the first speaker completed the task, changes were made and only after we presented to the second speaker the modified version of the task. The same method was followed for the second revision. In the end, a final pilot session was carried out with the first speaker. As for the training tasks, they were also firstly tested by this dissertation' author and after, only tasks for the 1st sessions (AX task) and the 5th sessions (AXB) for each group were piloted by one of the Hungarian speakers who piloted the tests. The tasks created for EP native speakers were piloted by the first supervisor of the dissertation and a master student of Linguistics. Piloting of these tasks followed an identical sequential procedure as the one conducted for the tasks for the Hungarian participants.

With the tasks tested and corrected, we then proceeded to our perceptual training program.

Once the names and email addresses of the recruited students were collected, two steps followed. First, a participant identification number (ID) was assigned by the author of the dissertation and associated with a given email of each participant. The list of email addresses and ID was uploaded to the tasks, in Gorilla. Second, an email was sent to the participants, which included the information given during the visit. It also included the instructions for the first session of the pretest, with an indication that they would receive an email from Gorilla with an access link and a participant ID.[93] Each email sent to Hungarian participants included instructions in writing and in a video. Written instructions were presented in Hungarian, translated by a native Hungarian speaker, while the video, made by the author of this dissertation, was in English.

During the training program, monitorization of students' progress in each task was based on the information provided by Gorilla: in each task, the indication of 'Pending',

[93] Upon activating the participant ID in each task, in the experiment platform, an automatic message was sent by the program to each student, with a unique link to access the task.

‘Live’, or ‘Completed’ was shown for each participant ID in the platform. When necessary, a reminder email was sent to the participants.

The tasks could only be accessed after participants provided their consent. After consenting, the first information displayed was always that the task consisted of an auditory exercise, and therefore use of headphones or earbuds was compulsory.

The participant's answers in the trials could only be given by left clicking the mouse or touch pad. To ensure each participant the possibility of reading the instructions at his/her own pace, no time limit was defined in the instruction screen. Instead, a ‘Continue’ button was provided: by clicking on it, participants could step to the next screen.

When judged necessary, we inserted a 30-second pause in a session. In these pauses, participants were asked not to leave the side of the computer or laptop, and not to interrupt the experiment. After the pauses, the trials always restarted automatically.

Next, we will detail the procedure in each phase of the program, and, in Appendix XI, we present computer screens of different moments of the tasks.

5.4.1. Pretest

After accessing the link provided by Gorilla and by entering their unique ID, participants were presented the consent form. Only by clicking ‘I accept the above terms and wish to participate in the experiment’, followed by ‘Next’, the participants accessed the first task: the background questionnaire.

Once this task was completed, participants were automatically redirected to the vowel discrimination test. This test included an introduction, a volume adjustment screen, the familiarization task, and the main task. The introduction included the structure of the test, and the information that an opportunity to adjust the volume would be provided, and that it was an important step to complete the test. It also included a request not to interrupt the test, and to answer as fast as possible, since no time limit was set for answering.

The next section was volume adjustment. Participants were shown three sounds, which they could hear repeatedly until they reached a comfortable sound volume. They were informed that once the volume was set, they should not change it throughout the experiment. After this, participants would step to the next section – the familiarization task –, by clicking ‘Continue’.

The first part of the familiarization task included interactive instructions, that is, the participants could try out each step of a trial, while reading the instructions. The second part included a sequence of eight oddity trials with feedback. After each answer, participants received a written message: ‘Correct!’ or ‘Incorrect… try again!’. If the answer was correct, they would automatically step onto the next trial. If the answer was incorrect, they were presented with the same trial immediately, and had as many tryouts as needed. Once the familiarization section was completed, a message appeared (‘End of the familiarization task. Please do not quit the experiment.’), and the main task began.

The main task started with written instructions. Participants were not explicitly informed that the sounds were produced by Portuguese native speakers, but since they were informed of the general topic of the study, they were expected to be aware of this. Instructions also warned participants that the test would end with the message ‘End of the experiment.’, and that until then, they could not interrupt the session. The instructions were followed by the oddity trials, without pause. In each trial, when participants clicked on the answer rectangle, the following trial would start automatically.

At the end of the main task, participants received the message ‘End of experiment. Thank you for participating.’ with an open question for optional comments. They would end the test by clicking the button ‘Conclusion’.

When participants were assigned a ‘Completed’ status in the vowel discrimination pretest, a new email with general instructions for the second session of the pretest was individually sent, along with an automatic email from Gorilla with a new link. Participant ID remained the same throughout the entire study.

Stress discrimination test followed a similar procedure as the vowel discrimination test.

Participants took around 15 minutes to complete the first session of the pretest (background questionnaire and vowel discrimination test), and 25 minutes to complete the second session of the pretest (stress discrimination).

Figure 5.1 (next page) displays a computer screen of one trial.

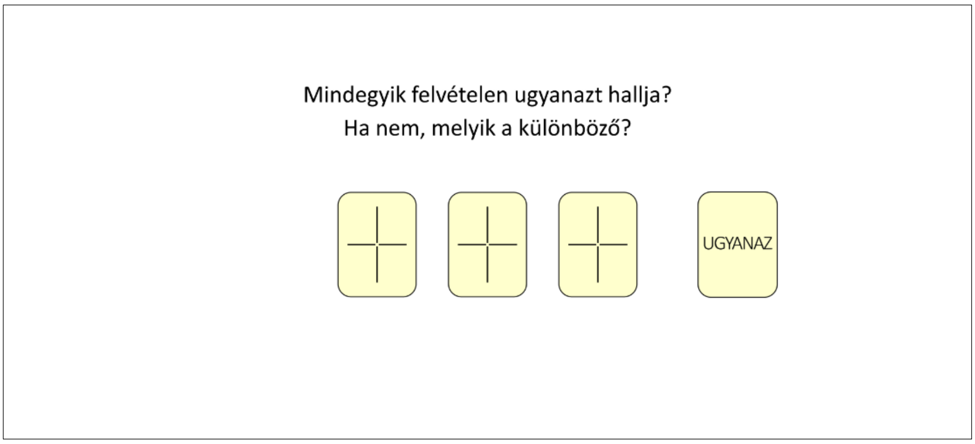


Figure 5.1. Computer screen of the main trials of the vowel discrimination pretest. ('Do you hear the same in each of the recordings? If not, which one is different?')

5.4.2. Training sessions

After participants completed the pretest phase, they were assigned to one of the three groups: Group *Vowels*, Group *Stress*, or Group *Vowels & Stress*. The assignment of students to intervention groups was only partly randomized. Specifically, in order to have approximately the same number of students from each affiliation in each training group, we randomly assigned students from each course into three groups.[94]

Once groups were formed, an email with instructions about the training was sent to each participant. This message consisted of an explicit explanation of the linguistic feature the training was focused on, which was presented in a language accessible to speakers with no or little knowledge in phonetics and phonology (see Appendix VIII). Along with the written instructions in Hungarian, a video with the same instructions in English, recorded by the author of the dissertation, was also included.

Immediately after receiving the email with the instructions for the training, participants also received an automatic email from Gorilla, with the link for the first session. When accessing the link, participants were asked to accept the informed consent, after which they started the training. All sessions followed the same procedure. Instructions were presented only in the beginning[95], and included the number of the session, a warning

[94] At the end of the study, this balance was not completely achieved, due to student attrition, as shown in Table 5.16.

[95] In the tests, instructions were presented before the familiarization tasks as well as before the main tasks.

to use headphones, and interactive examples of the tasks. Main training trials would start automatically after the examples.

Once the training trials were completed, participants were given two choices: exit the session, or continue to the bonus task.

As previously explained, participants had one week to complete each training session, and they would only receive the link to a new session, after completing the previous one.

Excluding the time for the bonus task, participants took 8 to 15 minutes to complete each of the first four sessions (the AX discrimination tasks), and 10 to 18 minutes in the case of sessions 5 and 6 (the AXB discrimination tasks).

Figure 5.2 displays a computer screen of one AX trial, for Group *Vowels*.

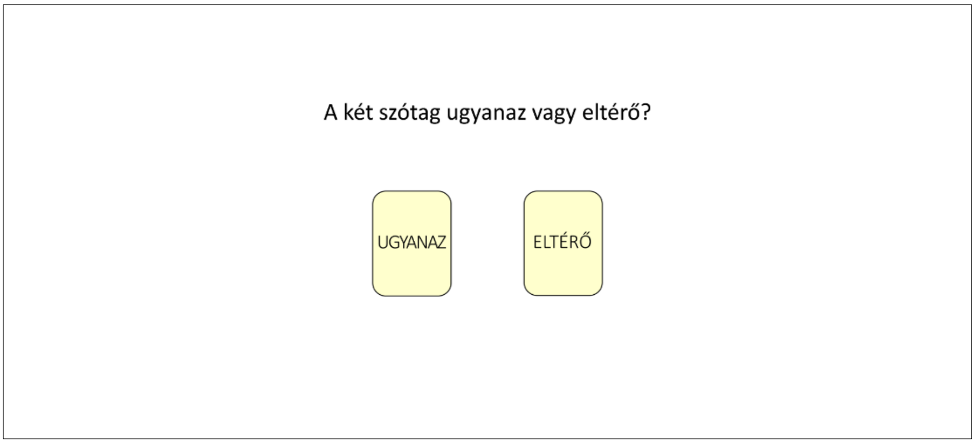


Figure 5.2. Computer screen of the main trials of an AX task for Group *Vowels*.
('Are the two syllables identical or different?')

5.4.3. Post-test

Once the six training sessions were completed, a one-week break was made, after which, a new email was sent to the students, with instructions for the post-test, similar to those sent in the pretest phase. The post-test procedure was also similar to the pretest, with the difference that instead of the background questionnaire, participants were presented with a follow-up questionnaire before the vowel discrimination test.

Participants took approximately 22 minutes to complete the follow-up questionnaire and the vowel discrimination test, and 20 minutes to complete the stress discrimination test.

Students who completed the training were given a certificate of participation, signed by the project host institutions, as mentioned previously. Additionally, after the conclusion of the training program (post-test phase included), they were given the opportunity to complete the other two trainings. For example, if a participant had completed the training for Group *Vowels*, he/she could receive a link from Gorilla for training sessions of Groups *Stress* and *Vowels & Stress*. This possibility, however, was only provided upon participants' request.

5.4.4. Stimuli validation

To validate the stimuli used in the study, we had Portuguese native speakers (the *Baseline* group) completing the same trials presented to the intervention groups. However, the procedure was different, hence we describe it here, separately. As explained in section 5.2.2, *Baseline* participants were divided into two groups: one to validate test stimuli, and the other to validate training stimuli.

The Portuguese participants who validated test stimuli had to completed four tasks in three sessions. The first session included a background questionnaire (see Appendix II) and the vowel discrimination test trials. This test consisted of the [gV] and [sV] trials (n = 96), equally divided into three blocks, separated by 30-second pauses. The second session included the stress discrimination test trials (n = 135). These trials were divided into 5 blocks, each containing 27 trials, separated by 30-second pauses. The third session included the trials for consonant discrimination (n = 48), divided into two blocks, and separated by a 30-second pause. All the trials in each session were randomized across blocks, and between participants, by Gorilla.

The procedure for the tasks for test stimuli validation was the following: after collecting the email addresses of the volunteers, an individual participant ID was created for each, and written instructions were sent, by email, as well as the link to access the first part. Once participants had completed one part, a new link for the following task was sent.

Validation of the training stimuli was more complex due to the elevated number of trials of the training tasks (n = 600). These trials were divided into five sets, and each set included one block with [zV] trials, one block with [zVzV] trials, and one block with trisyllabic training trials (for the structure, see Appendix IX). Each set was completed by six native speakers of EP, that is, in total 30 speakers were needed to complete the training stimuli validation. Similar to test stimuli validation, all trials were randomized across the blocks, and between participants. Other than the discrimination tasks, participants were also presented with a background questionnaire in the beginning.

The procedure for training stimuli validation was different than that for test stimuli validation in one aspect. While in the latter case three different links were provided for each of the three tests, in the training stimuli validation tests we provided one single link, with a pause between each section. Participants were able to quit the test and resume it later in these pauses, if they felt necessary, or they could complete two or three parts in one go.

The design for both test stimuli and training stimuli validation was similar to the tasks presented to the Hungarian participants, with the difference that all the written instructions and familiarization stimuli were presented in EP.

The structure of the task for test stimuli validation and for training stimuli validation are displayed in Appendix IX.

5.5. Results: exploratory data

In this section we present the exploratory analysis conducted on the data collected from the different experimental tasks – pretest, post-test, and training sessions – as well as from the questionnaires. This is an important step of the study, since exploring the data visually and with descriptive statistics allows us to have different perspectives of the results, which in turn, is essential to understand the outcomes and draw conclusions. The aim of this section is not to discuss the results (Discussion is presented in 5.6), but rather to investigate the data in order to find patterns in the results.

Data was downloaded from Gorilla in csv format. Raw data was processed and analyzed with the software R (R Core Team, 2014). Statistical analysis was performed by fitting linear mixed-effects logistic regression models with the GLMER function from the lme4 package (Bates et al. 2015), and pairwise comparisons of least-squared means were conducted with the LSMEANS function from the lsmeans package (Length, 2018). For the testing and training tasks, we collected participants' responses, the dependent variable, coded as 1 or 0 (1 = incorrect, 0 = correct). Regarding reaction times, we opted to exclude this parameter from the analysis due to the difficulties in controlling this variable on online tasks (Holden et al., 2019). The predictors of the model (the independent variables) were all effect-coded. These variables as well as the statistical models' formulas are reported for each analysis, in the beginning of each subsection. The independent variables were inserted in the models as fixed effects, with the exception of *participant*, which was inserted as random effect.

The statistical models were run by creating subsets of the main data. This procedure was necessary for two reasons. First, some predictors were only observed for some features. For example, *stress location* (*1st-2nd syllable* vs. *2nd-3rd syllable*) was a predictor for stress discrimination but not for vowel discrimination. By contrast, *vowel contrast type* (*perceptual overlap* vs. *stressed-unstressed*) was observed only for vowel discrimination, but not in stress discrimination. Accordingly, we analyzed a *vowel discrimination subset data* and a *stress discrimination subset data* separately. Second, in some cases the model would fail due to its complexity (e.g., *test* (3 levels) * *group* (4 levels) * *vowel contrast* (7 levels)).

Statistical summaries of the results – count, mean error rate, standard deviation, standard error of the mean, and confidence interval (at 95% significance level) – were calculated using the function summarySE (Wang, n. d.). These summaries were used for the graphs, in which we display mean error rate and standard error of the mean (represented in the error bars). Tables with the summaries are presented in Appendix X.

Our first step in exploring the data was to look at the results for the *Baseline* group (5.5.1). After, we looked at the results for consonant contrasts and catch trials for the Hungarian participants (5.5.2 and 5.5.3), both operating as control conditions for discrimination abilities. We then proceeded to the results for vowel discrimination (5.5.4 and 5.5.5) and for stress discrimination (5.5.6 and 5.5.7). These sections are organized as follows. First, we analyzed the results for each test – pretest, post-test^R and post-test^N, comparing the behavior between intervention groups, as well as their performance with that of the *Baseline* group. In these analyses we first plotted individual mean error rates, to observe individual variability. Other than investigating the overall results in each test, we also analyzed the results considering the conditions (e.g., *perceptual overlap* vowel contrasts vs. *stressed-unstressed* vowel contrasts). Having investigated the results for each test, we then looked at progress of the intervention groups, from the pretest to the post-test^R and from the pretest to the post-test^N. In the analyses of progress, *the Baseline* group was not included. We also analyzed the results of the training sessions, for each intervention group (5.5.8). Finally, we concluded by briefly observing the effect of nonlinguistic factors in individual progress achieved with the training program (5.5.9).

5.5.1. Stimuli validation

Previous studies approach validation of stimuli by native speakers in two ways. Some researchers set a minimum threshold of accuracy – usually 90% or 95% – in order to include the stimuli (Rato, 2013; Ylinen et al., 2010). In other studies, no threshold is defined, and researchers rather compare the behavior of native speakers with the behavior of non-native speakers, by looking at the differences in the results for each group (Correia et al., 2015; Grenon et al., 2019). We opted for the last method, on the grounds of the following reasons. First, in the case of stimuli for stress discrimination without vowel reduction we could not expect a high accuracy by the Portuguese listeners since previous studies indicate that Portuguese native speakers show stress 'deafness', with error rate values of around 20% (Correia et al., 2015).[96] Second, we have no data on EP vowel perception by Portuguese native speakers that allow us to determine the threshold for excluding stimuli. Comparisons between the intervention groups' results and the *Baseline* group results are presented in sections 5.5.4 and 5.5.6.

Other than the above-mentioned comparisons, we also looked separately at the results obtained for the *Baseline* group, for test stimuli validation and training stimuli

[96] This value is for AXB trials that included disyllabic and trisyllabic non-words. In the second experiment of this study, a sequence recall task, mean error rate was above 75%.

validation. The aim of this analysis was to investigate if there were significant differences in discrimination within each type of stimuli included in test and training tasks: monosyllabic stimuli, disyllabic stimuli, and trisyllabic stimuli. Specifically, we aimed at confirming that differences observed in the results for the Hungarian participants between, for example, trials with stimuli used in the post-test^{R} and trials with stimuli used in the post-test^{N} were not due to deviant stimuli.

For each type of stimuli, we looked separately at testing stimuli (presented in oddity tasks) and training stimuli (presented in AX and AXB tasks).

5.5.1.1. Monosyllabic stimuli validation

Three sets of monosyllabic stimuli were used in the study: [gV] (in oddity tasks, in the pretest and the post-test^{R}), [sV] (also in oddity tasks, in the post-test^{N}), and [zV] (in AX and AXB tasks, in training). Table 5.20 displays the independent variables and the statistical models' formulas for the analysis of the monosyllabic stimuli used in the tests.

Table 5.20. Independent variables and mixed-effect models' formulas for the analysis of the monosyllabic testing stimuli validation

<table>
<tr><th>independent variables</th><th></th><th>levels</th></tr>
<tr><td>participant</td><td>13 levels:</td><td>PT003', 'PT005', 'PT006', etc.</td></tr>
<tr><td>stimuli set</td><td>2 levels:</td><td>'[gV]', '[sV]'</td></tr>
<tr><td>trial type</td><td>2 levels:</td><td>'change', 'catch'</td></tr>
<tr><td>vowel contrast</td><td>13 levels:</td><td>'[ɐ]-[ɛ]', '[ɛ]-[e]', '[e]-[i]', '[ɐ]-[ɨ]', '[a]-[ɐ]', '[ɛ]-[ɨ]', [e]-[ɨ]', '[a]', '[ɐ]', '[ɛ]', '[e]', '[ɨ]', '[i]'</td></tr>
<tr><th></th><th colspan="2">GLMER models (family = binominal)</th></tr>
<tr><td>null model</td><td colspan="2">answer ~ 1 + (1|participant)</td></tr>
<tr><td>model 1</td><td colspan="2">answer ~ stimuli set + (1|participant)</td></tr>
<tr><td>model 2</td><td colspan="2">answer ~ trial type + (1|participant)</td></tr>
<tr><td>model 3</td><td colspan="2">answer ~ vowel contrast + (1|participant)</td></tr>
</table>

Results from the ANOVA between the null model and each of the models displayed in Table 5.19 (next page) were as follows. *Stimuli set* ([gV] vs. [sV]) did not affect results, nor *trial type* (*change* vs. *catch*). However, we found a significant effect of *vowel contrast* in the discrimination (χ^2(14, N = 12) = 107.99, p < .001; Figure 5.3).

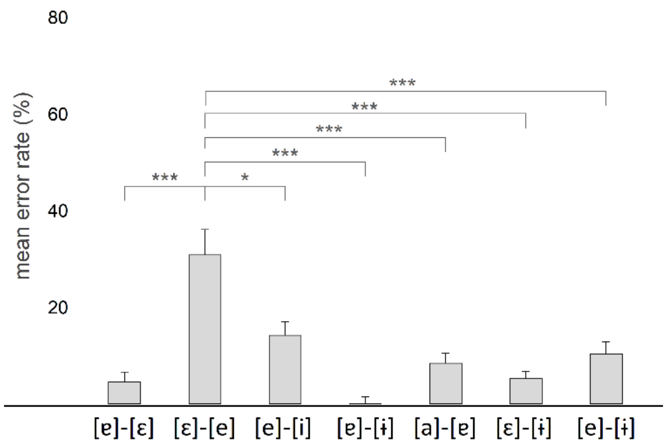


Figure 5.3. Mean error rates as a function of vowel contrast, for the *Baseline* group, for trials with monosyllabic testing stimuli

Pairwise comparisons showed that discrimination of [ɛ]-[e] and [e]-[i] posed more difficulty than the other contrasts ([ɛ]-[e] vs. [ɐ]-[ɛ], p < .001; [ɛ]-[e] vs. [e]-[i], p = .02; [ɛ]-[e] vs. [ɐ]-[ɨ], p < .001; [ɛ]-[e] vs. [a]-[ɐ], p < .001; [ɛ]-[e] vs. [ɛ]-[ɨ], p < .001; [ɛ]-[e] vs. [e]-[ɨ], p < .001).

This difficulty was also observed in the trials with the monosyllabic training stimuli ([zV]). The models and variables for the analysis of the results for these trials are displayed in Table 5.21.

Table 5.21. Independent variables and mixed-effect models' formulas for analysis of the monosyllabic training stimuli validation

independent variables		**levels**
participant	30 levels:	PT033', 'PT034', 'PT035', etc.
task type	2 levels:	'AX', 'AXB'
vowel contrast	7 levels:	'[ɐ]-[ɛ], '[ɛ]-[e]', '[e]-[i]', '[ɐ]-[ɨ]', '[a]-[ɐ]', '[ɛ]-[ɨ]', '[e]-[ɨ]'

	GLMER models (family = binominal)
null model	answer ~ 1 + (1|participant)
model 1	answer ~ task + (1|participant)
model 2	answer ~ vowel contrast + (1|participant)

In these trials, *vowel contrast* also affected significantly the results (χ^2(6, N = 8) = 71.73, *p* < .001; [ɛ]-[e] vs. [ɐ]-[ɛ], *p* = .004; [ɛ]-[e] vs. [ɐ]-[ɨ], *p* < .001; [ɛ]-[e] vs. [a]-[ɐ], *p* = .001; [ɛ]-[e] vs. [ɛ]-[ɨ], *p* < .001; [ɛ]-[e] vs. [e]-[ɨ], *p* = .01; [e]-[i] vs. [ɐ]-[ɛ], *p* < .001; [e]-[i] vs. [ɐ]-[ɨ], *p* < .001; [e]-[i] vs. [a]-[ɐ], *p* < .001; [e]-[i] vs. [ɛ]-[ɨ], *p* < .001; [e]-[i] vs. [e]-[ɨ], *p* < .001; [ɛ]-[e] vs. [e]-[i], *p* = .003). No significant effect of *task type* (*AX* vs. *AXB*) was found. Figure 5.4 displays the results by vowel contrast.

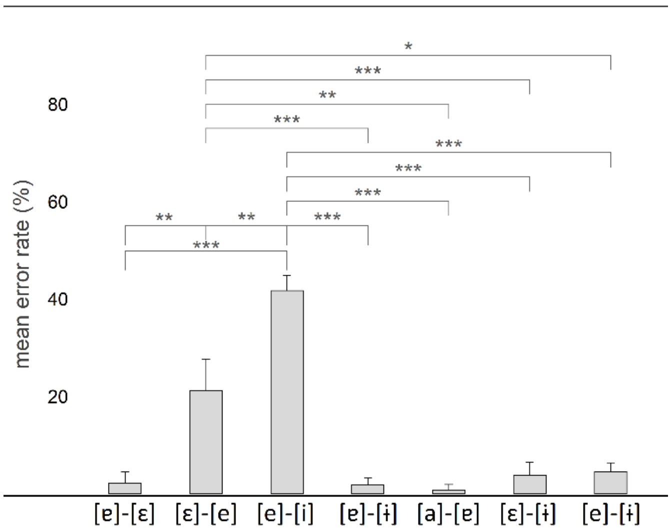


Figure 5.4. Mean error rates as a function of vowel contrast, for the *Baseline* group, for trials with training stimuli

5.5.1.2. Disyllabic stimuli validation

Disyllabic stimuli consisted of the [zVzV] pseudowords presented to Group *Vowels & Stress* in the training trials. Recall that these stimuli included three conditions: *vowel reduction on the 1st syllable* (e.g., [ˈza̲zi]-[zɐ̲ˈzi]), *vowel reduction on the 2nd syllable* (e.g., [ziˈza̲]-[ˈzizɐ̲]), and *vowel reduction on both syllables* (e.g., [ˈza̲zɐ̲]-[zɐ̲ˈza̲]). Table 5.22 presents the independent variables and models run to analyze trials with these stimuli.

Table 5.22. Independent variables and mixed-effect models' formulas for analysis of the disyllabic training stimuli validation

independent variables		**levels**
participant	30 levels:	PT033', 'PT034', 'PT035', etc.
task type	2 levels:	'AX', 'AXB'
vowel reduction	3 levels:	'reduction on the 1st syllable', 'reduction on the 2nd syllable' 'reduction on both syllables'
	GLMER models (family = binominal)	
null model	answer ~ 1 + (1|participant)	
model 1	answer ~ task + (1|participant)	
model 2	answer ~ vowel reduction + (1|participant)	

Portuguese native speakers achieved very low mean error rates (Figure 5.5), and neither *task type* nor *vowel reduction* placement affected the results.

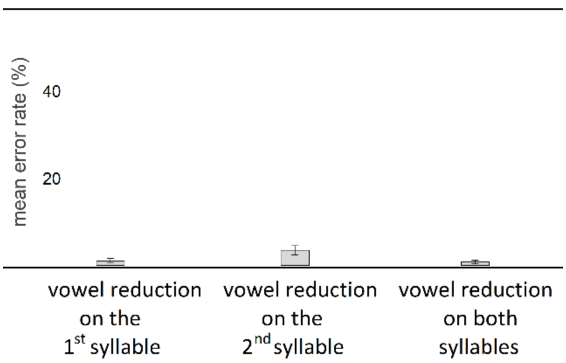


Figure 5.5. Mean error rates as a function of vowel reduction location, for the *Baseline* group

5.5.1.3. Trisyllabic stimuli validation

Trisyllabic stimuli consisted of CVCVCV pseudowords, contrasting in stress location or in consonant (e.g., [ˈtudiki]-[tuˈdiki] or [kuˈfitu]-[kuˈmitu], respectively). All the pseudowords included only the vowels [i] and [u], to exclude vowel reduction.

Stimuli contrasting in consonant was used only as a control condition. Thus, we looked only into the mean error rate, which was 7.4% across contrasts ([p]-[ʃ]: 4.5%, [m]-[f]: 5.1%, [b]-[ʃ]: 6.1%, [f]-[v]: 7.1%, [n]-[m]: 9.1%, [p]-[b]: 12.6%).

Regarding discrimination of stress contrasts, similar to the monosyllabic stimuli validation (5.5.1.1), we analyzed separately testing stimuli and training stimuli. Table 5.23 presents the independent variables and mixed-effect models used for the testing stimuli analysis.

Table 5.23. Independent variables and mixed-effect models' formulas for analysis of the trisyllabic testing stimuli validation

independent variables		**levels**
participant	12 levels:	PT005', 'PT006', 'PT007', etc.
stimuli set	2 levels:	'pretest', 'novel'
trial type	2 levels:	'catch', 'change'
contrast location	5 levels:	'1^{st}-2^{nd} syllable', '2^{nd}-3^{rd} syllable', '1^{st} syllable', '2^{nd} syllable', '3^{rd} syllable'
	GLMER models (family = binominal)	
null model	answer ~ 1 + (1|participant)	
model 1	answer ~ stimuli set + (1|participant)	
model 2	answer ~ trial type + (1|participant)	
model 3	answer ~ contrast location + (1|participant)	

No significant differences were found between results for trials with pretest (and post-test^{R}) stimuli and results that included post-test^{N} stimuli, but change trials and catch trials' results differed significantly (Figure 5.6, next page).

As we can see in Figure 5.6, there were significant differences between the results achieved for *1^{st}-2^{nd} syllable* contrast trials and the other conditions (*1^{st}-2^{nd} syllable* vs. *2^{nd}-3^{rd} syllable*, $p < .001$; *1^{st}-2^{nd} syllable* vs. *1^{st} syllable*, $p = .003$; *1^{st}-2^{nd} syllable* vs. *2^{nd} syllable*, $p = .02$; *1^{st}-2^{nd} syllable* vs. *3^{rd} syllable*, $p < .001$; Figure 5.6).

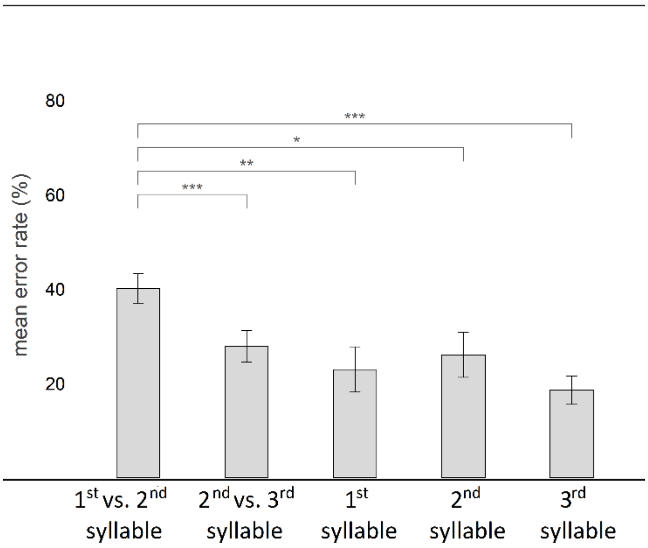


Figure 5.6. Mean error rates as a function of stress contrast location, for the *Baseline* group, for trials with trisyllabic testing stimuli

With respect to training stimuli validation, the independent variables and models used in the analysis are displayed below, in Table 5.24.

Table 5.24. Independent variables and mixed-effect models' formulas for analysis of the trisyllabic training stimuli validation

independent variables		**levels**
participant	30 levels:	PT033', 'PT034', 'PT035', etc.
task type	2 levels:	'AX', 'AXB'
contrast location	2 levels:	'1st-2nd syllable', '2nd-3rd syllable'
	GLMER models (family = binominal)	
null model	answer ~ 1 + (1|participant)	
model 1	answer ~ task + (1|participant)	
model 2	answer ~ contrast location + (1|participant)	

The higher mean error rate in the *1st-2nd syllable* contrasts observed for testing stimuli was also present in the results of the training trials. In these trials, *contrast location* was a significant factor (χ^2(1, N = 3) = 40.267, $p < .001$; Figure 5.7, next page). *Task type* did not affect significantly the results.

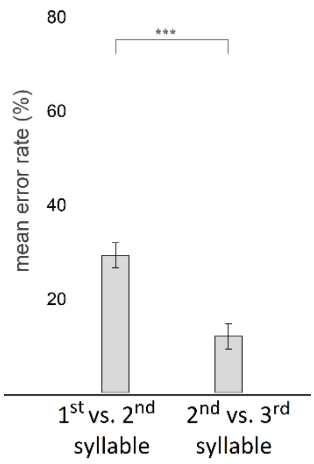


Figure 5.7. Mean error rates as a function of stress contrast location, for the *Baseline* group, for trials with trisyllabic training stimuli

Having analyzed the results for the *Baseline* group, we will now turn our attention to the intervention groups.

5.5.2. Consonant *versus* stress discrimination by the intervention groups

In the pretest session, we presented the participants a task for consonant contrast discrimination (e.g., [kuˈfitu]-[kuˈmitu]). These trials were included as control condition for stress perception. Overall, discrimination of consonant contrast did not pose a problem to the participants (mean error rate across groups: 10.7%). To compare consonant and stress discrimination, we run the models displayed in Table 5.25 (next page).

Table 5.25. Independent variables and mixed-effect models' formulas for analysis of consonant *versus* stress discrimination

independent variables		**levels**
participant	66 levels:	'BGE03', 'ELTEBA12', 'CORV04', etc.
group	3 levels:	'Vowels', 'Stress', 'Vowels & Stress'
condition	2 levels:	'stress discrimination', 'consonant discrimination'

	GLMER models (family = binominal)
null model	answer ~ 1 + (1|participant)
model 1	answer ~ condition + (1|participant)
model 2	answer ~ condition*group + (1|participant)

As expected, we observed a significant difference between discrimination of *consonant contrasts* and *stress contrast* (χ^2(1, N = 4) = 207.23, $p < .001$; Figure 5.8). Moreover, this effect was observed across intervention groups.

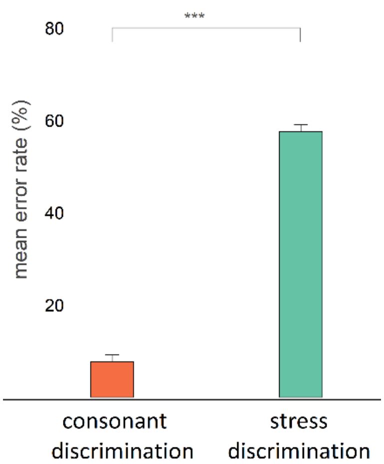


Figure 5.8. Mean error rates for stress and consonant discrimination in pretest, for the intervention groups

5.5.3. Catch trials' results for the intervention groups

Catch trials were included also as a control condition. With the results from these trials, we aimed at observing if Hungarian learners could identify tokens as belonging to the same category. Table 5.26 displays the independent variables and mixed-effect models for the analysis we conducted on the results of the catch trials for vowel discrimination, in the pretest.

Table 5.26. Independent variables and mixed-effect models' formulas for analysis of the catch trials for vowel discrimination in pretest

independent variables		**levels**
participant	66 levels:	'BGE03', 'ELTEBA12', 'CORV04', etc.
group	3 levels:	'Vowels', 'Stress', 'Vowels & Stress'
vowel	6 levels:	'[a]', '[ɐ]', '[ɛ]', '[e]', '[ɨ]', '[i]'
	GLMER models (family = binominal)	
null model	answer ~ 1 + (1|participant)	
model 1	answer ~ vowel + (1|participant)	
model 2	answer ~ vowel*group + (1|participant)	

In vowel discrimination, we found an effect of *vowel* (χ^2(5, N = 7) = 92.56, $p < .001$), but not of *vowel**group*, that is, the three groups exhibited a similar behavior in the perception of the vowels. As Figure 5.9 shows, discrimination of the unknown vowels [ɐ] and [ɨ] posed more problems than other EP vowels, specially [ɨ] ([ɐ]-[a]: $p = .027$; [ɨ] vs. all other contrasts: $p < .001$).

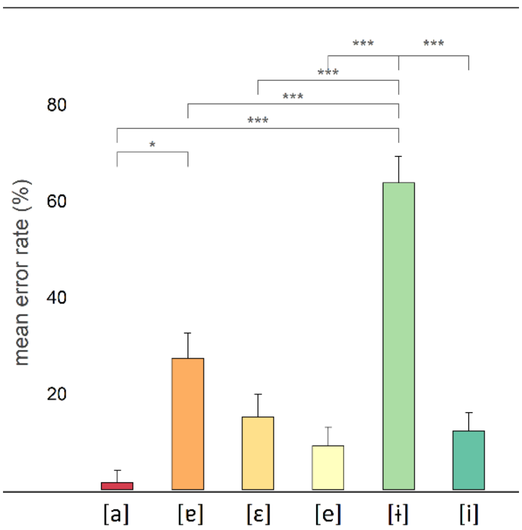


Figure 5.9. Mean error rate for the catch trials for vowel discrimination in pretest, for the intervention groups

This suggests that Hungarian learners were able to identify sequences of tokens as belonging to the same category, although phonetically different, with exception for the two unknown vowels, [ɐ] and [ɨ].

Regarding the results of catch trials for stress discrimination, in the pretest, Table 5.27 displays the independent variables and fixed-effect models used for the analysis of these trials.

Table 5.27. Independent variables and mixed-effect models' formulas for analysis of the catch trials for stress discrimination in pretest

independent variables		levels
participant	66 levels:	'BGE03', 'ELTEBA12', 'CORV04', etc.
group	3 levels:	'Vowels', 'Stress', 'Vowels & Stress'
stress location	6 levels:	'1st syllable', '2nd syllable', '3rd syllable'
	GLMER models (family = binominal)	
null model	answer ~ 1 + (1|participant)	
model 1	answer ~ stress location + (1|participant)	
model 2	answer ~ stress location*group + (1|participant)	

We found an effect of *stress location* (χ^2(2, N = 4) = 84.01, p < .001; Figure 5.10), but not of *stress location*group*, which suggests that the three groups displayed a similar behavior on the discrimination of stress, in these trials.

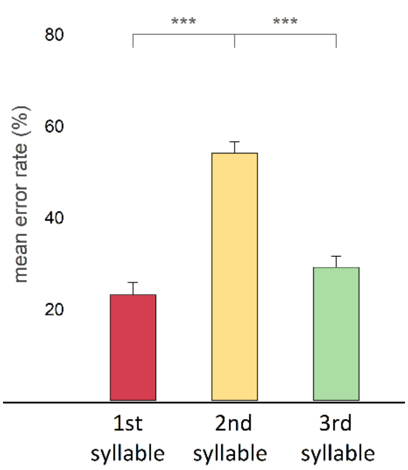


Figure 5.10. Mean error rate for the catch trials for stress discrimination in pretest, for the intervention groups

As we can see in Figure 5.10, discrimination of pseudowords with stress located in the 2nd syllable was more problematic than the other conditions (*1st syllable* vs. *2nd syllable*: $p < .001$; *2nd syllable* vs. *3rd syllable*: $p < .001$). No significant differences were found between discrimination of pseudowords stressed on the 1st syllable and pseudowords stressed on the 3rd syllable.

Having looked at our control conditions – consonant discrimination and catch trials results – we proceed to investigate the results for the change trials, for vowel discrimination and stress discrimination.

5.5.4. Vowel discrimination: tests' results

In this section, we look at the results for each test separately (pretest, post-test^R, and post-test^N), comparing the results between the *Baseline* group and the intervention groups, as well as observing differences between the intervention groups. Recall that in each test we observed discrimination of 7 contrasts, 4 of them in the *perceptual overlap* condition, and 3 of them in the *stressed-unstressed* condition (Figure 5.11).

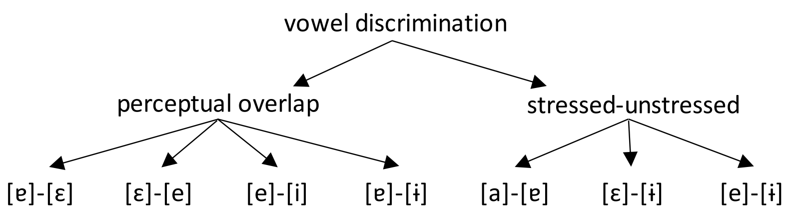


Figure 5.11. Vowel contrasts and respective conditions observed in vowel discrimination

For the analyses, we created three subsets of the data, each including the results for each test (pretest, post-test^R, post-test^N). Table 5.28 displays the independent variables and fixed-effect models run for each data subset.

Table 5.28. Independent variables and mixed-effect models' formulas for analysis of the change trials for vowel discrimination

independent variables		levels
participant	79 levels:	'BGE03', 'ELTEBA12', 'PT014', etc.
group	4 levels:	'Vowels', 'Stress', 'Vowels & Stress', 'Baseline'
contrast type	2 levels:	'perceptual overlap', 'stressed-unstressed'
vowel contrast	7 levels:	'[ɐ]-[ɛ]', '[ɛ]-[e]', '[e]-[i]', '[ɐ]-[ɨ]', '[a]-[ɐ]', '[ɛ]-[ɨ]', '[e]-[ɨ]'
	GLMER models (family = binominal)	
null model	answer ~ 1 + (1|participant)	
model 1	answer ~ group + (1|participant)	
model 2	answer ~ contrast type + (1|participant)	
model 3	answer ~ contrast type*group + (1|participant)	
model 4	answer ~ vowel contrast + (1|participant)	
model 5	answer ~ vowel contrast*group + (1|participant)	

5.5.4.1. Vowel discrimination results in the pretest

In the pretest, each participant completed 42 change trials, six for each contrast ([ɐ]-[ɛ], [ɛ]-[e], [e]-[i], [ɐ]-[ɨ], [a]-[ɐ], [ɛ]-[ɨ], and [e]-[ɨ]). These contrasts were included in one of two conditions: *perceptual overlap* or *stressed-unstressed contrasts*.

Figure 5.12 (next page) displays individual mean error rates (dots) and group mean error rates (dashed lines). The three intervention groups showed a similar individual dispersion (28.2% for Group *Vowels* and Group *Vowels & Stress*; 26.2% for Group *Stress*).[97] Moreover, the dispersion in the three intervention groups ranged within approximately similar lowest and highest values. The *Baseline* group also showed individual dispersion (23.8%), but ranging within lower values.

We found an effect of *group* ($\chi^2(3, N = 5) = 84.204$, $p < .001$), and pairwise comparisons showed that all the intervention groups performed significantly worse than the *Baseline* group, but no significant differences were found between the intervention groups (Group *Vowels* vs. *Baseline* group, $p < .001$; Group *Stress* vs. *Baseline* group, $p < .001$; Group *Vowels & Stress* vs. *Baseline* group, $p < .001$).

[97] These dispersion values were calculated by subtracting the lowest individual mean error rate from the highest individual mean error rate.

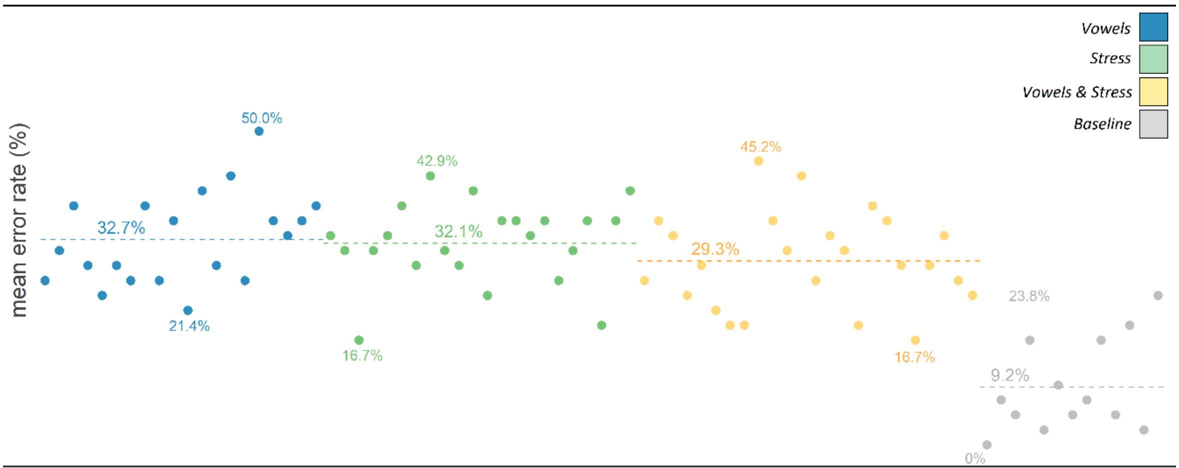


Figure 5.12. Individual and groups' mean error rates for vowel discrimination in pretest

We then investigate differences considering the type of contrast: *perceptual overlap* contrasts and *stressed-unstressed* contrasts (Figure 5.13).

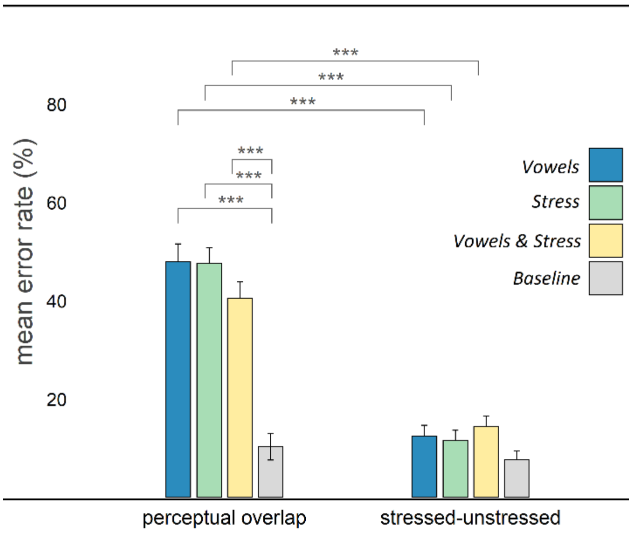


Figure 5.13. Groups' mean error rates for vowel discrimination in pretest, as a function of contrast type

We found a significant effect of the interaction *contrast type*group* (χ^2(6, N = 9) = 104.99, $p < .001$), with pairwise comparisons showing the following results. First, Hungarian students performed significantly worse in the perceptual overlap contrasts than in the stressed-unstressed contrasts (all *p* values of the comparisons between *stressed-unstressed* and *perceptual overlap* scores, in each intervention group, were lower than .001). Second, the three intervention groups displayed significantly higher mean error rates than the *Baseline* group in *perceptual overlap* contrasts ($p < .001$ in each comparison). However, this was not observed in the *stressed-unstressed* contrasts. Third, differences between the intervention groups did not reach the level of significance, not in the case of *perceptual overlap* contrasts, nor for *stressed-unstressed* contrasts.

Finally, we analyzed differences considering *vowel contrast* (Figure 5.14). We found a significant effect of the interaction *vowel contrast*group* ($\chi^2(21, N = 29) = 147.39$, $p < .001$).

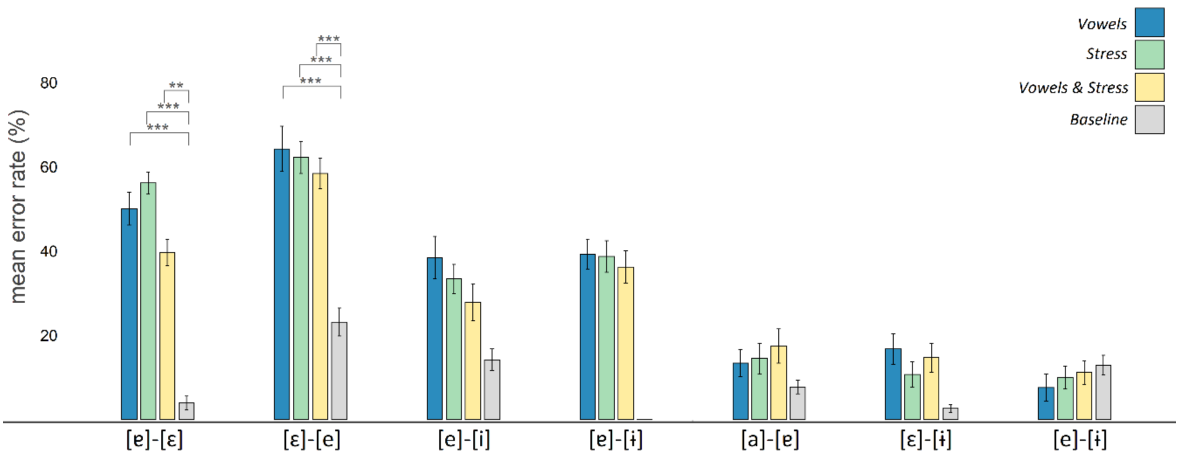


Figure 5.14. Groups' mean error rates for vowel discrimination in pretest, as a function of vowel contrast

Pairwise comparisons showed no significant differences between the three intervention groups. Regarding comparisons with the *Baseline* group, differences in the *perceptual overlap* contrasts were more prominent than for *stressed-unstressed* contrasts, although they reached significance levels only for the contrasts [ɐ]-[ɛ] and [ɛ]-[e] (in the first case, Group *Vowels* vs. *Baseline* group, $p < .001$; Group *Stress* vs. *Baseline* group, $p < .001$; Group *Vowels & Stress* vs. *Baseline* group, $p = .002$; in the second case: Group *Vowels* vs. *Baseline* group, $p < .001$; Group *Stress* vs. *Baseline* group, $p < .001$; Group *Vowels & Stress* vs. *Baseline* group, $p < .001$).

5.5.4.2. Vowel discrimination results in the post-test^R

In the post-test^R, dispersion of individual mean error rates was higher than in the pretest (35.7%, 31.0%, and 42.9%, for Group *Vowels*, Group *Stress*, and Group *Vowels & Stress*, respectively), although the lowest values were lower than in the pretest (Figure 5.15[98]).

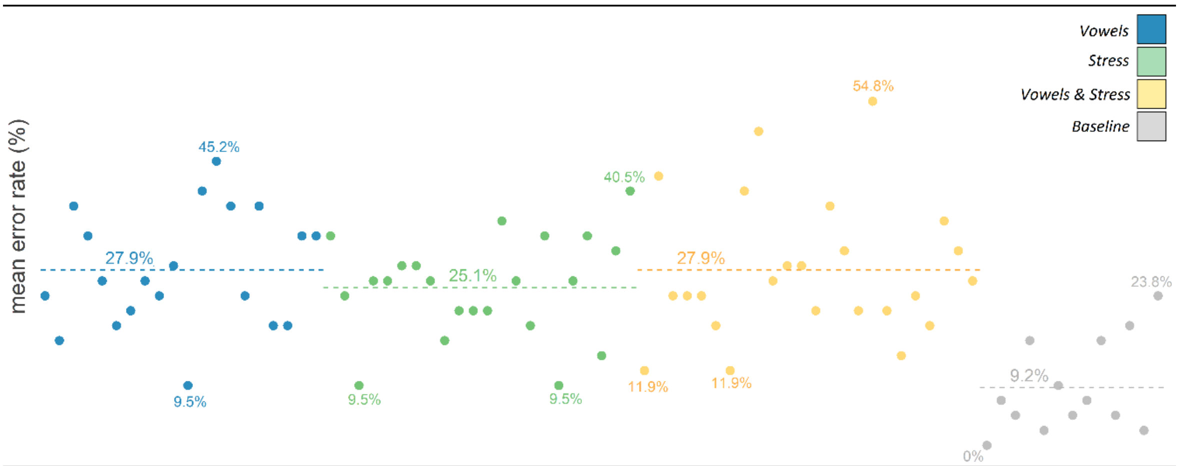


Figure 5.15. Individual and groups' mean error rates for vowel discrimination in post-test^R

Group was a significant factor ($\chi^2(3, N = 5) = 45.175$, $p < .001$). Pairwise comparisons showed that error rates for the three intervention groups performance were significantly higher compared to the *Baseline* group (in the three cases, $p < .001$), but no significant differences between the intervention groups were found.

The interaction *contrast type*group* was also significant ($\chi^2(6, N = 9) = 71.266$, $p < .001$; Figure 5.16, next page). Pairwise comparisons showed a similar scenario as in the pretest. In the *perceptual overlap* contrasts, the *Baseline* group outperformed significantly each of the three intervention groups ($p < .001$ in each case), and the intervention groups displayed mean error rates significantly higher in the *perceptual overlap* condition than in the *stressed-unstressed* condition ($p < .001$ for each intervention group).

[98] Results here displayed for the *Baseline* Group are those for the pretest, since the post-test^N consisted of the repetition of the pretest trials.

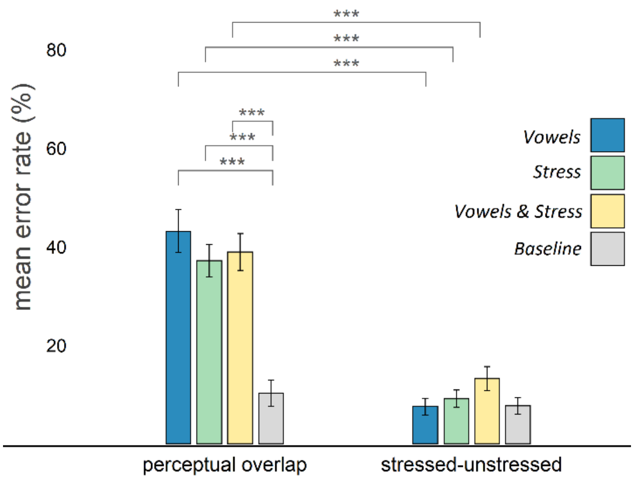


Figure 5.16. Groups' mean error rates for vowel discrimination in post-test[R], as a function of contrast type

Similar as in the pretest, we found a significant effect of the interaction *vowel contrast*group* (χ^2(21, N = 29) = 118.56, *p* < .001; Figure 5.17).

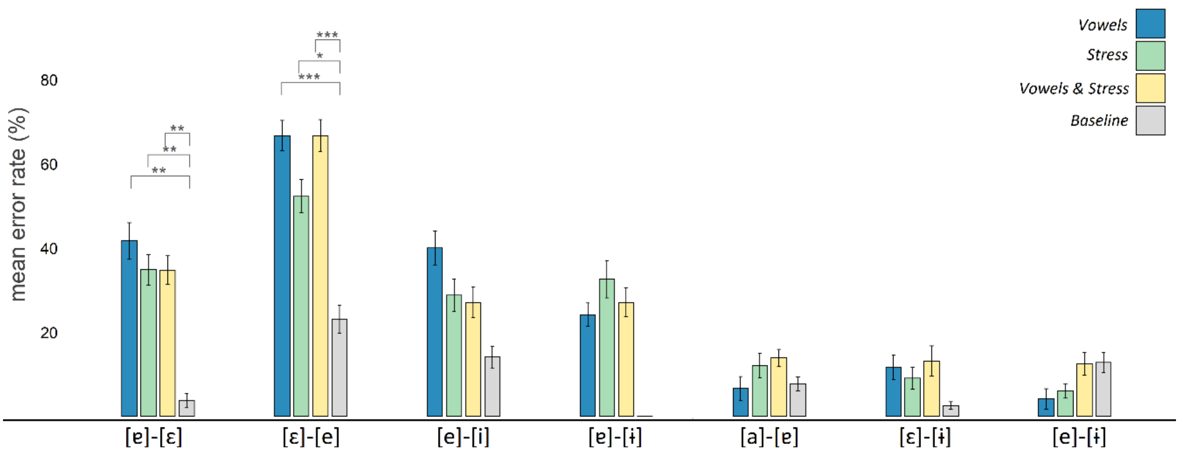


Figure 5.17. Groups' mean error rates for vowel discrimination in post-test[R], as a function of vowel contrast

Although, differences were less prominent in the post-test[R] than in the pretest, the three intervention groups still exhibit significantly higher mean error rates than the *Baseline* group in the contrasts [ɐ]-[ɛ] (Group *Vowels* vs. *Baseline* group, *p* = .001; Group *Stress* vs. *Baseline* group, *p* = .008; Group *Vowels & Stress* vs. *Baseline* group, *p* = .008) and [ɛ]-[e] (Group *Vowels* vs. *Baseline* group, *p* < .001; Group *Stress* vs. *Baseline* group, *p* = .05;

Group *Vowels & Stress* vs. *Baseline* group, $p < .001$). No significant differences were found between the three intervention groups' results.

5.5.4.3. Vowel discrimination results in the post-test^{N}

The novel trials for vowel discrimination included the same contrasts tested in the previous tests, but in a novel context [sV], produced by novel talkers. In this test, Hungarian participants mean error means ranged within higher lowest and highest values, although dispersion values were similar as the previous tests (31,0% for Group *Vowels*, 33.3% for Group *Stress*, and 35.7% for Group *Vowels & Stress*).

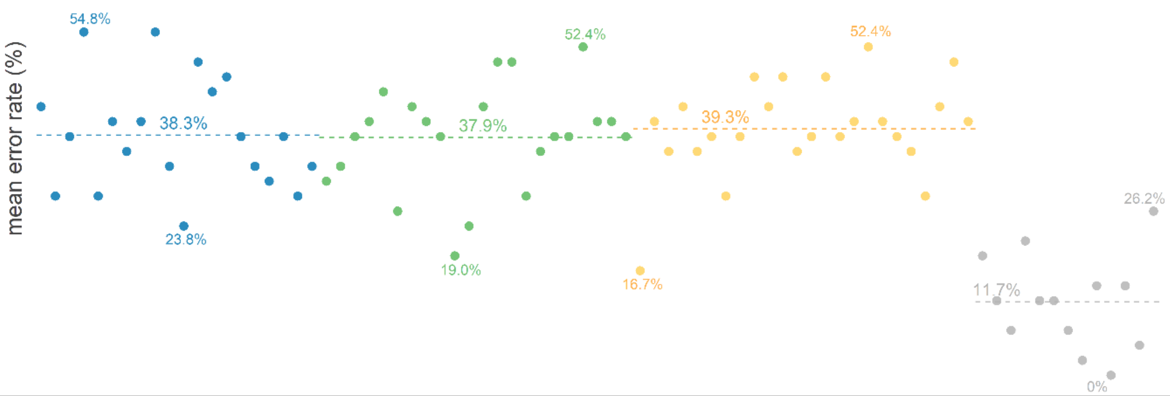


Figure 5.18. Individual and groups' mean error rates for vowel discrimination in post-test^{N}

As we can see in Figure 5.18, the gap between native speakers and students increased for all intervention groups. The mixed-effect model showed that *group* was a significant factor in the results (χ^2(3, N = 5) = 88.314, $p < .001$). Pairwise comparisons showed significant differences between each intervention group and the *Baseline* group ($p < .001$ in the three cases). No significant differences between intervention groups were found.

The interaction *contrast type*group* (χ^2(6, N = 9) = 90.30, $p < .001$; Figure 5.19, next page) was significant. However, we observed some differences between the novel trials' results and the pretest and the repeated trials' results. Looking at the intervention groups, the gap between discrimination of *perceptual overlap* contrasts and *stressed-unstressed* contrasts decreased, due to the higher mean error rates in the three intervention groups, in trials for *stressed-unstressed* condition. In the case of Group *Vowels & Stress*, differences are no longer significant (for Group *Vowels* and Group *Stress*, *p* values in *perceptual overlap* vs. *stressed-unstressed* contrasts were < .001). The higher mean error rate in the *stressed-unstressed* conditions compared to the pretest and the post-test[R] entailed a bigger difference between intervention groups and *Baseline* group in this condition. The pairwise comparisons showed that in these contrasts, differences between each intervention group and the *Baseline* group were significant (*p* < .001 in each case).

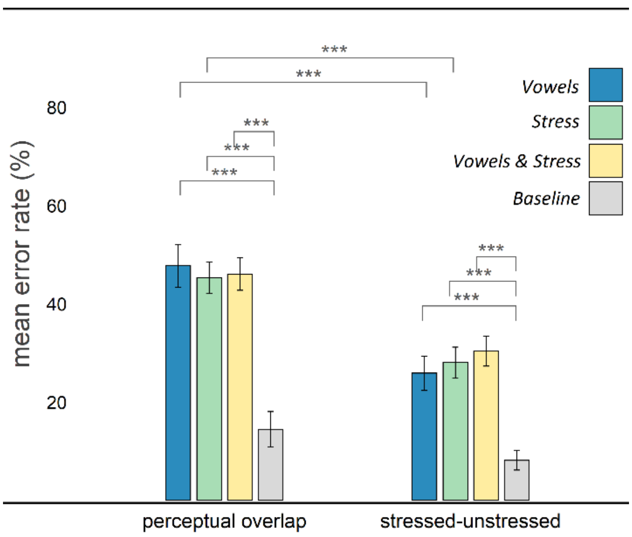


Figure 5.19. Groups' mean error rates for vowel discrimination in post-test[N], as a function of contrast type

*Vowel contrast*group* interaction was significant (χ^2(21, N = 29) = 153.05, *p* < .001). Pairwise comparisons showed that differences between each intervention groups and *Baseline* group in contrasts [ɛ]-[ɛ] and [ɛ]-[e] were not significant with exception of Group *Vowels & Stress,* for the latter (*p* = .001). Additionally, in the contrasts [e]-[i] and [e]-[ɨ] the gap between the intervention groups and the *Baseline* group increased, and mean error rate for each intervention group was significantly higher than for the *Baseline* group (in the first contrast, *p* values between the *Baseline* group and each intervention group were < .001; in the second contrast: Group *Vowels* vs. *Baseline* group, *p* < .001; Group *Stress* vs. *Baseline* group, *p* = .014; Group *Vowels & Stress* vs. *Baseline* group, *p* = .001). Moreover, we found a significant difference between Group *Vowels* and Group *Stress* in the contrast [a]-[ɐ] (p = .044). Figure 5.20 displays these results.

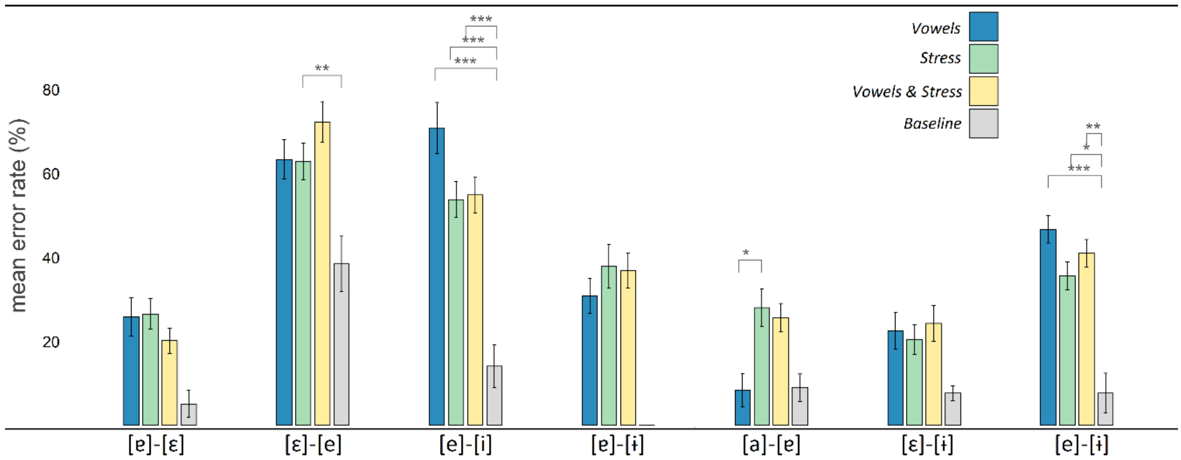


Figure 5.20. Groups' mean error rates for vowel discrimination in post-test^N, as a function of vowel contrast

5.5.5. Vowel discrimination: progress from the pretest to the post-test

Having looked at results in each test and compared performances between intervention groups, and between these and the *Baseline* group, we then analyzed the progress from the pretest to the post-test phase, thus considering only the three intervention groups. We examine progress from the pretest to the post-test^R and progress from the pretest to the post-test^N separately. To this purpose, we created two subsets of the data collected for vowel discrimination, one with the results for the pretest and the post-test^R and another with the results for the pretest and the post-test^N. For the statistical analysis, we used the independent variables and models displayed in Table 5.29 (next page).

Table 5.29. Independent variables and mixed-effect models' formulas for analysis of progress in vowel discrimination, from pretest to post-test

independent variables		**levels**
participant	66 levels:	'BGE03', 'ELTEBA12', 'CORV01', etc.
group	3 levels:	'Vowels', 'Stress', 'Vowels & Stress'
contrast type	2 levels:	'perceptual overlap', 'stressed-unstressed'
vowel contrast	7 levels:	'[ɐ]-[ɛ]', '[ɛ]-[e]', '[e]-[i]', '[ɐ]-[ɨ]', '[a]-[ɐ]', '[ɛ]-[ɨ]', '[e]-[ɨ]'
test	2 levels:	'pretest', 'post-test repeated trials'
		or
		'pretest', 'post-test novel trials'

	GLMER models (family = binominal)
null model	answer ~ 1 + (1|participant)
model 1	answer ~ test + (1|participant)
model 2	answer ~ test*group + (1|participant)
model 3	answer ~ contrast type*test + (1|participant)
model 4	answer ~ contrast type*test*group + (1|participant)
model 5	answer ~ vowel contrast*test + (1|participant)
model 6	answer ~ vowel contrast*test*group + (1|participant)

5.5.5.1. Vowel discrimination: progress from the pretest to the post-test^R

Although in general there was a significant improvement[99] from the pretest to the repeated trials (that is, a significant effect of *test*: $\chi^2(1, N = 3) = 12.78$, $p < .001$), no significant interaction *test**group* was found (Figure 5.21, next page, on the left side of the dashed line). We found a significant *test*contrast type*group* effect ($\chi^2(8, N = 13) = 19.78$, $p = .01$) (Figure 5.21, on the right side of the dashed line). However, pairwise comparisons showed that only Group *Stress* improvement in *perceptual overlap* contrasts reached a significant level ($p = .03$).

[99] We use 'improvement' to refer to decrease in mean error rates from the pretest to the post-test^R or the post-test^N. Conversely, we use 'decline' when we observe an increase in mean error rates from the pretest to the post-test^R or the post-test^N.

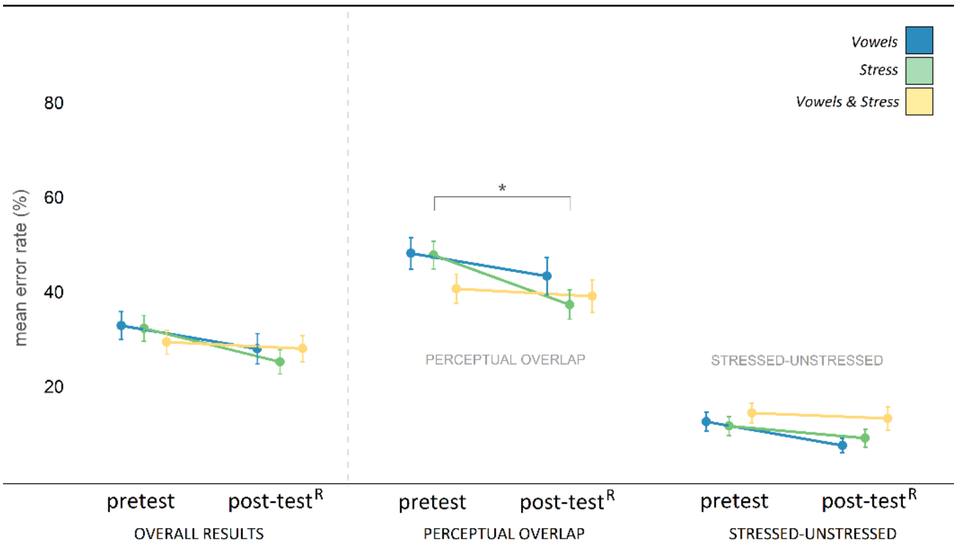


Figure 5.21. Progress in vowel discrimination from pretest to post-test^R, by intervention group: overall results, and results as a function of contrast type

We decided to investigate if, at vowel contrast level, the scenario would be consistent with the one presented above. Although we found a significant interaction effect of *test*vowel contrast*group* (χ^2(28, N = 43) = 45.29, *p* = .02), pairwise comparisons showed no significant differences between groups' results for each vowel contrast. Nevertheless, when visualizing the results in graphs, we found different scenarios worth mentioning.

In the case of *stressed-unstressed* contrasts (Figure 5.22, next page) the picture for each contrast was consistent with that seen in Figure 5.21, that is, low mean error rates and mild differences between the pretest and the post-test^R.

As for the *perceptual overlap* contrasts, however, as Figure 5.23 (next page) shows, two scenarios were observed: in the contrasts [ɐ]-[ɛ] and [ɐ]-[ɨ] the three groups displayed an improvement, but in the remaining two contrasts no similar behavior is observed between groups. In the contrast [ɛ]-[e], Group *Vowels* presented a very slight decline in the results, Group *Stress*, improvement, and Group *Vowels & Stress* also a decline, more pronounced than for Group *Vowels*. As for [e]-[i], Groups *Vowels* and *Vowels & Stress* displayed a very slight decline, and Group *Stress*, a slight improvement.

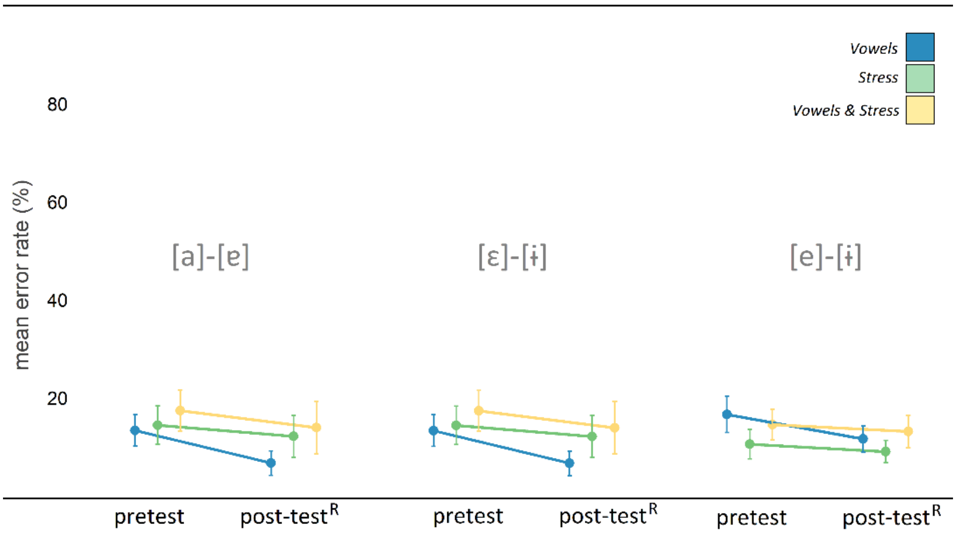


Figure 5.22. Progress from pretest to post-test^R for stressed-unstressed contrasts, by intervention group

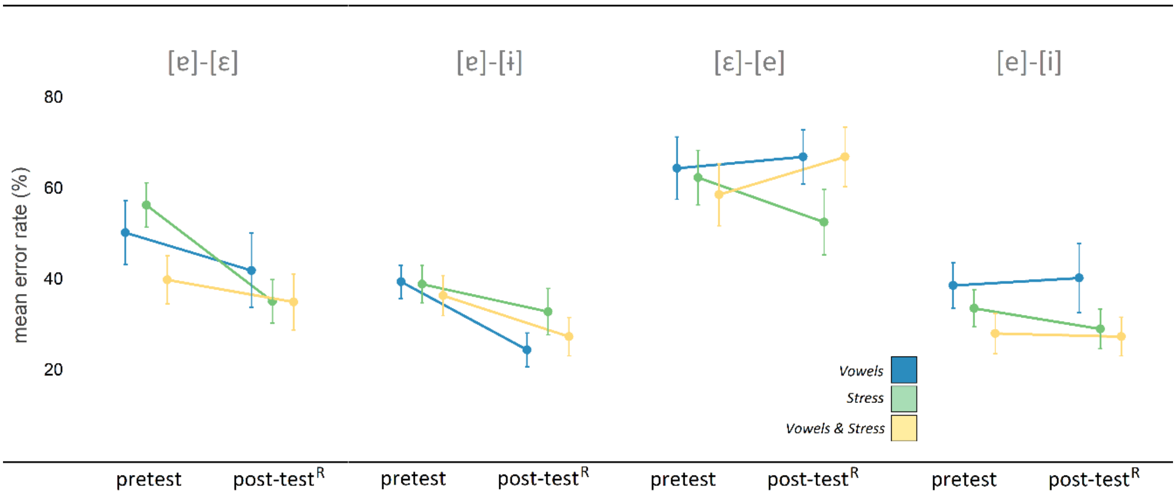


Figure 5.23. Progress from pretest to post-test^R for [ɐ]-[ɛ], [ɐ]-[ɨ], [ɛ]-[e], and [e]-[i], by intervention group

5.5.5.2. Vowel discrimination: progress from the pretest to the post-test^N

Contrary to the progress from the pretest to the repeated trials in the post-test, in the novel trials all groups presented a general decline in the results (Figure 5.24, next page, on the left side of the dashed line). Although a significant effect of *test* was found (χ^2(1, N = 3) = 32.20, p < .001), no significant *test**group* interaction effect was found, that is, the three group displayed the same behavior.

Contrast type had a significant effect on the results ($\chi^2(2, N = 5) = 447.9$, $p < .001$), but the interaction effect of *contrast type*group* was not significant. As we can see in Figure 5.24, the decline in the results in the novel trials was due to problems in the *stressed-unstressed* contrasts, that were not observed in the post-test^R, nor in the pretest. This difficulty was present in the three intervention groups (all relevant pairwise comparisons showed $p < .001$). As for *perceptual overlap* contrasts, no progress or only mild progress was observed.

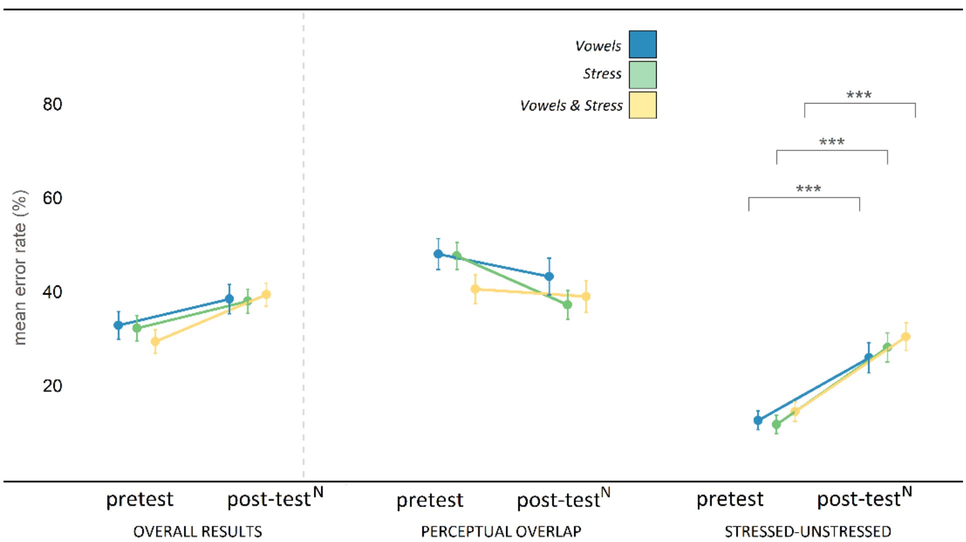


Figure 5.24. Progress in vowel discrimination from pretest to post-test^N, by intervention group: overall results, and results as a function of contrast type

We found a significant interaction *test*vowel contrast*group* ($\chi^2(28, N = 43) = 56.56$, $p = .001$). Looking at the pairwise comparisons and the graphs for each vowel contrast, we observed a chaotic picture, with three different scenarios. First, in the contrast [ɐ]-[ɛ] (Figure 5.25, next page), all groups displayed an expressive improvement. However, improvement reaches significance only for Group *Stress* ($p = .001$).

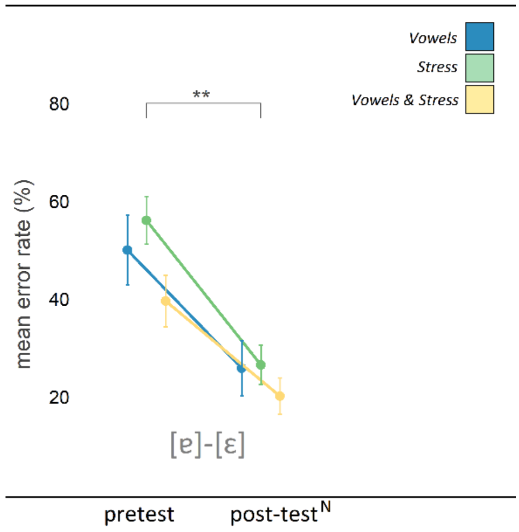


Figure 5.25. Progress from pretest to post-test[N] for [ɐ]-[ɛ], by intervention group

Second, in the contrasts [e]-[i], [e]-[ɨ], and [ɛ]-[ɨ] (Figure 5.26), mean error rate increased in the three groups.

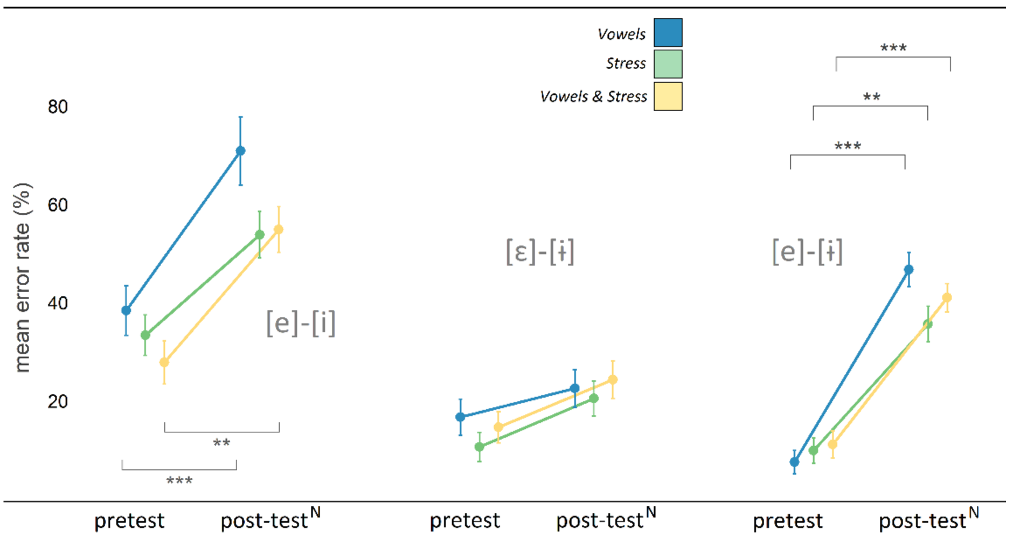


Figure 5.26. Progress from pretest to post-test[N] for [e]-[i], [e]-[ɨ], and [ɛ]-[ɨ], by intervention group

Pairwise comparisons showed the following results when comparing the pretest and the post-test[N] results. For [e]-[i], for Group *Vowels*, $p < .001$, and for Group *Vowels & Stress*, $p = .003$. In the case of the contrast for [e]-[ɨ], for Groups *Vowels* and *Vowels &*

Stress, $p < .001$, and for Group *Stress*, $p = .002$. In the contrast [ɛ]-[ɨ], the three groups exhibited a similar decline in the results.

Third, in the remaining contrasts the three groups did not evolve in the same way (Figure 5.27). In the contrast [a]-[ɐ], Groups *Stress* and *Vowels & Stress* showed a decline in the results and Group *Vowels* improved. In the contrast [ɐ]-[ɨ], we observe that Groups *Stress* and *Vowels & Stress* performed very similar as in the pretest, while Group *Stress* displayed a slight improvement in the results. Finally, in the case of the contrast [ɛ]-[e], Group *Vowels* and Group *Stress* maintained the same level of performance as in the pretest, whereas Group *Vowels & Stress* showed a decline in the results.

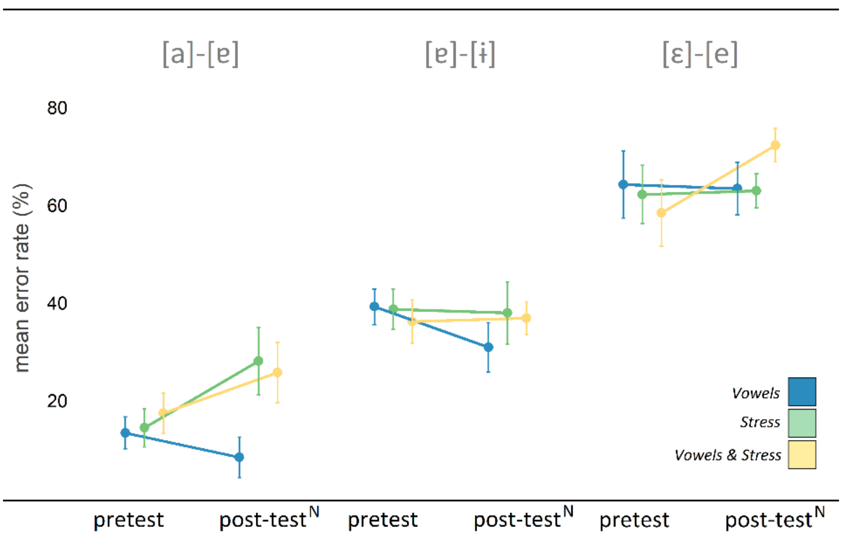


Figure 5.27. Progress from pretest to post-test^N for [a]-[ɐ], [ɐ]-[ɨ], and [ɛ]-[ɨ], by intervention group

5.5.6. Stress discrimination: tests' results

In this section, we look at the results of stress discrimination tests. Similar to the analysis of vowel discrimination, we started by analyzing the results for each test separately (pretest, post-test^R, and post-test^N). For each test, we compared the results between the *Baseline* group and the intervention groups, and between intervention groups.

Three subsets of the main database were created, one for each test. The analysis conducted included the independent variables and mixed-effect models displayed in Table 5.30 (next page).

Table 5.30. Independent variables and mixed-effect models' formulas for analysis of the change trials for stress discrimination

independent variables		levels
participant	79 levels:	'BGE03', 'ELTEBA12', 'PT014', etc.
group	4 levels:	'Vowels', 'Stress', 'Vowels & Stress', 'Baseline'
contrast location	2 levels:	'1st-2nd syllable', '2nd-3rd syllable'
	GLMER models (family = binominal)	
null model	answer ~ 1 + (1|participant)	
model 1	answer ~ group + (1|participant)	
model 2	answer ~ contrast location + (1|participant)	
model 3	answer ~ contrast location*group + (1|participant)	

5.5.6.1. Stress discrimination results in the pretest

In the pretest, each participant completed 60 change trials, and two conditions were tested: *1st-2nd syllable contrasts* (30 trials) and *2nd-3rd syllable contrasts* (30 trials). With respect to individual performances dispersion (Figure 5.28), all the groups displayed a similar high dispersion regarding individual mean error rates, slightly lower for Group *Vowels & Stress* (71.6% for Group *Vowels*, 70.0% for Group *Stress*, and 71.7% for *Baseline* group; Group *Vowels & Stress* dispersion was 60%). However, Hungarian participants dispersion ranged within higher lowest and highest values than the *Baseline* participants.

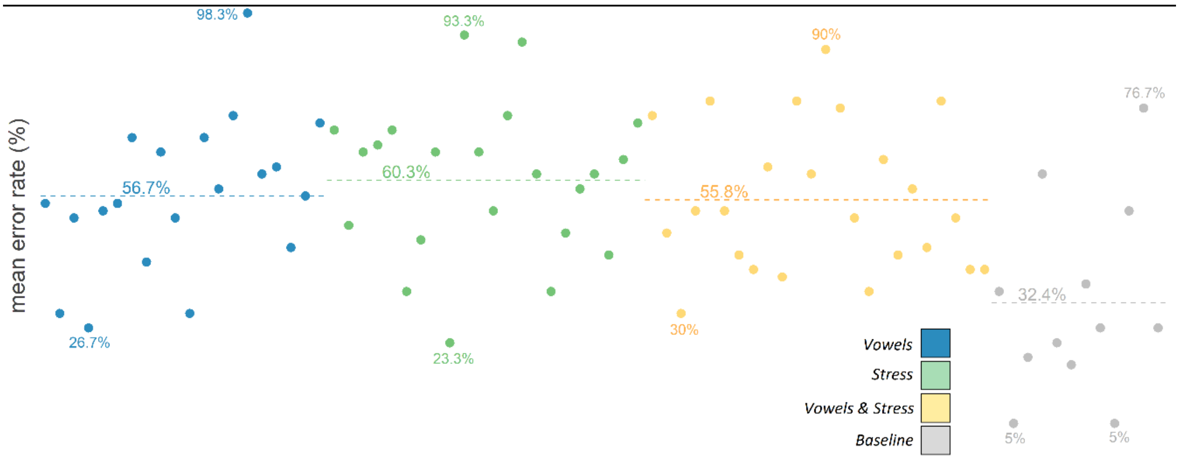


Figure 5.28. Individual and groups' mean error rate for stress discrimination in pretest

We found a significant effect of *group* ($\chi^2(3, N = 5) = 19.68$, $p < .001$) and pairwise comparisons showed that each intervention group performed significantly different than the *Baseline* group (in the three cases, $p < .001$), but they did not differ from each other.

Regarding the location of the stress contrast (Figure 5.29), the interaction *contrast location*group* was significant ($\chi^2(6, N = 9) = 22.39$, $p = .001$). Pairwise comparisons showed that results in *1st-2nd syllable* and in *2nd-3rd syllable* contrasts for each intervention group were significantly different (worse) than results for the *Baseline* group, but not between each other (*1st-2nd syllable*: Group *Vowels* vs. *Baseline* group, $p = .001$; Group *Stress* vs. *Baseline* group, $p = .002$; Group *Vowels & Stress* vs. *Baseline* group, $p = .03$; *2nd-3rd syllable*: Group *Vowels* vs. *Baseline* group, $p < .001$; Group *Stress* vs. *Baseline* group, $p = .001$; Group *Vowels & Stress* vs. *Baseline* group, $p = .004$). Furthermore, in all groups we found significant differences between *1st-2nd syllable* and *2nd-3rd syllable* contrasts, even in the *Baseline* group (*1st-2nd syllable* vs. *2nd-3rd syllable*, for Group *Vowels* and for Group *Stress*, $p < .001$; for Group *Vowel & Stress*, $p = .005$; for *Baseline* group, $p = .001$).

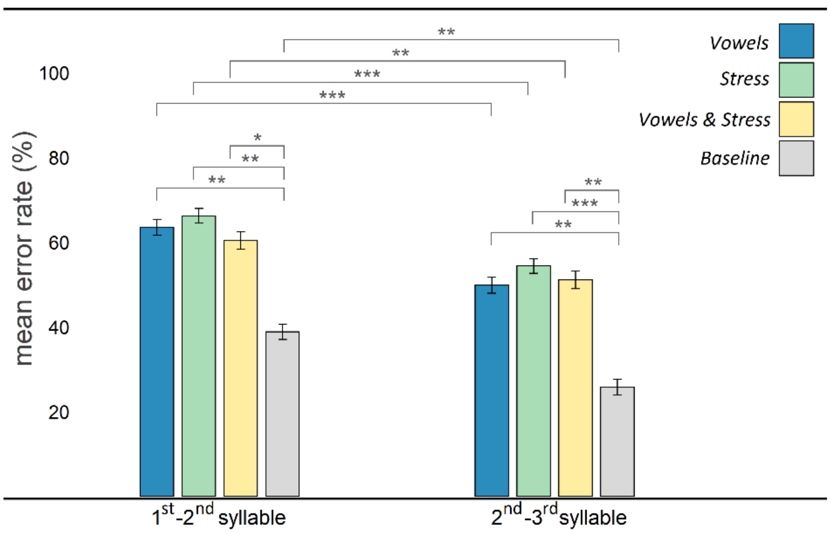


Figure 5.29. Groups' mean error rates for stress discrimination in pretest, as a function of stress contrast location

5.5.6.2. Stress discrimination results in the post-test[R]

Repetition of the pretest trials for stress discrimination included only 48 of the 60 trials from the pretest, 24 for *1st-2nd syllable* contrasts, and 24 for *2nd-3rd syllable* contrasts.

Similar to the pretest, in the post-test[R] we also found high individual dispersion (79.2% for Group *Vowels*, 68.8% for Group *Stress*, 68.8% for Group *Vowels & Stress*, and 75% for the *Baseline* group; Figure 5.30).

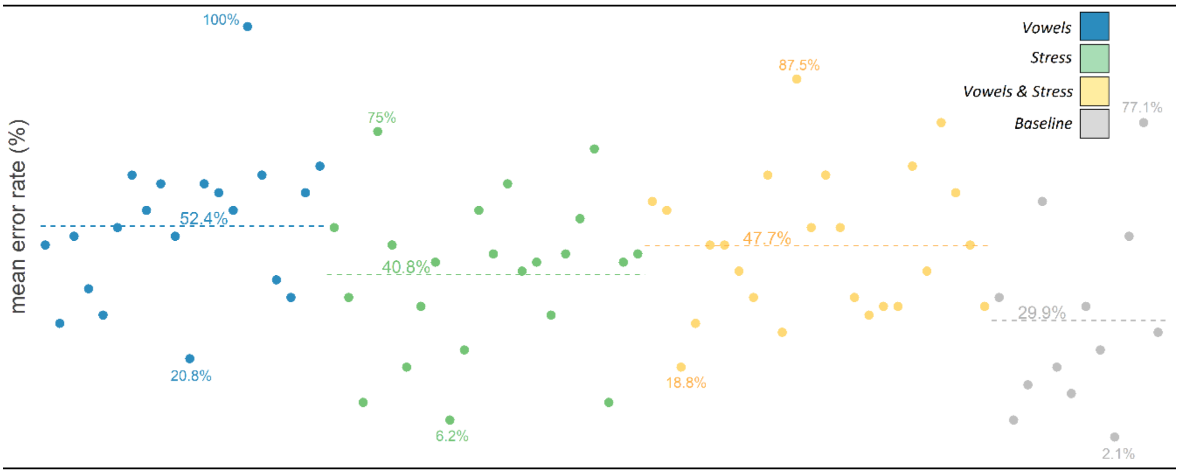


Figure 5.30. Individual and groups' mean error rates for stress discrimination in post-test[R]

Although *group* significantly affected the results, (χ^2(3, N = 5) = 13.67, *p* = .003), pairwise comparisons showed that differences between mean error rates for the *Baseline* group and Group *Stress* were not significant. As for Groups *Vowels* and *Vowels & Stress*, although these groups' mean error rate decreased compared to the pretest, they still differed significantly from the *Baseline* group (*p* = .002 and *p* = .02, respectively).

The interaction *contrast location**group* was significant (χ^2(6, N = 9) = 24.10, *p* < .001). However, we observed some changes, comparing to pretest results. First, as we can see in Figure 5.31 (next page), only in *2nd-3rd syllable* contrasts significant differences emerged between intervention groups and the *Baseline* group, and only for two groups: *Vowels* and *Vowels & Stress* (*p* = .003 and *p* = .03, respectively). Second, when comparing discrimination of *1st-2nd syllable* with *2nd-3rd syllable* contrasts, only for Group *Stress*, Group *Vowels & Stress*, and *Baseline* group differences reached the level of significance (*p* < .001, *p* = .014, *p* = .007, respectively). In Group *Vowels*, the difference in results between *1st-2nd syllable* and *2nd-3rd syllable* contrasts was not significant.

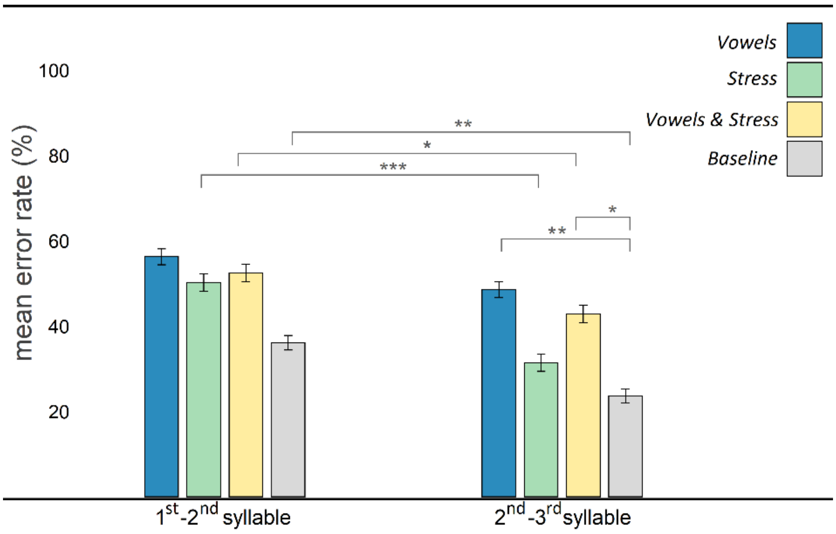


Figure 5.31. Groups' mean error rates for stress discrimination in post-test^R, as a function of stress contrast location

5.5.6.3. Stress discrimination results in the post-test^N

Finally, we investigated the results for the novel trials in the post-test. When looking at individual dispersion and group means (Figure 5.32), we observe that this was high in the four groups (70.8% for Group *Vowels*, 62.4% for Group *Stress*, 64.6% for Group *Vowels & Stress*, and 79.2% for the *Baseline* group), although *Baseline* group ranged at lower values.

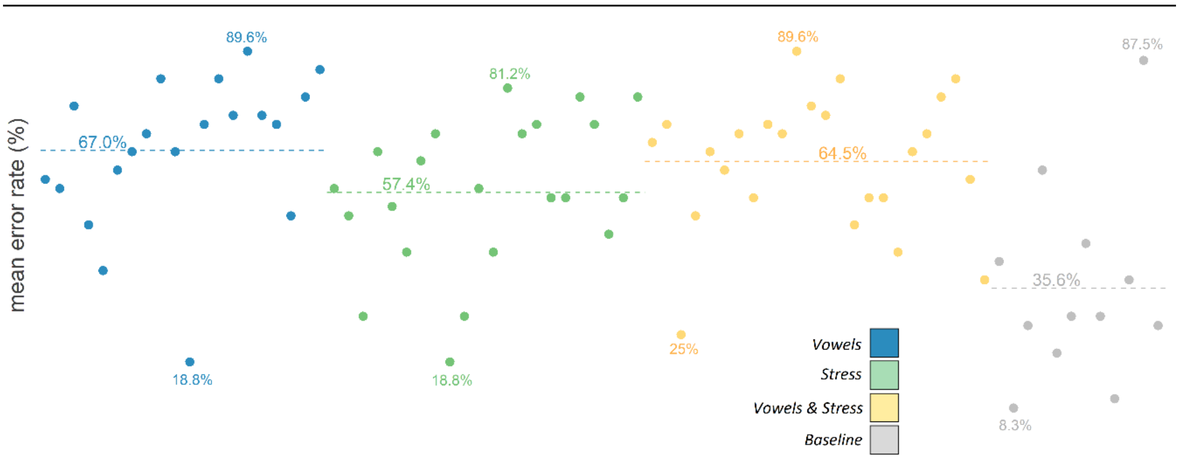


Figure 5.32. Individual and groups' mean error rates for stress discrimination in post-test^N

Pairwise comparisons between groups showed that each intervention group were significantly outperformed by the *Baseline* group (Group *Vowels* vs. *Baseline* group, $p < .001$; Group *Stress* vs. *Baseline* group, $p = .003$; Group *Vowels & Stress* vs. *Baseline* group, $p < .001$), but no significant differences were found between the intervention groups.

Similar to the previous tests, we found a significant effect of *contrast location*group* on the novel trials of the post-test ($\chi^2(6, N = 9) = 25.54$, $p < .001$). Pairwise comparisons showed similarities to pretest results, since all intervention groups performed significantly different from the *Baseline* group, and in both *contrast locations* (*1st-2nd syllable*: Group *Vowels* vs. *Baseline* group, $p < .001$; Group *Stress* vs. *Baseline* group, $p = .05$; Group *Vowels & Stress* vs. *Baseline* group, $p = .002$; *2nd-3rd syllable*: Group *Vowels* vs. *Baseline* group, $p < .001$; Group *Stress* vs. *Baseline* group, $p = .02$; Group *Vowels & Stress* vs. *Baseline* group, $p < .001$). However, only for Group *Stress* and the *Baseline* group differences between *1st-2nd syllable* and *2nd-3rd syllable* contrasts reached the level of significance ($p = .006$ and $p = .04$, respectively). Figure 5.33 displays these results.

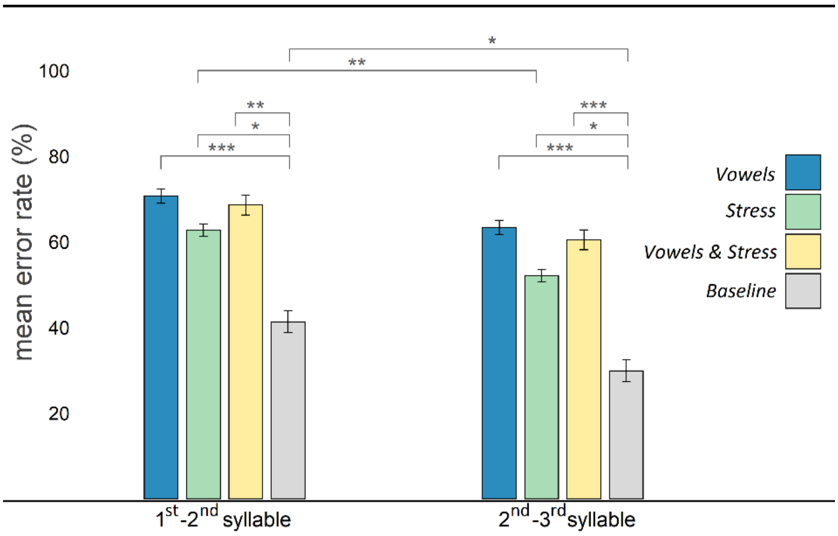


Figure 5.33. Groups' mean error rates for stress discrimination in post-test^N, as a function of stress contrast location

5.5.7. Stress discrimination: progress from the pretest to the post-test

Once we had looked at results for each testing phase for stress discrimination, we then investigated progress between these phases, specifically, between the pretest and the post-test^R and the pretest and the post-test^N. Similar to the analysis of progress for vowel discrimination, we excluded comparisons with the *Baseline* group. The independent variables and mixed-effect models' formulas used for these analyses are displayed in Table 5.31 below.

Table 5.31. Independent variables and mixed-effect models' formulas for analysis of progress in stress discrimination, from pretest to post-test

independent variables		**levels**
participant	66 levels:	'BGE03', 'ELTEBA12', 'CORV01', etc.
group	3 levels:	'Vowels', 'Stress', 'Vowels & Stress'
contrast location	2 levels:	'1st-2nd syllable', '2nd-3rd syllable'
test	2 levels:	'pretest', 'post-test repeated trials' or 'pretest', 'post-test novel trials'
	GLMER models (family = binominal)	
null model	answer ~ 1 + (1|participant)	
model 1	answer ~ test + (1|participant)	
model 2	answer ~ test*group + (1|participant)	
model 3	answer ~ contrast location*test + (1|participant)	
model 4	answer ~ contrast location*test*group + (1|participant)	

5.5.7.1. Stress discrimination: progress from the pretest to the post-test^R

We found a significant interaction of *test***group* (χ^2(4, N = 7) = 32.07, *p* < .001). Pairwise comparisons showed that improvement reach significance level only in the case of Groups *Stress* and *Vowels & Stress* (*p* < .001 in both cases; Figure 5.34, next page, on the left side of the dashed line). Regarding the effect of stress contrast location (Figure 5.34, on the right), the interaction *contrast location***group* was significant (χ^2(8, N = 13) = 42.80, *p* < .001). Pairwise comparisons showed that only Group *Stress* presented significant improvement, and in both contrasts (*p* < .001 in both cases).

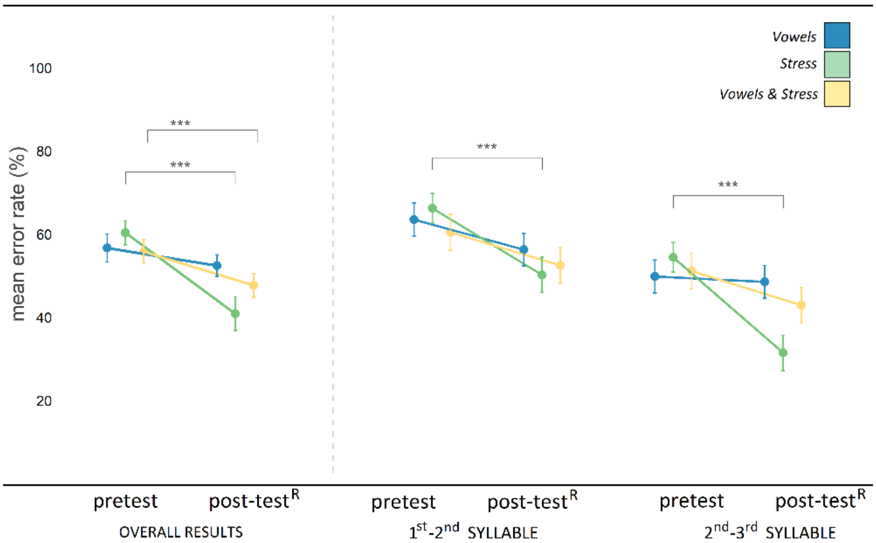


Figure 5.34. Progress in stress discrimination from pretest to post-test^R, by intervention group: overall results, and results as a function of stress contrast location

5.5.7.2. Stress discrimination: progress from the pretest to the post-test^N

Contrary to progress to the post-test^R, with the exception of Group *Stress*, in the novel trials, we observed a decline in the results (Figure 5.35, next page, left side of the dashed line). *Test*group* (χ^2(4, N = 7) = 28.41, $p < .001$) was significant. The pairwise comparisons showed that differences between the pretest and the post-test^N results were significant for Groups *Vowels* and *Vowels & Stress* ($p < .001$ in both cases). As for Group *Stress*, although a mild improvement was found, it did not reach significance level. Regarding *contrast location* effect, the interaction *contrast location*group* was significant (χ^2(8, N = 13) = 30.69, $p < .001$). The pairwise comparisons showed that the increase of error rate was significant for *2nd-3rd syllable* contrasts, for Groups *Vowels* and *Vowels & Stress* ($p < .001$ and $p = .025$, respectively). Figure 5.35 (next page) displays these results.

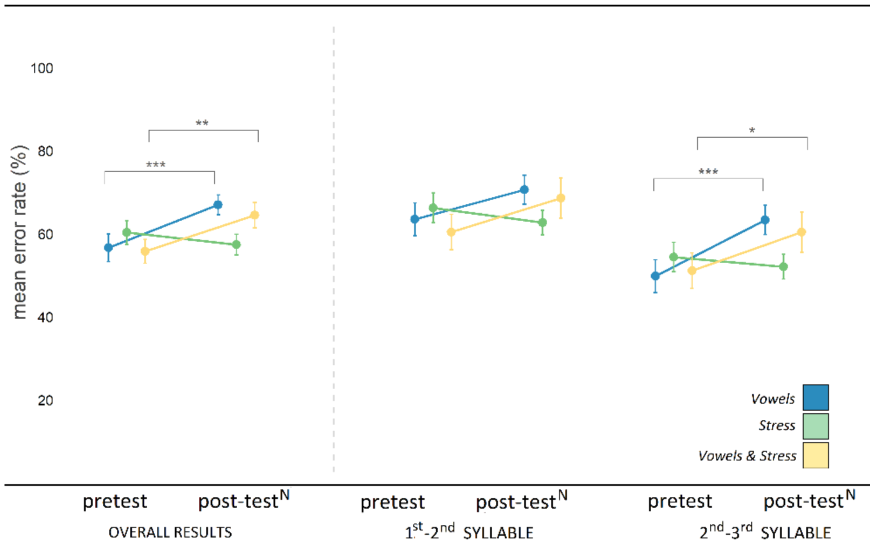


Figure 5.35. Progress in stress discrimination from pretest to post-test^N, by intervention group: overall results, and results as a function of stress contrast location

5.5.8. Training sessions' results

In order to investigate the results of the training sessions, we were compelled to use a different method from the tests, since in training trials there was no limit to tryouts. Here, we calculated the percentage of correct answers in the first attempt, considering the total number of trials. For example, in session 1 of Group *Vowels*' training, in 960 trials (48 trials × 20 participants) we calculated that in 93.6% of the trials, students answered correctly at the first attempt, and they only needed to repeat the trial in 6.4% of the times. Although in some cases participants needed to repeat more than once the same trial, this was a rare situation; therefore, we simplified the results to two outcomes: percentages of correct answers at the first attempt vs. percentages of correct answers with repeated trials. In the graphs for training progress, we represent in axis Y the percentage of correct answers at the first attempt.

Recall that in each training regimen, each contrast or condition was first trained in AX tasks (in sessions 1, 2, 3, and 4), and reviewed in two AXB tasks (in sessions 5 and 6). In the AXB of session 5, contrasts were trained in separated blocks, and in the AXB task of session 6, trials were randomized across contrasts. We analyzed differences in results between each of the 6 sessions, for each trained contrast. To this purpose, we created subsets of the data collected for each contrast. For example, we created a subset data for the contrast [ɐ]-[ɛ], and then analyzed differences between each session results for this contrast. Analysis for each feature was run with the formula *answer ~ session + (1|participant).* Fillers were excluded from the analyses.

5.5.8.1. Training sessions' results for Group *Vowels*

Group *Vowels* training was focused on seven vowel contrasts, from which four were *perceptual overlap* contrasts ([ɐ]-[ɛ], [ɐ]-[ɨ], [ɛ]-[e], and [e]-[i]) and three, *stressed-unstressed* contrasts ([a]-[ɐ], [ɛ]-[ɨ], and [e]-[ɨ]). Each vowel contrast was trained three times, once in an AX task and twice in AXB tasks, in the 5th and 6th sessions (see Table 5.9). We analyzed each vowel contrast, and found significant effects of training only in the contrasts [ɐ]-[ɨ] (χ^2(2, N = 4) = 7.90, *p* = .019) and [e]-[ɨ] (χ^2(2, N = 4) = 8.49, *p* = .014). Figure 5.36 and Figure 5.37 (next page) display the results for the *perceptual overlap* contrasts and the *stressed-unstressed* contrasts, respectively.

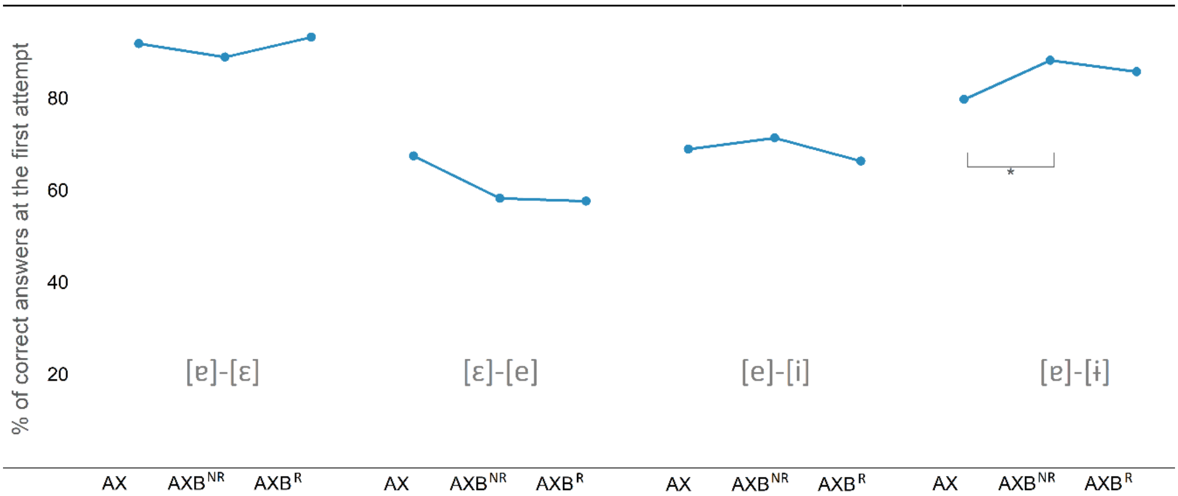


Figure 5.36. Progress observed between training sessions in *perceptual overlap* contrasts, for Group *Vowels*. NR = not randomized trials R = randomized trials

Pairwise comparisons showed that learners exhibited a significant improvement for [ɐ]-[ɨ], from the AX session to the first AXB session (*p* = .04).

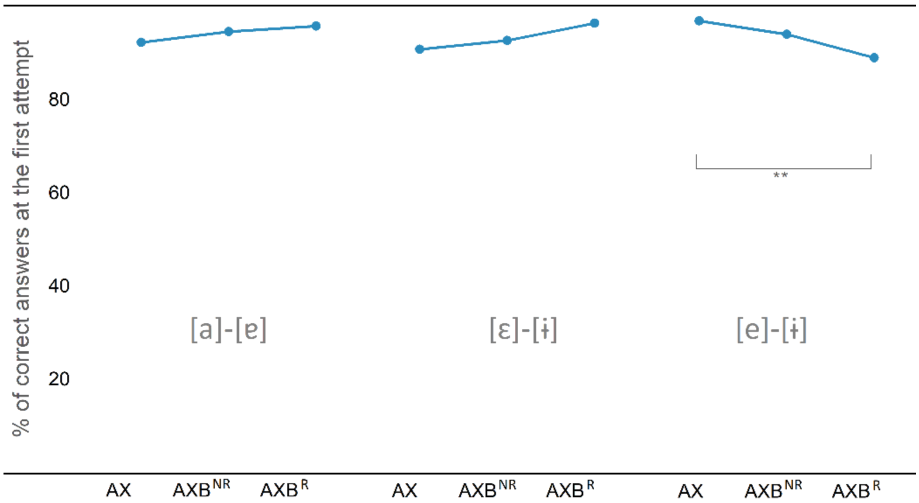


Figure 5.37. Progress observed between training sessions in *stressed-unstressed* contrasts, for Group *Vowels*. ^{NR} = not randomized trials ^{R} = randomized trials

As for the contrast [e]-[ɨ], the analysis showed that learners had significantly worse results in the second AXB session than in the AX task (*p* = .007).

5.5.8.2. Training sessions' results for Group *Stress*

Regarding Group *Stress*, participants trained stress discrimination in two conditions: *1st-2nd syllable contrast* and *2nd-3rd syllable contrast*. Each condition was trained in five sessions: three with AX tasks and two with AXB tasks (see Table 5.12). We analyzed progress in each condition (Figure 5.38, next page) and found that *training sessions* was a significant factor for *1st-2nd syllable* trials (χ^2(4, N = 6) = 11.10, *p* = .025) and for *2nd-3rd syllable* trials (χ^2(4, N = 6) = 51.82, *p* < .001). However, pairwise comparisons showed significant differences only between sessions of the *2nd-3rd syllable* trials (1st AX session vs. 1st AXB session, *p* < .001; 1st AX session vs. 2nd AXB session, *p* < .001; 2nd AX session vs. 1st AXB session, *p* < .001; 2nd AX session vs. 2nd AXB session, *p* < .001).

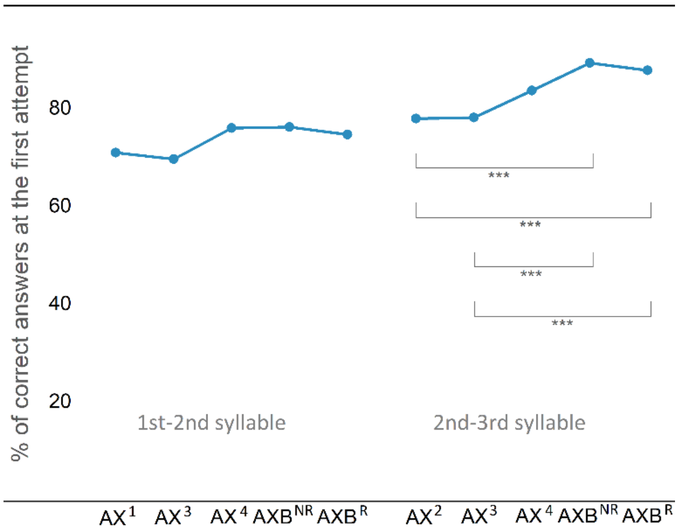


Figure 5.38. Progress observed in Group *Stress* between training sessions, as a function of stress contrast location. $^{1/2/3/4}$ = session number; ^{NR} = not randomized trials ^{R} = randomized trials

5.5.8.3. Training sessions' results for Group *Vowels & Stress*

Finally, training for Group *Vowels & Stress* included contrasts in three conditions: *vowel reduction on the 2nd syllable*, *vowel reduction on the 1st syllable*, and *vowel reduction on both syllables*. *Vowel reduction on the 2nd syllable* and *vowel reduction on the 1st syllable* were each trained in one session with AX trials and two sessions with AXB trials. *Vowel reduction on both syllables* was trained in two sessions with AX trials and two sessions with AXB trials (see Table 5.15). Figure 5.39 (next page) displays the progress between sessions for each condition. In the three cases we can see an improvement from the beginning of the training to the last sessions. The statistical model for the trials with *vowel reduction on the 1st syllable* showed a significant effect of session (χ^2(2, N = 4) = 8.87, p = .012). However, pairwise comparisons did not reveal significant differences between each of the three sessions. In the case of *vowel reduction on the 2nd syllable*, training was also a significant factor (χ^2(2, N = 4) = 43.976, $p < .001$), and pairwise comparisons showed that between the AX session and the 1st AXB session, and between the AX session and the 2nd AXB session, participants significantly improved (for both cases, $p < .001$). Finally, the training of *vowel reduction on both syllables* had a significant effect (χ^2(3, N = 5) = 53.782, $p < .001$), with a significant improvement from the 3rd AX session to the 4th AX session ($p < .001$), from the

3rd AX session to the 1st AXB session ($p < .001$), and from the 3rd AX session to the 2nd AXB session ($p < .001$).

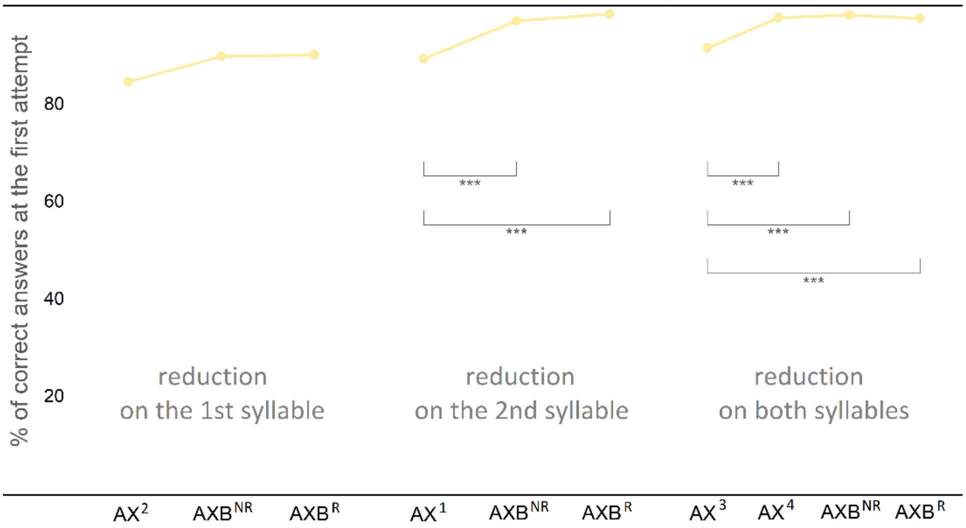


Figure 5.39. Progress observed in Group *Vowels & Stress,* between training sessions, as a function of vowel reduction location. $^{1/2/3/4}$ = session number; ^{NR} = not randomized trials ^{R} = randomized trials

5.5.9. Nonlinguistic factors and individual progress

In the background and follow-up questionnaires, we collected data that allowed us to observe factors that are not related to the L2 linguistic features but may affect the results at the individual level. Considering that an analysis of individual variability is a complex task, and it is not directly the target of our study, we will look only at some factors.

Before approaching individual progress, we investigated if previous contact with EP that participants may had have affected the results at the pretest. Based on the description provided by the participants (see Appendix II), we categorized previous contact with EP in three levels: *none*, *weak*, or *strong*. We ran two logistic mixed-effect analyses, with the formula *answer ~ previous contact with EP + (1|participant)*, one for the results in vowel discrimination and the other for the results in stress discrimination. Both models showed that previous contact with Portuguese did not significantly affect pretest results, neither for vowel perception, nor for stress perception.

In the follow-up questionnaire we collected individual information about aspects such as motivation or self-evaluation of the progress in perception. Table 5.32 presents a summary of the questions presented to the participants for evaluation of some aspects of the training program.

Table 5.32. Medians of responses to the evaluation of training

how interesting the task was	**duration of tasks**	**fatigue**
1 = not at all interesting	1 = too long	1 = not tired at all
2= not very interesting	2 = adequate duration	2 = a little bit tired
3 = interesting	3 = too short	3 = very tired
4 = very interesting		
median = 3	median = 2	median = 2

motivation	**concentration**	**lengh of training**
1 = not motivated at all	1 = not concentrated at all	1 = not at all sufficient
2 = not feel very motivated	2 = not very concentrated	2 = more or less sufficient
3 = motivated	3 = concentrated	3 = sufficient
4 = very motivated	4 = very concentrated	4 = more than sufficient
median = 3	median = 3	median = 2

In general, participants had a positive opinion of the training, with the exception of the length, in which most of the students found that it was not enough long. Two other important close-ended questions were presented to the participants: "Do you think the training contributed to improve your auditory perception of Portuguese language?" and "Do you think the training contributed to improve your pronunciation of Portuguese language?" In the first question, only four students replied "no", the remaining 62 answered positively. As for the second question, in this case 22 participants answered "yes" and 44 answered "no". Appendix II presents further commentaries from the participants collected in the follow-up questionnaire.

Next, we looked for factors (other than the training) that could have affected individual progress from the pretest to the post-test, such as the contact with EP outside the classroom, the affiliation (university/course in which they are enrolled), and courseload – which varied between institutions. Other than these factors, one question particularly interested us: is there a relation between discrimination abilities exhibited in training and in the tests? That is, the learners who showed improvement during the training sessions were also the learners who showed improvement from the pretest to the post-test? If so, then improvement may be related to a more general learning ability, as suggested in Schwab & Llisterri (2014). To answer this question, we plotted the progress between the pretest and the post-test, and combined this visually with the progress in training. Regarding progress from the pretest to the post-test, considering that in general no significant improvement was observed from the pretest to the post-test[N], neither for vowel nor for stress discrimination, we looked only at progress from the pretest to the repeated trials of the post-test. As for progress in training, we looked only into differences from the 5th session to the 6th sessions, since these were the only sessions in which all contrasts/conditions were trained, in the three groups.

In Figure 5.40, in the columns we represented three possibilities of progress: (1) general improvement (i.e., improvement in both vowel and stress discrimination), (2) partial improvement (i.e., in vowel or in stress discrimination), or (3) no improvement.

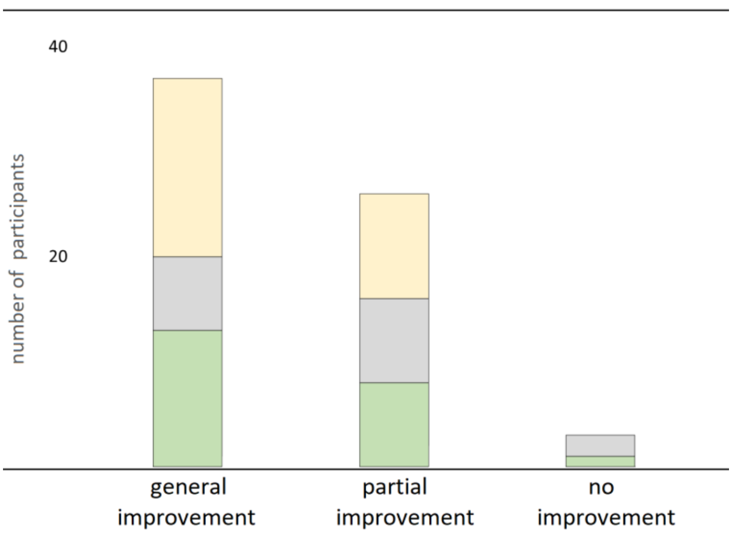


Figure 5.40. Relationship between progress in discrimination from pretest to post-test[R] and progress from the 5th to the 6th training sessions

Thirty-seven students showed general improvement, 26 students showed partial improvement, and three students showed no improvement. The three colors – green, grey, and yellow –, represent progress in the training. The green color indicates the number of students who improved in the training, grey signals the students who did not improve or receded, and yellow signals participants who exhibited regression in training results. As visible in the figure, no relation between improvement during the training sessions and improvement from the pretest to the post-test^{R} was revealed. In the three columns in Figure 5.40 we observe that there were students who improved in the training, students who maintained their abilities, and students who showed a regression during the training.

We then ran fixed-effect logistic regressions, to confirm the above-mentioned observations, as well as investigating if there was an effect of contact with EP, affiliation, or course load in general progress. General progress was calculated as 'result in the post-test^{R} minus results in the pretest'. If positive, it was coded *positive*, otherwise it was coded as *none or negative*. Progress in training was calculated as 'result of the 6th session minus result of the 5th session', and it was also coded as *positive* or *none or negative*. Affiliation consisted of ten levels, corresponding to the ten EP courses. Regarding courseload, the EP courses included three different courseloads: 90 minutes/week, 180 minutes/week, and 270 minutes/week. As for the contact with Portuguese, this information was collected in the follow-up questionnaire, in a similar way as the information about previous contact with Portuguese (see Appendix II). We grouped students into two levels: students with *no contact* with EP outside the classroom (n = 45) and students with *normal* (expected) contact with EP outside the classroom (n = 21).

For these analyses, we used the function GLM (R Core Team, 2014). Levels and models' formulas are displayed in Table 5.33.

Table 5.33. Independent variables and fixed-effect models' formulas for analysis of effects of nonlinguistic factors

independent variables		**levels**
participant	66 levels:	'BGE03', 'ELTEBA12', 'MET06', etc.
progress in training	3 levels:	'positive', 'none or negative'
affiliation	10 levels:	'BGE', 'CORV', 'ELETEBA', etc.
contact with EP	2 levels:	'none', 'normal'
weekly course load	3 levels:	'270', 180', '90'

	GLM models (family = binominal)
null model	improvement ~ 1
model 1	improvement ~ progress in training
model 2	improvement ~ affiliation
model 3	improvement ~ contact with EP
model 4	improvement ~ weekly course load

The outcome of the statistical models showed that none of these factors affected improvement from the pretest to the post-test[R] significantly.

The analyses reported in this section of exploratory data will be reviewed and commented on in the next section.

5.6. Summary and discussion of the results

In this chapter, we reported a study conducted with Hungarian learners of L2 Portuguese, designed to investigate if changes in perception can be promoted by a perceptual training. A ten-week program was created, which included the pretest and the post-test, and in between, three training regimens, separately conducted with three intervention groups. In the previous section we explored the results collected from the experimental tasks completed by the participants. In this section we will summarize and discuss those results.

Before approaching the results for the learners, we must comment on the performance of the *Baseline* group.

In the vowel discrimination tasks, we observed that Portuguese listeners exhibited an unexpected difficulty with the discrimination of [ɛ]-[e] and [e]-[i]. This difficulty occurred across the different tasks (oddity, AXB and AX) and across contexts ([gV], [sV], and [zV]). Correia et al. (2023) observed a similar problem in the discrimination of [ɛ]-[e]. In their study, the researchers designed a perceptual training with English native speakers with no contact with Portuguese. The L2 target segments were the EP contrasts /ɛ/-/e/, /ɔ/-/o/, /n/-/ɲ/, and /l/-/ʎ/. Preliminary results for the baseline group, with Portuguese native speakers, showed a 71.8% of accuracy for the contrast /ɛ/-/e/, by the Portuguese listeners. This result is in line with our results for the contrast [ɛ]-[e] (30.7% mean error rate). We suggest two possible explanations for this difficulty. First, it may be the case that Portuguese listeners need to access lexical representations to discriminate this contrast (e.g., *sede* [ˈsɛdɨ] 'headquarters' vs. *sede* [sedɨ] 'thirst'). Recall that in the experiment reported in Chapter 4, Portuguese participants did not display difficulties when they were asked to identify [gɛ] and [ge] (the same tokens as the ones we used in the present perceptual training study) in real words. Second, it is possible that orthography interfered with the response for discrimination, since <e> is the only orthographic form for both [ɛ] and [e].

Regarding stress discrimination, the results obtained with our *Baseline* group are in line with the results in Correia et al. (2015), confirming that Portuguese native speakers exhibit stress 'deafness' in the absence of vowel reduction. The importance of vowel reduction in stress perception to EP listeners is also attested by the lower error rates in the perception of the [zVzV] tokens, observed in the disyllabic stimuli validation trials. As for the higher difficulty in the perception of *1st-2nd syllable* contrasts compared to *2nd-3rd syllable* contrasts, a possible explanation is the pattern of prominence of the stressed syllable in EP. Assuming that the stimuli we used in our study were not deviant from the standard EP productions, our CVˈCVCV pseudowords complied with the regular EP stress pattern (the unmarked pattern), consequently, the stressed syllable (the 2nd) does not need to be signaled with significantly more prominent acoustic cues than the unstressed syllables. On the contrary, in the ˈCVCVCV and CVCVˈCV pseudowords, stressed syllables (1st or 3rd) were signaled with extra prominence, since these pseudowords do not obey the regular EP stress pattern. However, the prominence is not equal for the 1st and 3rd stressed syllables, due to the fact that the last syllable of a word in citation is usually longer. In Table 5.34 (next page) we provide an example, with the duration for the syllables of the token /filuvi/ produced by talker T1^{F}. As we can see in this table, the last syllable is always longer than the others, even when unstressed.

Table 5.34. Syllable duration for the tokens created for /filuvi/ produced by talker T1[F]

	fi	lu	vi
['filuvi]	245 ms	170 ms	293 ms
[fi'luvi]	240 ms	219 ms	312 ms
[filu'vi]	218 ms	177 ms	360 ms

We now turn our attention to the results for the learners.

At the onset of training, which coincided with the onset of L2 learning in the case of the participants of the present study, we observed homogeneity between the three interventional groups. This homogeneity was assessed by comparing pretest results: for both stress and vowel discrimination, mean error rates in the pretest were not significantly different between the three groups. Furthermore, all experimental groups exhibited significantly higher mean error rates than the Baseline group. Given the fact that some learners reported a previous contact with EP, we also investigated if this factor had significantly affected participants' performances in the pretest, but found no evidence for this effect.

It must be pointed out that homogeneity between groups does not necessarily entail homogeneity between learners. In Figures 5.12 and 5.28, we reported individual mean error rates for vowel and stress discrimination in the pretest, respectively. In these images, individual variability is clearly visible, within and across groups. It is also present in the *Baseline* group results, which suggests that individual variability occurs independently of language knowledge. Moreover, this variability was present throughout the study. Individual variability is difficult to assess, due to the multitude of factors involved (see section 1.2, in Chapter 1), and such a complex analysis was not in the scope of our work. Nevertheless, in section 5.5.9, we investigated some nonlinguistic factors which might have contributed to the variability between learners in the progress during the study. Specifically, we investigated if the ability to improve during training sessions was related to the ability to improve from the pretest to the post-test. We also looked at the effects of (1) affiliation (EP course), (2) contact with EP outside the classroom, and (3) course load, in individual improvement from the pretest to the post-test^{R}. The results of the statistical models showed that none of these factors did explain the variability of the results.

Before discussing in detail vowel discrimination and stress discrimination results, we need to make a note on generalization. As described in 5.1 (see Table 5.1), post-test trials involved two degrees of generalization, when comparing to training items. The *post-test*^{R} consisted of the repetition of the pretest trials. These included untrained stimuli, but produced by the same talkers as the stimuli in the training sessions. The *post-test*^{N} included untrained stimuli, produced by novel talkers. Post-test results showed a similar picture for both vowel and stress discrimination. With rare exceptions, from the pretest to the post-test^{R} there was a general improvement in the results. However, from the pretest to the post test^{N}, we observed a decline in the results. Even in the case of progress in stress acquisition for Group *Stress*, which was the only group presenting a significant improvement from the pretest to the post-test^{R}, between the pretest and the post-test^{N} this group exhibited only a mild improvement (considering the three groups progress). We can thus conclude that there was some generalization of learning – since learners were able to improve to untrained tokens – but this generalization was limited, since it was not observed when we included more phonetic variability (novel talkers).

Next, we review and discuss the results for vowel and stress discrimination (separately). In both cases, we start by overviewing the results at the onset of the training program (in the pretest), followed by the discussion of the progress from the pretest to the post-test.

5.6.1. Vowel discrimination

The results in the pretest for vowel discrimination showed that Hungarian learners exhibited significantly more difficulties in discriminating EP vowel contrasts than Portuguese speakers, particularly in the contrasts were the two segments perceptually overlap each other: [ɐ]-[ɛ], [ɛ]-[e], [e]-[i], and [ɐ]-[ɨ] (see Chapter 4). The effect of perceptual overlap in contrast discrimination was attested in previous L2 perceptual studies (Elvin et al., 2021; Faris et al., 2018; Flege & Mackay, 2004; Tyler et al., 2014). Additionally, based on the results for the catch trials in the pretest, we also observed that

Hungarian learners were able to identify tokens from the same category but phonetically different as belonging to one category only for the familiar vowels, that is, the vowels existing in their L1 ([a], [ɛ], [e] and [i]). In trials with sequences of [ɐ] or [ɨ] – vowels absent from the Hungarian inventory – participants showed high error rates.

Regarding progress from the pretest to the post-test^{R}, results across contrasts showed a mild improvement in the three groups. The improvement was observed in both *perceptual overlap* contrasts and *stressed-unstressed* contrasts, but was slightly more visible in the *perceptual overlap* condition. In *stressed-unstressed* contrasts ([a]-[ɐ], [ɛ]-[ɨ], and [e]-[ɨ]), we observed only slight differences between the pretest and the post-test^{R} and, in both tests, error rates were low. In contrasts with perceptual overlap ([ɛ]-[e], [ɐ]-[ɛ], [ɐ]-[ɨ], and [e]-[i]), we observed two outcomes. In the contrasts with the unfamiliar vowels, [ɐ]-[ɛ] and [ɐ]-[ɨ], participants exhibited improvement. As for the contrasts [ɛ]-[e] and [e]-[i], the groups which trainings included these vowels (Group *Vowels* and Group *Vowels & Stress*) exhibit a slight decline in the results. By contrast, Group *Stress* (whose training included only the vowels [i] and [u]) displayed improvement in the discrimination of both contrasts. These results suggest two conclusions. First, learners had less problems in improving their perception of contrasts with unfamiliar vowels ([ɐ]-[ɛ] and [ɐ]-[ɨ]). Second, training may had had a negative effect on the perceptual adjustment of the contrasts involving the mid-front vowels ([ɛ]-[e] and [e]-[i]).

Regarding progress from the pretest to the post-test^{N}, overall, we observed an increase in the error rates. This increase is more pronounced, and significant for all groups, in *stressed-unstressed* contrasts. When we investigated the results for each vowel contrast, we obtained a chaotic picture. We must first highlight the case of the contrast [ɐ]-[ɛ], which presented a substantial improvement for the three groups (although, comparatively, only significant for Group *Stress*). This result is contrary to the changes we observed for the remaining contrasts, among which, the more problematic cases were the contrasts [e]-[i] and [e]-[ɨ]. In these two contrasts, the three groups exhibited a very pronounced decline in discrimination. An increase in mean error rates was also observed for [ɛ]-[ɨ], but less prominent. As for [ɛ]-[e], no changes in the results were observed for Group *Vowels*, nor for Group *Stress*, but in Group *Vowels & Stress* error rate increased from the pretest to the post-test^N. Finally, in the contrast [a]-[ɐ], with exception of Group *Vowels*, which presented some improvement, we also observed a decline in the results. It must also be noted, that in the contrast [ɐ]-[ɨ], Group *Vowels* also outstands form the others with its improvement.

The different results between the contrasts in progress from the pretest to the post-test suggest that the learning path is not similar between those contrasts. The improvement observed in the contrast [ɐ]-[ɛ] stands out from the other contrasts, suggesting that in this contrast, learners were more successful in overcoming discrimination problems. Although less expressive, improvement in discrimination of [ɐ]-[ɨ] was also achieved.

By contrast, problems with the vowels [ɛ], [e], and [i] seem to persist, or even aggravate. One possible explanation is that the acoustic/articulatory proximity between these phones – all front vowels – may curtail perceptual adjustments. This may explain, for example, the decline in the results observed from the pretest to the post-test^N for the contrast [e]-[i], a contrast between a mid-close front and a close front vowel. The three groups also showed a decline from the pretest to the post-test^N for the contrast [ɛ]-[e] (mid-open front vs. a close front vowel). However, the increase of error rate for this contrast was not so high as for [e]-[i]. Interestingly, from the pretest to the post-test^N we also observed a pronounced decline in discrimination for the contrast [e]-[ɨ] (mid-close front vs. close central vowels). This suggests that perceptual tunning between mid-close and close vowels was more difficult than between mid-open and close vowels. As for the contrast [ɐ]-[ɨ], although this configures a contrast between two central vowels (mid and close, respectively), the fact that both sounds are unfamiliar to the learners, may have drawn more their attention to the differences between these two vowels.

5.6.2. Stress discrimination

Results for the pretest confirmed the previously reported high degree of stress 'deafness' in Hungarian speakers (Dupoux et al., 2008; Honbolygó, Kóbor, et al., 2017; Peperkamp & Dupoux, 2002). We obtained a mean error rate of 57.6% in discrimination of stress contrasts across the intervention groups, 25.2% higher than the *Baseline* group. Additionally, the mean error rates for stress contrast perception for the Hungarian listeners were significantly higher than for consonant contrasts, which is also in line with the above-mentioned studies. These studies, however, used CVCV pseudowords exclusively and only *1^{st}-2^{nd} syllable* contrasts were investigated. In our study, other than *1^{st}-2^{nd} syllable* and *2^{nd}-3^{rd} syllable* contrasts, we also obtained data for perception of stress located on the *1^{st} syllable*, on the *2^{nd} syllable*, and on the *3^{rd} syllable*. The stress contrast conditions (*1^{s} - 2^{nd} syllable* and *2^{nd}-3rd syllable*) were investigated in the change trials, and perception of stress on the *1^{st} syllable*, *2^{nd} syllable*, and *3^{rd} syllable* was tested in the catch trials. Regarding the results of catch trials, Hungarian participants displayed the lowest error rate in discrimination of sequences with pseudowords stressed on the *1^{st} syllable*, as expected, since this is consistent with the Hungarian stress pattern. However, this value was not significantly different from the mean error rate for perception of stimuli exhibiting stress on their *3^{rd} syllable*. By contrast, performance in trials with stimuli stressed on the *2^{nd} syllable* was significantly higher than in the conditions *1^{st} syllable* and *2^{nd} syllable*. As for the stress contrast conditions, we expected that Hungarian participants display better discrimination in the condition *1^{st}-2^{nd} syllable* than in the condition *2^{nd}-3^{rd} syllable*, since the first consists of a contrast between a familiar and an unfamiliar stress location, whereas *2^{nd}-3^{rd} syllable* stimuli includes two unfamiliar stress locations. However, results showed that the Hungarian learners exhibited more difficulties in the *1^{st}-2^{nd} syllable* condition than in the *2^{nd}-3^{rd} syllable* condition, and the difference between these two conditions was significant. We must point out that a similar result was observed with the *Baseline* group. Although this group, in general, exhibited lower error rates than Hungarian learners, it also displayed more difficulties in the discrimination of *1^{st}-2^{nd} syllable* contrasts than in the condition *2^{nd}-3^{rd} syllable*. The explanation proposed for the outcome of trials with the Portuguese listeners can be applied to the results of the Hungarian learners, both in the case of catch trials as well as for the change trials: the EP stress pattern assigns stress prominence in marked cases ('CVCVCV or CVCV'CV) but not in unmarked cases (CV'CVCV). Consequently, considering that the last syllable of the word (in isolation and falling intonation) is naturally prolonged, durational differences in *2^{nd}-3^{rd} syllable* contrasts are more prominent than in *1^{st}-2^{nd} syllable* contrasts.

Regarding changes in stress contrasts' perception, here we observe a different picture than that for the EP vowel contrasts.

From the pretest to the post-test^{R}, overall, the three groups exhibited improvement in the results, although not significant for Group *Vowels*. When we look at each stress contrast location (*1st-2nd syllable* and *2nd-3rd syllable*), we have a similar picture, although the statistical analysis showed that only Group *Stress* improved significantly in both conditions. Furthermore, in the pretest, differences between perception of *1st-2nd syllable* contrasts and perception of *2nd-3rd syllable* contrasts were significant for all intervention groups. However, this was not the case in the post-test^{R}, for Group *Stress*. Moreover, in this test, in the *1st-2nd syllable* contrast trials, results of Group *Stress* were not significantly different from results of the *Baseline* group.

Regarding progress in discrimination from the pretest to the post-test^{N}, we observed that Groups *Vowels* and *Vowels & Stress* exhibited more difficulties in the post-test^{N} than in the pretest. By contrast, Group *Stress* displayed a mild improvement, although it did not reach significance.

These results suggest that both trainings focused on stress perception (Groups *Stress* and *Vowels & Stress*) had a positive effect, although more visible for the training focused solely on stress contrasts (Group *Stress*).

Overall, the study we reported in this chapter provided us with important insights on changes in the discrimination of EP vowels and stress that occur during the first moments of L2 acquisition. The results and issues presented in this last section will be resumed and further discussed, in light of our *Hypotheses* and considering some literature presented in Chapter 1, next, in General Discussion.

General Discussion

The present work aimed at investigating how L2 learners acquire non-native vowels absent from the L1 considering that (1) those vowels are unstressed and reduced, and (2) those learners are insensitive ('deaf') to word stress contrasts. Thus, this dissertation explored the acquisition of both segmental and suprasegmental L2 features.

Research in L2 phonology is a growing area, and the last two decades have witnessed an increasing number of studies investigating a variety of factors – linguistic and nonlinguistic – that contribute to L2 speech perception. This dissertation approaches L2 speech perception from a cross-linguistic point of view and observes the influence of the relation between L1 and L2 in the acquisition of L2 Portuguese.

From the review of studies undertaken in Chapter 1, two observations stood out. First, L2 speech research has been more focused on investigating the perception (and production) of segmental features than of suprasegmental aspects. Second, results from studies on the L2 acquisition of vowels point to the benefit of perceptual training, whereas results on L2 stress acquisition are conflicting regarding the effectiveness of such training. Suprasegmental features such as word stress inspire more caution among researchers for several reasons. While segments involve fine-grained spectral changes, word stress depends on the interaction of diverse global acoustic cues – pitch, duration, and intensity – which are used not only for stress but also for other prosodic functions (Dupoux et al., 2008). Moreover, that interaction is not straightforward. For example, in EP word stress, Portuguese speakers resort to modulations in duration and energy only in the absence of vowel reduction and only if the word does not comply with the regular stress pattern (Delgado-Martins, 1986). Specifically, in words that do not comply with the EP regular stress pattern (and do not include reduced vowels) stressed syllables are more prominent than unstressed syllables. This is not observed in words that obey the regular stress pattern, in which stressed syllables are not significantly more salient than unstressed syllables. Such language-specific patterns are not widely studied among languages, and even less so in L2 acquisition (Chrabaszcz et al., 2014). To these aspects related to word stress, there are conflicting results in studies on L2 stress perception/acquisition.

Specifically, while some studies indicate that stress 'deafness' cannot be surpassed, others show that this problem is not necessarily persistent. The consequence to this is that, although researchers point out the importance of studying L2 acquisition of word stress (Field, 2005; Hahn, 2004), to date, very few studies have focused on this topic.

The present dissertation investigated the acquisition of reduced EP vowels – [ɐ] and [ɨ] –, as well as EP word stress by Hungarian learners of L2 Portuguese at the onset of learning. To this purpose, we conducted two studies.

The first study (Chapter 4) was designed to observe how Hungarian speakers map EP vowels into their L1 system. The aim of this study was to identify the difficulties Hungarian learners may encounter in the perception of [ɐ] and [ɨ] when they start learning Portuguese, and thus design a perceptual training aimed at those difficulties. The study consisted of a forced-choice identification task with goodness-of-fit ratings. Based on comparisons between the Hungarian and the EP F1×F2 vowel spaces, we hypothesized that *Hungarian learners of L2 Portuguese, at the onset of learning, categorize the EP [ɐ] into /ɛ/, /eː/ or /ø/, and the EP [ɨ] into /y/, /eː/, /ɛ/ or /ø/.*

The results partially support this hypothesis. Hungarian listeners identified [ɐ] firstly as /ɛ/ and secondly as /ø/, and [ɨ] firstly as /y/ and secondly as /ø/. However, the vowel [ɐ] was never identified as /eː/, [ɨ] was never identified as /ɛ/ and its identification as /eː/ was below chance level.

This entails a perceptual problem for [ɐ] and [ɛ], since both are perceived as /ɛ/ in most of the cases, which means that Hungarian speakers are not able to perceive the contrast [ɐ]-[ɛ]. Furthermore, Hungarian speakers rated the EP [ɐ] as a better fit to the Hungarian /ɛ/ than the EP [ɛ]. These results support and explain the observation in the classroom about [ɐ] reported in the Introduction: Hungarian learners exhibit a strong tendency to transcribe [ɐ] as <e>, which in Hungarian corresponds unequivocally to /ɛ/. As for [ɨ], results showed that no perceptual identification with EP vowels occurs.

Although the aim of the identification experiment was to observe the perception of [ɐ] and [ɨ], we tested the nine EP oral vowels: [a], [ɐ], [ɛ], [e], [ɨ], [i], [ɔ], [o], and [u]. The results revealed several cases of perceptual overlap involving directly or indirectly our target vowels (Figure 6.1, next page).

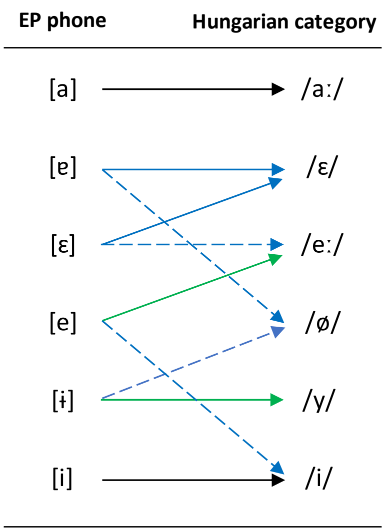


Figure 6.1. Perception of the EP front and central vowels by Hungarian native listeners (Chapter 4). Continuous lines: top-ranked choice; dashed lines: second-ranked choice; black lines: identifications above 90%; green lines: identifications above 70%; blue lines: identifications below 70%

As shown in Figure 6.1, the perception of [ɐ] and [ɨ] is entangled with the perception of the EP mid-front vowels [ɛ] and [e], and in turn, a perceptual overlap between [e] and [i] is observed. Consequently, in order to acquire [ɐ] and [ɨ], learners must be provided with tasks that help them adjusting the perceptual boundaries between [ɐ], [ɨ], [ɛ], [e], and [i], which is the aim of the study reported in Chapter 5.

Portuguese listeners perceive their L1 vowels as stressed ([a], [ɛ], [e], [i], [ɔ], [o], [u]) or unstressed ([ɐ], [ɨ], [i], [u]). The full-reduced pattern emerges from the presence or absence of variable word stress. Acquisition of reduced EP vowels is, thus, dependent on the acquisition of stress. Crucially, stress in Hungarian is fixed in the 1st syllable, not lexically contrastive, and Hungarian listeners exhibit stress 'deafness' (Dupoux et al., 2008; Honbolygó et al., 2017; Peperkamp & Dupoux, 2002). In order to acquire and use the EP vowel system, Hungarian learners must therefore overcome stress 'deafness' and acquire EP word stress.

Considering the important relation between vowels and word stress in EP, the difficulties with EP vowels exhibited by Hungarian speakers and the 'stress' deafness attested in those speakers, we developed an online high variability phonetic training to investigate (1) if perceptual training promotes the acquisition of the vowels [ɐ] and [ɨ], and word stress, and (2) which training focus is more effective: one that focuses on vowels and word stress separately, or one that combines vowels and stress. Furthermore, with the results of the study we also aim at observing if there is a hierarchy in acquisition between segmental features and suprasegmental features in L2.

The Hungarian learners of L2 Portuguese who took part in the study were enrolled in a Portuguese language Beginner course. Although some participants reported having had previous contact with Portuguese, the statistical analysis revealed that this contact did not significantly affect the results in the pretest. The participants were divided into three intervention groups. Group *Vowels* (n = 20) received training in [zV] monosyllables which included the contrasts with *perceptual overlap*: [ɐ]-[ɛ], [ɛ]-[e], [e]-[i], and [ɐ]-[ɨ] (Figure 6.1.) To these contrasts, we added [a]-[ɐ], [ɛ]-[ɨ], and [e]-[ɨ]. The inclusion of these three contrasts aimed at promoting the distinction between *stressed-unstressed* for our target vowels. Group *Stress* (n = 22) received training in word stress contrasts, with trisyllabic pseudowords contrasting in *1st vs. 2nd syllable* (e.g., [ˈlutinu]-[luˈtinu]) and in *2nd vs. 3rd syllable* ([luˈtinu]-[lutiˈnu). The pseudowords included the vowels [i] and [u] only, to exclude vowel reduction. Finally, Group *Vowels & Stress* (n = 24) was trained with disyllabic pseudowords ([zVzV]) that included *1st* vs. *2nd syllable* stress contrasts, and contrasts of vowel reduction on the first syllable (e.g., [ˈzazi]-[zɐˈzi]), on the second syllable (e.g., [ziˈza]-[ˈzizɐ]), or on both syllables (e.g., [ˈzazɐ]-[zɐˈza]). These pseudowords included the vowels trained by Group *Vowels*.

Training lasted for six weeks, with one session of about 10-15 minutes per week, in which participants were presented with AX or AXB tasks. Before and after this six-week training, we assessed vowel and stress discrimination, separately. Students completed a pretest and a post-test, both consisting of oddity discrimination tasks with catch trials. Pseudowords in the tests were different but complied with the structure of those used in the training sessions and were produced by the same talkers used in the training tasks. Besides the repetition of the pretest trials after training (post-test^{R}), the post-test included trials with novel stimuli, that is, new tokens with the same structure of the remaining stimuli, but produced by novel talkers (post-test^{N}). To comply with the HVPT method, stimuli were produced by different Portuguese talkers.

The results for vowel perception showed that Hungarian learners exhibit more difficulties in the perception of EP vowel contrasts than the *Baseline* group of Portuguese native speakers (31.4% against 9.2%). Additionally, we confirmed the prediction established with the results of the forced-choice identification task, according to which the higher the *perceptual overlap* the more discrimination difficulties are exhibited. In the pretest, mean error rates for contrasts with vowels with perceptual overlap were higher than for the *stressed-unstressed* contrasts (45.3% against 12.8%). This result is in line with previous studies, in which discrimination difficulties in L2 vowel contrast have been associated to perceptual overlap (Elvin et al., 2021; Faris et al., 2018; Flege & Mackay, 2004; Tyler et al., 2014).

When we compared the results between the pretest and the post-test for vowel discrimination, we observed a general improvement, but only from the pretest to the post-test^{R}. This improvement was similar across the three intervention groups. The results from the pretest to the post-test^{N} show a decline in the performance and, as from the pretest to the post-test^{R}, similarity across groups.

However, a closer look at the results reveals that participants exhibit different progress in learning for each contrast. The contrast [ɐ]-[ɛ] stands out among the others, since it showed improvement not only from the pretest to the post-test^{R} but also to the post-test^{N}. The contrast [ɐ]-[ɨ] also showed a similar improvement from the pretest to the post-test^{R}, but not to the post-test^{N}. In the remaining contrasts, we observed no improvement or only a small improvement, from the pretest to the post-test^{R}, and mainly a decline in the results from the pretest to the post-test^{N}. To account for these results, we discuss the L2 crosslinguistic models described in Chapter 1 – SLM, PAM, and the L2LP model.

PAM was originally targeted at observing non-native perception and not L2 learning. Although this model was later adapted to L2, studies on PAM-L2 focus on the effect of L2 vocabulary in perceptual tunning. Specifically, those studies observe how the increase in the L2 vocabulary size may benefit learning for some L2 contrasting segments while inhibiting learning in other contrasts (Bundgaard-Nielsen et al., 2011). These studies observe differences in perception between participants with different levels of L2 vocabulary knowledge, therefore, they are not in line with our research. Contrary to PAM, SLM was originally proposed for L2 learning. Studies on this model observe acquisition of *new* categories in contrast with acquisition of *similar* categories.[100] Considering that both target vowels in this study, [ɐ] and [ɨ], are *new* to Hungarian learners, this model does not provide much insight on possible differences between the acquisition of these two *new* categories. As for the L2LP model, although not restricted to formal learning, research on the model has been more focused on this context. Moreover, this model provides a more detailed acquisition path for L2 segments, proposing four *learning scenarios*: *new*, *subset easy*, *subset difficult*, and *similar*. Considering this, we investigated if this model provides explanations for our results.

When Hungarian listeners are exposed to the contrast [ɐ]-[ɛ], the results of the identification task (Chapter 4) revealed that, in most of the cases, both vowels are identified as /ɛ/. This means that the listeners do not perceive a contrast. According to the L2LP model, this contrast configures a *new scenario*, since learners will have to create a new category (or split the existing one) for the [ɐ] vowel. The learning path to change this situation is assumed to be the most difficult one, since it entails learning tasks of different nature: perceptual and representational. Also in line with the L2LP model, the contrasts [ɛ]-[e] and [ɐ]-[ɨ] configure a *subset scenario*, since the vowels in each contrast were identified as more than two L1 vowels. The segments [ɛ] and [e] were identified by Hungarian listeners as three different Hungarian categories: /ɛ/, /eː/, and /eː/, /i/, respectively. A similar situation was observed in the case of the contrast [ɐ]-[ɨ], as [ɐ] and [ɨ] were respectively identified as /ɛ/, /ø/, and /y/, /ø/, to varying degrees. However, according to Elvin et al. (2021), *subset scenarios* include a wide range of situations, with varying degrees of difficulty. To determine the degree of difficulty of these scenarios, we must take into account the degree of perceptual overlap. Considering that in the contrast [ɛ]-[e] there was a 41.0% of perceptual overlap, and in the perception of [ɐ]-[ɨ] there was only 25.2%, the first is presumably considered a *subset difficult* and the latter a *subset easy scenario*. The L2LP model predicts that for *subset scenarios*, learners also have perceptual

[100] Studies on SLM investigate a wide range of factors other than the issue of (dis)similarity between L1 and L2 segments, such as age of acquisition, age of arrival, and input.

and representational learning tasks; however, difficulty level is expected to be lower than in the *new scenario*. With respect to the EP vowels [e] and [i], these were perceived, in most of the cases, as the Hungarian /eː/ and /i/, respectively. According to the L2LP model, this configures a *similar scenario*. The model predicts that in such a situation, learners only have a perceptual task, since they only need to adjust the perceptual boundaries of each category. Accordingly, this is the easiest learning scenario.

Figure 6.2 summarizes the learning scenarios implicated in the acquisition of EP [ɐ] and [ɨ] in Hungarian learners of EP, at the onset of learning, as predicted by the L2LP model.

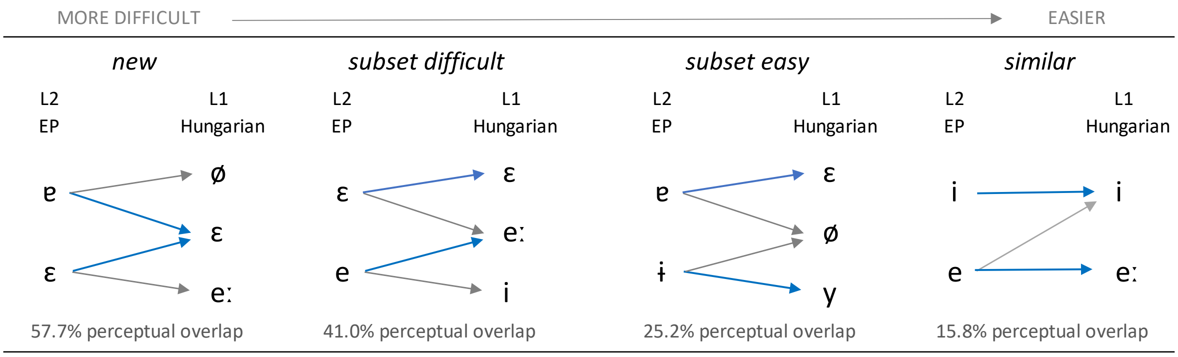


Figure 6.2. Learning scenarios for the EP contrasts [ɐ]-[ɛ], [ɛ]-[e], [ɐ]-[ɨ], and [e]-[i] in Hungarian learners of EP, according to the L2LP model. Blue lines: top-ranked choice; grey lines: second-ranked choice

In sum, applying the L2LP model framework to our case, we observe four different *scenarios* with varying difficulty (from more difficult to easier): *New:* [ɐ]-[ɛ] > *subset difficult:* [ɛ]-[e] > *subset easy:* [ɐ]-[ɨ] > *similar scenario:* [e]-[i]. However, our results were not consistent with these predictions, since learners exhibit greater improvement exactly for the contrast [ɐ]-[ɛ]. Additionally, [e]-[i] was the most problematic contrast, with learners exhibiting more difficulties in the post-test than in the pretest. According to our results, the sequence of learning difficulty was (from more difficult to easier): [e]-[i] > [ɛ]-[e] > [ɐ]-[ɨ] > [ɐ]-[ɛ]. The outcomes in our study suggest that learners adjust first their perception of contrast with unfamiliar vowels – [ɐ]-[ɨ] and [ɐ]-[ɛ] –, while they have more difficulties in contrasts with vowels existing in their L1 – [e]-[i] and [ɛ]-[e]. This is in line with the SLM framework, which posits that the greater the perceptual dissimilarity between L1 and L2 sounds, the higher the probabilities of a new category creation. However, comparisons between our results and results of studies on SLM must be interpreted with caution, since our participants are acquiring the L2 in a formal learning context and SLM observes learners in naturalistic learning context and with higher control over the L2. Moreover, SLM does not observe (dis)similarities between contrasts, but rather segments. In other words, we do not posit that our results support the SLM predictions.

Two other factors may have hindered changes in perception for the EP [ɛ], [e] and [i]. First, the high frequency of [ɛ] in Hungarian speech may have paradoxically contributed to the difficulties with the EP contrast [ɛ]-[e]. The learners from our study are in contact with the target language mostly in the classroom (in Hungary), and are thus much more exposed to the Hungarian productions of [ɛ], than to EP productions of the same vowel. Moreover, despite the fact that both EP and Hungarian [ɛ] are represented by the same symbol of the International Phonetic Alphabet, their quality is not necessarily identical. In the identification experiment reported in Chapter 4, the EP [ɛ] was identified as a poor fit to the Hungarian /ɛ/ or /eː/ (in both cases, mean goodness-of-fit was 2 = *bad example*).[101] It is possible that the improvement observed for [ɐ], which indicates the successful perceptual adjustment for this EP vowel, may have been partially at the expense of a regression in the perception of the EP [ɛ]. Second, orthography may have also difficulted the tunning in perception of contrasts with [ɛ], [e] and [i]. Hungarian orthography is transparent, and in this language, these vowels correspond to three distinct graphemes: <e>, <é> and <i>, respectively. However, in EP, both [ɛ] and [e] have only one orthographic form, <e>, and [i] is also frequently represented by <e>.[102]

Other than the *perceptual overlap* contrasts above discussed, we also observed *stressed-unstressed* minimal pairs: [a]-[ɐ], [ɛ]-[ɨ], and [e]-[ɨ]. The purpose of including these contrasts was to draw learners' attention to the EP stressed-unstressed patterns, complementing the perceptual and representational tasks involved in the acquisition of contrasts with *perceptual overlap*. In other words, we aimed at promoting the task of matching [ɐ] with /a/, and [ɨ] with /ɛ/ or /e/. Discrimination in these three contrasts was

[101] Recall that the EP [ɐ] was considered a better fit to the Hungarian /ɛ/ than the EP [ɛ].
[102] For example: *selo* [ˈsɛlu] 'I stamp', *selo* [ˈselu] 'stamp', and *exame* [iˈzɐmɨ] 'exam'.

not expected to be problematic, since no perceptual overlap was observed between these vowels in the identification experiment reported in Chapter 4. Although in the pretest and the post-test[R] we obtained low mean error rates for these vowels, in the post-test[N] we observed significantly higher error rates than in the pretest, for the three groups. This decline in perception was more pronounced for [e]-[ɨ]. This may be due to, first, the learning process, that is, learners are in the process of adjusting their perception and representations of the L2 vowels, but they still did not reach the end-state and, thus, were not able to generalize learning to the post-test[N], as observed in Ou (2011). Second, the difficulties observed for the contrasts [ɛ]-[ɨ], [e]-[ɨ], and [e]-[i] in the post-test[N] may be related with acoustic similarity/proximity. When more phonetic variability was introduced in the novel trials, which included stimuli produced by unfamiliar talkers, we observed more problems between the mid-close front and the close central vowel ([e]-[ɨ]) than between the mid-open front and the close central vowels ([ɛ]-[ɨ]). Moreover, the more problematic contrast was [e]-[i], a contrast between mid-close and close front vowels.

With respect to the effectiveness of training, as previously mentioned, when comparing the progress of the groups whose training included EP vowel contrasts – Group *Vowels* and Group *Vowels & Stress* – with the results for Group *Stress*, we did not find significant improvement in EP vowel discrimination. In general, we observed a similar learning path in the three groups. With this in consideration, *Hypothesis 2: A perceptual training focusing on vowel discrimination promotes the acquisition of the EP vowels [ɐ] and [ɨ] by Hungarian learners, although to different extent for each of the vowels* is refuted.

Two comments must be added. First, the differences in progress observed between each vowel contrast shows that vowel learning is not a homogenous process. Second, although the results point to the unsuccess of perceptual training focused on vowels, in the contrasts [a]-[ɐ] and [ɐ]-[ɨ] Group *Vowels* displayed a visible improvement, contrary to the other groups. The fact that this improvement did not reach significance level suggests that training length or intensity was perhaps insufficient.

With respect to stress discrimination at the onset of learning, the Hungarian students displayed 'deafness', as attested in previous studies (Dupoux et al., 2008; Honbolygó et al., 2017; Peperkamp & Dupoux, 2002). Interestingly, participants performed significantly better in the contrasts between the 2nd and the 3rd syllables (e.g., [duˈtiku]-[dutiˈku]) than in the contrasts between the 1st and the 2nd syllables (e.g., [ˈdutiku]-[duˈtiku]). The explanation for this result may be found in the relative prominence of acoustic cues in EP stress pattern. As explained in 5.6.2, the last syllable of a word in citation form is by nature longer, whereas the second syllable, even if stressed, does not receive extra durational prominence.[103] The durational difference between the 2nd and the 3rd stressed syllables is, in normal cases, higher than between the 1st and the 2nd stressed syllables. Salsignac (1998) observed the influence of prominence of stress, concluding that when stress is more prominent, listeners relied first on acoustic cues, and, by contrast, when it is less salient, they resorted to their L1 stress patterns.

As for the progress in learning, we observed an effect of training from the pretest to the post-test^{R}, but not to the post-test^{N}. Specifically, results showed that from the pretest to the post-test^{R}, both Groups *Stress* and *Vowels & Stress* improved significantly, but not Group *Vowels*. These results support *Hypothesis 3: Hungarian learners benefit from perceptual training focusing on stress discrimination to overcome stress 'deafness'*. However, we must point out that improvement was observed only to some extent by Group *Stress* (since progress from the pretest to the post-test^{N} did not reach the significance level) and by Group *Vowels & Stress* (since it displayed a decline from the pretest to the post-test^{N}).

The improvement observed in the discrimination of stress contrasts contradicts the assumption of stress deafness being a persistent state, as observed by Dupoux et al. (2008) and Honbolygó et al. (2017), and is in line with the results of the studies of Brawerman-Albini (2012) and Ou (2011). An important difference between the four studies is that, while the first two studies test stress discrimination in learners with different L2 proficiency, Brawerman-Albini (2012) and Ou (2011) conducted perceptual interventions. This supports the idea that perceptual training is effective to overcome problems in stress discrimination and, eventually, other suprasegmental features.

[103] We assume, as previously mentioned, that our stimuli were not deviant from the standard EP speech, and accordingly, presented the suprasegmental patterns described in Delgado-Martins (1986).

The comparison of the results collected in the three intervention groups, for vowel and stress discrimination, allows us to answer the important question of whether auditory tasks should focus on vowel and word stress separately or combine both features. In vowel discrimination, we observed that neither training focused on vowels, nor the training focused on vowels and stress combined had a significant effect on perceptual learning. Consequently, our data did not provide us with the necessary information to conclude which training is more beneficial for L2 vowel discrimination: a training focusing only on vowels, or a training focusing on vowels and stress. However, we must point out that in the post-test^{N}, in the contrasts [a]-[ɐ], [ɐ]-[ɨ] and [ɛ]-[e], Group *Vowels & Stress* displayed worse results than Group *Vowels*, although differences between the groups' progress did not reach significance level.

As for stress discrimination, we observed that, overall, both Group *Stress* and Group *Vowels & Stress* exhibited significant improvement (from the pretest to the post-test^{R}). However, in general, Group *Stress* displayed better results than Group *Vowels & Stress*[104]. The difference between the two groups performance is observable when we analyzed separately each condition (*1st-2nd syllable* contrast and *2nd-3rd syllable* contrast): Group *Stress* significantly improved from the pretest to the post-test^{R}, in both conditions, whereas *Group Vowels & Stress* improvement did not reach significance level. Moreover, Group *Stress* improved from the pretest to the post-test^{N} (although not significantly), but Group *Vowels & Stress* displayed an increase of mean error rate. This fact suggests that a training focused on stress discrimination may be more effective for stress acquisition than a training that combines stress and vowel contrasts.

With the above observations in consideration, *Hypothesis 4: Hungarian learners of L2 Portuguese benefit more from a perceptual training that combines stress and vowel reduction than from a training focused on stress or vowels only* was refuted.

A comment must be added on the possible effect of the stimuli on the results. Regarding vowel discrimination, in the contrast [ɛ]-[e], for example, from the pretest to the post-test^{N} we observed more difficulties in Group *Vowels & Stress* than in Group *Vowels* (or in Group *Stress*). The explanation may lie in the stimuli provided in the training sessions.

[104] Group *Stress* improved from the pretest to the post-test^{N}, but Group *Vowels & Stress* no.

While Group *Vowels* was presented with the contrast [ɛ]-[e] in the same sequences (e.g., [zɛ]-[ze]), this was not the case for the stimuli presented to Group *Vowels & Stress*. In this group, the vowels [ɛ] and [e] were presented in different sequences, that is, they did not contrast with each other (e.g., [ˈzɛzi]-[zɨˈzi] in a sequence, and [ˈzezi]-[zɨˈzi] in a different sequence).

In respect to stress discrimination, Group *Stress* displayed better results than Group *Vowels & Stress.* However, it is worth mentioning that while Group *Stress* was trained in two conditions (*1st-2nd syllable* and *2nd-3rd syllable* contrasts), Group *Vowels & Stress* was only presented with stimuli contrasting in *1st-2nd syllable*. The inclusion of two contrasts (*1st-2nd* and *2nd-3rd syllables*) in the regimen for Group *Stress* may explain the better results for this group, as compared to the performance of Group *Vowels & Stress*. Additionally, the stimuli presented to Group *Stress* in the training sessions included only two vowels ([i] and [u]) but it contained varied consonants, all present in the learners' L1, whereas stimuli for the regimen of Group *Vowels & Stress* included only the consonant [z]. According to Michaux (2016), absence of phonotactic information can negatively impact improvement in stress perception.

The last hypothesis put forth in this dissertation aimed at testing whether there is a hierarchy in L2 acquisition between segmental and suprasegmental properties. As previously reported, training regimens for Groups *Vowels* and *Vowels & Stress* had no effect in vowel discrimination, whereas training for Group *Stress* significantly contributed to improvement in stress discrimination. These results support *Hypothesis 5: Hungarian learners of L2 Portuguese at the onset of learning focus first on the L2 suprasegmental features. Accordingly, stress precedes vowels in L2 acquisition*.

In order words, results suggest a similar learning sequence for L1 and L2, with a focus on prosodic features first, in order to extract the boundaries between words and thus begin vocabulary acquisition. If so, the differences between L1 and L2 acquisition seem to be fewer than previously thought, and the L1 learning mechanisms are also available for L2 acquisition in adult learners, as posited by the *full access hypothesis* (Flege 2005; Escudero & Boersma, 2004). However, the interpretation of these results must be taken with caution.

The effectiveness of the training regimen focusing on stress contrasts may be related to the design of the training itself. The six training sessions for Group *Stress* included only two contrasts (*1st-2nd syllable* and *2nd-3rd syllable* stress contrasts) whereas Group *Vowels* trained seven vowel contrasts.[105] As for Group *Vowels & Stress*, it combined three conditions of vowel reduction (*on the 1st syllable*, *on the 2nd syllable*, and *on both syllables*), with *1st-2nd syllable* contrast, and with several vowel contrasts. In other words, the training of Group *Stress* was less varied than the other two trainings. The fact that this group was asked to focus on fewer conditions may have led to its better results.

One must also mention the fact that training for Group *Stress* focused on a perceptual cue that exists in the L1 of the participants – duration. Vowel length is contrastive in Hungarian; therefore, Hungarian listeners use durational contrasts at the segmental level. Stimuli used in the training for Group *Stress* included contrasts of prominence achieved by modulation of duration and energy at suprasegmental level. One may suggest that learners in Group *Stress* had only to 'adjust' perception of durational contrasts from segmental to suprasegmental level.

Finally, one should also mention that the previous knowledge of L2 English by the participants of our study, mostly at B2 or C1 level, may have had an impact on the results obtained in Group *Stress*. English is also a lexically contrastive stress language, and learners may have transferred the word stress knowledge from English to EP. However, although English displays vowel reduction (thus, reduced vowels), neither [ɐ] nor [ɨ] are present in the English vowel inventory.

Therefore, the conclusion that suprasegmental precedes segmental features in L2 acquisition and perception must be approached with caution, since the results may have been influenced by the training design, language-specific features of the L1, and previous knowledge of L2 English.

In the next chapter we will present a summary of the studies conducted in this dissertation and discuss the implications of our findings to the development of materials for L2 Portuguese, as well as make propositions for future research.

[105] Furthermore, in Group *Stress* regimen, each contrast was trained in 96 AX trials and 48 AXB trials. As for Group *Vowels*, each contrast was trained in 24 AX trials and in 8 or 16 AXB trials.

Conclusion

In this dissertation, we investigated the acquisition of two features of EP – the reduced vowels [ɐ] and [ɨ], and word stress – by Hungarian learners.

As far as word accent is concerned, native Hungarian speakers exhibit stress 'deafness', that is, insensitivity in discriminating word accent contrasts (Dupoux et al., 2008; Honbolygó et al., 2017; Peperkamp & Dupoux, 2002). Regarding the perception of the two target vowels, to the best of our knowledge, no study on this topic has been carried out with native Hungarian speakers. Consequently, a first experiment (Chapter 4) – a forced-choice identification task with goodness-of-fit rating – was carried out to investigate this issue. The results showed that, in most cases, Hungarian listeners categorize the Portuguese vowel [ɐ] as /ɛ/ and the vowel [ɨ] as /y/ in their L1. The data collected also revealed that, other than the perceptual conflicts directly involving the target vowels, cases of perceptual overlap were also displayed between [ɛ] and [e], and between [e] and [i].

Considering the results of this experiment, as well as the stress 'deafness' attested in previous studies, we designed a perceptual training program (Chapter 5) with the aim of helping Hungarian learners of L2 Portuguese overcome their difficulties with the target features.

Contrary to previous studies, we aimed at observing changes in perception of L2 speech in the early stages of contact with the target language. To this purpose, we recruited 66 Hungarian learners of L2 Portuguese at the onset of learning, and conducted a six-week training program, preceded by a pretest and followed by a post-test. The training aimed at observing the effectiveness of (1) a training focused on vowel discrimination ([ɐ] and [ɨ]), (2) a training focused on word stress discrimination, and (3) a training that combined both vowels and stress. The training for vowels included monosyllabic pseudowords (ˈCV), thus with no stress modulations. The training for stress included trisyllabic pseudowords with the vowels [i] and [u] only, to exclude reduced vowels. These stimuli presented two possible stress contrasts: 1st vs. 2nd syllable (ˈCVCVCV-CVˈCVCV) or 2nd vs. 3rd syllable (CVˈCVCV-CVCVˈCV). The third training comprised dissyllabic pseudowords which included the same vowels of the first training and 1st vs. 2nd syllable stress contrasts (ˈCVCV-CVˈCV).

The three trainings were administered to three intervention groups: Group *Vowels* (n = 20), Group *Stress* (n = 22), and Group *Vowels & Stress* (n = 24), respectively.

Before and after the training, learners were tested by means of an oddity discrimination task. The training sessions consisted of AX and AXB tasks and were administered to participants once a week, tooking about 10 or 15 minutes each. In the post-test, participants were presented with the same trials as in the pretest and additional trials with stimuli produced by novel talkers. The training program – questionnaires, tests, and training sessions – was conducted online with Gorilla Experimental Builder.

The results showed that, overall, the groups which followed the training regimens *Vowels* or *Vowels & Stress* did not improve significantly, when compared to group *Stress*, in the EP vowel contrasts. A more detailed analysis showed improvement in the performance by the three groups in the contrast [ɐ]-[ɛ], and partly in the contrast [ɐ]-[ɨ]. The remaining contrasts displayed conflicting results. In the acquisition of word stress, both Group *Stress* and Group *Vowels & Stress* exhibited a significant improvement, but it was not generalized to novel talkers. Furthermore, Group *Stress* displayed better results than Group *Vowels & Stress* in both post-tests.

Based on the results found, the following conclusions can be drawn. First, learners do not display a similar learning path in the acquisition of the EP vowels, since they improved in the discrimination of some contrasts, but exhibited difficulties in the remaining contrasts. Specifically, they exhibited a significant improvement in [ɐ]-[ɛ], a contrast particularly difficult for Hungarian learners since they perceive both EP vowels as a sole Hungarian category: /ɛ/. Contrarily, contrasts between the mid-front vowels [ɛ] and [e], and between these and the close-central [ɨ] and the close-front [i] were problematic. Moreover, we did not find evidence of effectiveness of training focused on vowels.

Second, improvement in stress contrast discrimination suggests that stress ‘deafness’ is not persistent in Hungarian learners, and that with adequate perceptual training it can be overcome. Effectiveness of training was observed both for stress contrasts between 1st and 2nd syllables, as for stress contrasts between 2nd and 3rd syllables.

Third, a perceptual training that combines vowels and word stress is not more effective in the acquisition of these features than a training program that focuses on vowels and word stress separately. However, we did not find evidence that the training regimen that combined vowels and stress was significantly less effective than the training focusing on vowels, or than a training focusing on stress.

Lastly, although we found evidence that Hungarian learners at the onset of the L2 acquisition are able to benefit from training perception at a suprasegmental level (word stress) but not at a segmental level (vowels), suggesting that similar to L1, in L2 suprasegmental properties are first acquired, this conclusion must be interpreted with caution. The reason is twofold. First, the design of the training sessions may have been more adequate in the training focusing on stress perception than on the training focusing on vowel perception or on vowel and stress perception. Second, suprasegmental cues of EP stress involve an acoustic cue familiar to Hungarian learners – duration –, whereas the target segmental features – the vowels [ɐ] and [ɨ] – are absent from the Hungarian system.

In the next paragraphs, the implications of our findings to development of materials for L2 Portuguese will be discussed, considering the target features of this research: [ɐ] and [ɨ], and word stress.

Pedagogical implications

The perceptual study reported in this dissertation provided us some experience for future research, as well as for the design of didactic online material, to be implemented in the classroom or for autonomous learning for the L2 Portuguese learner. We must, however, point out some methodological issues to be taken into consideration in the design of such material.

The results for vowel discrimination suggest that the input received in the classroom promotes discrimination between the EP vowels [ɐ] and [ɛ] in Hungarian learners. We infer this from the fact that the three groups displayed a similar and considerable improvement for [ɐ]-[ɛ]. Recall that, at the onset of learning, both vowels are perceived by Hungarian learners as /ɛ/, and in order to split this category into two different categories, learners have to adjust their perception of the two vowels. Improvement in the discrimination of [ɐ] and [ɛ] must be combined with tasks that help learners to match [ɐ] with the category /a/, since this learning step is important to form a long-term category for [ɐ]. Observations in class showed that this learning step occurs late in EP acquisition by Hungarian learners, since they show a persistent tendency to transcribe this vowel as <e>, even in more advanced proficiency levels. To this purpose, we also included discrimination between [a] and [ɐ] in the training, which proved relatively successful. With respect to [ɨ], the case is not as straightforward as for [ɐ], since that vowel is very frequently deleted in production and, in classroom, learners seem to capture this regularity of the language easily. However, error rates in the post-test[N] for the contrast [e]-[ɨ] were significantly higher than in the pretest, which indicates that learners may have problems with this vowel when it is phonetically realized. In the case of Hungarian learners, we observed some perceptual overlap within [ɛ] and [e]. Acquisition of [ɨ] entails adjustment of the perceptual boundaries between these vowels, while learning the stressed-unstressed link that connects [ɛ]-[ɨ] and [e]-[ɨ]. To fulfill this goal, our perceptual training was not effective.

In what concerns the acquisition of word stress, results showed that perceptual training focusing on stress contrasts at suprasegmental level is effective to overcome stress 'deafness'. However, one must point out that EP stress induces vowel reduction. In other words, for Hungarian learners, acquiring EP phonology entails overcoming stress 'deafness' and acquiring reduced vowels that are a result of the absence of stress. The training regimen of Group *Vowels & Stress* was not successful in the discrimination of EP vowels contrasts, nor it was more successful than Group *Stress* in the discrimination of word stress. However, as previously suggested, these results may have been due to the complexity of the training design and stimuli presented to Group *Vowels & Stress*, compared to those presented to Group *Stress*.

Finally, it must be pointed out that our perceptual training study was original in observing learners at the onset of the L2 acquisition process. As previously explained, this entails advantages and disadvantages. Specifically, the onset of learning may be the ideal moment for perceptual tunning, since it is not constrained by morphosyntax or vocabulary knowledge (Best & Tyler, 2007). In contrast, the lack of knowledge in the L2 necessarily restricts the methodology of the training, since only a few types of tasks can be presented to the learners.

With the above reflections in mind, we make the following suggestions for the design of auditory perceptual tasks for didactic purposes, for Portuguese L2 speech acquisition:

(i) perceptual tasks can be administered from the onset of learning. However, considering that the learning path is not identical to all L2 features, that is, some speech features are more difficult to acquire than others, this fact should be taken into consideration when designing auditory exercises for different features. For example, Hungarian learners need to devote more time to the perceptual adjustment of EP front vowels than of EP central vowels.

(ii) Reduced vowels and word stress should be trained both separately and in combination. Considering that Hungarian learners were able to improve in word stress perception after six short training sessions, we suggest starting a training with this feature. Once acquainted with the variable nature of EP stress, learners can be trained in vowels combined with stress, to promote the link between full and reduced vowels and the presence or absence of stress.

(iii) Discrimination tasks should be combined with other types of tasks, so that the boost in perceptual tunning of phonetic features is combined with the creation of representational categories. This can be achieved in different ways. For example, using not only pseudowords but also real words, such as the ones used in Oliveira (2006; e.g., *p<u>e</u>dra* [ˈp<u>ɛ</u>dɾɐ] 'stone' vs. *p<u>e</u>drinha* [p<u>ɨ</u>ˈdɾiɲɐ] 'small stone'). Other than promoting the creation of lexically functional representations and the link stress-reduced vowels, the use of real words, with the possibility of expanding L2 vocabulary, may be more appealing to learners. Also, identification tasks have been shown to promote L2 speech features at representational level (Carlet, 2017). When learners already control at least some L2 vocabulary, this type of task can be introduced. Moreover, tasks should include multimodal information, with auditory stimuli complemented with visual information too, as advised by some researchers (Hazan et al., 2005; Kuhl et al., 2008).

The proposals above mentioned are evidence-based and aimed at providing relevant information for the development of online tasks that can help learners in L2 Portuguese phonological acquisition. Considering that the EP features investigated in this dissertation – the vowels [ɐ] and [ɨ], and word stress – are absent in languages other than Hungarian, such material could be used for learners of other L1. Crucially, more research in L2 Portuguese speech perception, specifically in the features mentioned above, would decisively contribute to more learning developments.

Considerations for future research

Before concluding this dissertation, one would like to present suggestions for future research. These come out as a result of an honest reflection on what can be done in the future, based on what has been done in the past. It is useful (and motivating) that some aspects remained unexplored or that some steps carried out in this research could have been taken differently.

As mentioned in the Introduction, research in L2 Portuguese phonology is still scarce (at least when compared to the productivity of studies on L2 Portuguese syntax, morphosyntax and didactics, for instance), and so it is the research on L2 with Hungarian learners. In order words, the list of topics which need to be or can be investigated is extensive. Therefore, the suggestions made here will be limited to matters related to the two studies reported in this dissertation.

In the forced-choice identification experiment reported in Chapter 4, the EP oral vowels were presented to the Hungarian listeners, which they identified among nine Hungarian vowels. However, although Hungarian vowel inventory presents nine vowel qualities, it includes 14 vowels in total: [ɒ], [aː], [ɛ], [eː], [i], [iː], [o], [oː], [ø], [øː], [u], [uː], [y], and [yː]. Considering that vowel length is contrastive in Hungarian, a similar experiment should be conducted presenting Hungarian participants with those 14 vowels as options for the identification of the EP vowels. Furthermore, considering the effect of consonant contrast in vowel perception (Bohn & Steinlen, 2003), EP stimuli should include monosyllables with different consonants too.

One of the constraints experienced in the experiments reported in this dissertation was the scarcity of studies in Portuguese phonetics and phonology concerning vowels and stress production and perception by Portuguese native speakers. For example, for measurements of the acoustic features of EP stress, the only data available are the one collected and analyzed by Delgado-Martins (1986). In our study, we collected two types of results. We observed perception of EP vowels and stress by Portuguese, as well as conducted measurements for the stimuli recorded with Portuguese speakers. Both data is worth to be further analyzed.

The perceptual training study reported in Chapter 5 involved a great number of variables related to stress and to vowel discrimination, and would need further analyses. The training of Group *Vowels & Stress* included not only the three conditions analyzed (*vowel reduction on the 1st syllable*, *vowel reduction on the 2nd syllable*, *vowel reduction on both syllables*), but also different L2 vowel contrasts, which were not analyzed individually.

Regardless of the limitations described above, the present dissertation provides an insight on different issues in L2 speech acquisition, such as the perception of non-native vowels and the effectiveness of perceptual training in learners at the onset of the L2 acquisition. It also adds empirical evidence supporting the idea that stress 'deafness' is not necessarily persistent, and that there appears to be a hierarchy in L2 acquisition between segmentals and suprasegmentals, suggesting a similarity between L1 and L2.

With the results of the studies reported in this dissertation, we provided empirical evidence that we hope can contribute to a better understanding of L2 speech perception and acquisition. Furthermore, we aimed at developing research in L2 Portuguese and provide evidence-based knowledge that helps designing scientifically attested auditory tasks for L2 teaching and learning.

Conclusão [Portuguese]

Na presente tese, investigámos a aquisição de duas características do português europeu (PE) – as vogais reduzidas [ɐ] e [ɨ], e o acento de palavra – por aprendentes húngaros.

No que diz respeito ao acento de palavra, os falantes nativos húngaros exibem 'surdez acentual', ou seja, insensibilidade na discriminação de contrastes de acento de palavra (Dupoux et al., 2008; Honbolygó et al., 2017; Peperkamp & Dupoux, 2002). Relativamente à perceção das duas vogais-alvo, tanto quanto sabemos, não foi realizado qualquer estudo neste tópico com falantes nativos de húngaro. Consequentemente, realizámos uma primeira experiência (Capítulo 4) – um teste de identificação (*forced-choice identification task with goodness-of-fit rating*) – para investigar esta questão. Os resultados mostraram que, na maioria dos casos, os ouvintes húngaros categorizam a vogal portuguesa [ɐ] como /ɛ/ e a vogal [ɨ] como /y/, na sua L1. Os dados recolhidos mostraram ainda outros casos de sobreposição percetiva (*perceptual overlap*): entre [ɛ] e [e], e entre [e] e [i].

Tendo em conta os resultados desta experiência, assim como a 'surdez acentual' atestada em estudos anteriores, desenhámos um treino percetivo (Capítulo 5) com o objetivo de auxiliar os aprendentes húngaros de Português L2 (PL2) a ultrapassar as dificuldades na perceção das vogais reduzidas [ɐ] e [ɨ], e do acento de palavra.

Ao contrário de estudos anteriores, o nosso objetivo foi observar as mudanças na perceção de fala em L2 nas fases iniciais de contacto com a língua-alvo. Para o efeito, recrutámos 66 aprendentes húngaros de PL2 no início do processo de aprendizagem e realizámos um programa de treino percetivo de seis semanas, precedido de um pré-teste e seguido de um pós-teste. O programa visou observar a eficácia de (1) um treino focado na discriminação de vogais, (2) um treino focado na discriminação do acento de palavra, e (3) um treino que combinava vogais e acento de palavra. Para investigar estas questões, desenhámos três intervenções, conduzidas paralelamente em três grupos. A primeira, focada nas vogais orais do PE, incluiu monossílabos (ˈCV), ou seja, estímulos sem contrastes de acento de palavra. Os estímulos consistiram nas vogais vogais-alvo [ɐ] e [ɨ], bem como [ɛ], [e] e [i], na perceção das quais os ouvintes húngaros apresentam sobreposição percetiva. A segunda intervenção incluiu pseudopalavras trissilábicas (CVCVCV), exclusivamente com as vogais [i] e [u], as únicas vogais portuguesas que não sofrem redução em contexto átono. Estas pseudopalavras apresentavam dois contrastes de acento: 1ª vs. 2ª sílaba (ˈCVCVCV- CVˈCVCV), e 2ª vs. 3ª sílaba (CVˈCVCV- CVCVˈCV). A terceira intervenção incluiu pseudopalavras dissilábicas. Este estímulo incluiu as mesmas vogais presentes na primeira intervenção, e contrastes de acento entre a 1ª e a 2ª sílaba (ˈCVCV- CVˈCV).

Os participantes foram divididos em três grupos, um para cada intervenção, respetivamente: Grupo *Vowels* (n = 20), Grupo *Stress* (n = 22) e Grupo *Vowels & Stress* (n = 24). Antes e após o treino, testámos os participantes em discriminação de contrastes de vogais ([ɐ]-[ɛ], [ɐ]-[ɨ], [ɛ]-[e], [e]-[i], [a]-[ɐ], [ɛ]-[ɨ] e [e]-[ɨ]) e de acento de palavra (1ª vs. 2ª sílaba e 2ª vs. 3ª sílaba). O pré-teste e o pós-teste incluíram testes de intruso (*oddity tasks*). O pós-teste consistiu na repetição dos exercícios do pré-teste, assim como exercícios adicionais com novos estímulos produzidos por novos falantes. As sessões de treino consistiram em tarefas de discriminação auditiva AX e AXB. As intervenções foram administradas durante seis sessões, uma sessão por semana, com a duração de cerca de 10 ou 15 minutos cada. Todas as atividades do estudo foram criadas e apresentadas aos participantes em formato *online* com o programa Gorilla Experiment Builder (Anwyl-Irvine et al., 2018).

Os resultados mostraram que, no geral, os grupos que seguiram os regimes de treino *Vowels* ou *Vowels & Stress* não melhoraram significativamente, quando comparados com o Grupo *Stress*, nos contrastes de vogais do PE. Uma análise mais detalhada mostrou uma melhoria no desempenho dos três grupos no contraste [ɐ]-[ɛ], e parcialmente no contraste [ɐ]-[ɨ]. Os restantes contrastes apresentaram resultados contraditórios. Na aquisição do acento de palavra, tanto o Grupo *Stress* como o Grupo *Vowels & Stress* apresentaram uma melhoria significativa, comparando com o Grupo *Vowels*, mas esta não foi generalizada para os novos estímulos produzidos por novos falantes. É de referir também que o Grupo *Stress* apresentou melhores resultados do que o Grupo *Vowels & Stress* em ambos os pós-testes.

Com base nos resultados encontrados, apresentamos as seguintes conclusões.

Em primeiro lugar, os aprendentes não exibem uma trajetória de aprendizagem semelhante na aquisição das vogais do PE, uma vez que melhoraram na discriminação de alguns contrastes, mas exibiram dificuldades nos restantes contrastes. Especificamente, os aprendentes melhoraram significativamente na discriminação de [ɐ]-[ɛ], um contraste particularmente difícil, uma vez que os falantes húngaros percecionam ambas as vogais do PE como uma única categoria húngara: /ɛ/. Pelo contrário, os contrastes entre as vogais [ɛ] e [e], e entre estas e [ɨ] ou [i] revelaram-se problemáticos para os aprendentes. De referir ainda que não encontrámos evidência de eficácia de um treino percetivo focado em contrastes vocálicos.

Em segundo lugar, os grupos cujo treino focou a discriminação de contrastes de acento de palavra – Grupo *Stress* e Grupo *Vowels & Stress* – revelou-se eficaz. A eficácia do treino foi observada, no Grupo *Stress*, tanto para contrastes de acento entre a 1ª e a 2ª sílabas, como para contrastes entre a 2ª e a 3ª sílabas. Este resultado sugere que a 'surdez' na discriminação de contrastes de acento de palavra não é persistente nos aprendentes húngaros e que, com um treino percetivo adequado, pode ser ultrapassada.

Em terceiro lugar, um treino percetivo que combine vogais e acento de palavra (regime do Grupo *Vowels & Stress*) não é mais eficaz na aquisição destas características do que um programa de treino que se concentre em vogais e no acento de palavra separadamente. No entanto, não encontrámos evidências de que o regime de treino que combinou vogais e acento foi significativamente menos eficaz do que o treino centrado nas vogais, ou do que um treino centrado no acento de palavra.

Por último, embora os resultados indiquem que os aprendentes húngaros no estágio inicial de aprendizagem são capazes de beneficiar de treino percetivo a nível suprassegmental (acento de palavra), mas não a nível segmental (vogais), sugerindo que, à semelhança da L1, na L2 as propriedades suprassegmentais são adquiridas em primeiro lugar, esta conclusão deve ser interpretada com cautela. Em primeiro lugar, o desenho das sessões de treino pode ter sido mais adequado no regime de treino centrado na perceção do acento do que no regime de treino centrado na perceção das vogais ou na perceção das vogais e do acento. Em segundo lugar, as características suprassegmentais do acento de palavra do PE envolvem pistas acústicas familiares aos aprendentes húngaros – duração e energia –, ao passo que as características segmentais alvo – as vogais [ɐ] e [ɨ] – estão ausentes do sistema húngaro.

Nos parágrafos seguintes, serão discutidas as implicações dos resultados para o desenvolvimento de materiais para o PL2, considerando os traços-alvo desta pesquisa: [ɐ] e [ɨ], e o acento de palavra.

Implicações pedagógicas

O estudo percetivo relatado nesta tese proporcionou-nos alguma experiência para futuras investigações, bem como para a conceção de material didático online a ser implementado na sala de aula ou para a aprendizagem autónoma do aprendente de PL2. Devemos, no entanto, apontar algumas questões metodológicas a serem levadas em consideração na elaboração desse material.

Os resultados relativos à discriminação de vogais sugerem que o input recebido em sala de aula promove a discriminação entre as vogais do PE [ɐ] e [ɛ] em aprendentes de húngaro. Inferimos isto a partir de o facto de os três grupos terem apresentado uma melhoria semelhante para [ɐ]-[ɛ], mesmo o Grupo *Stress*, cujo treino não incluía estas vogais. Recorde-se que, no início da aprendizagem, ambas as vogais são percecionadas pelos aprendentes húngaros como /ɛ/, e, para dividir esta categoria em duas, terão de ajustar a sua perceção das duas vogais. A melhoria na discriminação de [ɐ]-[ɛ] deve ser combinada com tarefas que ajudem os aprendentes a fazer a correspondência entre o fone [ɐ] e o fonema /a/, uma vez que este passo de aprendizagem é importante para formar uma categoria de longo prazo para [ɐ]. Observações em sala de aula mostram que esta etapa ocorre tardiamente na aquisição do PE pelos aprendizes húngaros, uma vez que estes mostram uma tendência persistente em transcrever essa vogal como <e>, mesmo em níveis de proficiência mais avançados. Para promover a relação [a]-[ɐ], incluímos também no treino este contraste, o que se revelou relativamente bem-sucedido. Relativamente a [ɨ], o caso não é tão linear como o de [ɐ], uma vez que esta vogal é frequentemente apagada em produção e, em sala de aula, os aprendentes parecem captar facilmente esta regularidade da língua. No entanto, as taxas de erro nos exercícios do pós-teste com novos estímulos/falantes para o contraste [e]-[ɨ] foram significativamente mais elevadas do que no pré-teste, o que indica que os aprendentes podem ter problemas com esta vogal quando realizada foneticamente. No caso dos falantes de húngaro, na experiência de identificação foi observada alguma sobreposição percetiva entre [ɛ] e [e]. A aquisição de [ɨ] implica o ajustamento das fronteiras percetivas entre [ɛ] e [e], e paralelamente a aquisição das relações *tónico-átono* [ɛ]-[ɨ] e [e]-[ɨ]. Para tal, estes contrastes também foram incluídos no treino de vogais, à semelhança de [a]-[ɐ]. No entanto, neste caso o treino percetivo não se revelou eficaz.

No que diz respeito à aquisição do acento de palavra, os resultados mostraram que um treino percetivo centrado em contrastes de acento a nível suprassegmental é eficaz para ultrapassar a 'surdez acentual' inicialmente apresentada pelos aprendentes húngaros. No entanto, é preciso salientar que o acento do PE induz redução vocálica. Por outras palavras, para estes aprendentes, a aquisição da fonologia do PE implica que, além de terem de superar a 'surdez acentual', é necessário adquirirem as vogais reduzidas que resultam da ausência de acento. O regime de treino do Grupo *Vowels & Stress* não foi bem-sucedido na discriminação dos contrastes das vogais do PE, nem foi mais eficaz do que o Grupo *Stress* na discriminação do acento de palavra. No entanto, como sugerido no capítulo "General Discussion", esses resultados podem ter sido devidos à complexidade do desenho do treino e dos estímulos apresentados ao Grupo *Vowels & Stress*, em comparação com os apresentados ao Grupo *Stress*.

Por último, é de salientar que o nosso estudo de treino percetivo foi original ao observar aprendentes no início do processo de aquisição de uma L2. Como explicámos anteriormente, este facto apresenta vantagens e desvantagens. Especificamente, o início da aprendizagem pode ser o momento ideal para a afinação percetiva (*perceptual tunning*), uma vez que não é restringido por conhecimentos de morfossintaxe ou de vocabulário (Best & Tyler, 2007). Em contrapartida, a falta de domínio na L2 limita necessariamente a metodologia do treino, uma vez que apenas alguns tipos de tarefas podem ser apresentados aos aprendentes.

Tendo em conta as reflexões anteriores, apresentamos as seguintes sugestões para a conceção de tarefas percetivas auditivas com fins didáticos, para a aquisição da perceção de fala em PL2:

(i) as tarefas percetivas podem ser administradas desde o início da aprendizagem. No entanto, tendo em conta que o percurso de aprendizagem não é idêntico para todas as características da L2, ou seja, algumas características são mais difíceis de adquirir do que outras, este facto deve ser tido em consideração na elaboração de exercícios auditivos. Por exemplo, os aprendentes de húngaro precisam de dedicar mais tempo ao ajuste percetivo das vogais anteriores do que das vogais centrais do PE.

(ii) As vogais reduzidas e o acento de palavra devem ser treinados separadamente e em combinação. Tendo em conta que os aprendentes húngaros conseguiram melhorar a perceção do acento de palavra a nível suprassegmental após seis curtas sessões de treino, sugerimos que se inicie um treino focado nesta caraterística. Quando os aprendentes estiverem familiarizados com a natureza variável do acento do PE, podem ser treinados em vogais combinadas com acento, para promover a relação entre vogais tónicas e átonas e a presença ou ausência de acento de palavra.

(iii) Os exercícios de discriminação auditiva devem ser combinados com outros tipos de tarefas, de modo a promover não só a afinação percetiva das características fonéticas, mas também a criação de categorias representacionais. Tal objetivo pode ser conseguido de diferentes formas, por exemplo, recorrendo não só a pseudopalavras mas também a palavras reais, como as utilizadas em Oliveira (2006; e.g., *pedra* [ˈpɛdrɐ] vs. *pedrinha* [pɨˈdriɲɐ]). Para além de promover a criação de representações lexicalmente funcionais e a relação entre vogais tónicas e átonas, o uso de palavras reais, com a possibilidade de expandir o vocabulário da L2, pode ser mais apelativo para os aprendentes. Além disso, estudos anteriores demonstram que tarefas de identificação promovem a aquisição fonológica da L2 a nível representacional (Carlet, 2017). Quando os aprendentes já controlarem algum vocabulário da L2, este tipo de tarefa pode ser introduzido. Por último, as tarefas devem incluir informação multimodal, com estímulos auditivos complementados com informação visual, tal como recomendado por alguns investigadores (Hazan et al., 2005; Kuhl et al., 2008).

As propostas acima mencionadas são baseadas em evidências empíricas e têm como objetivo fornecer informações relevantes para o desenvolvimento de tarefas online que possam auxiliar os aprendentes na aquisição fonológica do PL2. Tendo em conta que as características do PE investigadas nesta tese – as vogais [ɐ] e [ɨ], e o acento de palavra – estão ausentes noutras línguas para além do húngaro, este material poderá ser utilizado por aprendentes de outras L1. Mais investigação sobre a perceção da fala em PL2, especificamente sobre as características acima mencionadas contribuiria decisivamente para mais desenvolvimentos no ensino/aprendizagem.

Considerações para futura investigação

Antes de concluir esta tese, gostaríamos de apresentar sugestões para investigação futura. Estas surgem como resultado de uma reflexão honesta sobre o que pode ser feito no futuro, com base no que foi feito no passado. É útil (e motivador) que alguns aspetos tenham ficado por explorar ou que alguns passos dados nesta investigação pudessem ter sido dados de forma diferente.

Como foi referido na Introdução, a investigação em fonologia do PL2 é ainda escassa (quando comparada com a produtividade dos estudos sobre sintaxe, morfossintaxe e didática do PL2, por exemplo), e o mesmo acontece com a investigação em L2 com aprendentes húngaros. Por outras palavras, a lista de tópicos que precisam de ser ou podem ser investigados é extensa. Tendo este facto em consideração, as sugestões aqui apresentadas limitar-se-ão a questões relacionadas com os dois estudos relatados nesta tese.

Na experiência de identificação relatada no Capítulo 4, as vogais orais do PE foram apresentadas aos ouvintes húngaros, que as identificaram entre nove vogais húngaras. No entanto, apesar de o inventário de vogais do húngaro apresentar nove qualidades vocálicas, inclui 14 vogais no total: [ɒ], [aː], [ɛ], [eː], [i], [iː], [o], [oː], [ø], [øː], [u], [uː], [y] e [yː]. Considerando que a duração vocálica é contrastiva em húngaro, uma experiência semelhante deveria ser realizada apresentando aos participantes húngaros as 14 vogais da L1 como opções para a identificação das vogais do PE. Além disso, tendo em conta o efeito do contexto consonântico na perceção de vogais (Bohn & Steinlen, 2003), os estímulos auditivos deveriam incluir diferentes consoantes.

Um dos constrangimentos verificados nas experiências relatadas nesta tese foi a escassez de estudos em fonética e fonologia do PE relativos à produção e perceção das vogais e do acento de palavra por falantes nativos do português. Por exemplo, para a medição das características acústicas do acento de palavra em PE, os únicos dados disponíveis são os recolhidos e analisados por Delgado-Martins (1986). No nosso estudo, recolhemos dois tipos de resultados: observámos a perceção das vogais orais e do acento de palavra do PE por falantes nativos e realizámos medições para os estímulos gravados com falantes de português. Ambos os dados merecem uma análise mais aprofundada.

O estudo de treino percetivo relatado no Capítulo 5 envolveu um grande número de variáveis relacionadas com o acento de palavra e com a discriminação de vogais, e necessitaria de mais análises estatísticas. A título de exemplo: o treino do Grupo *Vowels & Stress* incluiu não só as três condições analisadas (redução vocálica na 1ª sílaba, redução vocálica na 2ª sílaba, redução vocálica em ambas as sílabas), mas também diferentes contrastes de vogais da L2, que não foram analisados individualmente.

Apesar das limitações descritas acima, a presente tese fornece uma visão sobre diferentes questões na aquisição fonológica em L2, tais como a perceção de vogais não nativas e a eficácia do treino percetivo em aprendizes no início da aquisição da L2. Também acrescenta evidências empíricas que apoiam a ideia de que a 'surdez acentual' não é necessariamente persistente, e que parece haver uma hierarquia na aquisição da L2 entre características segmentais e suprassegmentais, sugerindo uma semelhança entre a L1 e a L2.

Com os resultados dos estudos relatados nesta tese, fornecemos evidências empíricas que esperamos vir a contribuir para uma melhor compreensão da perceção e aquisição da fala em L2. Além disso, esperamos também promover a investigação em PL2 e fornecer conhecimentos baseados em evidências que ajudem a conceber tarefas auditivas cientificamente comprovadas para o ensino e a aprendizagem de L2.

References

Abrahamsson, N., & Hyltenstam, K. (2009). Age of Onset and Nativelikeness in a Second Language: Listener Perception Versus Linguistic Scrutiny. *Language Learning*, *59*(2), 249–306. https://doi.org/https://doi.org/10.1111/j.1467-9922.2009.00507.x

Akahane-Yamada, R., Tohkura, Y., Bradlow, A. R., & Pisoni, D. B. (1996). Does training in speech perception modify speech production? *Proceedings of the 4th International Conference on Spoken Language Processing (ICSLP 1996)*, 606–609. https://doi.org/10.21437/icslp.1996-153

Aliaga-García, C. (2017). *The effect of auditory and articulatory phonetic training on the perception and production of L2 vowels by Catalan-Spanish learners of English* [Unpublished doctoral dissertation]. University of Barcelona.

Aliaga-García, C., & Mora, J. C. (2009). Assessing the effects of phonetic training on L2 sound perception and production. In M. Watkins, A. Rauber, & B. Baptista (Eds.), *Recent Research in Second Language Phonetics/Phonology: Perception and Production* (pp. 2–31). Cambridge Scholars Publishing.

Alispahic, S., Mulak, K. E., & Escudero, P. (2017). Acoustic properties predict perception of unfamiliar dutch vowels by adult australian english and Peruvian Spanish listeners. *Frontiers in Psychology*, *8*, Article 52. https://doi.org/10.3389/fpsyg.2017.00052

Allegro, F. R. P. (2014). *Ensino de Pronúncia em Português Língua Estrangeira: análise de livros didáticos Ensino de Pronúncia em Português Língua Estrangeira: análise de livros didáticos* [Unpublished doctoral dissertation]. Pontifícia Universidade Católica de São Paulo.

Alves, U. K., & Luchini, P. L. (2017). Effects of perceptual training on the identification and production of word-initial voiceless stops by Argentinean learners of English. *Ilha Do Desterro*, *70*(3), 15–32. https://doi.org/10.5007/2175-8026.2017v70n3p15

Andrade, A. (2020). Vocalismo. In Eduardo B. P. Raposo, M. F. B. do Nascimento, M. A. C. da Mota, L. Segura, A. Mendes, & A. Andrade (Eds.), *Gramática do Português - Volume III* (pp. 3241–3330). Fundação Calouste Gulbenkian.

Anwyl-Irvine, A., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. (2018). Gorilla in our Midst: An online behavioral experiment builder. *Behavior Research Methods*, 1–20. https://doi.org/10.3758/s13428-019–01237-x

Audacity Team. (2020). *Audacity(R): Free Audio Editor and Recorder* [Computer application]. *Version 2.4.2*. https://audacityteam.org/

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language speech research* (pp. 171–204). York Press.

Best, C. T., & Strange, W. (1992). Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics*, *20*(3), 305–330. https://doi.org/10.1016/s0095-4470(19)30637-0

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception. Commonalities and complementaries. In M. J. Munro & O.-S. Bohn (Eds.), *Language Experience in Second Language Speech Learning: In honor of James Emil Flege* (pp. 13–34). John Benjamins.

Bettoni-Techio, M. (2008). *Perceptual training and word-initial /s/-clusters in Brazilian Portuguese/English interphonology* [Unpublished doctoral dissertation]. Universidade Federal de Santa Catarina.

Boersma, P., & Weenink, D. (2020). *Praat: doing phonetics by computer* [Computer program]. *Version 6.1.16*. http://www.praat.org/

Bohn, O.-S. (1995). Cross-language speech perception in adults. First language transfer doesn't tell it all. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross Language Research* (pp. 275–300). York Press.

Bohn, O.-S., & Bundgaard-Nielsen, R. L. (2008). Second Language Speech Learning with Diverse Inputs. In T. Piske & M. Young-Scholten (Eds.), *Input Matters in SLA* (pp. 207–218). Multilingual Matters. https://doi.org/10.21832/9781847691118-014

Bohn, O.-S., & Flege, J. E. (1990). Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics*, *11*(3), 303–328. https://doi.org/https://doi.org/10.1017/S0142716400008912

Bohn, O.-S., & Steinlen, A. (2003). Consonantal Context affects Cross-Language Perception of Vowels. *Proceedings of the XVth International Congress of Phonetic Sciences*, *August*, 2289–2292. http://www.hum.au.dk/engelsk/engosb/pdf/ICPhS226bostei.pdf

Borgwaldt, S. R., Hellwig, F. M., & Groot, A. M. B. De. (2004). Word-initial entropy in five languages: Letter to sound, and sound to letter. *Written Language & Literacy,7*(2), 165–184. https://doi.org/10.1075/wll.7.2.03bor

Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., & Tohkura, Y. (1999). Training Japanese listeners to identify english /r/and /1/: Long-term retention of learning in perception and production. *Perception and Psychophysics*, *61*(5), 977–985. https://doi.org/10.3758/BF03206911

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. (1997). Training Japanese listeners to identify English / r / and / l /: IV. Some effects of perceptual learning on speech production. *The Journal of the Acoustical Society of America*, *101*(4), 2299–2310. https://doi.org/10.1121/1.418276

Braun, B., Lemhöfer, K., & Cutler, A. (2008). English word stress as produced by English and Dutch speakers: The role of segmental and suprasegmental differences. *Proceedings of the 9th Annual Conference of the International Speech Communication Association (INTERSPEECH 2008), Brisbane*, 1953.

Brawerman-Albini, A. (2012). *Os efeitos de um treinamento de percepção na aquisição de pré-proparoxítonas inglesas por brasileiros* [Unpublished doctoral dissertation]. Universidade Federal do Paraná.

Brawerman-Albini, A., Kluge, D. C., Silva, A. H. P., & Albini, L. C. P. (2017). Perceptual Training Effects on the Acquisition of English Stress by Brazilian Learners. *International Journal of English Linguistics*, *7*(6), 1–15. https://doi.org/10.5539/ijel.v7n6p1

Brekelmans, G., Lavan, N., Saito, H., Clayards, M., & Wonnacott, E. (2022). Does high variability training improve the learning of non-native phoneme contrasts over low variability training? A replication. *Journal of Memory and Language*, *126*, Article 104352. https://doi.org/10.1016/j.jml.2022.104352

Brown, C. (2000). The Interrelation between Speech Perception and Phonological Acquisition from Infant to Adult. In J. Archibald (Ed.), *Second Language Acquisition and Linguistic Theory* (pp. 4–63). Blackwell.

Bundgaard-Nielsen, R. L., Best, C. T., & Tyler, M. D. (2011). Vocabulary size is associated with second-language vowel perception performance in adult learners. *Studies in Second Language Acquisition*, *33*, 433–461. https://doi.org/10.1017/S0272263111000040

Camões – Instituto da Cooperação e da Lingua (2016a). Aprender no Ensino Superior. https://www.instituto-camoes.pt/activity/o-que-fazemos/aprender-portugues/presencial/no-ensino-superior

Camões – Instituto da Cooperação e da Lingua (2016b). *Referencial Camões PLE*. https://www.instituto-camoes.pt/activity/centro-virtual/referencial-camoes-ple

Carlet, A. (2017). *L2 perception and production of English consonants and vowels by Catalan speakers: The effects of attention and training task in a cross-training study* [Unpublished doctoral dissertation]. Universitat Autònoma de Barcelona.

Castelo, A. (2005). The perception of word primary stress by European Portuguese speakers. In S. Frota, M. Vigário, & M. J. Freitas (Eds.), *Prosodies. With Special Reference to Iberian Languages* (pp. 159–176). De Gruyter Mouton. https://doi.org/10.1515/9783110197587.1.159

Castelo, A., & Santos, R. N. (2017). As vogais do português entre aprendentes chineses e suas implicações no desenvolvimento de um programa de português. In *Português como língua estrangeira, de herança e materna: abordagens, contextos e práticas* (pp. 123–136). Boavista Press.

Cebrian, J. (2006). Experience and the use of non-native duration in L2 vowel categorization. *Journal of Phonetics*, *34*(3), 372–387. https://doi.org/10.1016/j.wocn.2005.08.003

Cebrian, J. (2010). Cross-linguistic perception of Catalan and English diphthongs. In K. Dziubalska-Kołaczyk, M. Wrembel, & M. Kul (Eds.), *Proceedings of the 6th International Symposium on the Acquisition of Second Language Speech (New Sounds 2010), Poznań,* 73–76.

Chan, A. Y. W. (2014). The perception and production of English speech sounds by Cantonese ESL learners in Hong Kong. *Linguistics*, *52*(1), 35–72. https://doi.org/10.1515/ling-2013-0056

Chao, S. C., Ochoa, D., & Daliri, A. (2019). Production variability and categorical perception of vowels are strongly linked. *Frontiers in Human Neuroscience*, *13,* Article 96. https://doi.org/10.3389/fnhum.2019.00096

Chládková, K., & Podlipský, V. J. (2011). Native dialect matters: Perceptual assimilation of Dutch vowels by Czech listeners. *The Journal of the Acoustical Society of America*, *130*(4), EL186–EL192. https://doi.org/10.1121/1.3629135

Chrabaszcz, A., Winn, M., Lin, C. Y., & Idsardia, W. J. (2014). Acoustic Cues to Perception of Word Stress by English, Mandarin, and Russian Speakers. *Journal of Speech, Language, and Hearing Research*, *57*(4), 1468–1479. https://doi.org/10.1044/2014_JSLHR-L-13-0279

Colantoni, L., Steele, J., & Escudero, P. (2015). *Second Language Speech: Theory and Practice*. Cambridge University Press.

Conselho da Europa (2001). Quadro Europeu Comum de Referência para as Línguas – Aprendizagem, ensino, avaliação [Portuguese edition of the Common European Framework of Reference for languages: Learning, Teaching, Assessment (QECR). Council of Europe, 2001]. Edições Asa.

Cooper, N., Cutler, A., & Wales, R. (2002). Constraints of lexical stress on lexical access in English: Evidence from native and non-native listeners. *Language and Speech*, *45*(3), 207–228. https://doi.org/10.1177/00238309020450030101

Correia, S., Butler, J., Vigário, M., & Frota, S. (2015). A stress "deafness" effect in European Portuguese. *Language and Speech*, *58*(1), 48–67. https://doi.org/10.1177/0023830914565193

Correia, S., Ge, Y., Dinis Fernandes, J., Rato, A., Kachlicka, M., Saito, K., & Rebuschat, P. (2023, July 13–14). *Effects of perceptual training and cognitive aptitudes in the development of non-native speech perception abilities: Evidence from L2 Portuguese* [Conference presentation]. L2 Portuguese @ Lancaster (L2P@L 2023), U.K. https://www.lancaster.ac.uk/media/lancaster-university/content-assets/documents/heritage-language/L2PL2023programv3July12.pdf

Council of Europe (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*. Council of Europe Publishing, http://www.coe.int/lang-cefr

Couper-Kuhlen, E. (1986). *An Introduction to English Prosody*. Edward Arnold.

Cruz-Ferreira, M. (1999). Portuguese (European). *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet*. Cambridge University Press.

d'Andrade, E., & Viana, M. do C. (1989). Ainda sobre o acento e o ritmo em português. *Atas do 4º Encontro Nacional da Associação Portuguese de Linguística, Lisboa*, 3–16.

Delgado-Martins, M. R. (1986). *Sept etudes sur la perception*. Instituto Nacional de Investigação Científica.

Deme, A. (2014). Intelligibility of sung vowels: The effect of consonantal context and the onset of voicing. *Journal of Voice*, *28*(4), 523.e19-523.e25. https://doi.org/10.1016/j.jvoice.2014.01.003

Derwing, T. M., & Munro, M. J. (2005). Second Language Accent and Pronunciation Teaching: A Research-Based Approach. *TESOL Quarterly*, *39*(3), 379–397. https://doi.org/10.2307/3588485

Dörnyei, Z. (2014). The Psychology of the Language Learner. Individual Differences in Second Language Acquisition. In S. M. Gass, J. Schachter, & A. Mackey (Eds.), *The Psychology of the Language Learner*. Routledge. https://doi.org/10.1006/jmla.1996.2500

Dupoux, E., Peperkamp, S., & Sebastián-Gallés, N. (2001). A robust method to study stress "deafness." *Journal of Acoustical Society of America*, *110*(3), 1606–1618. https://doi.org/10.1121/1.1380437

Dupoux, E., Sebastián-Gallés, N., Navarrete, E., & Peperkamp, S. (2008). Persistent stress "deafness": The case of French learners of Spanish. *Cognition*, *106*(2), 682–706. https://doi.org/10.1016/j.cognition.2007.04.001

Dupoux, E., Sebastian, N., Pallier, C., & Mehler, J. (1997). A Destressing "Deafness" in French? *Journal of Memory and Language*, *36*(3), 406–421.

Eckman, F. R. (2008). Typological markedness and second language phonology. In J. G. H. Edwards & M. L. Zampini (Eds.), *Phonology and Second Language Acquisition* (pp. 95–115). John Benjamins. https://doi.org/10.1075/sibil.36.06eck

Edwards, J. G. H., & Zampini, M. L. (2008). Introduction. In Jette G. Hansend Edwards & M. L. Zampini (Eds.), *Phonology and Second Language Acquisition* (pp. 1–11). John Benjamins. https://doi.org/10.1075/sibil.36

Elvin, J., Escudero, P., & Vasiliev, P. (2014). Spanish is better than English for discriminating Portuguese vowels: Acoustic similarity versus vowel inventory. *Frontiers in Psychology*, *5,* Article 1188. https://doi.org/10.3389/fpsyg.2014.01188

Elvin, J., Williams, D., Shaw, J. A., Best, C. T., & Escudero, P. (2021). The role of acoustic similarity and non-native categorisation in predicting non-native discrimination: Brazilian Portuguese vowels by English vs. Spanish listeners. *Languages*, *6*(1), Article 44. https://doi.org/10.3390/languages6010044

Escudero, P. (2005). *Linguistic Perception and Second Language Acquisition: Explaining the attainment of optimal phonological categorization* (Publication No. 113) [Doctoral dissertation, Ultrecht University]. *LOT, Dissertation Series*.

Escudero, P., & Boersma, P. (2004). Bridging the Gap Between L2 Speech Perception Research and Phonological Theory. *Studies in Second Language Acquisition*, *26*(4), 551–585. https://doi.org/10.1017/S0272263104040021

Escudero, P., & Boersma, P. (2002). The subset problem in L2 perceptual development: Multiple-category assimilation by Dutch learners of Spanish. In B. Skarabela, S. Fish, & A. H.-J. Do (Eds.), *Proceedings of the 26th Annual Boston University Conference on Language Development*, 208–219. Cascadilla Press.

Escudero, P., Boersma, P., Rauber, A. S., & Bion, R. A. H. (2009). A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. *The Journal of the Acoustical Society of America*, *126*(3), 1379–1393. https://doi.org/10.1121/1.3180321

Escudero, P., & Vasiliev, P. (2011). Cross-language acoustic similarity predicts perceptual assimilation of Canadian English and Canadian French vowels. *The Journal of the Acoustical Society of America*, *130*(5), EL277–EL283. https://doi.org/10.1121/1.3632043

Escudero, P., & Wanrooij, K. (2010). The Effect of L1 Orthography on Non-native Vowel Perception. *Language and Speech*, *53*(3), 343–365. https://doi.org/10.1177/0023830910371447

Faris, M. M., Best, C. T., & Tyler, M. D. (2016). An examination of the different ways that non-native phones may be perceptually assimilated as uncategorized. *The Journal of the Acoustical Society of America*, *139*(1), EL1–EL5. https://doi.org/10.1121/1.4939608

Faris, M. M., Best, C. T., & Tyler, M. D. (2018). Discrimination of uncategorised non-native vowel contrasts is modulated by perceived overlap with native phonological categories. *Journal of Phonetics*, *70*, 1–19. https://doi.org/10.1016/j.wocn.2018.05.003

Field, J. (2005). Intelligibility and the Listener: The Role of Lexical Stress. *TESOL Quarterly*, *39*(3), 399–424. https://doi.org/https://doi.org/10.2307/3588485

Fitch, W. T. (1997). Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *Journal of Acoustical Society of America*, *102*(2), 1213–1222. https://doi.org/10.1121/1.421048

Flege, J. E. (1987). Effects of equivalence classification on the production of foreign language speech sounds. In A. James & J. Leather (Eds.), *Sound patterns in second language acquisition* (pp. 9–40). De Gruyter Mouton. https://doi.org/10.1515/9783110878486-003

Flege, J. E. (1988). Factors affecting degree of perceived foreign accent in English sentences. *Journal of the Acoustical Society of America*, *84*(1), 70–79. https://doi.org/10.1121/1.396876

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–276). York Press.

Flege, J. E. (1999). The Relation between Production and Distribution. In J. J. Ohala, Y. Hasegawa, M. Ohala, G. Daniel, & A. C. Bailey (Eds.), *Proceedings of the XVIth International Congress of Phonetics Sciences, San Francisco, 2*, 1273–1276. https://doi.org/10.1007/978-1-349-01255-8_12

Flege, J. E. (2008). Give Input a Chance! In T. Piske & M. Young-Scholten (Eds.), *Input Matters in SLA* (pp. 175–190). Multilingual Matters. https://doi.org/10.21832/9781847691118-012

Flege, J. E. (2018). It's input that matters most, not age. *Bilingualism: Language and Cognition*, *21*(5), 919–920. https://doi.org/10.1017/S136672891800010X

Flege, J. E., & Bohn, O. (2021). The Revised Speech Learning Model (SLM-r). In R. Wayland (Ed.), *Second Language Speech Learning. Theoretical and Empirical Progress*. Cambridge University Press. https://doi.org/10.1017/9781108886901.002

Flege, J. E., Fox, R. A., & Munro, M. J. (1994). Auditory and categorical effects on cross-language vowel perception. *Journal of the Acoustical Society of America*, *95*(6), 3623–3641. https://doi.org/10.1121/1.409931

Flege, J. E., & Mackay, I. R. A. (2004). Perceiving vowels in a second language. *Studies in Second Language Acquisition*, *26*(1), 1–34. https://doi.org/10.1017/S0272263104026117

Flege, J. E., MacKay, I. R. A., & Meador, D. (1999). Native Italian speakers' perception and production of English vowels. *The Journal of the Acoustical Society of America*, *106*(5), 2973–2987. https://doi.org/10.1121/1.428116

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age Constraints on Second-Language Acquisition. *Journal of Memory and Language*, *41*(1), 78–104. https://doi.org/10.1006/jmla.1999.2638

Fónagy, I. (1958). A hangsúlyról. In P. Fábián (Ed.), *Nyelvtudományi Értekezések* (18). Akadémiai Kiadó.

Fowler, C. A. (1986). An event approach to the study of speech perception from a direct-realist perspective. *Journal of Phonetics*, *14*(1), 3–28. https://doi.org/10.1016/s0095-4470(19)30607-2

Frota, S., Vigário, M., Martins, F., & Cruz, M. (2010). *FrePOP (version 1.0)*. Laboratório de Fonética (CLUL), Faculdade de Letras da Universidade de Lisboa. http://frepop.fl.ul.pt/

Gervain, J. (2014). Neonatal speech perception. In P. J. Brooks & V. Kempe (Eds.), *Encyclopedia of language development* (pp. 402–404). Sage.

Gósy, M. (1997). Semleges magánhangzók a magyar beszédben. *Magyar Nyelvőr*, *121*, 9–19.

Gósy, M. (2004). *Fonetika, a beszéd tudománya*. Osiris Kiadó.
Gottfried, T. L. (1984). Effects of consonant context on the perception of French vowels. *Journal of Phonetics*, *12*(2), 91–114. https://doi.org/10.1016/s0095-4470(19)30858-7
Grenon, I., Kubota, M., & Sheppard, C. (2019). The creation of a new vowel category by adult learners after adaptive phonetic training. *Journal of Phonetics*, *72*, 17–34. https://doi.org/10.1016/j.wocn.2018.10.005
Guedes, A. M. (1998). A identidade, a propaganda e o nacionalismo: o projecto de leitorados de língua e cultura portuguesas, 1921-97. *Lusotopie. Des protestantismes en lusophonie catholique* (5), 107–132.
Guion, S. G., & Pederson, E. (2007). Investigating the role of attention in phonetic learning. In O.-S. Bohn & M. J. Munro (Eds.), *Language Experience in Second Language Speech Learning: In honor of James Emil Flege* (pp. 57–77). John Benjamins. https://doi.org/10.1075/lllt.17.09gui
Hahn, L. D. (2004). Primary Stress and Intelligibility: Research to Motivate the Teaching of Suprasegmentals. *Tesol Quarterly*, *38*(2), 201–223. https://doi.org/10.2307/3588378
Halácsy, P., Kornai, A., Németh, L., Rung, A., Szakadát, I., & Trón, V. (2003). A szószablya projekt. In Z. Alexin & D. Csendes (Eds.), *I. Magyar Számítógépes Nyelvészeti Konferencia előadásai* [Proceedings], 298. Szegedi Tudományegyetem Informatikai Tanszék. http://acta.bibl.u-szeged.hu/id/eprint/59176
Hammarberg, B. (2007). Roles of L1 and L2 in L3 Production and Acquisition. In J. Cenoz, B. Hufeisen & U. Jessner (Ed.), *Cross-Linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives* (pp. 21-41). Bristol: Multilingual Matters. https://doi.org/10.21832/9781853595509-003
Harnsberger, J. D. (2001). On the relationship between identification and discrimination of non-native nasal consonants. *The Journal of the Acoustical Society of America*, *110*(1), 489–503. https://doi.org/10.1121/1.1371758
Hayes, B. (1995). *Metrical Stress Theory. Principles and Case Studies*. University of Chicago Press.
Hayes, B. (2009). *Introductory Phonology*. Wiley-Blackwell.
Hazan, V., Sennema, A., Iba, M., & Faulkner, A. (2005). Effect of audiovisual perceptual training on the perception and production of consonants by Japanese learners of English. *Speech Communication*, *47*(3), 360–378. https://doi.org/10.1016/j.specom.2005.04.007
Holden, J., Francisco, E., Lensch, R., Tommerdahl, A., Kirsch, B., Zai, L., Dennis, R., & Tommerdahl, M. (2019). Accuracy of different modalities of reaction time testing: Implications for online cognitive assessment tools. *BioRxiv* [Preprint], 1–21. https://doi.org/10.1101/726364
Honbolygó, F., Kóbor, A., & Csépe, V. (2017). Cognitive components of foreign word stress processing difficulty in speakers of a native language with non-contrastive stress. *International Journal of Bilingualism*, *23*(2), 1–15. https://doi.org/10.1177/1367006917728393
Hyltenstam, K., & Abrahamsson, N. (2003). Maturational Constraints in SLA. In C. J. Doughty & M. H. Long (Eds.), *The Handbook of Second Language Acquisition* (pp. 539–589). Blackwell.

Ilari, R. (2013). O português no contexto das línguas românicas. In Eduardo Buzaglo Paiva Raposo, M. F. B. do Nascimento, M. A. C. da Mota, L. Segura, & A. Mendes (Eds.), *Gramática do Português - Volume I*. Fundação Calouste Gulbenkian.

Iverson, P., & Evans, B. G. (2009). Learning English vowels with different first-language vowel systems II: Auditory training for native Spanish and German speakers. *The Journal of the Acoustical Society of America*, *126*(2), 866–877. https://doi.org/10.1121/1.3148196

Iverson, P., Hazan, V., & Bannister, K. (2005). Phonetic training with acoustic cue manipulations: A comparison of methods for teaching English /r/-/l/ to Japanese adults. *The Journal of the Acoustical Society of America*, *118*(5), 3267–3278. https://doi.org/10.1121/1.2062307

Justi, F. R. dos R., de Oliveira, B. S. F., & Justi, C. N. G. (2023). The relationship between morphological awareness and word reading in Brazilian Portuguese: a longitudinal study. *Psicologia: Reflexao e Critica*, *36*(1). https://doi.org/10.1186/s41155-022-00245-9

Kiss, G. (1985). A magyar magánhangzók első két formánsának meghatározása szintetizált hangmintákat felhasználó percepciós kísérlet segítségével. *Nyelvtudományi Közlemények*, *87*(1), 160–174.

Kissling, E. M. (2015). Phonetics instruction improves learners' perception of L2 sounds. *Language Teaching Research*, *19*(3), 254–275. https://doi.org/10.1177/1362168814541735

Kuhl, P. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences of the United States of America*, *87*(22), 11850–11857. https://doi.org/10.1073/pnas.97.22.11850

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1493), 979–1000. https://doi.org/10.1098/rstb.2007.2154

Kuhl, P., & Iverson, P. (1995). Linguistic experience and the "perceptual magnet effect." In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 121–124). York Press.

Kuhl, P., & Miller, J. D. (1978). Speech perception by the chinchilla: Identification functions for synthetic VOT stimuli. *The Journal of the Acoustical Society of America*, *63*(3), 905–917. https://doi.org/10.1121/1.381770

Lengeris, A. (2009). Perceptual assimilation and L2 learning: Evidence from the perception of Southern British english vowels by native speakers of Greek and Japanese. *Phonetica*, *66*(3), 169–187. https://doi.org/10.1159/000235659

Lengeris, A., & Hazan, V. (2010). The effect of native vowel processing ability and frequency discrimination acuity on the phonetic training of English vowels for native speakers of Greek. *The Journal of the Acoustical Society of America*, *128*(6), 3757–3768. https://doi.org/10.1121/1.3506351

Lenneberg, E. H. (1967). *Biological Foundations of Language*. Wiley.

Levis, J. M., & Sonsaat, S. (2017). Pronunciation teaching in the early CLT era. In O. Kang, R. I. Thomson, & J. M. Murphy (Eds.), *The Routledge Handbook of Contemporary English Pronunciation* (pp. 267–283). Routledge.

Levy, E. S. (2009). On the assimilation-discrimination relationship in American English adults' French vowel learning. *The Journal of the Acoustical Society of America*, *126*(5), 2670–2682. https://doi.org/10.1121/1.3224715

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, *74*(6), 431–461. https://doi.org/10.1037/h0020279

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, *54*(5), 358–368. https://doi.org/10.1037/h0044417

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revisited. *Cognition*, *21*(1), 1–36. https://doi.org/10.1016/0010-0277(85)90021-6

Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. *The Journal of the Acoustical Society of America*, *94*(3), 1242–1255. https://doi.org/10.1121/1.408177

Lively, S. E., Pisoni, D. B., Yamada, R. A., Yoh'ichi, T., & Yamada, T. (1994). Training japanese listeners to identify english /r/ and /l/. III. Long-term retention of new phonetic categories. *Journal of the Acoustical Society of America*, *96*(4), 2076–2087. https://doi.org/10.1121/1.410149

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training japanese listeners to identify english /r/ and /l/: A first report. *Journal of Acoustical Society of America*, *89*(2), 874–886. https://doi.org/10.1121/1.1894649

Lotto, A. J., & Holt, L. L. (2016). Speech Perception. In G. Hickok & S. L. Small (Eds.), *Neurobiology of Language* (pp. 185–194). Academic Press. https://doi.org/10.1007/978-3-319-16999-6_3334-2

Mackay, I. R. A., Flege, J. E., & Imai, S. (2006). Evaluating the effects of chronological age and sentence duration on degree of perceived foreign accent. *Applied Psycholinguistics*, *27*(2), 157–183. https://doi.org/10.1017/s0142716406060231

Major, R. C. (2001). *Foreign Accent. The Ontogeny and Phylogeny of Second Language Phonology*. Routledge. https://doi.org/10.4324/9781410604293

Major, R. C. (2008). Transfer in second language phonology: A review. In Jette G. Hansen Edwards & M. L. Zampini (Eds.), *Phonology and Second Language Acquisition* (pp. 63–94). John Benjamins. https://doi.org/10.1075/sibil.36.05maj

Markó, A. (2017). Hangtan. In A. Imrényi, N. Kugler, M. Ladányi, A. Markó, S. Tátrai, & G. Tolcsvai Nagy (Eds.), *Nyelvtan* (pp. 75–203). Osiris Kiadó.

Markó, A., Bartók, M., Csapó, T. G., Deme, A., & Gráczi, T. E. (2019). The effect of focal accent on vowels in Hungarian: Articulatory and acoustic data. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia 2019* (pp. 2715–2719). Australasian Speech Science and Technology Association Inc. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2019/papers/ICPhS_2764.pdf

Martins, P., Carbone, I., Pinto, A., Silva, A., & Teixeira, A. (2008). European Portuguese MRI based speech production studies. *Speech Communication*, *50*(11–12), 925–952. https://doi.org/10.1016/j.specom.2008.05.019

Mateus, M. H. M. (2003). Fonologia. In M. H. M. Mateus, A. M. Brito, I. Duarte, I. H. Faria, S. Frota, G. Matos, F. Oliveira, M. Vigário, & A. Villalva (Eds.), *Gramática da Língua Portuguesa - 5ª ed., revista e aumentada*. Caminho.

Mateus, M. H. M., Brito, A. M., Duarte, I., Faria, I. H., Frota, S., Matos, G., Oliveira, F., Vigário, M., & Villalva, A. (2003). *Gramática da Língua Portuguesa - 5ª ed., revista e aumentada*. Caminho.

Mateus, M. H. M., & d'Andrade, E. (2000). *The Phonology of Portuguese*. Oxford University Press.

Mayr, R., & Escudero, P. (2010). Explaining individual variation in L2 perception: Rounded vowels in English learners of German. *Bilingualism*, *13*(3), 279–297. https://doi.org/10.1017/S1366728909990022

Michaux, M. C. (2016). La perception de l'accent lexical néerlandais par les apprenants francophones. *Langages*, *202*, 47–74. https://doi.org/10.3917/lang.202.0047

Moyer, A. (2013). *Foreign Accent: The Phenomenon of Non-Native Speech*. Cambridge University Press.

Munson, B., & Solomon, N. P. (2004). The effect of phonological neighborhood density on vowel articulation. *Journal of Speech, Language, and Hearing Research*, *47*(5), 1048–1058. https://doi.org/10.1044/1092-4388(2004/078)

Nobre-Oliveira, D. (2007). *The effect of perceptual training on the learning of English vowels by Brazilian Portuguese speakers* [Unpublished doctoral dissertation]. Universidade Federal de Santa Catarina.

Oliveira, I. G. (2006). *A aquisição do sistema vocálico por falantes de português europeu como língua não materna* [Unpublished master's thesis]. Universidade de Lisboa.

Oravecz, C., Váradi, T., & Sass, B. (2014). The Hungarian Gigaword Corpus. In N. Calzolari, K. Choukri, T. Declerck, H. Loftsson, B. Maegaard, J. Mariani, A. Moreno, J. Odijk, & S. Piperidis (Eds.), *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14), Reykjavik* (pp. 1719– 1723). European Language Resources Association (ELRA). https://aclanthology.org/L14-1

Ou, S. (2011). Training Taiwanese EFL learners to perceive English lexical stress contrast: a pilot study. *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS XVII), Hong Kong*, 1550–1553. http://www.icphs2011.hk/resources/OnlineProceedings/RegularSession/Ou/Ou.pdf

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Peltola, T., Palo, P., & Aaltonen, O. (2014). A short term study of Hungarians learning Finnish vowels. In K. Jähi & L. Taimi (Eds.), *XXVIII Fonetiikan päivät, Turku* [Proceedings], 65–71. Turun Yliopisto.

Peperkamp, S., & Dupoux, E. (2002). A typological study of stress ' deafness .' In C. Gussenhoven & N. Warner (Eds.), *Laboratory Phonology 7* (pp. 203–240). Mouton de Gruyter. https://doi.org/10.1515/9783110197105.1.203

Peperkamp, S., Vendelin, I., & Dupoux, E. (2010). Perception of predictable stress: A cross-linguistic investigation. *Journal of Phonetics*, *38*(3), 422–430. https://doi.org/10.1016/j.wocn.2010.04.001

Pereira, I. (2020). Acento de Palavra. In Eduardo B. P. Raposo, M. F. B. do Nascimento, M. A. C. da Mota, L. Segura, A. Mendes, & A. Andrade (Eds.), *Gramática do Português - Volume III* (pp. 3397–3425). Fundação Calouste Gulbenkian.

Perkell, J. S., Guenther, F. H., Lane, H., Matthies, M. L., Stockmann, E., Tiede, M., & Zandipour, M. (2004). The distinctness of speakers' productions of vowel contrasts is related to their discrimination of the contrasts. *The Journal of the Acoustical Society of America*, *116*(4), 2338–2344. https://doi.org/10.1121/1.1787524

Pisoni, D. B. (1977). Identification and discrimination of the relative onset time of two component tones: Implications for voicing perception in stops. *Journal of the Acoustical Society of America*, *61*(5), 1352–1361. https://doi.org/10.1121/1.381409

Pisoni, D. B., & Lively, S. E. (1995). Variability and invariance in speech perception: A new look at some old problems in perceptual learning. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross Language Research* (pp. 433–459). York Press.

Poeppel, D., Idsardi, W. J., & Van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. *Philosophical Transactions of the Royal Society B*, *363*(1493), 1071–1086. https://doi.org/10.1098/rstb.2007.2160

Polka, L., & Werker, J. F. (1994). Developmental Changes in Perception of Nonnative Vowel Contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(2), 421–435. https://doi.org/10.1037/0096-1523.20.2.421

R Core Team. (2014). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Rakerd, B. (1984). Vowels in consonantal context are perceived more linguistically than are isolated vowels: Evidence from an individual differences scaling study. *Perception & Psychophysics*, *35*(2), 123–136. https://doi.org/10.3758/BF03203892

Raposo, Eduardo Buzaglo Paiva, do Nascimento, M. F. B., da Mota, M. A. C., Segura, L., & Mendes, A. (Eds.). (2013a). *Gramática do Português - Volume I: Vol. I*. Fundação Calouste Gulbenkian.

Raposo, Eduardo Buzaglo Paiva, do Nascimento, M. F. B., da Mota, M. A. C., Segura, L., & Mendes, A. (Eds.). (2013b). *Gramática do Português - Volume II: Vol. II*. Fundação Calouste Gulbenkian.

Raposo, Eduardo Buzaglo Paiva, do Nascimento, M. F. B., da Mota, M. A. C., Segura, L., & Mendes, A. (Eds.). (2020). *Gramática do Português - Volume III*. Fundação Calouste Gulbenkian.

Rato, A. (2013). *Cross-language Perception and Production of English Vowels by Portuguese Learners: The Effects of Perceptual Training* [Unpublished doctoral dissertation]. Universidade do Minho.

Rauber, A. S., Escudero, P., Bion, R. A. H., & Baptista, B. O. (2005). The Interrelation between the perception and production of English vowels by native speakers of Brazilian Portuguese. *Proceedings of the 9th European Conference on Speech Communication and Technology (INTERSPEECH 2005), Lisbon*, 2913–2916. https://doi.org/10.21437/interspeech.2005-769

Rauber, A. S., Rato, A., & Silva, A. L. (2010). Percepção e produção de vogais anteriores do inglês por falantes nativos de mandarim. *Diacrítica*, *24*(1), 5–23. http://repositorium.sdum.uminho.pt/bitstream/1822/23229/1/Diacritica24-1.pdf

Reed, M., & Levis, J. L. (Eds.). (2015). *The handbook of English Pronunciation*. John Wiley & Sons.

Reiterer, S. M., Hu, X., Sumathi, T. A., & Singh, N. C. (2013). Are you a good mimic? Neuro-acoustic signatures for speech imitation ability. *Frontiers in Psychology*, *4*(Article 782), 1–13. https://doi.org/10.3389/fpsyg.2013.00782

Rice, K. (2007). Markedness in phonology. In P. de Lacy (Ed.), *The Cambridge Handbook of Phonology*. Cambridge University Press.

Rosa, E. A. B. O. da R. (2022). *The effect of perception training with synthetic and natural stimuli on BP learners' ability to identify the English vowels /æ-ɛ/* [Unpublished master's thesis]. Universidade Federal de Santa Catarina.

Rost, M. (2011). *Teaching and Researching Listening* (2nd ed.). Pearson Education Limited.

Salsignac, J. (1998). Une étude sur la perception de l'accent primaire de langues étrangères. *Letras de Hoje*, *33*(4), 81–107.

Schwab, S., & Llisterri, J. (2014). Does training make French speakers more able to identify lexical stress? *Proceedings of the International Symposium on the Acquisition of Second Language Speech. Concordia Working Papers in Applied Linguistics*, *5*, 624–636.

Seymour, P. H. K., Aro, M., Erskine, J. M., Wimmer, H., Leybaert, J., Elbro, C., Lyytinen, H., Gombert, J. E., Le Normand, M. T., Schneider, W., Porpodas, C., Ragnarsdottir, H., Tressoldi, P., Vio, C., De Groot, A., Licht, R., Iønnessen, F. E., Castro, S. L., Cary, L., … Olofsson, Å. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, *94*(2), 143–174. https://doi.org/10.1348/000712603321661859

Sheldon, A., & Strange, W. (1982). The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics*, *3*(3), 243–261. https://doi.org/10.1017/S0142716400001417

Siptár, P., & Törkenczy, M. (2000). *The Phonology of Hungarian*. Oxford University Press.

Szabó, C., Miranda, L., Gráczi, T. E., Csapó, T. G., Juhász, K., Deme, A., & Markó, A. (2023). Você disse "avô" ou "avó"? Análise acústica das vogais médias posteriores na fala de aprendizes húngaros de PLA. *Anais Do I e Do II Conecta Leitores*.

Szalontai, A., Wagner, P., Mády, K., & Windmann, A. (2016). Teasing apart lexical stress and sentence accent in Hungarian and German. *12. Tagung Phonetik Und Phonologie Im Deutschsprachigen Raum, München* [Proccedings], 215–218. https://epub.ub.uni-muenchen.de/29405/

Szende, T. (1990). Hungarian. *Journal of the International Phonetic Association*, *24*(2), 91–94. https://doi.org/https://doi.org/10.1017/S0025100300005090

Sturgis, P., Roberts, C., & Smith, P. (2014). Middle Alternatives Revisited: How the neither/nor Response Acts as a Way of Saying "I Don't Know"? *Sociological Methods and Research*, *43*(1), 15–38. https://doi.org/10.1177/0049124112452527

Tavares, G., Deme, A., & Correia, S. (2022). Phonological acquisition of European Portuguese oral vowels: A forced-choice identification study with Hungarian native speakers. *Toronto Working Papers in Linguistics*, *44*, 1–19. https://doi.org/10.33137/twpl.v44i1.36758

Teixidó, M., François, C., Bosch, L., & Männel, C. (2018). The role of prosody in early speech segmentation and word-referent mapping: Electrophysiological evidence. In P. Prieto & N. Esteve-Gibert (Eds.), *The development of prosody in first language acquisition* (pp. 79–100). John Benjamins.

Traunmüller, H. (1981). Perceptual dimension of openness in vowels. *Journal of the Acoustical Society of America*, *69*(5), 1465–1475. https://doi.org/10.1121/1.385780

Trubetzkoy, N. S. (1969). *Principles of Phonology* (Baltaxe, C. A. M., Trans.). University of California. (Original work published 1958)

Tyler, M. D., Best, C. T., Faber, A., & Levitt, A. G. (2014). Perceptual assimilation and discrimination of non-native vowel contrasts. *Phonetica*, *71*(1), 4–21. https://doi.org/10.1159/000356237

van Heugten, M., Dautriche, I., & Christophe, A. (2014). Phonological and Prosodic Bootstrapping. In P. J. Brooks & V. Kempe (Eds.), *Encyclopedia of language development*. Sage.

van Heuven, V. J. (2019). Acoustic Correlates and Perceptual Cues of Word and Sentence Stress. Towards a Cross-Linguistic Perspective. In R. Goedemans, J. Heinz, & H. van der Hulst (Eds.), *The study of word stress and accent. Theories, methods and data* (pp. 15–52). Cambridge University Press.

VanPatten, B., & Williams, J. (Eds.). (2015). *Theories in Second Language Acquisition: An Introduction* (2nd ed.). Routledge.

Veloso, J. (2012). Vogais centrais do português europeu contemporâneo: uma proposta de análise à luz da fonologia dos elementos. *Letras de Hoje*, *47*(3), 234–243. http://revistaseletronicas.pucrs.br/ojs/index.php/fale/article/view/11853

Vigário, M. (2003). *The Prosodic Word in European Portuguese*. De Gruyter Mouton. https://doi.org/10.1515/9783110900927

Wade, T., Jongman, A., & Sereno, J. (2007). Effects of acoustic variability in the perceptual learning of non-native-accented speech sounds. *Phonetica*, *64*(2–3), 122–144. https://doi.org/10.1159/000107913

Wang, Y., Jongman, A., & Sereno, J. A. (2003). Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *The Journal of the Acoustical Society of America*, *113*(2), 1033–1043. https://doi.org/10.1121/1.1531176

Wang, W. (n.d.). Cookbook for R. Plotting means and error bars (ggplot2). Retrieved January 2021, from http://www.cookbook-r.com/Graphs/Plotting_means_and_error_bars_(ggplot2)/

White, L., & Mády, K. (2008). The long and the short and the final: Phonological vowel length and prosodic timing in Hungarian. *Proceedings of the 4th International Conference on Speech Prosody, Campinas*, 363–366. https://www.isca-speech.org/archive/speechprosody_2008/index.html

Whiteside, S. P. (2001). Sex-specific fundamental and formant frequency patterns in a cross-sectional study. *The Journal of the Acoustical Society of America*, *110*(1), 464–478. https://doi.org/10.1121/1.1379087

Wilson, S. M., Saygin, A. P., Sereno, M. I., & Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nature Neuroscience*, *7*(7), 701–702. https://doi.org/10.1038/nn1263

Wong, J. W. S. (2012). The effects of high and low variability phonetic training on the perception and production of english vowels /e/-/æ/ by cantonese ESL learners with high and low l2 proficiency levels. In F. Cox, K. Demuth, S. Lin, K. Miles, S. Palethorpe, J. Shaw, & I. Yuen (Eds.), *Proceedings of the 14th Australasian International Conference on Speech Science and Technology, Sydney* (pp. 37–40). Australasian Speech Science and Technology Association (ASSTA).

Yazawa, K., Whang, J., Kondo, M., & Escudero, P. (2020). Language-dependent cue weighting: An investigation of perception modes in L2 learning. *Second Language Research*, *36*(4), 557–581. https://doi.org/10.1177/0267658319832645

Ylinen, S., Uther, M., Latvala, A., Vepsäläinen, S., Iverson, P., Akahane-Yamada, R., & Näätänen, R. (2010). Training the brain to weight speech cues differently: A study of finnish second-language users of english. *Journal of Cognitive Neuroscience*, *22*(6), 1319–1332. https://doi.org/10.1162/jocn.2009.21272

UNIVERSIDADE ABERTA
UNIVERSIDADE NOVA DE LISBOA

Perceção de acento e redução vocálica na aquisição fonológica em PL2. Implicações para a prática pedagógica

Word stress and reduced vowels perception and acquisition in L2 Portuguese. Pedagogical implications

Gabriela Pereira Tavares

APPENDICES

Tese de Doutoramento em Didática das Línguas – Multilinguismo e Educação para a Cidadania Global

PhD dissertation in Language Teaching - Multilingualism and Education for Global Citizenship

2024

UNIVERSIDADE ABERTA
UNIVERSIDADE NOVA DE LISBOA

Perceção de acento e redução vocálica na aquisição fonológica em PL2. Implicações para a prática pedagógica

Word stress and reduced vowels perception and acquisition in L2 Portuguese. Pedagogical implications

Gabriela Pereira Tavares

APPENDICES

Tese de Doutoramento em Didática das Línguas – Multilinguismo e Educação para a Cidadania Global

PhD dissertation in Language Teaching - Multilingualism and Education for Global Citizenship

Supervision:

Professor Susana Correia (NOVA University Lisbon)
Professor Andrea Deme (Eötvös Loránd University)

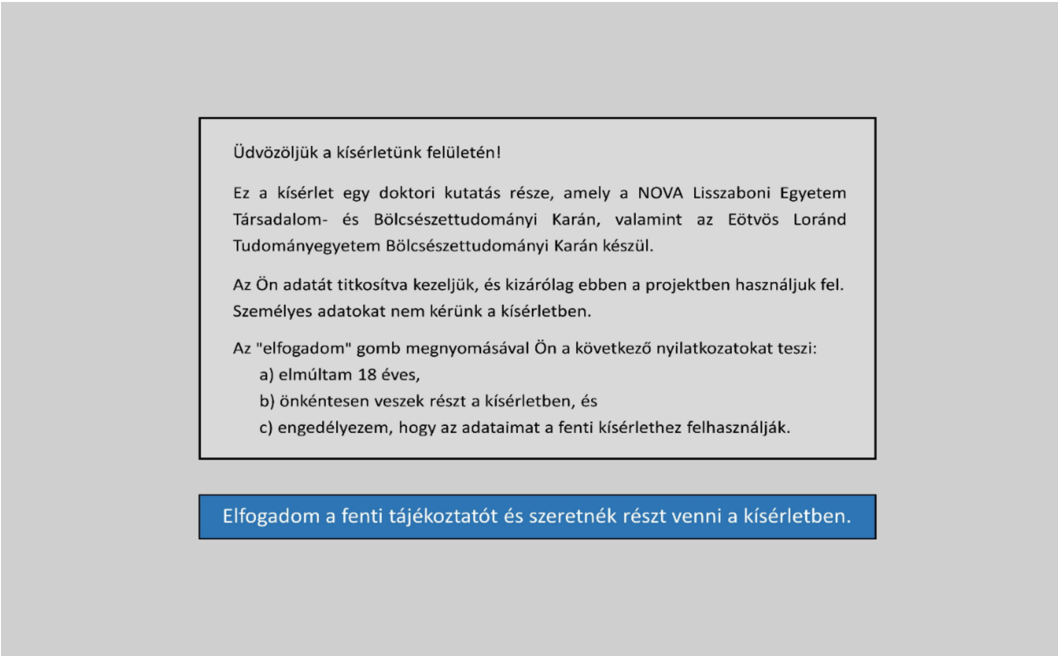
Janeiro 2024

This work is licensed under a Creative Commons
NonCommercial-NoDerivatives 4.0 International licence (CC BY-NC-ND 4.0)

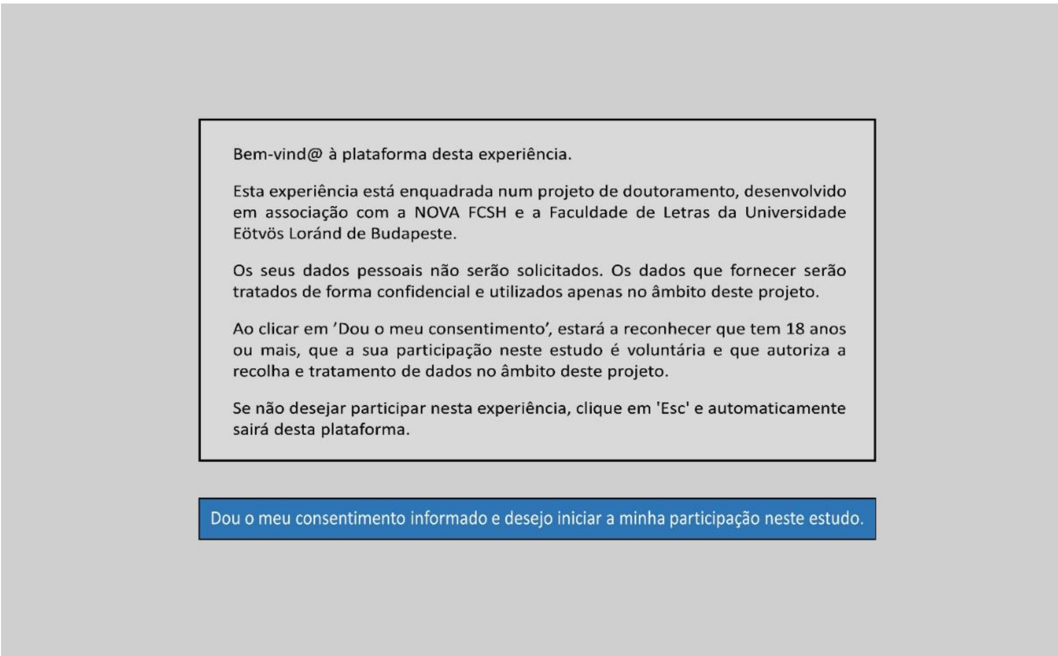
TABLE OF CONTENTS

Appendix I
Consent forms

A. Consent screen presented to the Hungarian participants in the identification experiment (Chapter 4)



B. Consent screen presented to the Portuguese participants in the identification experiment (Chapter 4)



C. Consent screen presented to the Hungarian participants in the beginning of each task of the perceptual training study (Chapter 5)

Hozzájárulás

Üdvözöljük a kísérletünk felületén! Ez a kísérlet egy doktori kutatás része, amely a NOVA Lisszaboni Egyetem Társadalom- és Bölcsészettudományi Karán, valamint az Eötvös Loránd Tudományegyetem Bölcsészettudományi Karán készül.

Az Ön adatait titkosítva kezeljük, és kizárólag ebben a projektben használjuk fel.

Kérjük **pipálja ki az „Elfogadom..."** jelölőnégyzetet, és **nyomja meg a "Next" gombot** a továbblépéshez. Ezzel Ön a következő nyilatkozatokat teszi:

a) elmúltam 18 éves,

b) önkéntesen veszek részt a kísérletben, és

c) engedélyezem, hogy az adataimat a fenti kísérlethez felhasználják.

☐ **Elfogadom a fenti tájékoztatót és szeretnék részt venni a kísérletben.**

Nem fogadom el és kilépek a kísérletből.

Next

C. Consent screen presented to the Portuguese participants in the beginning of each task of the perceptual training study (Chapter 5)

Consentimento

Bem-vind@ à plataforma desta experiência.

Esta experiência está enquadrada num projeto de doutoramento, desenvolvido em associação com a NOVA FCSH e a Faculdade de Letras da Universidade Eötvös Loránd de Budapeste.

Os dados que fornecer serão tratados de forma confidencial e utilizados apenas no âmbito deste projeto. Ao clicar em "Dou o meu consentimento", estará a reconhecer que tem 18 anos ou mais, que a sua participação neste estudo é voluntária e que autoriza a recolha e tratamento de dados no âmbito deste projeto.

Se deseja participar, clique em **"Dou o meu consentimento"** e depois em **'Continuar'**.

☐ **Dou o meu consentimento informado e desejo iniciar a minha participação neste estudo.**

Não dou o meu consentimento.

Next

Appendix II
Questionnaires

A. Translation of the background questionnaire presented to the Hungarian and Portuguese participants in the identification experiment (Chapter 4)

Choose an option:

a) Age
18-25 / 26-30 / 31-40 / 41-45

b) Gender
female / male / other

c) Academic studies completed
secondary studies bachelor / master / doctorate

Is your mother language Hungarian/Portuguese?[1]
yes / no

Choose one language:

(Description: To the Hungarian participants, seven languages were presented (English, Germain, Spanish, French, Italian, Portuguese, and Russian), each with four levels (native, basic, intermediate, advanced). To the Portuguese participants, Russian was not included, and instead of Portuguese, Hungarian was included. In both cases, the option “neither one” was presented. The options were presented in a grid (see below)).

B. Computer screen of the grid for L2 knowledge, presented to the Hungarian participants in the identification experiment (Chapter 4)

[1] The first option was presented to the Hungarian participants, and the second, to the Portuguese participants.

C. Translation of the background questionnaire presented to the Hungarian participants in the perceptual training study (Chapter 5)

1. Name (Surname, First name)
2. Age
3. Gender (Female, Male, Other)
4. Academic studies completed (Secondary studies, Bachelor, Master, Doctorate)
5. In which university are you studying? (Write the name of the university/faculty, course, and year)
6. Do you have any acute or chronic hearing problems (example: cold, asthma, Reinke's edema, vocal cord nodules, polyp)?
   a) Yes. Please, describe it. If you received treatment, describe it also, briefly.
   b) No
7. First language(s)[2]
8. Second language (If you don't speak (other) second languages, click on the 'Continue' button)[3]
   8.1 Language
   8.2. Level (If you are uncertain of your level, consult the following page, where the levels are described:
   https://hu.wikipedia.org/wiki/K%C3%B6z%C3%B6s_Eur%C3%B3pai_Referenciakeret
      a) Breakthrough - A1
      b) Waystage - A2
      c) Threshold - B1
      d) Vantage - B2
      e) Advanced - C1
      f) Mastery - C2
   8.3. How frequently do you use this language?
      a) daily
      b) weekly
      c) monthly
      d) annually
      e) rarely
      f) never
9. Previous contact with the Portuguese language: Think about where you met/hear? the Portuguese language before (e.g., tv programs, films, meeting with Portuguese native speakers, a Portuguese course in the past, etc.) Did you have contact with the Portuguese language before starting the Portuguese course you presently attend?

[2] Three options were given, the first was compulsory.

[3] This section was presented five times, that is, participants were given the opportunity to report five L2.

a) Yes. Briefly describe under which conditions you had contact with Portuguese (when, where, for how long, etc.).
b) No

10. Presently, do you have contact with the Portuguese language outside the Portuguese classes?
    a) Yes. Briefly describe under which conditions (when, where, for how long, etc.).
    b) No
11. Why do study Portuguese (motivation/reason)?

D. Translation of the follow-up questionnaire presented to the Hungarian participants in the perceptual training study (Chapter 5)

1. Do you have any acute or chronic hearing problems (example: cold, asthma, Reinke's edema, vocal cord nodules, polyp)?
   a) Yes. Please, describe it. If you received treatment, describe it also, briefly.
   b) No
2. Approximately, how many 60 minutes classes have you attended since September?
3. Since September, the nature and amount of contact with the Portuguese language
   a) didn't change
   b) change (describe briefly how)
4. Which of the following statements is true for you? In your opinion, the training
   a) was not at all interesting
   b) was not very interesting
   c) was interesting
   d) was very interesting
5. Which of the following statements is true for you? In your opinion, the training tasks
   a) were too long
   b) had the adequate duration
   c) were too short
6. Which of the following statements is true for you? In your opinion, in the end of the training tasks
   a) you didn't feel tired at all
   b) you felt a little bit tired
   c) you felt very tired
7. Which of the following statements is true for you? In your opinion, during the training tasks
   a) you didn't feel motivated at all
   b) you didn't feel very motivated
   c) you felt motivated
   d) you felt very motivated

8. Which of the following statements is true for you? In your opinion, during the training tasks
    a) you didn’t feel concentrated at all
    b) you didn’t feel very concentrated
    c) you felt concentrated
    d) you felt very concentrated
9. Which of the following statements is true for you? In your opinion, the six training sessions,
    a) were not at all sufficient
    b) were more or less sufficient
    c) were sufficient
    d) were more than sufficient
10. Do you think the training contributed to improve your auditory perception of Portuguese language?
    a) Yes. Describe briefly why.
    b) No
11. Do you think the training contributed to improve your pronunciation of Portuguese language?
    Yes
    No
    Describe briefly why.
12. Other comments (optional)

E. Translation of the background questionnaire presented to the Portuguese participants in the perceptual training study (Chapter 5)

1. Age
2. Gender Gender (Female, Male, Other)
3. Academic studies completed (Secondary studies, Bachelor, Master, Doctorate)
4. Occupation (if you are a student, indicate the university/faculty, course, and year)
5. Do you have any problem, temporary or chronic, that affects your hearing abilities?
    a) yes
    b) no
    If you answer 'yes', indicate which (example: cold, otitis, influenza, asthma)

6.1. Is European Portuguese your mother language?
    a) yes
    b) no

6.2. Do you speak another variety of Portuguese as your mother language?
    a) yes
    b) no

If you speak another variety of Portuguese, indicate which (example: Mozambican Portuguese)

6.3. Write where did you have lived in the last 10 years and in the present.

6.4. Do you speak Hungarian?

a) yes
b) no

6.5. Do you have other mother languages besides Portuguese?

a) yes
b) no
If you answer 'yes', indicate which.

7. Second language (If you don't speak (other) second languages, click on the 'Continue' button)[4]

7.1. Language

7.2 Level (If you are uncertain of your level, consult the following page, where the levels are described:
https://hu.wikipedia.org/wiki/K%C3%B6z%C3%B6s_Eur%C3%B3pai_Referenciakeret

a) Breakthrough - A1
b) Waystage - A2
c) Threshold - B1
d) Vantage - B2
e) Advanced - C1
f) Mastery - C2

How frequently do you use this language?

a) daily
b) weekly
c) monthly
d) annually
e) rarely

F. Translation of participants responses to the question "Briefly describe under which conditions you had contact with Portuguese (when, where, for how long, etc.)", presented in the background questionnaire in the perceptual training experiment (Chapter 5)

participant ID	response
BGE04	Listening to Latin-American songs, for 2-3 years constantly.
BGE08	Listening to Portuguese/Brazilian music.
CORV02	Travelling abroad.
CORV04	I saw a movie in Portuguese a few weeks ago, and a video of a girl who speaks several languages talking in Portuguese.

[4] This section was presented five times.

CORV05 Brazil in 2014 and Portugal in 2019.

CORV06 During travelling in 2019.

CORV08 In movies, TV shows, in interviews, once a week or month.

CORV14 During a 4- or 5-days visit to Portugal. Moreover, I listened to a few Portuguese videos.

CORV16 I spent last semester in Spain on Erasmus, where I got some Portuguese friends.

CORV19 On Erasmus + on an exchange program, and when I was on vacation in Portugal.

CORV32 On vacation in Portugal 5 years ago.

DEB06 My husband speaks Portuguese, also in the summer of 2021 and 2022 I spent some weeks in Portugal.

ELTEBA02 2-3 years ago, in high school I was in a Portuguese elective coursefor about half a year.

ELTEBA04 Series that I've watched in Portuguese with English subtitles.

ELTEBA06 In Portugal, 4 years ago.

ELTEBA18 In Portugal, in the summer of 2022.

ELTEBA24 Yes, I've had previous contact with the Portuguese language before. In the last few months, I've became friends with Brazilians studying here and with whom I practiced Portuguese.

ELTEBA26 Autonomous language learning in the past few years.

ELTEBA27 This year I've studied by myself, online, on a basic level.

ELTEMI02 By listening to Fado.

ELTEMI07 TV series (1 year ago), Brazilian friend (currently).

ELTEMI08 I like to listen and to sing Portuguese songs.

ELTEMI09 Through my partner who is Portuguese, I was in Portugal (Algarve) for 2 weeks about 3 months ago, where they spoke in Portuguese in their family.

ELTEMI14 In 2014 I was in Portugal.

ELTEMI16 In the university as my minor.

ELTEMI17 In high school I had it as my elective course, until half a year ago.

ELTEMI36 Travelling to Portugal. In the hotel where I work, I meet Portuguese guests as well.

ELTEMI38 I've watched a Portuguese series in Portuguese, with Hungarian subtitles. On Instagram I've seen several times Portuguese content.

KAROL10 In 2022 spring I attended a two weeklong intensive language course (40 hours in total) in Lisbon.

KAROL13 For a few months I followed the program called "Portuguese with Carla - The Journey" and also watched Portuguese videos on YouTube.

MET01 I lived for two months in Portugal, this year's April and May. I communicated in English, but I heard Portuguese on a daily basis.

MET02 Listening to music.

MET03 I spent my Erasmus in Lisbon from 25.09.2021 until 02.07.2022. I've learned Portuguese there a little (beginning of A1).

MET06	Brazilian-Portuguese course half a year ago.
NKE01	I've watched YouTube videos during the summer.
NKE08	Brazilian movie on Netflix.
PECS13	In the summer of 2021, I spent 3,5 weeks in Portugal and although I communicated in English, I heard people in my surrounding talking in Portuguese.
SZEGED01	I was on vacation in Portugal for a month, my Hungarian friend taught me some things.
SZEGED02	I've had minimal previous contact with Portuguese, my friends talk in Portuguese between themselves.
SZEGED04	During a Portuguese journey in 2016.
SZEGED 07	During my university studies, my department organized different programs with the Portuguese lecturer and the students.

G. Translation of participants responses to the question "Briefly describe under which conditions (when, where, for how long, etc.) [you have presently contact with the Portuguese language outside the Portuguese classes", presented in the background questionnaire in the perceptual training experiment (Chapter 5)

participant ID	response
BGE03	By watching series and listening to music.
BGE08	By listening to music.
CORV02	I watch movies in Portuguese or with Portuguese subtitles.
CORV08	On the internet.
CORV14	Sometimes I watch or listen to Portuguese videos. As the course goes, I want to do this even more often.
CORV18	On the internet, watching videos.
CORV19	Series on Netflix (although I think that is Brazilian accent).
CORV32	Music and series.
DEB04	In the university, in the class.
ELTEBA03	Watching series, using web applications.
ELTEBA07	I'm following more and more Portuguese sites.
ELTEBA24	Yes, I have present contact with Portuguese. I try to meet up at least once a week with my Brazilian friends, moreover I watch Portuguese content on the internet every day, on different sites (Youtube, Google, Instagram, etc.)
ELTEMI08	Also with music.
ELTEMI09	We practice it with my partner, which mainly helps with the pronunciation.
ELTEMI16	Also in the university, twice a week on the lesson.
KAROL13	I listen to "Mia Esmeriz Academy – Destination: Portugal" course, also Portuguese language learning videos on YouTube.

MET01	In the university's class, Duolingo (Brazilian Portuguese).
MET02	Music.
MET06	With Brazilian friends.
NKE08	Listening to Portuguese music (also mostly Brazilian).
SZEGED04	By listening to music.

H. Translation of participants responses to the question "Why do study Portuguese (motivation/reason)?", presented in the background questionnaire in the perceptual training experiment (Chapter 5)

participant ID	translation
BGE03	I've always wanted to learn Portuguese or a similar language. I like it a lot, it is completely different than the languages I've learnt so far.
BGE04	I like to study unique languages.
BGE08	To move to Brazil.
CORV01	I wanted to learn a Latin language and I liked Portuguese the most.
CORV02	Because of Erasmus.
CORV04	I'm learning Portuguese because not a lot of people study it as foreign language, therefore it could be very useful in the future if I have this language skill.
CORV05	I like the language and the culture; I would like to learn something new.
CORV06	I want to be able to speak Portuguese when I go back to Portugal.
CORV07	For Erasmus, improving my language knowledge.
CORV08	I find it to be one of the most beautiful languages, it has been my goal for a long time to study Portuguese after Spanish.
CORV09	I will be in Portugal in the next semester, on Erasmus.
CORV10	I'm going to live in Portugal, and I would like to learn the language more or less before that.
CORV12	Next semester I will be in Lisbon on Erasmus.
CORV14	Firstly, because I'm interested in the language, the country, the culture, secondly, it will be useful in my career.
CORV16	It's similar to Spanish, therefore I expect that I can learn it easily. Moreover, I like new Latin languages and cultures.
CORV17	I find it a beautiful language and I would like to travel to Portugal, where it is useful to speak the language.
CORV18	For fun.
CORV19	Because I like the language, Portugal and Brazil are both wonderful, and because I am interested in the culture. I would like to live there, even just for a bit.
CORV32	I like the language and the culture.
CORV34	I would like to learn as many languages as possible and Portuguese was the one I liked the most from the options.

DEB06 To expand my knowledge with this language, and also because I like the country and the culture.

DEB12 Exploring a new language area, gaining more interest in it.

ELTEBA01 Next to Spanish learning another new Latin language, I like the Portuguese language and the culture.

ELTEBA02 I've always liked learning languages and I was already interested in portuguese in high school. Since I've been in Portugal my interest in the culture and the language grew even more and that is the reason why I chose this course in university.

ELTEBA03 I like to study languages; I want to learn as many as possible including Portuguese.

ELTEBA04 I like the language.

ELTEBA06 I'm interested in new Latin languages.

ELTEBA07 I wanted to learn a Latin language, moreover I want to learn as many languages as possible.

ELTEBA12 I'm interested in it.

ELTEBA18 Because it is a unique language which is in part of a language family with a lot of others, and I want to learn as many as possible.

ELTEBA24 I'm obsessed with different sounds, and I tend to think it is possible to learn them well. I became interested in Portuguese by the Portuguese phonology. I find it very nice and it is like an instrument I want to learn to play.

ELTEBA26 I'm interested in the new Latin linguistics since high school.

ELTEBA27 It sounds very nice, and I would like to be able to speak it. The culture of Portuguese speaking countries stands very close to me.

ELTEBA28 The language is very interesting.

ELTEMI01 The Portuguese language and Portugal's particularity.

ELTEMI02 I started to like Fado and its atmosphere, and I would like to speak the language, to get to know the people who can convey these emotions through the music.

ELTEMI04 I like how it sounds.

ELTEMI07 I want to learn a language that's similar to Spanish and I'm also interested in the Portuguese/Brazilian culture, I want to visit the country.

ELTEMI08 I want to sing the Portuguese songs with also a good pronunciation and understanding the lyrics. The country and Lisbon is really appealing.

ELTEMI09 I really like it, I like how it sounds, a lot of people use it, and I want to understand my partner's surroundings.

ELTEMI10 I've heard stories from my high school lecturer of the time she lived in Brazil and Portugal. She suggested I have Portuguese as a minor in university.

ELTEMI13 I like the language.

ELTEMI14 I wanted to learn language as a minor and Portuguese was the most appealing to me of the languages that we can learn from the basics.

ELTEMI16 I was advised to learn it next to Spanish, advantageous combination.

ELTEMI17 I found the culture interesting and now I can imagine moving there in the future, or going on Erasmus there and for that learning the language is essential.

ELTEMI36 In any case I wanted to choose a new Latin language as a minor, and this was the language which didn't require previous knowledge.

ELTEMI38 From the new Latin languages, I want to learn as many as possible. I've already had Italian and Spanish, so now it's time for French and Portuguese.

KAROL08 I wanted to get to know another language.

KAROL10 I find Portugal stunningly beautiful, as a country and the people living there. It is hard for me to learn languages, but I feel strongly intrigued by Portuguese language.

KAROL13 I have some Portuguese friends and by learning the language I want to develop our friendship, for travelling and self-improvement, and in the hope of future job opportunities (translation, interpreting).

MET01 I would like to go back on Erasmus.

MET02 I was interested in this language.

MET03 Because I started and want to keep up/improve the level of my knowledge. Also, because I miss Lisbon and it takes me back to that wonderful period of my life that I spent there.

MET06 It could be useful in the future, sport journalism.

NKE01 After Spanish I wanted to learn another new Latin language and I found them to be similar, so it could be useful.

NKE02 I'm interested in the countries of Latin America and since Brazil is an important country among them, I want to learn Portuguese and after that Brazilian Portuguese.

NKE04 In the future I would like to live in a Portuguese speech country, therefore I find it important to know the language.

NKE08 The priority for me was to study a Latin language and it was not possible to have Spanish as my course.

PECS11 Learning languages is my passion and I want to take advantage of the opportunity of Portuguese learning courses in the university and obtain a language exam.

PECS13 I decided I want to learn Portuguese during my stay in Algarve in the summer, because I fell in love with the country, and I really want to go back there in the future with the ability to communicate in Portuguese.

PECS30 To broaden my horizons, plus after Spanish it's easier.

SZEGED01 I want to move to Portugal.

SZEGED02 My best friend's wife and child is Portuguese, I want to keep in touch with them.

SZEGED04 I've been interested in Portuguese culture and history for a long time, moreover, I have several friends living in Portugal, and I'd like to keep in touch with them in their mother tongue.

SZEGED06 It was possible to fit the lesson in my timetable, it sounds interesting.

SZEGED07	Since Portuguese is also a new Latin language, my knowledge of Spanish and Catalan helps me to learn this language. I like to study languages, get to know cultures, use the languages I've learnt in general and during travelling.

I. Translation of participants responses to the question "Do you think the training contributed to improve your auditory perception of Portuguese language? Describe briefly why", presented in the follow-questionnaire, in the perceptual training experiment (Chapter 5)

participant ID	translation
	comments for the answer "no"
CORV18	I do not hear out better the differences than before the training.
ELTEBA18	I don't know. It depends on the person talking how obvious is the difference in the sounds.
ELTEMI04	No.
	comments for the answer "yes"
BGE03	The exercises after listening are getting easier for me, also I understand more words from the film than before.
BGE04	I hear out the differences easier.
BGE08	The training helped but because of the short time it could not help as much as I expected. 1 test per week which are 10 minutes long, therefore it did not have much success for me.
CORV01	I hear out the word stress in the words better.
CORV02	I pay more attention how to pronounce each sound.
CORV04	Differentiating the sounds helped me to write down even words I did not know in Portuguese.
CORV05	I can recognize easier the syllables which sound similar.
CORV06	It helped me to distinguish the different sounds.
CORV07	I feel like I can tell apart easier similar sounds that I have had thought to be the same in the beginning of the training.
CORV08	It helped a lot in recognizing, differentiating the sounds.
CORV09	I realized the significance of the difference in sounds.
CORV10	I memorized the pronunciation of a lot of Portuguese sounds.
CORV12	I hear out the word stress a bit better now.
CORV14	It enabled me to differentiate better on which syllable is the word stress in a certain word.
CORV16	Yes, I think it will help in the future so I can recognize better the words in the listening exercises.
CORV17	Yes.
CORV19	I recognized that by the end I could differentiate much better the two versions of pronouncing "a". In the beginning I could not hear any difference.

CORV32	I can differentiate better sounds.
CORV34	The words do not sound so unfamiliar, and I can distinguish them easier in my head.
DEB06	It is easier now to tell apart similar sounding words.
DEB12	I got to know better the Portuguese sound system.
ELTEBA01	I can tell apart easier the different Portuguese vowels.
ELTEBA02	I feel that I improved in the listening exercises. I can understand better what they are talking about.
ELTEBA03	I understand better the texts that I am reading.
ELTEBA04	I cannot really tell in what way it helped exactly but the exercises are easier now for me.
ELTEBA06	I can pay more attention to the differences in pronunciation.
ELTEBA07	I paid more attention to the sounds.
ELTEBA12	I got more used to the language.
ELTEBA24	Yes, it contributed because it helped to differentiate the different phonemes.
ELTEBA26	To listen several times from several people the same thing helped to distinguish similar syllables in other languages too.
ELTEBA27	It helped to recognize the words.
ELTEBA28	I could distinguish easier the differences in the sounds.
ELTEMI01	I can hear out better the small differences.
ELTEMI02	It helped to recognize the word stress learned on the course in the text we were listening to.
ELTEMI07	I can hear out the small differences as well.
ELTEMI08	I felt that I can tell apart easier the differences during the exercise.
ELTEMI09	It contributed in a way that I can hear now how close certain sounds are.
ELTEMI10	The training exercises helped in a way that I could recognize sounds better and I could differentiate them.
ELTEMI13	I feel it is a bit better now.
ELTEMI14	It helped me to concentrate better on the text we are listening to.
ELTEMI16	Understanding the text after listening is easier now.
ELTEMI17	My pronunciation and understanding after listening is better.
ELTEMI36	It made me realize to be more open to Portuguese language and to be able to associate to words.
ELTEMI38	I don't know, but I've improved for sure - the exercises after listening might be easier now for me.
KAROL08	We just listened to combinations of sounds, no full words.
KAROL10	I can understand a bit better Portuguese texts.
KAROL13	With the bonus tasks I got to know numerous useful words, expressions.
MET01	Maybe yes but I cannot tell actually. I have been hearing Portuguese from other places as well, I don't know how much exactly the training helped in it.

MET02	Yes, it was good to hear how they pronounce the words and what is the correct word stress.
MET03	Maybe I can hear out better the small differences.
MET06	It helped to improve.
NKE01	I can recognize the sounds better; it does not seem all so incomprehensible.
NKE02	The word stress has great importance in the difference between words, therefore the training was useful for me.
NKE04	From my point of view, the biggest difficulty in Portuguese language is the pronunciation, and I could improve exactly in that during the training.
NKE08	I can hear out the similar pronunciation.
PECS11	I think I can tell apart easier the sounds and understand the authentic Portuguese language.
PECS13	It helped in differentiating certain sounds.
PECS30	Yes, I can recognize easier the words.
SZEGED01	I know much better where the word stress in a word is.
SZEGED02	It was easier to distinguish the differences which are not recognizable in Hungarian language.
SZEGED04	It made the differentiating between vowels in Portuguese easier.
SZEGED06	I've learned the importance of the word stress in a certain word.
SZEGED07	For sure. It's very difficult to separate words from each other, and that I understand the different consonants and syllables.

J. Translation of participants responses to the question "Do you think the training contributed to improve your pronunciation in Portuguese language? Describe briefly why", presented in the follow-questionnaire, in the perceptual training experiment (Chapter 5)

participant ID	translation
	comments for the answer "no"
BGE08	Because to learn to pronunciation after listening is quite difficult and for that the exercises were not enough.
CORV01	It did not use real words and the same words came up repeatedly over the weeks.
CORV02	It mostly helped in the listening; I felt the bigger difference in that.
CORV04	In my opinion, listening is not enough for the pronunciation, we should repeat the words out loud.
CORV05	As I did not see the syllables written, it did not help a lot in the pronunciation.
CORV07	There were only listening tasks and I feel like I could not use the sounds that appeared in those correctly. As the recordings were short, I could not really connect them to exact words that they could appear in.
CORV12	I was listening to non-existent words, so it didn't help in my pronunciation.

CORV18	Because we did not have to talk during the training, only understand what we are listening to.
CORV19	Until now I cannot feel that it helped. On the Portuguese lessons the teacher still has to correct me, it helped mostly in understanding after listening.
CORV32	I do not feel that I improved in my pronunciation greatly.
CORV34	As I only listened, I don't think my pronunciation got better.
ELTEBA07	Because there were sounds.
ELTEMI02	Forming the words that do not exist in Hungarian is mostly the difficulty for me in pronunciation of Portuguese, which the course did not help.
ELTEMI04	No.
ELTEMI07	Because they showed non-existent words.
ELTEMI09	As the syllables we were hearing were not written, I could not associate the sounds to anything to learn.
ELTEMI16	I don't have a broad vocabulary yet, so the improvement is not visible.
ELTEMI36	I cannot tell that yet.
ELTEMI38	Unfortunately, Portuguese pronunciation is still quite hard for me.
MET06	As there was only listening and understanding, I don't think so.
NKE02	No, because I cannot write down the words by only listening to them, even if they are not real words.
SZEGED06	In listening I can tell the words apart easier, but the pronunciation is harder.
	comments for the answer "yes"
BGE03	I can read and talk smoother and with less mistakes.
BGE04	It is easier to pronounce the words.
CORV06	I can memorize easier the pronunciation of certain sounds after listening to them more times.
CORV08	Yes, the listening of the pronunciation of the syllables definitely helped.
CORV09	I started to pay more attention on the lessons to learn the pronunciation more properly.
CORV10	After listening to it repeatedly it was easier to imitate.
CORV14	Probably only a bit. I learned to pronounce better the words that appeared in the exercises.
CORV16	Yes, because I've heard a lot of times syllables and sounds that do not exist in Hungarian.
CORV17	My pronunciation improved.
DEB06	I pay more attention in the pronunciation of the sounds.
DEB12	During the bonus exercises I got to know various new language elements, which extended my knowledge.
ELTEBA01	I pay more attention to the pronunciation of Portuguese syllables.
ELTEBA02	I observed a bit of change in my pronunciation. And I pronounce the words more bravely now.
ELTEBA03	I try to distinguish where the word stress is in each word.

ELTEBA04 The pronunciation is a critical point in Portuguese language, how I speak evolved a lot during these few months, the words come more naturally for me.

ELTEBA06 I can understand the intonation/pronunciation much better and the differences in these.

ELTEBA12 I could distinguish easier the differences in the sounds.

ELTEBA18 I am more conscious about emphasizing the differences in the pronunciation.

ELTEBA24 Although it helped me to differentiate/recognize the different phonemes, it was not that helpful to reproduce these. It would have been great if we would have got more practical tips.

ELTEBA26 As in other languages more times, from different people hearing the same helped to distinguish the similar syllables, and it got more and more fixed in my mind.

ELTEBA27 It became helpful in the pronunciation of words.

ELTEBA28 I can pronounce the words better.

ELTEMI01 Words that were hard for me until now also appeared in the exercises. After multiple tasks, hearing them multiple times it is easier now to pronounce them.

ELTEMI08 A bit, but yes, all practicing helps and improves.

ELTEMI10 The pronunciation for me is easier after listening. Although we did not hear existent words in the exercises, pronouncing the syllables we listened to helped to improve my pronunciation overall.

ELTEMI13 I feel it is a bit better.

ELTEMI14 I pay more attention to the word stress since then.

ELTEMI17 I can tell the pronunciation of "e" and "i" sounds apart, and how to use the word stress.

KAROL08 I can already distinguish sometimes the similar sounding sound scales.

KAROL10 I hope so.

KAROL13 The "word" that we listened to I repeated to myself.

MET01 I can feel better where to put the word stress.

MET02 Yes, it contributed, it helped a lot in the more playful part when the word was written and there was a recording to listen to as well.

MET03 I see more clearly the particularities of the pronunciation and by that I can possibly recognize better in the future the mass of sounds coming to my ear.

NKE01 It helped to see more into the language, what are its characteristics.

NKE04 It definitely helped to differentiate between the various ways of pronunciation.

NKE08 By seeing and hearing the words in the bonus exercises.

PECS11 I can distinguish the sounds better and I can probably even repeat them.

PECS13 It helped in differentiating the sounds.

PECS30 It is easier to acquire the pronunciation after listening.

SZEGED01 It made me conscious about the importance of word stress, therefore I also try to do that well now.

SZEGED02	After starting to understand the differences between the sounds, learning the words and the pronunciation got less hard than in the beginning.
SZEGED04	We could notice well the pronunciation in the bonus exercises.
SZEGED07	Yes, definitely. As I knew already that the use of word stress is especially important in the neo-Latin languages, I reckon that the training helped that my pronunciation in Portuguese can be even better in the future.

K. Translation of participants responses to the question "Other comments", presented in the follow-questionnaire, in the perceptual training experiment (Chapter 5)

participant ID	**translation**
BGE03	I really liked this work together, would be glad to participate another time as well. :)
CORV07	The bonus tasks were amusing and interesting.
CORV18	The bonus tasks were interesting.
CORV19	Although its much clearer now to me how certain sounds are in Portuguese, I think the training could be a bit longer. What I mean is that there could be more parts and modules for the improvement, not that these modules should be longer.
DEB06	The bonus tasks were very interesting and useful. :)
DEB12	The assignments were really interesting, I hope the investigation will be successful.
ELTEBA24	I am satisfied with it. :)
ELTEBA26	In my opinion there could have been twice this number of assignments too.
ELTEMI38	When I made a mistake in an assignment, it was too obvious for me that if for instance, if the square in the right is not the right answer, then it is the left one.
MET01	The bonus tasks are really good. :)
NKE04	Although the length of the assignments is okay to solve, towards the end my concentration started to ease down, and I started to get confused in the sense that I had difficulty in make difference between the words.

Appendix III

Certificates for the perceptual training study (Chapter 5)

Department of
Applied Linguistics and Phonetics
Eötvös Loránd University, Faculty of Humanities
H-1088 Budapest, Múzeum krt. 4/A
(36-1) 411-6500 / 5099, 5353 | mnyfi@btk.elte.hu

CLUNL
Linguistics Research Centre of NOVA University Lisbon
NOVA University Lisbon, School of Social Sciences and Humanities
Colégio Almada Negreiros, gabinete 343, 1099-032 Lisboa
(+351) 217908309 | Ext.: 40325 | sec.clunl@fcsh.unl.pt

CERTIFICATE OF PARTICIPATION

We hereby certify that **NAME has completed an auditory perceptual training in Portuguese as Foreign Language**, in the context of a research conducted by CLUNL - Linguistics Research Centre of NOVA University Lisbon, in collaboration with the Department of Applied Linguistics and Phonetics of the Faculty of Humanities of Eötvös Loránd University of Budapest. The specific focus of the training was European Portuguese phonetics and phonology.

The training took place between (month) and (month) (year).

CLUNL - Linguistics Research Centre of NOVA University Lisbon

Department of Applied Linguistics and Phonetics of ELTE University

Department of
Applied Linguistics and Phonetics
Eötvös Loránd University, Faculty of Humanities
H-1088 Budapest, Múzeum krt. 4/A
(36-1) 411-6500 / 5099, 5353 | mnyfi@btk.elte.hu

Linguistics Research Centre of NOVA University Lisbon
NOVA University Lisbon, School of Social Sciences and Humanities
Colégio Almada Negreiros, gabinete 343, 1099-032 Lisboa
(+351) 217908309 | Ext.: 40325 | sec.clunl@fcsh.unl.pt

CERTIFICATE OF PARTICIPATION

We hereby certify that **NAME has participated in the project "Perceção de acento e redução vocálica na aquisição fonológica em PL2. Implicações para a prática pedagógica com aprendentes húngaros"**. This research is conducted by CLUNL - Linguistics Research Centre of NOVA University Lisbon, in collaboration with the Department of Applied Linguistics and Phonetics of the Faculty of Humanities of Eötvös Loránd University of Budapest. This research project has received a grant from the Portuguese Foundation for Science and Technology (2020.05740.BD).

Date,

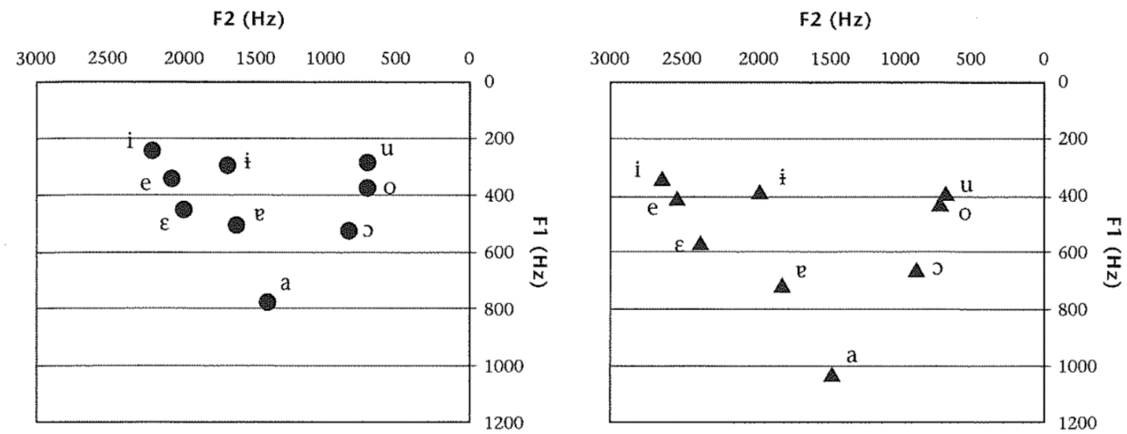
CLUNL - Linguistics Research Centre of NOVA University Lisbon

Department of Applied Linguistics and Phonetics of ELTE University

Appendix IV

Acoustic values for standard EP vowels

A. EP oral vowel space, for native male productions (on the left) and native female productions (on the right), retrived from Andrade (2020, p. 3251)



B. F1 and F2 values for the EP oral vowels, calculated from the graphics above

MALE SPEAKERS

	F1	F2
[a]	772	1369
[ɐ]	508	1630
[ɛ]	440	1931
[e]	332	2011
[ɨ]	292	1687
[i]	232	2137
[ɔ]	519	825
[o]	365	697
[u]	282	697

FEMALE SPEAKERS

	F1	F2
[a]	1023	1468
[ɐ]	725	1816
[ɛ]	566	2376
[e]	420	2550
[ɨ]	395	1974
[i]	348	2645
[ɔ]	663	884
[o]	478	726
[u]	389	687

Appendix V

Carrier sentences for stimuli recordings

A. Carrier sentences for monosyllabic stimuli

(GATO) **[ga]** . Diga **[ga]** , por favor: **[ga]** .
(GAVETA) **[gɐ]** . Diga **[gɐ]** , por favor: **[gɐ]** .
(GUERRA) **[gɛ]** . Diga **[gɛ]** , por favor: **[gɛ]** .
(GUE) **[ge]** . Diga **[ge]** , por favor: **[ge]** .
(GUERREAR) **[gɨ]** . Diga **[gɨ]** , por favor: **[gɨ]** .
(GUITO) **[gi]** . Diga **[gi]** , por favor: **[gi]** .
(GOLA) **[gɔ]** . Diga **[gɔ]** , por favor: **[gɔ]** .
(GOLO) **[go]** . Diga **[go]** , por favor: **[go]** .
(GULA) **[gu]** . Diga **[gu]** , por favor: **[gu]** .

(SATO) **[sa]** . Diga **[sa]** , por favor: **[sa]** .
(SAPATO) **[sɐ]** . Diga **[sɐ]** , por favor: **[sɐ]** .
(SETA) **[sɛ]** . Diga **[sɛ]** , por favor: **[sɛ]** .
(SEDA) **[se]** . Diga **[se]** , por favor: **[se]** .
(SENÃO) **[sɨ]** . Diga **[sɨ]** , por favor: **[sɨ]** .
(SICAL) **[si]** . Diga **[si]** , por favor: **[si]** .

(CASACO) **[za]** . Diga **[za]** , por favor: **[za]** .
(CASA) **[zɐ]** . Diga **[zɐ]** , por favor: **[zɐ]** .
(JOSÉ) **[zɛ]** . Diga **[zɛ]** , por favor: **[zɛ]** .
(AZEDO) **[ze]** . Diga **[ze]** , por favor: **[ze]** .
(ALISE) **[zɨ]** . Diga **[zɨ]** , por favor: **[zɨ]** .
(ASILO) **[zi]** . Diga **[zi]** , por favor: **[zi]** .
(CASOTA) **[zɔ]** . Diga **[zɔ]** , por favor: **[zɔ]** .
(ZONA) **[zo]** . Diga **[zo]** , por favor: **[zo]** .
(CASUAL) **[zu]** . Diga **[zu]** , por favor: **[zu]** .

(TAPA) **[ta]** . Diga **[ta]** , por favor: **[ta]** .
(TAREFA) **[tɐ]** . Diga **[tɐ]** , por favor: **[tɐ]** .
(TELA) **[tɛ]** . Diga **[tɛ]** , por favor: **[tɛ]** .
(TEMA) **[te]** . Diga **[te]** , por favor: **[te]** .
(TERROR) **[tɨ]** . Diga **[tɨ]** , por favor: **[tɨ]** .
(TINA) **[ti]** . Diga **[ti]** , por favor: **[ti]** .
(TOMA) **[tɔ]** . Diga **[tɔ]** , por favor: **[tɔ]** .
(TODA) **[to]** . Diga **[to]** , por favor: **[to]** .
(TUDO) **[tu]** . Diga **[tu]** , por favor: **[tu]** .

B. Carrier sentences for dissyllabic stimuli

(ZAZÁ)	**[ˈzazɐ]**	. Diga	**[ˈzazɐ]**	, por favor:	**[ˈzazɐ]**	.
(ZÁZA)	**[zɐˈza]**	. Diga	**[zɐˈza]**	, por favor:	**[zɐˈza]**	.
(ZÁZE)	**[ˈzazɨ]**	. Diga	**[ˈzazɨ]**	, por favor:	**[ˈzazɨ]**	.
(ZAZÉ)	**[zɐˈzɛ]**	. Diga	**[zɐˈzɛ]**	, por favor:	**[zɐˈzɛ]**	.
(ZAZÊ)	**[zɐˈze]**	. Diga	**[zɐˈze]**	, por favor:	**[zɐˈze]**	.
(ZÁZI)	**[zɐˈzi]**	. Diga	**[zɐˈzi]**	, por favor:	**[zɐˈzi]**	.
(ZAZI)	**[ˈzazi]**	. Diga	**[ˈzazi]**	, por favor:	**[ˈzazi]**	.
(ZÁZO)	**[ˈzazu]**	. Diga	**[ˈzazu]**	, por favor:	**[ˈzazu]**	.
(ZAZÓ)	**[zɐˈzɔ]**	. Diga	**[zɐˈzɔ]**	, por favor:	**[zɐˈzɔ]**	.
(ZÉZA)	**[ˈzɛzɐ]**	. Diga	**[ˈzɛzɐ]**	, por favor:	**[ˈzɛzɐ]**	.
(ZEZÁ)	**[zɨˈza]**	. Diga	**[zɨˈza]**	, por favor:	**[zɨˈza]**	.
(ZÉZE)	**[ˈzɛzɨ]**	. Diga	**[ˈzɛzɨ]**	, por favor:	**[ˈzɛzɨ]**	.
(ZEZÉ)	**[zɨˈzɛ]**	. Diga	**[zɨˈzɛ]**	, por favor:	**[zɨˈzɛ]**	.
(ZÊZE)	**[ˈzezɨ]**	. Diga	**[ˈzezɨ]**	, por favor:	**[ˈzezɨ]**	.
(ZEZÊ)	**[zɨˈze]**	. Diga	**[zɨˈze]**	, por favor:	**[zɨˈze]**	.
(ZÉZI)	**[ˈzɛzi]**	. Diga	**[ˈzɛzi]**	, por favor:	**[ˈzɛzi]**	.
(ZÊZI)	**[ˈzezi]**	. Diga	**[ˈzezi]**	, por favor:	**[ˈzezi]**	.
(ZEZI)	**[zɨˈzi]**	. Diga	**[zɨˈzi]**	, por favor:	**[zɨˈzi]**	.
(ZÉZU)	**[ˈzɛzu]**	. Diga	**[ˈzɛzu]**	, por favor:	**[ˈzɛzu]**	.
(ZEZÓ)	**[zɨˈzɔ]**	. Diga	**[zɨˈzɔ]**	, por favor:	**[zɨˈzɔ]**	.
(ZIZA)	**[ˈzizɐ]**	. Diga	**[ˈzizɐ]**	, por favor:	**[ˈzizɐ]**	.
(ZIZÁ)	**[ziˈza]**	. Diga	**[ziˈza]**	, por favor:	**[ziˈza]**	.
(ZIZE)	**[ˈzizɨ]**	. Diga	**[ˈzizɨ]**	, por favor:	**[ˈzizɨ]**	.
(ZIZÉ)	**[ziˈzɛ]**	. Diga	**[ziˈzɛ]**	, por favor:	**[ziˈzɛ]**	.
(ZIZÊ)	**[ziˈze]**	. Diga	**[ziˈze]**	, por favor:	**[ziˈze]**	.
(ZIZO)	**[ˈzizu]**	. Diga	**[ˈzizu]**	, por favor:	**[ˈzizu]**	.
(ZIZÓ)	**[ziˈzɔ]**	. Diga	**[ziˈzɔ]**	, por favor:	**[ziˈzɔ]**	.
(ZÓZI)	**[ˈzɔzi]**	. Diga	**[ˈzɔzi]**	, por favor:	**[ˈzɔzi]**	.
(ZOZI)	**[zuˈzi]**	. Diga	**[zuˈzi]**	, por favor:	**[zuˈzi]**	.

C. Carrier sentences for trisyllabic stimuli

[ˈbikulu]	**BICULO**	. Diga	**BICULO**	, por favor:	**BICULO**	.
[biˈkulu]	**BICULO**	. Diga	**BICULO**	, por favor:	**BICULO**	.
[bikuˈlu]	**BICULU**	. Diga	**BICULU**	, por favor:	**BICULU**	.
[ˈbifuʃi]	**BIFUCHI**	. Diga	**BIFUCHI**	, por favor:	**BIFUCHI**	.

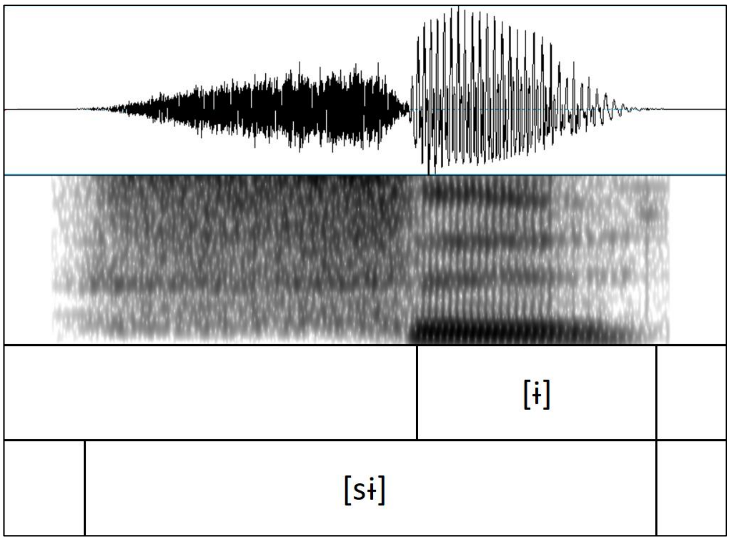
[biˈfuʃi]	**BIFUCHI**	. Diga	**BIFUCHI**	, por favor:	**BIFUCHI**	.
[bifuˈʃi]	**BIFUCHI**	. Diga	**BIFUCHI**	, por favor:	**BIFUCHI**	.
[ˈdutiku]	**DUTICO**	. Diga	**DUTICO**	, por favor:	**DUTICO**	.
[duˈtiku]	**DUTICO**	. Diga	**DUTICO**	, por favor:	**DUTICO**	.
[dutiˈku]	**DUTICU**	. Diga	**DUTICU**	, por favor:	**DUTICU**	.
[ˈfiluvi]	**FILUVI**	. Diga	**FILUVI**	, por favor:	**FILUVI**	.
[fiˈluvi]	**FILUVI**	. Diga	**FILUVI**	, por favor:	**FILUVI**	.
[filuˈvi]	**FILUVI**	. Diga	**FILUVI**	, por favor:	**FILUVI**	.
[ˈʒurigu]	**JURIGO**	. Diga	**JURIGO**	, por favor:	**JURIGO**	.
[ʒuˈrigu]	**JURIGO**	. Diga	**JURIGO**	, por favor:	**JURIGO**	.
[ʒuriˈgu]	**JURIGU**	. Diga	**JURIGU**	, por favor:	**JURIGU**	.
[ˈkitumi]	**QUITUMI**	. Diga	**QUITUMI**	, por favor:	**QUITUMI**	.
[kiˈtumi]	**QUITUMI**	. Diga	**QUITUMI**	, por favor:	**QUITUMI**	.
[kituˈmi]	**QUITUMI**	. Diga	**QUITUMI**	, por favor:	**QUITUMI**	.
[ˈlutinu]	**LUTINO**	. Diga	**LUTINO**	, por favor:	**LUTINO**	.
[luˈtinu]	**LUTINO**	. Diga	**LUTINO**	, por favor:	**LUTINO**	.
[lutiˈnu]	**LUTINU**	. Diga	**LUTINU**	, por favor:	**LUTINU**	.
[ˈmudini]	**MUDINI**	. Diga	**MUDINI**	, por favor:	**MUDINI**	.
[muˈdini]	**MUDINI**	. Diga	**MUDINI**	, por favor:	**MUDINI**	.
[mudiˈni]	**MUDINI**	. Diga	**MUDINI**	, por favor:	**MUDINI**	.
[ˈmupiʃu]	**MUPICHO**	. Diga	**MUPICHO**	, por favor:	**MUPICHO**	.
[muˈpiʃu]	**MUPICHO**	. Diga	**MUPICHO**	, por favor:	**MUPICHO**	.
[mupiˈʃu]	**MUPICHU**	. Diga	**MUPICHU**	, por favor:	**MUPICHU**	.
[ˈkiruʃi]	**QUIRUCHI**	. Diga	**QUIRUCHI**	, por favor:	**QUIRUCHI**	.
[kiˈruʃi]	**QUIRUCHI**	. Diga	**QUIRUCHI**	, por favor:	**QUIRUCHI**	.
[kiruˈʃi]	**QUIRUCHI**	. Diga	**QUIRUCHI**	, por favor:	**QUIRUCHI**	.
[ˈtikuru]	**TICURO**	. Diga	**TICURO**	, por favor:	**TICURO**	.
[tiˈkuru]	**TICURO**	. Diga	**TICURO**	, por favor:	**TICURO**	.
[tikuˈru]	**TICURU**	. Diga	**TICURU**	, por favor:	**TICURU**	.
[ˈtudiki]	**TUDIQUI**	. Diga	**TUDIQUI**	, por favor:	**TUDIQUI**	.
[tuˈdiki]	**TUDIQUI**	. Diga	**TUDIQUI**	, por favor:	**TUDIQUI**	.
[tudiˈki]	**TUDIQUI**	. Diga	**TUDIQUI**	, por favor:	**TUDIQUI**	.
[ˈtuduri]	**TUDURI**	. Diga	**TUDURI**	, por favor:	**TUDURI**	.
[tuˈduri]	**TUDURI**	. Diga	**TUDURI**	, por favor:	**TUDURI**	.
[tuduˈri]	**TUDURI**	. Diga	**TUDURI**	, por favor:	**TUDURI**	.
[ˈvumipi]	**VUMIPI**	. Diga	**VUMIPI**	, por favor:	**VUMIPI**	.
[vuˈmipi]	**VUMIPI**	. Diga	**VUMIPI**	, por favor:	**VUMIPI**	.
[vumiˈpi]	**VUMIPI**	. Diga	**VUMIPI**	, por favor:	**VUMIPI**	.
[ˈzituli]	**ZITULI**	. Diga	**ZITULI**	, por favor:	**ZITULI**	.

[ziˈtuli]	**ZITULI**	. Diga	**ZITULI**	, por favor:	**ZITULI**	.
[zituˈli]	**ZITULI**	. Diga	**ZITULI**	, por favor:	**ZITULI**	.
[buˈpili]	**BUPILI**	. Diga	**BUPILI**	, por favor:	**BUPILI**	.
[buˈʃili]	**BUCHILI**	. Diga	**BUCHILI**	, por favor:	**BUCHILI**	.
[tiˈfubi]	**TIFUBI**	. Diga	**TIFUBI**	, por favor:	**TIFUBI**	.
[tiˈvubi]	**TIVUBI**	. Diga	**TIVUBI**	, por favor:	**TIVUBI**	.
[tiˈburi]	**TIBURI**	. Diga	**TIBURI**	, por favor:	**TIBURI**	.
[tiˈʃuri]	**TICHURI**	. Diga	**TICHURI**	, por favor:	**TICHURI**	.
[kuˈmitu]	**CUMITO**	. Diga	**CUMITO**	, por favor:	**CUMITO**	.
[kuˈfitu]	**CUFITO**	. Diga	**CUFITO**	, por favor:	**CUFITO**	.
[diˈpuki]	**DIPUQUI**	. Diga	**DIPUQUI**	, por favor:	**DIPUQUI**	.
[diˈbuki]	**DIBUQUI**	. Diga	**DIBUQUI**	, por favor:	**DIBUQUI**	.
[zuˈmitu]	**ZUMITO**	. Diga	**ZUMITO**	, por favor:	**ZUMITO**	.
[zuˈnitu]	**ZUNITO**	. Diga	**ZUNITO**	, por favor:	**ZUNITO**	.

Appendix VI

Acoustic measurements

A. Screenshot of the labbeling in Praat for the token [sɨ] produced by talker T5^{F}



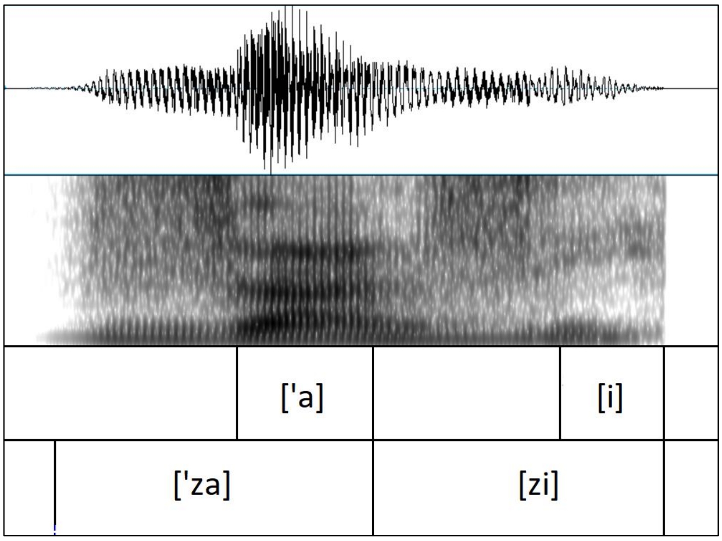
B. Acoustic measurements for the monosyllabic stimuli

	syllable	vowel						
tokens	duration (ms)	duration (ms)	Pitch (Hz)	F1 (Hz)	F2 (Hz)	F3 (Hz)	F4 (Hz)	Intensity (db)
[ga] T1^{F}	314	266	154	853	1645	2646	3593	83
[ga] T2^{F}	304	266	166	856	1487	2568	4388	87
[ga] T3^{F}	299	252	201	840	1497	2864	4227	85
[gɐ] T1^{F}	339	273	162	593	1960	2187	2940	85
[gɐ] T2^{F}	334	262	176	743	2040	2866	4249	88
[gɐ] T3^{F}	293	243	202	573	1901	2907	4304	91
[gɛ] T1^{F}	308	249	155	499	2373	2869	3954	88
[gɛ] T2^{F}	334	292	169	574	2601	3097	4506	87
[gɛ] T3^{F}	298	254	211	510	2375	2728	4282	90
[ge] T1^{F}	347	259	188	399	2692	3300	4560	88
[ge] T2^{F}	402	305	157	485	2756	3117	4586	88
[ge] T3^{F}	336	289	199	398	2798	3173	4529	90
[gɨ] T1^{F}	318	255	192	302	2256	2635	4227	91
[gɨ] T2^{F}	374	250	176	437	2358	2789	4312	88
[gɨ] T3^{F}	282	229	225	404	2093	2815	4728	89
[gi] T1^{F}	327	250	200	310	2718	3540	4563	93
[gi] T2^{F}	333	214	188	449	2941	3460	4806	88
[gi] T3^{F}	293	256	232	393	2863	3664	4784	91
[gɔ] T1^{F}	325	277	169	569	995	2616	3804	88

[gɔ] T2^F	352	295	189	609	1037	2869	3840	90
[gɔ] T3^F	294	252	200	593	1041	2751	4029	89
[go] T1^F	312	220	207	299	692	2571	3719	90
[go] T2^F	334	264	190	490	850	2777	3924	88
[go] T3^F	323	258	188	384	739	2820	4327	91
[gu] T1^F	321	254	197	276	612	2222	3653	89
[gu] T2^F	342	271	210	393	799	2618	3863	91
[gu] T3^F	292	258	203	380	742	2920	4196	89
[sa] T5^F	556	260	159	714	1758	2736	4239	74
[sa] T6^F	415	195	147	660	1846	3116	4417	73
[sa] T7^M	376	169	99	545	1607	2274	3869	73
[sɐ] T5^F	492	274	153	802	1506	2615	4038	74
[sɐ] T6^F	405	203	148	879	1509	2610	4105	73
[sɐ] T7^M	449	232	103	788	1391	2337	3775	77
[sɛ] T5^F	520	267	157	549	2260	2821	4635	75
[sɛ] T6^F	389	187	142	536	2406	2968	4452	72
[sɛ] T7^M	430	207	96	354	2058	2467	3748	76
[se] T5^F	556	300	156	431	2470	2896	4714	74
[se] T6^F	405	194	141	407	2672	3089	4346	74
[se] T7^M	466	240	113	344	2309	2936	4193	75
[sɨ] T5^F	582	329	159	404	1861	2682	4655	72
[sɨ] T6^F	457	192	161	371	2049	3181	4353	75
[sɨ] T7^M	500	238	118	284	2116	2586	4035	74
[si] T5^F	528	312	164	367	2729	3241	4789	72
[si] T6^F	396	169	147	321	2959	3430	4401	72
[si] T7^M	431	217	111	266	2200	3228	3994	73
[za] T1^F	452	270	138	854	1563	2670	4083	80
[za] T2^F	535	348	158	894	1565	2760	3924	79
[za] T3^F	474	261	157	760	1496	2854	4024	76
[zɐ] T1^F	455	248	141	644	1854	1930	2972	82
[zɐ] T2^F	561	369	161	738	1896	2921	4243	78
[zɐ] T3^F	501	276	170	507	1728	3012	4640	79
[zɛ] T1^F	466	254	136	503	2382	2892	4216	77
[zɛ] T2^F	539	339	159	544	2476	2980	4421	79
[zɛ] T3^F	489	233	160	419	2365	2847	4042	79
[ze] T1^F	434	247	141	388	2637	3203	4147	78
[ze] T2^F	535	341	155	476	2494	2860	4397	78
[ze] T3^F	472	254	175	382	2778	3246	4821	79
[zɨ] T1^F	527	268	147	387	1973	2818	4511	77

[zɨ] T2^F	587	361	162	459	2001	2768	4485	78
[zɨ] T3^F	557	280	173	372	1920	3028	4806	78
[zi] T1^F	450	264	145	331	2756	3495	4634	77
[zi] T2^F	580	358	161	397	2883	3213	4840	77
[zi] T3^F	504	253	186	378	2881	3757	4861	76
[zɔ] T1^F	419	237	145	555	1080	2702	3576	80
[zɔ] T2^F	485	321	152	631	1114	2731	3891	78
[zɔ] T3^F	524	251	174	518	1163	3033	3876	78
[zo] T1^F	517	257	141	364	799	2672	3926	76
[zo] T2^F	607	401	158	459	947	2697	3643	76
[zo] T3^F	506	242	160	412	826	3026	4040	76
[zu] T1^F	474	245	157	296	602	2681	3527	79
[zu] T2^F	531	316	166	418	861	2690	4026	79
[zu] T3^F	498	209	164	386	793	3018	4260	79

C. Screenshot of the labbeling in Praat for the token ['zazi] produced by talker T1^F



D. Acoustic measurements for the disyllabic stimuli

token	location of the measured vowel	syllable duration (ms)	vowel						
			duration (ms)	Pitch (Hz)	F1 (Hz)	F2 (Hz)	F3 (Hz)	F4 (Hz)	Intensity (db)
['zazɐ] T1^F	1st syllable	369	164	163	751	1597	2854	4329	79
['zazɐ] T1^F	2nd syllable	378	236	124	517	1736	2987	4147	70
['zazɐ] T2^F	1st syllable	380	209	198	891	1442	2647	2790	80
['zazɐ] T2^F	2nd syllable	502	357	160	746	1696	2948	4115	69
['zazɐ] T3^F	1st syllable	428	195	206	818	1462	2931	4564	78

['zazɐ] T3F	2nd syllable	363	200	148	299	1770	2933	4317	67
['zazɨ] T1F	1st syllable	369	172	154	821	1591	2889	4637	78
['zazɨ] T1F	2nd syllable	298	169	123	178	1507	2451	3625	69
['zazɨ] T2F	1st syllable	374	198	198	897	1606	2472	2933	79
['zazɨ] T2F	2nd syllable	428	311	148	444	1563	2681	4330	67
['zazɨ] T3F	1st syllable	472	224	199	675	1544	2974	4774	77
['zazɨ] T3F	2nd syllable	403	202	undefined	843	1770	3143	4096	56
['zazi] T1F	1st syllable	295	128	152	702	1588	2716	3466	80
['zazi] T1F	2nd syllable	263	96	171	264	2658	3168	4784	70
['zazi] T2F	1st syllable	334	181	198	866	1608	2449	3920	79
['zazi] T2F	2nd syllable	424	259	149	424	2651	2996	4743	69
['zazi] T3F	1st syllable	448	217	199	763	1411	2785	4055	76
['zazi] T3F	2nd syllable	333	173	149	349	2853	3314	4629	68
[zɐ'za] T1F	1st syllable	305	93	172	484	1768	3017	4599	77
[zɐ'za] T1F	2nd syllable	390	277	126	885	1639	2855	4176	75
[zɐ'za] T2F	1st syllable	302	140	214	609	1503	2679	2860	78
[zɐ'za] T2F	2nd syllable	479	384	159	911	1486	2669	3850	76
[zɐ'za] T3F	1st syllable	300	92	244	463	1701	3151	4794	77
[zɐ'za] T3F	2nd syllable	381	254	149	544	1445	2935	4071	72
[zɐ'zɛ] T1F	1st syllable	252	87	217	441	1736	2275	3035	80
[zɐ'zɛ] T1F	2nd syllable	457	331	132	491	2318	2845	4164	75
[zɐ'zɛ] T2F	1st syllable	299	155	230	527	1831	2878	4256	79
[zɐ'zɛ] T2F	2nd syllable	382	288	154	577	2328	2862	4387	75
[zɐ'zɛ] T3F	1st syllable	324	136	245	479	1816	3255	4884	76
[zɐ'zɛ] T3F	2nd syllable	396	258	162	453	2407	2953	4341	74
[zɐ'ze] T1F	1st syllable	320	92	181	469	1953	2947	4740	79
[zɐ'ze] T1F	2nd syllable	397	276	131	376	2621	3168	4449	75
[zɐ'ze] T2F	1st syllable	334	123	227	508	1829	2836	4246	78
[zɐ'ze] T2F	2nd syllable	544	401	151	451	2571	2873	4495	73
[zɐ'ze] T3F	1st syllable	314	120	274	531	1832	3097	4744	76
[zɐ'ze] T3F	2nd syllable	372	241	157	406	2729	3115	4163	76
[zɐ'zi] T1F	1st syllable	303	88	207	414	1856	2765	2988	81
[zɐ'zi] T1F	2nd syllable	403	261	130	304	2766	3560	4734	68
[zɐ'zi] T2F	1st syllable	286	120	208	516	1864	2843	4259	79
[zɐ'zi] T2F	2nd syllable	511	398	160	374	2891	3235	4708	72
[zɐ'zi] T3F	1st syllable	333	115	241	479	1705	3084	4902	79
[zɐ'zi] T3F	2nd syllable	366	208	154	340	2894	3428	4143	71
['zezɨ] T1F	1st syllable	347	128	182	432	2364	3038	4627	79
['zezɨ] T1F	2nd syllable	325	159	127	467	1825	3032	4605	68

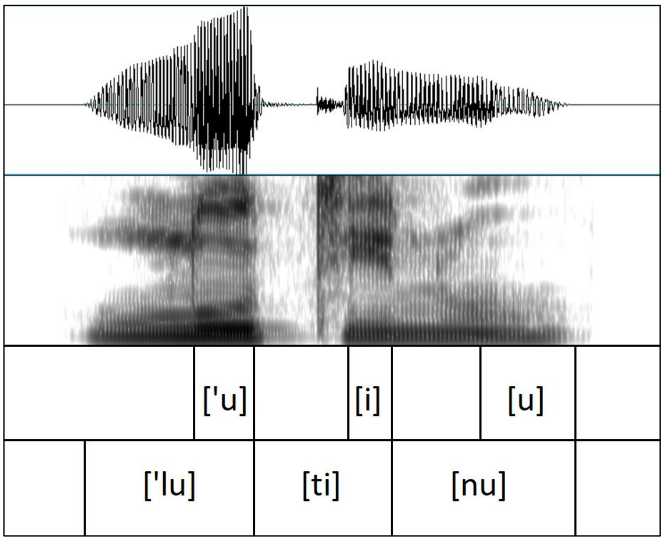
['zezɨ] T2^F	1^st syllable	346	176	191	453	2250	2691	4316	79
['zezɨ] T2^F	2^nd syllable	425	244	157	333	1669	2911	3856	68
['zezɨ] T3^F	1^st syllable	445	193	225	445	2568	3121	4359	76
['zezɨ] T3^F	2^nd syllable	357	180	152	248	1879	3166	4972	65
['zezi] T1^F	1^st syllable	339	141	204	402	2469	3005	4724	78
['zezi] T1^F	2^nd syllable	328	185	129	237	2641	3336	4679	66
['zezi] T2^F	1^st syllable	401	208	198	455	2363	2895	4466	81
['zezi] T2^F	2^nd syllable	408	297	171	415	2652	3094	4664	69
['zezi] T3^F	1^st syllable	435	213	233	465	2605	3104	4630	77
['zezi] T3^F	2^nd syllable	355	179	167	371	2800	3251	4590	68
['zɛzɐ] T1^F	1^st syllable	330	137	160	535	2077	2832	4471	79
['zɛzɐ] T1^F	2^nd syllable	332	214	123	1556	1756	3230	4207	70
['zɛzɐ] T2^F	1^st syllable	366	208	188	596	2024	2233	2800	78
['zɛzɐ] T2^F	2^nd syllable	465	372	158	632	1680	2907	4084	73
['zɛzɐ] T3^F	1^st syllable	454	240	205	570	2227	2939	4585	76
['zɛzɐ] T3^F	2^nd syllable	359	211	156	352	1787	3026	4911	68
['zɛzɨ] T1^F	1^st syllable	343	160	178	527	2124	2707	4359	79
['zɛzɨ] T1^F	2^nd syllable	311	165	131	231	1733	3085	4437	69
['zɛzɨ] T2^F	1^st syllable	422	225	189	604	2124	2466	2826	78
['zɛzɨ] T2^F	2^nd syllable	499	325	158	191	1717	2922	4066	68
['zɛzɨ] T3^F	1^st syllable	438	227	210	619	2210	3084	4444	75
['zɛzɨ] T3^F	2^nd syllable	313	159	152	418	1851	3172	5113	67
['zɛzi] T1^F	1^st syllable	321	144	191	549	2189	2947	4348	78
['zɛzi] T1^F	2^nd syllable	342	195	135	251	2629	3281	4845	66
['zɛzi] T2^F	1^st syllable	293	177	181	562	2285	2780	3084	79
['zɛzi] T2^F	2^nd syllable	380	273	157	426	2689	3018	4690	71
['zɛzi] T3^F	1^st syllable	447	247	211	519	2333	2972	4635	77
['zɛzi] T3^F	2^nd syllable	347	199	160	415	2702	3168	4468	69
['zɛzu] T1^F	1^st syllable	319	133	188	551	2059	2819	4273	79
['zɛzu] T1^F	2^nd syllable	349	208	137	333	1875	3185	4007	70
['zɛzu] T2^F	1^st syllable	374	208	194	580	2237	2798	4199	79
['zɛzu] T2^F	2^nd syllable	456	333	161	247	962	2634	4116	71
['zɛzu] T3^F	1^st syllable	487	233	215	501	2248	2962	4756	78
['zɛzu] T3^F	2^nd syllable	356	187	159	326	1152	3308	3732	68
[zɨ'za] T1^F	1^st syllable	333	70	215	390	1801	3025	4583	78
[zɨ'za] T1^F	2^nd syllable	371	249	130	1008	1613	2952	4237	75
[zɨ'za] T2^F	1^st syllable	332	104	219	437	1751	2843	4380	78
[zɨ'za] T2^F	2^nd syllable	554	430	159	896	1555	2740	4050	75
[zɨ'za] T3^F	1^st syllable	345	105	242	451	1703	3094	4791	76

[zɨ'za] T3^F	2nd syllable	439	248	154	482	1417	2839	4308	72
[zɨ'zɛ] T1^F	1st syllable	282	62	196	300	1899	2983	4577	77
[zɨ'zɛ] T1^F	2nd syllable	465	307	128	506	2401	2914	4239	78
[zɨ'zɛ] T2^F	1st syllable	312	110	235	453	1798	2007	2751	78
[zɨ'zɛ] T2^F	2nd syllable	497	370	147	541	2330	2821	4292	75
[zɨ'zɛ] T3^F	1st syllable	351	126	254	496	1793	3088	4974	73
[zɨ'zɛ] T3^F	2nd syllable	427	240	162	438	2453	2855	4259	76
[zɨ'ze] T1^F	1st syllable	286	61	226	292	1832	2926	4494	78
[zɨ'ze] T1^F	2nd syllable	400	243	125	376	2642	3183	4633	74
[zɨ'ze] T2^F	1st syllable	305	123	253	481	1844	2420	2785	79
[zɨ'ze] T2^F	2nd syllable	387	284	158	461	2389	2817	4467	76
[zɨ'ze] T3^F	1st syllable	408	163	253	308	1734	3052	4517	73
[zɨ'ze] T3^F	2nd syllable	423	252	156	413	2674	3110	4361	76
[zɨ'zi] T1^F	1st syllable	313	66	229	292	1870	2934	4567	77
[zɨ'zi] T1^F	2nd syllable	367	204	123	258	2804	3501	4704	73
[zɨ'zi] T2^F	1st syllable	312	138	247	470	1933	2746	2792	79
[zɨ'zi] T2^F	2nd syllable	521	359	160	368	2867	3248	4662	75
[zɨ'zi] T3^F	1st syllable	375	105	262	328	1737	3028	4782	75
[zɨ'zi] T3^F	2nd syllable	410	213	161	338	2854	3334	4738	74
['zizɐ] T1^F	1st syllable	327	96	198	354	2575	3202	4739	78
['zizɐ] T1^F	2nd syllable	354	203	127	886	1862	3030	4470	70
['zizɐ] T2^F	1st syllable	309	138	222	440	2637	3038	4587	78
['zizɐ] T2^F	2nd syllable	500	355	158	661	1797	2976	4120	73
['zizɐ] T3^F	1st syllable	460	177	208	356	2757	3251	4980	76
['zizɐ] T3^F	2nd syllable	403	231	148	569	1791	3025	4323	70
['zizɨ] T1^F	1st syllable	359	108	204	335	2642	3054	4704	77
['zizɨ] T1^F	2nd syllable	323	162	126	359	1891	3280	4784	72
['zizɨ] T2^F	1st syllable	375	182	229	427	2595	2973	4538	77
['zizɨ] T2^F	2nd syllable	434	284	155	457	1903	2730	4307	73
['zizɨ] T3^F	1st syllable	472	182	213	291	2795	3512	4941	76
['zizɨ] T3^F	2nd syllable	402	227	159	427	1928	3236	5167	71
[zi'za] T1^F	1st syllable	267	72	190	318	2282	3019	4744	77
[zi'za] T1^F	2nd syllable	379	269	140	987	1738	2769	4143	77
[zi'za] T2^F	1st syllable	288	135	225	440	2613	3074	4721	77
[zi'za] T2^F	2nd syllable	462	355	144	843	1624	2583	3982	74
[zi'za] T3^F	1st syllable	336	110	231	430	2717	3177	4991	75
[zi'za] T3^F	2nd syllable	418	267	154	863	1659	2971	4343	73
[zi'zɛ] T1^F	1st syllable	324	98	206	355	2665	3151	4629	78
[zi'zɛ] T1^F	2nd syllable	369	251	137	494	2424	2943	4220	77

[zi'zɛ] T2^{F}	1st syllable	287	116	229	444	2590	2991	4405	77
[zi'zɛ] T2^{F}	2nd syllable	435	333	147	538	2434	2913	4296	75
[zi'zɛ] T3^{F}	1st syllable	357	120	245	437	2695	3248	4944	73
[zi'zɛ] T3^{F}	2nd syllable	412	252	163	459	2496	3003	4346	76
[zi'ze] T1^{F}	1st syllable	317	82	218	250	2623	3207	4776	79
[zi'ze] T1^{F}	2nd syllable	359	199	124	361	2641	3210	4316	74
[zi'ze] T2^{F}	1st syllable	318	151	227	443	2594	3137	4520	76
[zi'ze] T2^{F}	2nd syllable	484	361	156	457	2477	2881	4394	75
[zi'ze] T3^{F}	1st syllable	360	155	243	384	2695	3312	4749	75
[zi'ze] T3^{F}	2nd syllable	416	232	157	431	2736	3114	4318	74
['zazu] T1^{F}	1st syllable	369	176	156	842	1657	2798	4170	80
['zazu] T1^{F}	2nd syllable	327	172	126	332	1585	3297	4175	66
['zazu] T2^{F}	1st syllable	362	224	201	863	1717	2915	3865	79
['zazu] T2^{F}	2nd syllable	509	374	159	201	1028	2859	4321	68
['zazu] T3^{F}	1st syllable	465	264	202	769	1484	2889	4691	77
['zazu] T3^{F}	2nd syllable	310	169	153	237	1143	3196	3721	68
[zɐ'zɔ] T1^{F}	1st syllable	295	115	158	470	1747	2937	4651	78
[zɐ'zɔ] T1^{F}	2nd syllable	326	224	129	569	1038	2836	3823	76
[zɐ'zɔ] T2^{F}	1st syllable	301	154	227	603	1815	2899	4125	77
[zɐ'zɔ] T2^{F}	2nd syllable	439	329	155	656	1076	2791	3771	74
[zɐ'zɔ] T3^{F}	1st syllable	275	102	245	463	1647	3102	4809	77
[zɐ'zɔ] T3^{F}	2nd syllable	381	232	148	417	1125	2834	3719	71
[zɨ'zɔ] T1^{F}	1st syllable	316	84	233	365	1791	3049	4633	76
[zɨ'zɔ] T1^{F}	2nd syllable	435	262	127	600	1060	2648	3622	75
[zɨ'zɔ] T2^{F}	1st syllable	328	119	232	457	1724	2459	2747	79
[zɨ'zɔ] T2^{F}	2nd syllable	495	376	149	607	1073	2827	3785	75
[zɨ'zɔ] T3^{F}	1st syllable	367	125	253	354	1634	3206	4787	74
[zɨ'zɔ] T3^{F}	2nd syllable	471	271	155	512	1174	3005	3772	74
[zi'zɔ] T1^{F}	1st syllable	277	85	197	344	2607	3124	4706	77
[zi'zɔ] T1^{F}	2nd syllable	401	268	135	536	1012	2810	3687	77
[zi'zɔ] T2^{F}	1st syllable	274	119	223	441	2431	2815	4497	79
[zi'zɔ] T2^{F}	2nd syllable	470	360	152	630	1103	2722	3807	74
[zi'zɔ] T3^{F}	1st syllable	316	164	238	444	2592	3215	4805	75
[zi'zɔ] T3^{F}	2nd syllable	389	253	155	487	1205	3022	3727	74
['zizu] T1^{F}	1st syllable	330	109	214	264	2599	3108	4543	78
['zizu] T1^{F}	2nd syllable	327	184	123	364	1582	3636	3969	68
['zizu] T2^{F}	1st syllable	371	176	224	434	2568	2924	4580	78
['zizu] T2^{F}	2nd syllable	460	331	160	457	977	2959	4183	72
['zizu] T3^{F}	1st syllable	383	142	201	368	2758	3278	5096	76

['zizu] T3^F	2nd syllable	406	217	154	240	1076	2842	4098	71
['zɔzi] T1^F	1st syllable	313	133	171	586	1140	2707	3516	80
['zɔzi] T1^F	2nd syllable	258	143	147	325	2567	3198	4631	71
['zɔzi] T2^F	1st syllable	348	179	217	727	1318	2877	3921	79
['zɔzi] T2^F	2nd syllable	358	235	176	419	2685	3057	4653	68
['zɔzi] T3^F	1st syllable	417	213	205	568	1203	2981	3825	76
['zɔzi] T3^F	2nd syllable	318	174	160	388	2696	3207	4255	70
[zu'zi] T1^F	1st syllable	272	137	204	257	1399	3056	3585	78
[zu'zi] T1^F	2nd syllable	325	190	131	267	2666	3377	4606	72
[zu'zi] T2^F	1st syllable	276	101	239	415	1100	2817	3647	78
[zu'zi] T2^F	2nd syllable	406	298	161	354	2749	3249	4374	74
[zu'zi] T3^F	1st syllable	285	115	304	318	1430	3175	3498	76
[zu'zi] T3^F	2nd syllable	407	251	171	332	2848	3639	4727	74

E. Screenshot of the labbeling in Praat for the token ['lutinu] produced by talker T3^F



F. Acoustic measurements for the trisyllabic stimuli

	syllable		vowels							
		duratio n (ms)		duratio n (ms)	Pitch (Hz)	F1 (Hz)	F2 (Hz)	F3 (Hz)	F4 (Hz)	Intensit y (db)
['bifuʃi]T1^F	['bi]	233	['i]	95	225	233	2523	3229	4460	88
['bifuʃi]T1^F	[fu]	180	[u]	97	221	1458	2343	3534	3912	67
['bifuʃi]T1^F	[ʃi]	269	[i]	90	undefined	255	2286	2990	3779	63
[bi'fuʃi]T1^F	[bi]	141	[i]	67	195	304	2647	3093	4460	80
[bi'fuʃi]T1^F	['fu]	238	['u]	69	184	348	1323	2491	4107	78

[bi'fuʃi]T1^F	[ʃi]	425	[i]	234	208	239	2408	3016	3911	70
[bifu'ʃi]T1^F	[bi]	105	[i]	50	214	238	2487	3007	4577	81
[bifu'ʃi]T1^F	[fu]	175	[u]	46	216	229	1552	2720	3893	77
[bifu'ʃi]T1^F	['ʃi]	401	['i]	220	259	273	2607	3425	4487	75
['bifuʃi]T4^M	['bi]	172	['i]	100	130	274	2249	2749	3485	81
['bifuʃi]T4^M	[fu]	176	[u]	54	111	296	1109	2483	3367	73
['bifuʃi]T4^M	[ʃi]	322	[i]	143	97	259	2249	2669	3489	69
[bi'fuʃi]T4^M	[bi]	109	[i]	61	109	270	2186	2716	3456	74
[bi'fuʃi]T4^M	['fu]	242	['u]	88	137	278	990	2347	3375	79
[bi'fuʃi]T4^M	[ʃi]	375	[i]	191	103	281	2213	3111	3603	72
[bifu'ʃi]T4^M	[bi]	92	[i]	58	105	289	2124	2584	3572	78
[bifu'ʃi]T4^M	[fu]	165	[u]	38	109	285	1115	2514	3240	73
[bifu'ʃi]T4^M	['ʃi]	388	['i]	222	99	266	2255	3032	3796	76
['bifuʃi]T2^F	['bi]	235	['i]	105	238	449	2816	3419	4604	81
['bifuʃi]T2^F	[fu]	216	[u]	48	162	463	1107	2615	3810	68
['bifuʃi]T2^F	[ʃi]	391	[i]	214	138	451	2790	3085	4427	67
[bi'fuʃi]T2^F	[bi]	204	[i]	79	239	463	2731	3207	4423	81
[bi'fuʃi]T2^F	['fu]	268	['u]	96	199	389	951	2683	4188	77
[bi'fuʃi]T2^F	[ʃi]	465	[i]	266	139	328	2828	2961	4192	69
[bifu'ʃi]T2^F	[bi]	193	[i]	82	253	488	2673	3044	4405	81
[bifu'ʃi]T2^F	[fu]	186	[u]	54	253	454	1262	2733	3844	77
[bifu'ʃi]T2^F	['ʃi]	462	['i]	277	145	423	2855	3104	4428	70
['filuvi]T1^F	['fi]	245	['i]	54	214	236	2376	2480	4075	81
['filuvi]T1^F	[lu]	170	[u]	71	150	318	1152	2579	3970	71
['filuvi]T1^F	[vi]	293	[i]	190	281	295	2596	3192	4384	67
[fi'luvi]T1^F	[fi]	240	[i]	49	199	259	2367	2772	4261	76
[fi'luvi]T1^F	['lu]	219	['u]	113	211	238	720	2546	3696	82
[fi'luvi]T1^F	[vi]	312	[i]	185	230	273	2633	3034	4356	71
[filu'vi]T1^F	[fi]	218	[i]	62	198	309	1473	2423	4092	74
[filu'vi]T1^F	[lu]	177	[u]	106	206	217	1119	2461	3971	77
[filu'vi]T1^F	['vi]	360	['i]	248	227	245	2556	3543	4378	71
['filuvi]T4^M	['fi]	255	['i]	79	145	294	2087	2468	3332	79
['filuvi]T4^M	[lu]	189	[u]	96	107	297	938	2369	3643	73
['filuvi]T4^M	[vi]	239	[i]	131	91	254	2029	2815	3633	67
[fi'luvi]T4^M	[fi]	230	[i]	82	110	289	2092	2421	3534	75
[fi'luvi]T4^M	['lu]	275	['u]	165	120	292	775	2485	3432	79
[fi'luvi]T4^M	[vi]	259	[i]	155	97	251	2107	2971	3750	73
[filu'vi]T4^M	[fi]	224	[i]	81	117	300	1959	2374	3508	76
[filu'vi]T4^M	[lu]	170	[u]	102	112	285	878	2490	3641	73

[filu'vi]T4M	['vi]	422	['i]	320	111	230	2330	3111	3664	77
['filuvi]T2F	['fi]	244	['i]	95	244	481	2571	3057	4533	80
['filuvi]T2F	[lu]	201	[u]	103	177	420	973	2905	4234	71
['filuvi]T2F	[vi]	343	[i]	258	156	465	2798	3289	4507	67
[fi'luvi]T2F	[fi]	232	[i]	105	245	483	2631	3057	4555	77
[fi'luvi]T2F	['lu]	296	['u]	169	209	378	828	2970	4338	76
[fi'luvi]T2F	[vi]	492	[i]	393	157	417	2847	3129	4565	66
[filu'vi]T2F	[fi]	241	[i]	85	247	483	2538	2927	4300	78
[filu'vi]T2F	[lu]	192	[u]	101	232	438	902	2944	4271	77
[filu'vi]T2F	['vi]	414	['i]	287	158	389	2836	3803	4525	66
['ʒurigu]T1F	['ʒu]	353	['u]	127	238	292	1139	2496	3831	79
['ʒurigu]T1F	[ri]	185	[i]	144	148	303	2475	2536	3902	70
['ʒurigu]T1F	[gu]	222	[u]	198	229	323	984	2537	3920	63
[ʒu'rigu]T1F	[ʒu]	271	[u]	124	202	309	1418	2988	4060	78
[ʒu'rigu]T1F	['ri]	187	['i]	154	207	342	2564	3447	4479	79
[ʒu'rigu]T1F	[gu]	264	[u]	204	136	332	1116	2537	4080	62
[ʒuri'gu]T1F	[ʒu]	173	[u]	80	168	343	1662	2599	3600	75
[ʒuri'gu]T1F	[ri]	169	[i]	126	177	331	2486	2717	3815	73
[ʒuri'gu]T1F	['gu]	441	['u]	354	144	330	685	2589	3747	77
['ʒurigu]T4M	['ʒu]	219	['u]	72	118	305	1304	2324	3247	75
['ʒurigu]T4M	[ri]	136	[i]	94	120	264	2297	2503	3494	74
['ʒurigu]T4M	[gu]	526	[u]	419	120	129	758	2368	3601	78
[ʒu'rigu]T4M	[ʒu]	267	[u]	95	129	295	1221	2221	3296	78
[ʒu'rigu]T4M	['ri]	179	['i]	139	122	271	2289	2601	3465	76
[ʒu'rigu]T4M	[gu]	376	[u]	292	120	129	1355	2522	3822	65
[ʒuri'gu]T4M	[ʒu]	233	[u]	75	123	313	1356	2327	3362	77
[ʒuri'gu]T4M	[ri]	131	[i]	93	123	278	2203	2439	3482	74
[ʒuri'gu]T4M	['gu]	534	['u]	454	120	128	1089	2426	3639	74
['ʒurigu]T2F	['ʒu]	327	['u]	143	239	470	1092	2620	3816	80
['ʒurigu]T2F	[ri]	137	[i]	117	181	371	2630	2924	4231	72
['ʒurigu]T2F	[gu]	363	[u]	270	153	412	859	2605	4166	67
[ʒu'rigu]T2F	[ʒu]	261	[u]	122	258	459	1179	2577	3489	80
[ʒu'rigu]T2F	['ri]	202	['i]	174	203	395	2758	3410	4709	75
[ʒu'rigu]T2F	[gu]	347	[u]	274	156	445	876	2604	4058	69
[ʒuri'gu]T2F	[ʒu]	232	[u]	116	232	442	1178	2538	3240	77
[ʒuri'gu]T2F	[ri]	160	[i]	122	241	467	2672	2945	3955	76
[ʒuri'gu]T2F	['gu]	358	['u]	302	158	373	737	2667	4087	68
['mupiʃu]T1F	['mu]	285	['u]	142	211	255	633	3010	3638	87
['mupiʃu]T1F	['pi]	119	['i]	57	164	291	2319	2771	3905	73

['mupiʃu]T1F	['ʃu]	364	['u]	169	171	959	2116	2932	3893	66
[mu'piʃu]T1F	[mu]	217	[u]	168	180	531	2024	3110	4293	74
[mu'piʃu]T1F	['pi]	163	['i]	77	224	235	2579	3294	4470	80
[mu'piʃu]T1F	[ʃu]	310	[u]	111	undefined	282	1865	2644	3760	62
[mup'iʃu]T1F	[mu]	217	[u]	132	213	221	697	2627	4143	87
[mup'iʃu]T1F	[pi]	118	[i]	75	220	227	2511	2775	4048	86
[mup'iʃu]T1F	['ʃu]	378	['u]	194	236	295	2086	3103	3944	77
['mupiʃu]T4M	['mu]	263	['u]	158	133	283	820	2358	3464	79
['mupiʃu]T4M	['pi]	132	['i]	81	114	259	2119	2763	3479	72
['mupiʃu]T4M	['ʃu]	350	['u]	187	100	770	1289	2476	3737	68
[mu'piʃu]T4M	[mu]	184	[u]	164	103	333	742	2473	3504	76
[mu'piʃu]T4M	['pi]	179	['i]	117	131	248	2248	2911	3681	79
[mu'piʃu]T4M	[ʃu]	342	[u]	164	100	362	1018	2361	3644	72
[mup'iʃu]T4M	[mu]	182	[u]	132	126	299	789	2359	3437	76
[mup'iʃu]T4M	[pi]	92	[i]	49	130	252	2189	2814	3471	74
[mup'iʃu]T4M	['ʃu]	353	['u]	204	101	257	856	2502	3538	76
['mupiʃu]T2F	['mu]	296	['u]	185	217	403	881	2686	4176	81
['mupiʃu]T2F	['pi]	124	['i]	88	174	375	2791	3011	4500	67
['mupiʃu]T2F	['ʃu]	448	['u]	278	147	433	1005	2738	4047	70
[mu'piʃu]T2F	[mu]	251	[u]	174	265	398	812	2780	4239	82
[mu'piʃu]T2F	['pi]	178	['i]	136	187	407	2744	3046	4466	73
[mu'piʃu]T2F	[ʃu]	448	[u]	255	149	426	955	2637	3594	71
[mup'iʃu]T2F	[mu]	236	[u]	128	250	478	1011	2676	4296	81
[mup'iʃu]T2F	[pi]	130	[i]	73	242	468	2690	3184	4440	72
[mup'iʃu]T2F	['ʃu]	473	['u]	298	153	413	764	2903	4297	71
['vumipi]T1F	['vu]	225	['u]	97	222	269	726	2296	4049	81
['vumipi]T1F	[mi]	209	[i]	153	141	292	2627	3401	4699	72
['vumipi]T1F	[pi]	305	[i]	188	undefined	621	2583	3304	4002	51
[vu'mipi]T1F	[vu]	283	[u]	81	172	506	1348	2790	4203	79
[vu'mipi]T1F	['mi]	264	['i]	201	160	404	2636	3620	4373	77
[vu'mipi]T1F	[pi]	307	[i]	262	132	284	2595	3222	4389	70
[vumi'pi]T1F	[vu]	224	[u]	47	207	304	819	2623	4158	80
[vumi'pi]T1F	[mi]	245	[i]	195	189	445	2606	3430	4274	76
[vumi'pi]T1F	['pi]	340	['i]	293	136	289	2529	2906	3644	67
['vumipi]T4M	['vu]	245	['u]	100	152	328	777	2204	3446	77
['vumipi]T4M	[mi]	243	[i]	160	102	266	2078	2783	3419	72
['vumipi]T4M	[pi]	176	[i]	131	97	257	2176	2756	3425	70
[vu'mipi]T4M	[vu]	195	[u]	92	101	389	773	2178	3394	72
[vu'mipi]T4M	['mi]	306	['i]	208	114	255	2234	2941	3653	77

[vu'mipi]T4^{M}	[pi]	261	[i]	214	95	260	2243	3054	3516	74
[vumi'pi]T4^{M}	[vu]	188	[u]	71	109	324	833	2360	3507	74
[vumi'pi]T4^{M}	[mi]	213	[i]	146	105	250	2131	2929	3146	72
[vumi'pi]T4^{M}	['pi]	295	['i]	213	116	237	2315	3008	3670	76
['vumipi]T2^{F}	['vu]	309	['u]	152	236	458	932	2744	4114	80
['vumipi]T2^{F}	[mi]	209	[i]	158	168	486	2686	3114	4204	71
['vumipi]T2^{F}	[pi]	299	[i]	242	148	434	2837	3294	4457	64
[vu'mipi]T2^{F}	[vu]	224	[u]	105	255	464	930	2878	4308	78
[vu'mipi]T2^{F}	['mi]	272	['i]	199	198	452	2855	3293	4479	74
[vu'mipi]T2^{F}	[pi]	305	[i]	261	156	439	2900	3482	4643	65
[vumi'pi]T2^{F}	[vu]	265	[u]	116	241	464	948	2837	4199	80
[vumi'pi]T2^{F}	[mi]	228	[i]	152	241	492	2689	3016	4231	76
[vumi'pi]T2^{F}	['pi]	329	['i]	277	154	441	2969	3943	4705	62
['zituli]T1^{F}	['zi]	256	['i]	84	243	391	2640	3126	4576	81
['zituli]T1^{F}	[tu]	278	[u]	107	149	380	1364	3246	3655	67
['zituli]T1^{F}	[li]	244	[i]	202	136	333	2763	3302	4356	69
[zi'tuli]T1^{F}	[zi]	237	[i]	67	194	348	2603	3197	4566	78
[zi'tuli]T1^{F}	['tu]	263	['u]	98	219	255	1920	3203	3954	79
[zi'tuli]T1^{F}	[li]	230	[i]	139	125	313	2488	3222	4449	71
[zitu'li]T1^{F}	[zi]	153	[i]	126	151	341	2190	3214	4464	73
[zitu'li]T1^{F}	[tu]	220	[u]	77	154	329	1489	3048	3509	68
[zitu'li]T1^{F}	['li]	298	['i]	145	142	342	2725	3580	4534	73
['zituli]T4^{M}	['zi]	261	['i]	96	169	285	2205	2652	3374	79
['zituli]T4^{M}	[tu]	235	[u]	73	115	312	1021	2428	3528	74
['zituli]T4^{M}	[li]	206	[i]	130	96	275	2405	2762	3831	69
[zi'tuli]T4^{M}	[zi]	215	[i]	61	110	286	2305	2670	3624	69
[zi'tuli]T4^{M}	['tu]	330	['u]	142	143	329	982	2253	3547	76
[zi'tuli]T4^{M}	[li]	274	[i]	183	103	282	2305	2912	3984	74
[zitu'li]T4^{M}	[zi]	186	[i]	62	112	287	2181	2638	4250	65
[zitu'li]T4^{M}	[tu]	255	[u]	134	116	264	903	2404	3499	76
[zitu'li]T4^{M}	['li]	259	['i]	227	120	252	2363	2875	3629	77
['zituli]T2^{F}	['zi]	279	['i]	102	239	470	2720	3275	4774	81
['zituli]T2^{F}	[tu]	271	[u]	140	169	438	933	2901	3991	70
['zituli]T2^{F}	[li]	241	[i]	216	146	428	2857	3349	4645	70
[zi'tuli]T2^{F}	[zi]	263	[i]	88	251	481	2738	3196	4750	80
[zi'tuli]T2^{F}	['tu]	327	['u]	193	176	394	923	2977	4090	75
[zi'tuli]T2^{F}	[li]	260	[i]	238	154	436	2853	3173	4502	70
[zitu'li]T2^{F}	[zi]	235	[i]	91	234	468	2589	3219	4725	78
[zitu'li]T2^{F}	[tu]	256	[u]	129	230	439	955	3038	3941	76

[zitu'li]T2F	['li]	310	['i]	282	159	408	2897	3707	4529	73
['dutiku]T1F	['du]	270	['u]	87	214	267	1095	2559	3782	86
['dutiku]T1F	[ti]	213	[i]	41	147	339	2676	3300	4441	71
['dutiku]T1F	[ku]	301	[u]	129	138	318	1599	3032	3833	70
[du'tiku]T1F	[du]	165	[u]	73	160	324	1538	3001	3560	76
[du'tiku]T1F	['ti]	232	['i]	63	194	331	2671	3472	4455	82
[du'tiku]T1F	[ku]	346	[u]	142	182	843	1856	2951	4319	70
[duti'ku]T1F	[du]	131	[u]	75	205	252	1266	2651	3408	85
[duti'ku]T1F	[ti]	152	[i]	39	231	236	2601	2646	4067	84
[duti'ku]T1F	['ku]	348	['u]	161	150	308	643	2309	3945	70
['dutiku]T4M	['du]	211	['u]	113	159	309	1002	2207	3291	80
['dutiku]T4M	[ti]	196	[i]	67	115	267	2272	2675	3457	72
['dutiku]T4M	[ku]	283	[u]	136	97	711	882	2477	3568	67
[du'tiku]T4M	[du]	120	[u]	53	110	304	1259	2418	3471	73
[du'tiku]T4M	['ti]	267	['i]	97	140	256	2320	2780	3448	77
[du'tiku]T4M	[ku]	296	[u]	145	99	328	818	2428	3531	77
[duti'ku]T4M	[du]	177	[u]	55	121	320	1333	2531	3497	76
[duti'ku]T4M	[ti]	178	[i]	63	117	261	2328	2718	3486	71
[duti'ku]T4M	['ku]	317	['u]	178	121	283	772	2209	3451	78
['dutiku]T2F	['du]	235	['u]	109	238	438	990	2843	4059	81
['dutiku]T2F	[ti]	178	[i]	60	193	393	2750	3093	4472	74
['dutiku]T2F	[ku]	451	[u]	286	150	395	905	2505	4096	70
[du'tiku]T2F	[du]	206	[u]	86	253	434	1222	2927	4165	80
[du'tiku]T2F	['ti]	228	['i]	97	203	428	2695	3035	4469	77
[du'tiku]T2F	[ku]	441	[u]	270	148	435	892	2555	4293	70
[duti'ku]T2F	[du]	208	[u]	102	237	427	1143	2810	4045	79
[duti'ku]T2F	[ti]	169	[i]	72	254	491	2571	3052	4393	77
[duti'ku]T2F	['ku]	486	['u]	322	162	380	744	2737	3869	76
['kiɾuʃi]T1F	['ki]	176	['i]	94	238	285	2579	3177	4355	81
['kiɾuʃi]T1F	[ɾu]	137	[u]	97	161	321	1587	2665	3799	75
['kiɾuʃi]T1F	[ʃi]	379	[i]	110	200	267	2541	3180	3898	59
[ki'ɾuʃi]T1F	[ki]	161	[i]	85	180	299	2570	3265	4558	77
[ki'ɾuʃi]T1F	['ɾu]	156	['u]	115	192	312	1310	2334	3482	78
[ki'ɾuʃi]T1F	[ʃi]	370	[i]	168	201	228	2541	3078	3917	63
[kiɾu'ʃi]T1F	[ki]	95	[i]	37	242	264	2197	2718	4448	80
[kiɾu'ʃi]T1F	[ɾu]	108	[u]	85	254	276	1559	2530	3034	77
[kiɾu'ʃi]T1F	['ʃi]	441	['i]	227	264	271	2517	3231	4184	72
['kiɾuʃi]T4M	['ki]	174	['i]	121	173	285	2163	3118	3233	78
['kiɾuʃi]T4M	[ɾu]	118	[u]	72	115	304	1209	2286	3416	74

['kiruʃi]T4^{M}	[ʃi]	354	[i]	171	99	261	2266	3111	3682	66
[ki'ruʃi]T4^{M}	[ki]	132	[i]	81	115	285	2278	2782	3476	75
[ki'ruʃi]T4^{M}	['ru]	166	['u]	128	142	288	1011	2344	3328	77
[ki'ruʃi]T4^{M}	[ʃi]	369	[i]	195	96	267	2254	2928	3622	70
[kiru'ʃi]T4^{M}	[ki]	121	[i]	63	116	298	2263	2659	3516	77
[kiru'ʃi]T4^{M}	[ru]	98	[u]	63	114	292	1306	2425	3305	75
[kiru'ʃi]T4^{M}	['ʃi]	412	['i]	241	106	243	2277	2987	3571	75
['kiruʃi]T2^{F}	['ki]	195	['i]	126	231	474	2718	3334	4507	81
['kiruʃi]T2^{F}	[ru]	102	[u]	85	186	408	1272	2726	3123	74
['kiruʃi]T2^{F}	[ʃi]	493	[i]	313	160	434	2771	3022	4502	70
[ki'ruʃi]T2^{F}	[ki]	175	[i]	105	260	502	2695	3106	4388	78
[ki'ruʃi]T2^{F}	['ru]	163	['u]	135	210	399	1011	2745	4323	77
[ki'ruʃi]T2^{F}	[ʃi]	468	[i]	264	149	406	2740	3246	4565	67
[kiru'ʃi]T2^{F}	[ki]	163	[i]	101	261	493	2709	2992	4364	78
[kiru'ʃi]T2^{F}	[ru]	103	[u]	85	277	403	1471	2784	3679	78
[kiru'ʃi]T2^{F}	['ʃi]	504	['i]	309	164	397	2810	3366	4506	71
['tikuru]T1^{F}	['ti]	105	['i]	58	248	288	2529	3023	4490	83
['tikuru]T1^{F}	[ku]	196	[u]	43	147	347	1344	2606	4054	72
['tikuru]T1^{F}	[ru]	210	[u]	177	227	312	1211	2396	3743	66
[ti'kuru]T1^{F}	[ti]	75	[i]	20	undefined	undefined	undefined	undefined	undefined	73
[ti'kuru]T1^{F}	['ku]	278	['u]	104	208	234	987	2457	3839	80
[ti'kuru]T1^{F}	[ru]	245	[u]	207	238	261	964	2314	3682	67
[tiku'ru]T1^{F}	[ti]	81	[i]	35	198	248	2501	2997	4391	81
[tiku'ru]T1^{F}	[ku]	197	[u]	66	220	238	1100	2523	3986	86
[tiku'ru]T1^{F}	['ru]	273	['u]	241	159	304	877	2275	4005	76
['tikuru]T4^{M}	['ti]	108	['i]	68	158	295	2216	2606	3241	79
['tikuru]T4^{M}	[ku]	223	[u]	77	115	305	827	2407	3488	76
['tikuru]T4^{M}	[ru]	207	[u]	182	99	300	869	2432	3430	70
[ti'kuru]T4^{M}	[ti]	77	[i]	27	119	269	2310	2560	3428	69
[ti'kuru]T4^{M}	['ku]	310	['u]	142	143	294	765	2166	3424	78
[ti'kuru]T4^{M}	[ru]	194	[u]	152	106	311	939	2335	3362	75
[tiku'ru]T4^{M}	[ti]	92	[i]	56	205	224	2100	2501	3184	78
[tiku'ru]T4^{M}	[ku]	207	[u]	70	163	301	817	2133	3318	74
[tiku'ru]T4^{M}	['ru]	262	['u]	216	102	293	775	2335	3421	72
['tikuru]T2^{F}	['ti]	147	['i]	103	247	480	2728	3366	4654	80
['tikuru]T2^{F}	[ku]	241	[u]	101	198	406	1001	2811	4181	77
['tikuru]T2^{F}	[ru]	290	[u]	266	157	425	858	2761	4036	74
[ti'kuru]T2^{F}	[ti]	101	[i]	60	268	469	2636	3192	4411	81
[ti'kuru]T2^{F}	['ku]	319	['u]	165	209	375	854	2796	4078	75

[ti'kuru]T2F	[ru]	311	[u]	294	159	444	850	2800	3993	68
[tiku'ru]T2F	[ti]	100	[i]	68	272	447	2718	3326	4668	82
[tiku'ru]T2F	[ku]	219	[u]	87	260	382	978	2824	4379	76
[tiku'ru]T2F	['ru]	350	['u]	331	157	388	796	2769	4278	72
['tudiki]T1F	['tu]	135	['u]	95	239	292	1269	2832	3949	83
['tudiki]T1F	[di]	143	[i]	76	166	312	2566	3363	4440	71
['tudiki]T1F	[ki]	353	[i]	199	145	289	2543	3393	4322	67
[tu'diki]T1F	[tu]	132	[u]	85	178	349	1413	2680	3658	78
[tu'diki]T1F	['di]	163	['i]	87	195	390	2638	3441	4434	79
[tu'diki]T1F	[ki]	339	[i]	185	187	243	2520	3324	4133	69
[tudi'ki]T1F	[tu]	94	[u]	39	187	332	1423	2628	3571	80
[tudi'ki]T1F	[di]	146	[i]	87	188	362	2676	3370	4371	79
[tudi'ki]T1F	['ki]	379	['i]	246	136	303	2566	3060	3752	71
['tudiki]T4M	['tu]	135	['u]	96	173	341	1091	2159	3283	80
['tudiki]T4M	[di]	157	[i]	74	107	263	2262	2718	3548	76
['tudiki]T4M	[ki]	353	[i]	202	101	276	2352	2893	3812	69
[tu'diki]T4M	[tu]	108	[u]	52	109	297	1234	2375	3478	71
[tu'diki]T4M	['di]	221	['i]	127	140	282	2332	2763	3527	79
[tu'diki]T4M	[ki]	376	[i]	223	103	242	2335	2895	3846	72
[tudi'ki]T4M	[tu]	95	[u]	36	109	285	1417	2346	3395	74
[tudi'ki]T4M	[di]	168	[i]	86	111	246	2339	2765	3364	75
[tudi'ki]T4M	['ki]	563	['i]	411	119	131	2337	2842	3557	76
['tudiki]T2F	['tu]	148	['u]	103	234	432	1104	2797	4059	82
['tudiki]T2F	[di]	148	[i]	87	193	407	2804	3307	4439	75
['tudiki]T2F	[ki]	416	[i]	266	153	448	2931	3345	4570	69
[tu'diki]T2F	[tu]	130	[u]	95	231	407	1008	2722	4196	81
[tu'diki]T2F	['di]	221	['i]	159	179	413	2869	3467	4477	77
[tu'diki]T2F	[ki]	387	[i]	262	150	418	2911	3426	4657	70
[tudi'ki]T2F	[tu]	135	[u]	96	251	437	1184	2853	4012	81
[tudi'ki]T2F	[di]	136	[i]	78	228	446	2752	3661	4586	75
[tudi'ki]T2F	['ki]	515	['i]	345	160	429	3009	3571	4609	71
['tuduri]T1F	['tu]	146	['u]	115	279	294	1141	2588	3833	79
['tuduri]T1F	[du]	168	[u]	80	161	348	1410	2627	3749	69
['tuduri]T1F	[ri]	237	[i]	213	124	300	2589	2991	3860	65
[tu'duri]T1F	[tu]	121	[u]	92	185	289	1279	2618	3684	79
[tu'duri]T1F	['du]	176	['u]	99	221	229	1099	2641	3837	88
[tu'duri]T1F	[ri]	210	[i]	178	277	283	2617	3177	4333	79
[tudu'ri]T1F	[tu]	111	[u]	75	177	319	1455	2697	3625	77
[tudu'ri]T1F	[du]	170	[u]	106	172	302	1375	2607	3674	75

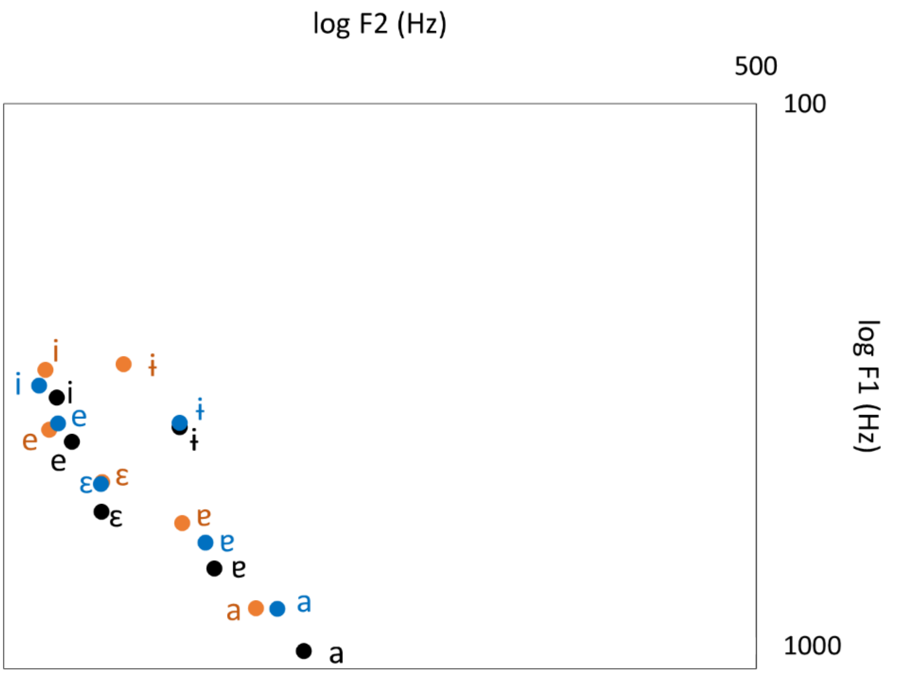
[tudu'ri]T1F	['ri]	345	['i]	297	149	314	2682	3620	4450	74
['tuduri]T4M	['tu]	142	['u]	96	159	276	986	2105	3164	78
['tuduri]T4M	[du]	163	[u]	87	109	299	1133	2284	3395	73
['tuduri]T4M	[ri]	197	[i]	152	97	250	2239	2951	3647	72
[tu'duri]T4M	[tu]	122	[u]	64	111	295	1149	2374	3483	69
[tu'duri]T4M	['du]	207	['u]	129	138	284	1060	2346	3372	78
[tu'duri]T4M	[ri]	192	[i]	146	99	250	2214	2902	3588	73
[tudu'ri]T4M	[tu]	100	[u]	50	112	309	1215	2386	3359	74
[tudu'ri]T4M	[du]	171	[u]	97	108	307	1126	2507	3487	74
[tudu'ri]T4M	['ri]	272	['i]	232	113	249	2358	3118	3592	76
['tuduri]T2F	['tu]	146	['u]	110	235	432	934	2828	4108	82
['tuduri]T2F	[du]	185	[u]	101	176	407	1094	2898	4161	72
['tuduri]T2F	[ri]	388	[i]	350	153	446	2822	3088	4582	66
[tu'duri]T2F	[tu]	119	[u]	88	285	336	885	2841	3838	82
[tu'duri]T2F	['du]	260	['u]	175	193	367	970	2775	3881	76
[tu'duri]T2F	[ri]	396	[i]	356	155	416	2801	3134	4585	70
[tudu'ri]T2F	[tu]	122	[u]	82	262	415	1021	2880	4061	78
[tudu'ri]T2F	[du]	186	[u]	122	235	432	1106	2797	4065	76
[tudu'ri]T2F	['ri]	378	['i]	344	162	385	2803	3394	4523	69
['bikulu]T6F	['bi]	233	['i]	95	220	325	2668	3211	3997	79
['bikulu]T6F	[ku]	283	[u]	132	165	323	818	3213	3631	74
['bikulu]T6F	[lu]	189	[u]	177	151	234	774	3306	4370	73
[bi'kulu]T6F	[bi]	257	[i]	91	259	268	2626	3147	4064	79
[bi'kulu]T6F	['ku]	360	['u]	195	184	365	827	3228	3788	74
[bi'kulu]T6F	[lu]	148	[u]	136	154	227	832	3365	3464	70
[biku'lu]T6F	[bi]	197	[i]	88	241	262	2593	2958	4075	78
[biku'lu]T6F	[ku]	255	[u]	112	223	306	768	3378	3814	73
[biku'lu]T6F	['lu]	190	['u]	180	161	347	806	3231	3517	72
['bikulu]T5F	['bi]	255	['i]	95	211	371	2503	3229	4461	80
['bikulu]T5F	[ku]	280	[u]	104	147	353	927	2606	3537	76
['bikulu]T5F	[lu]	298	[u]	280	145	455	1259	2910	3786	72
[bi'kulu]T5F	[bi]	194	[i]	83	203	334	2455	3103	4365	78
[bi'kulu]T5F	['ku]	339	['u]	174	175	391	826	2477	3703	77
[bi'kulu]T5F	[lu]	263	[u]	221	155	337	798	1958	3404	73
[biku'lu]T5F	[bi]	201	[i]	70	205	317	2455	2977	4310	77
[biku'lu]T5F	[ku]	243	[u]	105	202	363	880	2733	3714	78
[biku'lu]T5F	['lu]	305	['u]	264	158	358	733	2780	4241	76
['bikulu]T7M	['bi]	196	['i]	65	147	285	2296	3108	4035	81
['bikulu]T7M	[ku]	209	[u]	78	106	334	1047	2036	3644	65

['bikulu]T7M	[lu]	184	[u]	121	95	348	724	2805	3923	63
[bi'kulu]T7M	[bi]	176	[i]	67	173	276	2288	2936	3965	80
[bi'kulu]T7M	['ku]	270	['u]	132	117	325	686	2114	3671	77
[bi'kulu]T7M	[lu]	233	[u]	156	99	314	747	2897	3779	68
[biku'lu]T7M	[bi]	132	[i]	44	162	285	2214	2419	3853	79
[biku'lu]T7M	[ku]	224	[u]	99	134	317	645	2589	2988	73
[biku'lu]T7M	['lu]	284	['u]	235	106	294	709	2575	3736	74
['lutinu]T6F	['lu]	237	['u]	83	219	390	911	3048	4028	79
['lutinu]T6F	[ti]	190	[i]	59	175	377	2721	3138	4172	71
['lutinu]T6F	[nu]	254	[u]	129	155	231	1002	1395	3624	67
[lu'tinu]T6F	[lu]	210	[u]	72	263	293	1037	3109	3812	77
[lu'tinu]T6F	['ti]	246	['i]	96	187	426	2823	3178	4143	73
[lu'tinu]T6F	[nu]	219	[u]	146	149	293	1286	2345	3587	71
[luti'nu]T6F	[lu]	200	[u]	84	243	328	918	2986	4014	75
[luti'nu]T6F	[ti]	190	[i]	66	258	283	2570	2992	4097	76
[luti'nu]T6F	['nu]	287	['u]	148	148	250	760	1436	3472	68
['lutinu]T5F	['lu]	251	['u]	109	222	397	903	2630	3544	80
['lutinu]T5F	[ti]	212	[i]	44	179	432	2484	2924	4077	73
['lutinu]T5F	[nu]	300	[u]	153	159	181	1104	2184	3495	64
[lu'tinu]T5F	[lu]	193	[u]	76	205	338	1221	2825	4112	76
[lu'tinu]T5F	['ti]	292	['i]	136	201	291	1159	2765	3138	78
[lu'tinu]T5F	[nu]	329	[u]	244	150	695	1544	2646	3579	73
[luti'nu]T5F	[lu]	170	[u]	76	198	340	1165	2784	3800	75
[luti'nu]T5F	[ti]	214	[i]	80	199	299	2554	2935	4318	76
[luti'nu]T5F	['nu]	425	['u]	334	162	335	774	2401	3273	75
['lutinu]T7M	['lu]	137	['u]	56	157	328	1256	2590	3720	80
['lutinu]T7M	[ti]	212	[i]	42	104	233	2197	2667	3670	64
['lutinu]T7M	[nu]	212	[u]	106	76	420	1248	2258	3663	56
[lu'tinu]T7M	[lu]	138	[u]	61	191	300	1204	2816	3686	79
[lu'tinu]T7M	['ti]	247	['i]	88	118	240	2312	3173	3923	73
[lu'tinu]T7M	[nu]	226	[u]	113	91	308	1228	2270	3796	63
[luti'nu]T7M	[lu]	127	[u]	63	171	323	960	2521	3681	78
[luti'nu]T7M	[ti]	143	[i]	46	157	290	2208	2762	3771	74
[luti'nu]T7M	['nu]	230	['u]	114	102	364	820	2393	3484	67
['mudini]T6F	['mu]	258	['u]	108	216	346	921	2760	4100	74
['mudini]T6F	[di]	131	[i]	81	179	350	2631	3152	4159	69
['mudini]T6F	[ni]	286	[i]	214	159	273	822	2942	3426	67
[mu'dini]T6F	[mu]	275	[u]	105	259	314	1050	2635	4065	78
[mu'dini]T6F	['di]	197	['i]	124	200	382	2786	3248	3998	71

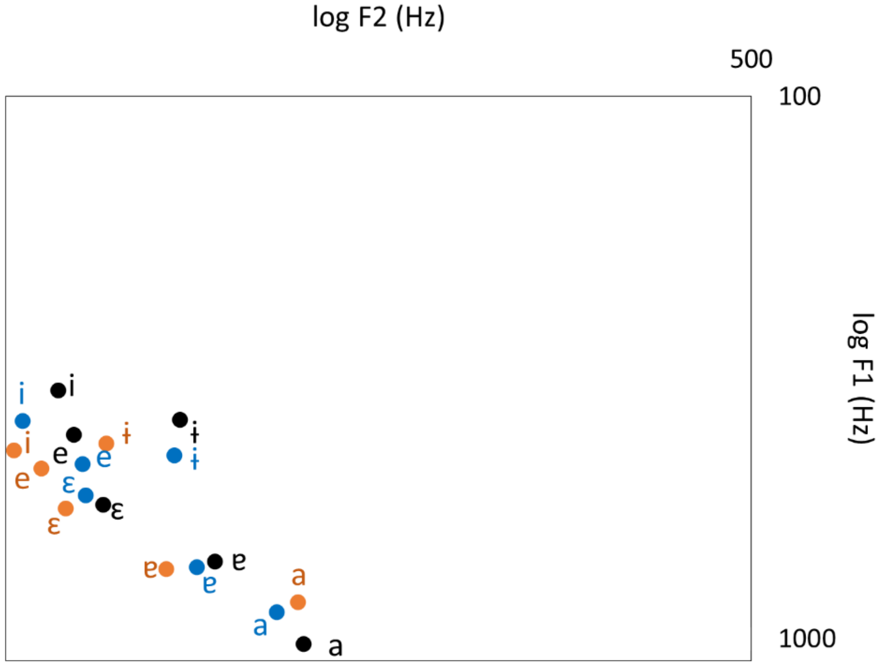
[mu'dini]T6^F	[ni]	266	[i]	192	155	281	981	2778	3188	68
[mudi'ni]T6^F	[mu]	248	[u]	94	242	324	1038	3063	4128	75
[mudi'ni]T6^F	[di]	137	[i]	92	234	295	2748	3214	4330	73
[mudi'ni]T6^F	['ni]	285	['i]	208	161	305	2994	3428	4179	70
['mudini]T5^F	['mu]	333	['u]	153	227	278	914	2239	3289	77
['mudini]T5^F	[di]	192	[i]	139	164	382	2581	3050	4466	71
['mudini]T5^F	[ni]	275	[i]	196	155	248	1271	2728	3197	65
[mu'dini]T5^F	[mu]	161	[u]	83	193	561	1363	2801	4336	77
[mu'dini]T5^F	['di]	249	['i]	151	212	422	2649	3159	4388	75
[mu'dini]T5^F	[ni]	321	[i]	259	156	264	1065	2754	3273	70
[mudi'ni]T5^F	[mu]	163	[u]	84	193	358	1232	2875	4286	76
[mudi'ni]T5^F	[di]	161	[i]	93	192	311	2578	3158	4414	75
[mudi'ni]T5^F	['ni]	373	['i]	301	164	323	2794	3321	4732	72
['mudini]T7^M	['mu]	188	['u]	80	156	347	877	2315	3798	79
['mudini]T7^M	[di]	159	[i]	78	110	260	2193	3169	3921	66
['mudini]T7^M	[ni]	228	[i]	155	90	302	1537	2371	3359	62
[mu'dini]T7^M	[mu]	146	[u]	71	169	314	1055	2347	3688	77
[mu'dini]T7^M	['di]	184	['i]	114	128	271	2300	3270	3919	74
[mu'dini]T7^M	[ni]	242	[i]	151	91	332	1109	2394	3375	65
[mudi'ni]T7^M	[mu]	132	[u]	66	152	354	1106	2272	3956	79
[mudi'ni]T7^M	[di]	143	[i]	78	136	273	2287	3303	3932	70
[mudi'ni]T7^M	['ni]	265	['i]	190	99	298	2363	3619	4183	69
['kitumi]T6^F	['ki]	115	['i]	58	235	285	2879	3306	4267	80
['kitumi]T6^F	[tu]	204	[u]	73	191	377	1053	3197	4220	73
['kitumi]T6^F	[mi]	249	[i]	176	163	334	2886	3362	4053	71
[ki'tumi]T6^F	[ki]	101	[i]	49	237	281	2817	3211	4281	79
[ki'tumi]T6^F	['tu]	248	['u]	124	199	387	940	3121	4069	77
[ki'tumi]T6^F	[mi]	230	[i]	183	159	294	1061	2881	3399	70
[kitu'mi]T6^F	[ki]	102	[i]	44	243	273	2642	3169	4171	78
[kitu'mi]T6^F	[tu]	194	[u]	88	236	288	1182	3173	4127	74
[kitu'mi]T6^F	['mi]	270	['i]	219	164	326	2933	3354	4087	72
['kitumi]T5^F	['ki]	194	['i]	78	225	393	2555	3088	4587	80
['kitumi]T5^F	[tu]	240	[u]	76	170	453	1399	2816	4163	76
['kitumi]T5^F	[mi]	292	[i]	214	155	280	1160	2573	3213	70
[ki'tumi]T5^F	[ki]	158	[i]	57	212	360	2472	3219	4510	76
[ki'tumi]T5^F	['tu]	329	['u]	161	226	304	934	2318	3559	78
[ki'tumi]T5^F	[mi]	317	[i]	255	169	310	1807	2877	3851	70
[kitu'mi]T5^F	[ki]	132	[i]	41	197	329	2485	3089	4545	75
[kitu'mi]T5^F	[tu]	240	[u]	118	195	313	1247	2658	4108	76

[kitu'mi]T5^F	['mi]	343	['i]	295	159	309	989	2799	3326	74
['kitumi]T7^M	['ki]	106	['i]	59	175	271	2222	3282	3946	80
['kitumi]T7^M	[tu]	194	[u]	54	106	244	1475	2306	3346	62
['kitumi]T7^M	[mi]	237	[i]	144	95	526	2367	3019	3839	60
[ki'tumi]T7^M	[ki]	97	[i]	58	177	287	2162	3022	3970	80
[ki'tumi]T7^M	['tu]	202	['u]	88	121	352	803	2501	3412	75
[ki'tumi]T7^M	[mi]	230	[i]	116	95	409	2357	2884	3594	62
[kitu'mi]T7^M	[ki]	103	[i]	52	163	252	2166	3041	4074	76
[kitu'mi]T7^M	[tu]	165	[u]	72	145	259	1943	2229	3700	71
[kitu'mi]T7^M	['mi]	202	['i]	105	105	325	2279	3543	4279	74

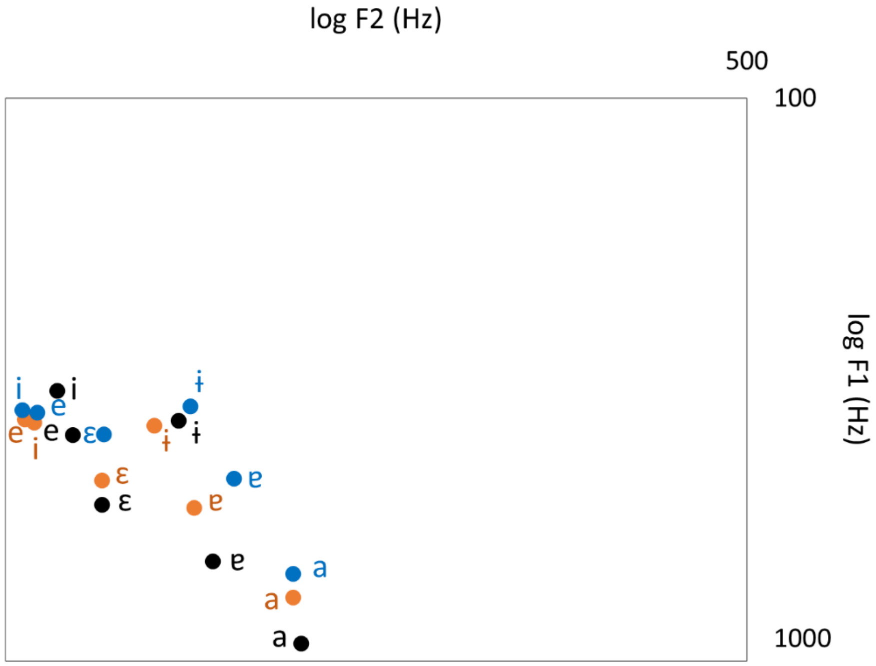
G. F1xF2 vowel space with values for productions of talker T1^F (in orange, inserted in [gV]; in blue, inserted in [zV]) and values for standard production of EP female speakers (in black, retrieved from Andrade 2020, p. 3251, produced in [sV] context)



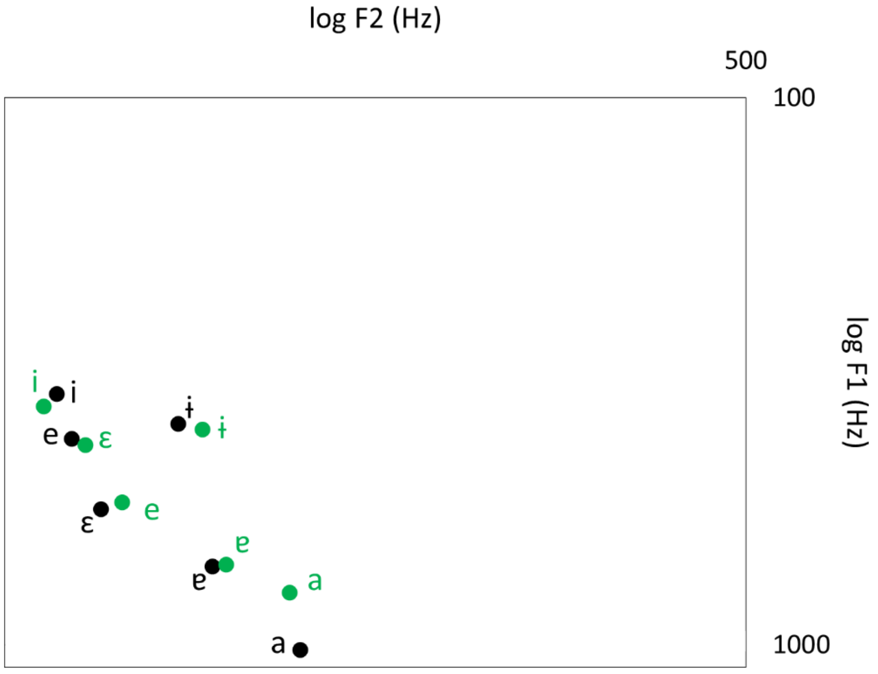
H. F1xF2 vowel space with values for productions of talker T2F (in orange, inserted in [gV]; in blue, inserted in [zV]) and values for standard production of EP female speakers (in black, retrieved from Andrade 2020, p. 3251, produced in [sV] context)



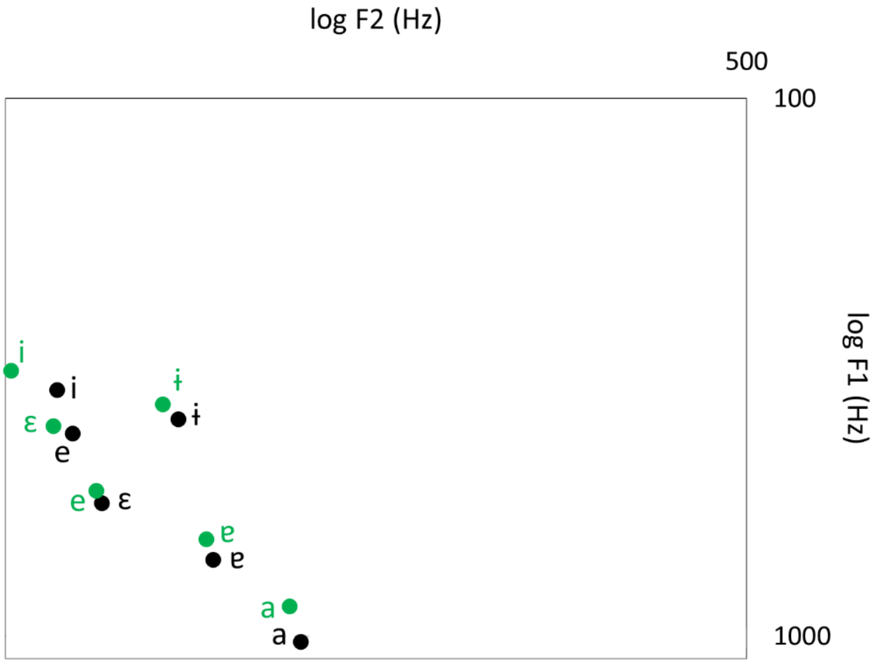
I. F1xF2 vowel space with values for productions of talker T3F (in orange, inserted in [gV]; in blue, inserted in [zV]) and values for standard production of EP female speakers (in black, retrieved from Andrade 2020, p. 3251, produced in [sV] context)



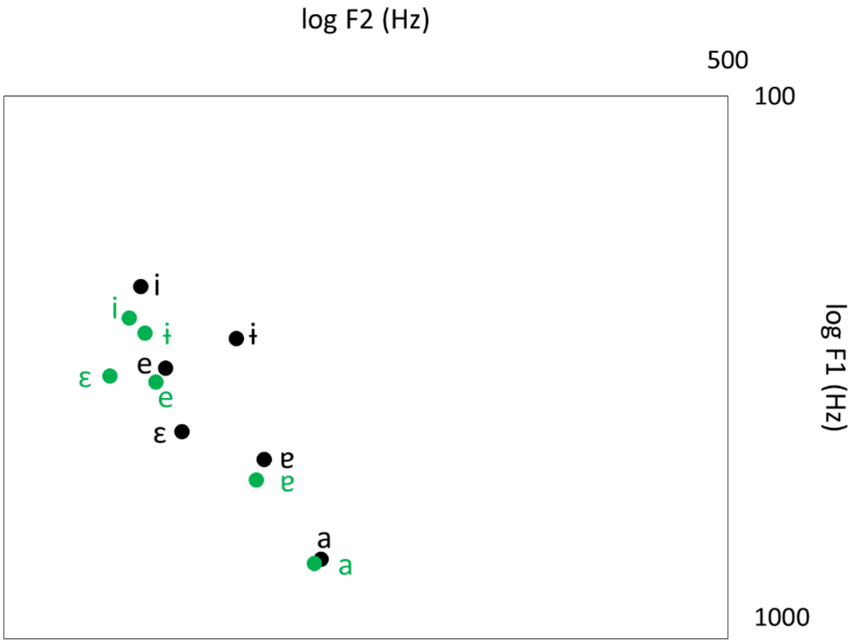
J. F1xF2 vowel space with values for productions of talker T5[F] (in green, inserted in [sV]) and values for standard production of EP female speakers (in black, retrieved from Andrade 2020, p. 3251, produced in [sV] context)



K. F1xF2 vowel space with values for productions of talker T6[F] (in green, inserted in [sV]) and values for standard production of EP female speakers (in black, retrieved from Andrade 2020, p. 3251, produced in [sV] context)



L. F1xF2 vowel space with values for productions of talker T7ᴹ (in green, inserted in [sV]) and values for standard production of EP male speakers (in black, retrieved from Andrade 2020, p. 3251, produced in [sV] context)



Appendix VII

Main trials' sequences for the experimental tasks, for the perceptual training study (Chapter 5)

A. Pretest trials for vowel discrimination (oddity task with catch trials)

1st token			2nd token			3rd token		answer
[gɐ]	T1^F	-	[gɐ]	T2^F	-	[gɛ]	T3^F	3rd token
[gɐ]	T1^F	-	[gɛ]	T2^F	-	[gɐ]	T3^F	2nd token
[gɐ]	T1^F	-	[gɛ]	T2^F	-	[gɛ]	T3^F	1st token
[gɛ]	T1^F	-	[gɐ]	T2^F	-	[gɐ]	T3^F	1st token
[gɛ]	T1^F	-	[gɐ]	T2^F	-	[gɛ]	T3^F	2nd token
[gɛ]	T1^F	-	[gɛ]	T2^F	-	[gɐ]	T3^F	3rd token
[gɛ]	T1^F	-	[gɛ]	T2^F	-	[ge]	T3^F	3rd token
[gɛ]	T1^F	-	[ge]	T2^F	-	[gɛ]	T3^F	2nd token
[gɛ]	T1^F	-	[ge]	T2^F	-	[ge]	T3^F	1st token
[ge]	T1^F	-	[gɛ]	T2^F	-	[gɛ]	T3^F	1st token
[ge]	T1^F	-	[gɛ]	T2^F	-	[ge]	T3^F	2nd token
[ge]	T1^F	-	[ge]	T2^F	-	[gɛ]	T3^F	3rd token
[ge]	T1^F	-	[ge]	T2^F	-	[gi]	T3^F	3rd token
[ge]	T1^F	-	[gi]	T2^F	-	[ge]	T3^F	2nd token
[ge]	T1^F	-	[gi]	T2^F	-	[gi]	T3^F	1st token
[gi]	T1^F	-	[ge]	T2^F	-	[ge]	T3^F	1st token
[gi]	T1^F	-	[ge]	T2^F	-	[gi]	T3^F	2nd token
[gi]	T1^F	-	[gi]	T2^F	-	[ge]	T3^F	3rd token
[gɐ]	T1^F	-	[gɐ]	T2^F	-	[gɨ]	T3^F	3rd token
[gɐ]	T1^F	-	[gɨ]	T2^F	-	[gɐ]	T3^F	2nd token
[gɐ]	T1^F	-	[gɨ]	T2^F	-	[gɨ]	T3^F	1st token
[gɨ]	T1^F	-	[gɐ]	T2^F	-	[gɐ]	T3^F	1st token
[gɨ]	T1^F	-	[gɐ]	T2^F	-	[gɨ]	T3^F	2nd token
[gɨ]	T1^F	-	[gɨ]	T2^F	-	[gɐ]	T3^F	3rd token
[gɐ]	T1^F	-	[gɐ]	T2^F	-	[ga]	T3^F	3rd token
[gɐ]	T1^F	-	[ga]	T2^F	-	[gɐ]	T3^F	2nd token
[gɐ]	T1^F	-	[ga]	T2^F	-	[ga]	T3^F	1st token
[ga]	T1^F	-	[gɐ]	T2^F	-	[gɐ]	T3^F	1st token
[ga]	T1^F	-	[gɐ]	T2^F	-	[ga]	T3^F	2nd token
[ga]	T1^F	-	[ga]	T2^F	-	[gɐ]	T3^F	3rd token
[gɛ]	T1^F	-	[gɛ]	T2^F	-	[gɨ]	T3^F	3rd token
[gɛ]	T1^F	-	[gɨ]	T2^F	-	[gɛ]	T3^F	2nd token
[gɛ]	T1^F	-	[gɨ]	T2^F	-	[gɨ]	T3^F	1st token
[gɨ]	T1^F	-	[gɛ]	T2^F	-	[gɛ]	T3^F	1st token
[gɨ]	T1^F	-	[gɛ]	T2^F	-	[gɨ]	T3^F	2nd token
[gɨ]	T1^F	-	[gɨ]	T2^F	-	[gɛ]	T3^F	3rd token

[ge]	$T1^F$	-	[ge]	$T2^F$	-	[gɨ]	$T3^F$	3rd token
[ge]	$T1^F$	-	[gɨ]	$T2^F$	-	[ge]	$T3^F$	2nd token
[ge]	$T1^F$	-	[gɨ]	$T2^F$	-	[gɨ]	$T3^F$	1st token
[gɨ]	$T1^F$	-	[ge]	$T2^F$	-	[ge]	$T3^F$	1st token
[gɨ]	$T1^F$	-	[ge]	$T2^F$	-	[gɨ]	$T3^F$	2nd token
[gɨ]	$T1^F$	-	[gɨ]	$T2^F$	-	[ge]	$T3^F$	3rd token
[ga]	$T1^F$	-	[ga]	$T2^F$	-	[ga]	$T3^F$	identical
[gɐ]	$T1^F$	-	[gɐ]	$T2^F$	-	[gɐ]	$T3^F$	identical
[gɛ]	$T1^F$	-	[gɛ]	$T2^F$	-	[gɛ]	$T3^F$	identical
[ge]	$T1^F$	-	[ge]	$T2^F$	-	[ge]	$T3^F$	identical
[gɨ]	$T1^F$	-	[gɨ]	$T2^F$	-	[gɨ]	$T3^F$	identical
[gi]	$T1^F$	-	[gi]	$T2^F$	-	[gi]	$T3^F$	identical

B. Novel trials for vowel discrimination (oddity task with catch trials)

1st token			2nd token			3rd token		answer
[sɐ]	$T5^F$	-	[sɐ]	$T6^F$	-	[sɛ]	$T7^M$	3rd token
[sɐ]	$T5^F$	-	[sɛ]	$T6^F$	-	[sɐ]	$T7^M$	2nd token
[sɐ]	$T5^F$	-	[sɛ]	$T6^F$	-	[sɛ]	$T7^M$	1st token
[sɛ]	$T5^F$	-	[sɐ]	$T6^F$	-	[sɐ]	$T7^M$	1st token
[sɛ]	$T5^F$	-	[sɐ]	$T6^F$	-	[sɛ]	$T7^M$	2nd token
[sɛ]	$T5^F$	-	[sɛ]	$T6^F$	-	[sɐ]	$T7^M$	3rd token
[sɛ]	$T5^F$	-	[sɛ]	$T6^F$	-	[se]	$T7^M$	3rd token
[sɛ]	$T5^F$	-	[se]	$T6^F$	-	[sɛ]	$T7^M$	2nd token
[sɛ]	$T5^F$	-	[se]	$T6^F$	-	[se]	$T7^M$	1st token
[se]	$T5^F$	-	[sɛ]	$T6^F$	-	[sɛ]	$T7^M$	1st token
[se]	$T5^F$	-	[sɛ]	$T6^F$	-	[se]	$T7^M$	2nd token
[se]	$T5^F$	-	[se]	$T6^F$	-	[sɛ]	$T7^M$	3rd token
[se]	$T5^F$	-	[se]	$T6^F$	-	[si]	$T7^M$	3rd token
[se]	$T5^F$	-	[si]	$T6^F$	-	[se]	$T7^M$	2nd token
[se]	$T5^F$	-	[si]	$T6^F$	-	[si]	$T7^M$	1st token
[si]	$T5^F$	-	[se]	$T6^F$	-	[se]	$T7^M$	1st token
[si]	$T5^F$	-	[se]	$T6^F$	-	[si]	$T7^M$	2nd token
[si]	$T5^F$	-	[si]	$T6^F$	-	[se]	$T7^M$	3rd token
[sɐ]	$T5^F$	-	[sɐ]	$T6^F$	-	[sɨ]	$T7^M$	3rd token
[sɐ]	$T5^F$	-	[sɨ]	$T6^F$	-	[sɐ]	$T7^M$	2nd token
[sɐ]	$T5^F$	-	[sɨ]	$T6^F$	-	[sɨ]	$T7^M$	1st token
[sɨ]	$T5^F$	-	[sɐ]	$T6^F$	-	[sɐ]	$T7^M$	1st token
[sɨ]	$T5^F$	-	[sɐ]	$T6^F$	-	[sɨ]	$T7^M$	2nd token

[sɨ]	T5^{F}	-	[sɨ]	T6^{F}	-	[sɐ]	T7^{M}	3rd token
[sɐ]	T5^{F}	-	[sɐ]	T6^{F}	-	[sa]	T7^{M}	3rd token
[sɐ]	T5^{F}	-	[sa]	T6^{F}	-	[sɐ]	T7^{M}	2nd token
[sɐ]	T5^{F}	-	[sa]	T6^{F}	-	[sa]	T7^{M}	1st token
[sa]	T5^{F}	-	[sɐ]	T6^{F}	-	[sɐ]	T7^{M}	1st token
[sa]	T5^{F}	-	[sɐ]	T6^{F}	-	[sa]	T7^{M}	2nd token
[sa]	T5^{F}	-	[sa]	T6^{F}	-	[sɐ]	T7^{M}	3rd token
[sɛ]	T5^{F}	-	[sɛ]	T6^{F}	-	[sɨ]	T7^{M}	3rd token
[sɛ]	T5^{F}	-	[sɨ]	T6^{F}	-	[sɛ]	T7^{M}	2nd token
[sɛ]	T5^{F}	-	[sɨ]	T6^{F}	-	[sɨ]	T7^{M}	1st token
[sɨ]	T5^{F}	-	[sɛ]	T6^{F}	-	[sɛ]	T7^{M}	1st token
[sɨ]	T5^{F}	-	[sɛ]	T6^{F}	-	[sɨ]	T7^{M}	2nd token
[sɨ]	T5^{F}	-	[sɨ]	T6^{F}	-	[sɛ]	T7^{M}	3rd token
[se]	T5^{F}	-	[se]	T6^{F}	-	[sɨ]	T7^{M}	3rd token
[se]	T5^{F}	-	[sɨ]	T6^{F}	-	[se]	T7^{M}	2nd token
[se]	T5^{F}	-	[sɨ]	T6^{F}	-	[sɨ]	T7^{M}	1st token
[sɨ]	T5^{F}	-	[se]	T6^{F}	-	[se]	T7^{M}	1st token
[sɨ]	T5^{F}	-	[se]	T6^{F}	-	[sɨ]	T7^{M}	2nd token
[sɨ]	T5^{F}	-	[sɨ]	T6^{F}	-	[se]	T7^{M}	3rd token
[sɐ]	T5^{F}	-	[sɐ]	T6^{F}	-	[sɐ]	T7^{M}	identical
[sɛ]	T5^{F}	-	[sɛ]	T6^{F}	-	[sɛ]	T7^{M}	identical
[se]	T5^{F}	-	[se]	T6^{F}	-	[se]	T7^{M}	identical
[si]	T5^{F}	-	[si]	T6^{F}	-	[si]	T7^{M}	identical
[sɨ]	T5^{F}	-	[sɨ]	T6^{F}	-	[sɨ]	T7^{M}	identical
[sa]	T5^{F}	-	[sa]	T6^{F}	-	[sa]	T7^{M}	identical

C. Pretest trials for stress discrimination (oddity task with catch trials)

(* indicates trials not repeated in the post-test[R])

1st token			2nd token			3rd token		answer
[ˈdutiku]	T1^{F}	-	[ˈdutiku]	T2^{F}	-	[duˈtiku]	T4^{M}	3rd token
[ˈdutiku]	T1^{F}	-	[duˈtiku]	T2^{F}	-	[ˈdutiku]	T4^{M}	2nd token
[ˈdutiku]	T1^{F}	-	[duˈtiku]	T2^{F}	-	[duˈtiku]	T4^{M}	1st token
[duˈtiku]	T1^{F}	-	[ˈdutiku]	T2^{F}	-	[ˈdutiku]	T4^{M}	1st token
[duˈtiku]	T1^{F}	-	[ˈdutiku]	T2^{F}	-	[duˈtiku]	T4^{M}	2nd token
[duˈtiku]	T1^{F}	-	[duˈtiku]	T2^{F}	-	[ˈdutiku]	T4^{M}	3rd token
[duˈtiku]	T1^{F}	-	[duˈtiku]	T2^{F}	-	[dutiˈku]	T4^{M}	3rd token
[duˈtiku]	T1^{F}	-	[dutiˈku]	T2^{F}	-	[duˈtiku]	T4^{M}	2nd token
[duˈtiku]	T1^{F}	-	[dutiˈku]	T2^{F}	-	[dutiˈku]	T4^{M}	1st token

[dutiˈku]	$T1^F$	-	[duˈtiku]	$T2^F$	-	[duˈtiku]	$T4^M$	1st token
[dutiˈku]	$T1^F$	-	[duˈtiku]	$T2^F$	-	[dutiˈku]	$T4^M$	2nd token
[dutiˈku]	$T1^F$	-	[dutiˈku]	$T2^F$	-	[duˈtiku]	$T4^M$	3rd token
[ˈtikuru]	$T1^F$	-	[ˈtikuru]	$T2^F$	-	[tiˈkuru]	$T4^M$	3rd token
[ˈtikuru]	$T1^F$	-	[tiˈkuru]	$T2^F$	-	[ˈtikuru]	$T4^M$	2nd token
[ˈtikuru]	$T1^F$	-	[tiˈkuru]	$T2^F$	-	[tiˈkuru]	$T4^M$	1st token
[tiˈkuru]	$T1^F$	-	[ˈtikuru]	$T2^F$	-	[ˈtikuru]	$T4^M$	1st token
[tiˈkuru]	$T1^F$	-	[ˈtikuru]	$T2^F$	-	[tiˈkuru]	$T4^M$	2nd token
[tiˈkuru]	$T1^F$	-	[tiˈkuru]	$T2^F$	-	[ˈtikuru]	$T4^M$	3rd token
[tiˈkuru]	$T1^F$	-	[tiˈkuru]	$T2^F$	-	[tikuˈru]	$T4^M$	3rd token
[tiˈkuru]	$T1^F$	-	[tikuˈru]	$T2^F$	-	[tiˈkuru]	$T4^M$	2nd token
[tiˈkuru]	$T1^F$	-	[tikuˈru]	$T2^F$	-	[tikuˈru]	$T4^M$	1st token
[tikuˈru]	$T1^F$	-	[tiˈkuru]	$T2^F$	-	[tiˈkuru]	$T4^M$	1st token
[tikuˈru]	$T1^F$	-	[tiˈkuru]	$T2^F$	-	[tikuˈru]	$T4^M$	2nd token
[tikuˈru]	$T1^F$	-	[tikuˈru]	$T2^F$	-	[tiˈkuru]	$T4^M$	3rd token
[ˈtuduri]	$T1^F$	-	[ˈtuduri]	$T2^F$	-	[tuˈduri]	$T4^M$	3rd token
[ˈtuduri]	$T1^F$	-	[tuˈduri]	$T2^F$	-	[ˈtuduri]	$T4^M$	2nd token
[ˈtuduri]	$T1^F$	-	[tuˈduri]	$T2^F$	-	[tuˈduri]	$T4^M$	1st token
[tuˈduri]	$T1^F$	-	[ˈtuduri]	$T2^F$	-	[ˈtuduri]	$T4^M$	1st token
[tuˈduri]	$T1^F$	-	[ˈtuduri]	$T2^F$	-	[tuˈduri]	$T4^M$	2nd token
[tuˈduri]	$T1^F$	-	[tuˈduri]	$T2^F$	-	[ˈtuduri]	$T4^M$	3rd token
[tuˈduri]	$T1^F$	-	[tuˈduri]	$T2^F$	-	[tuduˈri]	$T4^M$	3rd token
[tuˈduri]	$T1^F$	-	[tuduˈri]	$T2^F$	-	[tuˈduri]	$T4^M$	2nd token
[tuˈduri]	$T1^F$	-	[tuduˈri]	$T2^F$	-	[tuduˈri]	$T4^M$	1st token
[tuduˈri]	$T1^F$	-	[tuˈduri]	$T2^F$	-	[tuˈduri]	$T4^M$	1st token
[tuduˈri]	$T1^F$	-	[tuˈduri]	$T2^F$	-	[tuduˈri]	$T4^M$	2nd token
[tuduˈri]	$T1^F$	-	[tuduˈri]	$T2^F$	-	[tuˈduri]	$T4^M$	3rd token
[ˈkiruʃi]	$T1^F$	-	[ˈkiruʃi]	$T2^F$	-	[kiˈruʃi]	$T4^M$	3rd token
[ˈkiruʃi]	$T1^F$	-	[kiˈruʃi]	$T2^F$	-	[ˈkiruʃi]	$T4^M$	2nd token
[ˈkiruʃi]	$T1^F$	-	[kiˈruʃi]	$T2^F$	-	[kiˈruʃi]	$T4^M$	1st token
[kiˈruʃi]	$T1^F$	-	[ˈkiruʃi]	$T2^F$	-	[ˈkiruʃi]	$T4^M$	1st token
[kiˈruʃi]	$T1^F$	-	[ˈkiruʃi]	$T2^F$	-	[kiˈruʃi]	$T4^M$	2nd token
[kiˈruʃi]	$T1^F$	-	[kiˈruʃi]	$T2^F$	-	[ˈkiruʃi]	$T4^M$	3rd token
[kiˈruʃi]	$T1^F$	-	[kiˈruʃi]	$T2^F$	-	[kiruˈʃi]	$T4^M$	3rd token
[kiˈruʃi]	$T1^F$	-	[kiruˈʃi]	$T2^F$	-	[kiˈruʃi]	$T4^M$	2nd token
[kiˈruʃi]	$T1^F$	-	[kiruˈʃi]	$T2^F$	-	[kiruˈʃi]	$T4^M$	1st token
[kiruˈʃi]	$T1^F$	-	[kiˈruʃi]	$T2^F$	-	[kiˈruʃi]	$T4^M$	1st token
[kiruˈʃi]	$T1^F$	-	[kiˈruʃi]	$T2^F$	-	[kiruˈʃi]	$T4^M$	2nd token
[kiruˈʃi]	$T1^F$	-	[kiruˈʃi]	$T2^F$	-	[kiˈruʃi]	$T4^M$	3rd token

[ˈtudiki]	T1^{F}	-	[ˈtudiki]	T2^{F}	-	[tuˈdiki]	T4^{M}	3rd token
[ˈtudiki]	T1^{F}	-	[tuˈdiki]	T2^{F}	-	[ˈtudiki]	T4^{M}	2nd token
[ˈtudiki]	T1^{F}	-	[tuˈdiki]	T2^{F}	-	[tuˈdiki]	T4^{M}	1st token
[tuˈdiki]	T1^{F}	-	[ˈtudiki]	T2^{F}	-	[ˈtudiki]	T4^{M}	1st token
[tuˈdiki]	T1^{F}	-	[ˈtudiki]	T2^{F}	-	[tuˈdiki]	T4^{M}	2nd token
[tuˈdiki]	T1^{F}	-	[tuˈdiki]	T2^{F}	-	[ˈtudiki]	T4^{M}	3rd token
[tuˈdiki]	T1^{F}	-	[tuˈdiki]	T2^{F}	-	[tudiˈki]	T4^{M}	3rd token
[tuˈdiki]	T1^{F}	-	[tudiˈki]	T2^{F}	-	[tuˈdiki]	T4^{M}	2nd token
[tuˈdiki]	T1^{F}	-	[tudiˈki]	T2^{F}	-	[tudiˈki]	T4^{M}	1st token
[tudiˈki]	T1^{F}	-	[tuˈdiki]	T2^{F}	-	[tuˈdiki]	T4^{M}	1st token
[tudiˈki]	T1^{F}	-	[tuˈdiki]	T2^{F}	-	[tudiˈki]	T4^{M}	2nd token
[tudiˈki]	T1^{F}	-	[tudiˈki]	T2^{F}	-	[tuˈdiki]	T4^{M}	3rd token
[ˈdutiku]	T1^{F}	-	[ˈdutiku]	T2^{F}	-	[ˈdutiku]	T4^{M}	identical
[duˈtiku]	T1^{F}	-	[duˈtiku]	T2^{F}	-	[duˈtiku]	T4^{M}	identical
[dutiˈku]	T1^{F}	-	[dutiˈku]	T2^{F}	-	[dutiˈku]	T4^{M}	identical
[ˈtikuru]	T1^{F}	-	[ˈtikuru]	T2^{F}	-	[ˈtikuru]	T4^{M}	identical
[tiˈkuru]	T1^{F}	-	[tiˈkuru]	T2^{F}	-	[tiˈkuru]	T4^{M}	identical
[tikuˈru]	T1^{F}	-	[tikuˈru]	T2^{F}	-	[tikuˈru]	T4^{M}	identical
[ˈtuduri]	T1^{F}	-	[ˈtuduri]	T2^{F}	-	[ˈtuduri]	T4^{M}	identical
[tuˈduri]	T1^{F}	-	[tuˈduri]	T2^{F}	-	[tuˈduri]	T4^{M}	identical
[tuduˈri]	T1^{F}	-	[tuduˈri]	T2^{F}	-	[tuduˈri]	T4^{M}	identical
[ˈkiruʃi]	T1^{F}	-	[ˈkiruʃi]	T2^{F}	-	[ˈkiruʃi]	T4^{M}	identical
[kiˈruʃi]	T1^{F}	-	[kiˈruʃi]	T2^{F}	-	[kiˈruʃi]	T4^{M}	identical
[kiruˈʃi]	T1^{F}	-	[kiruˈʃi]	T2^{F}	-	[kiruˈʃi]	T4^{M}	identical
[ˈtudiki]	T1^{F}	-	[ˈtudiki]	T2^{F}	-	[ˈtudiki]	T4^{M}	identical
[tuˈdiki]	T1^{F}	-	[tuˈdiki]	T2^{F}	-	[tuˈdiki]	T4^{M}	identical
[tudiˈki]	T1^{F}	-	[tudiˈki]	T2^{F}	-	[tudiˈki]	T4^{M}	identical

D. Novel trials for stress discrimination (oddity task with catch trials)

1st token			2nd token			3rd token		answer
[ˈlutinu]	T5^{F}	-	[ˈlutinu]	T6^{F}	-	[luˈtinu]	T7^{M}	3rd token
[ˈlutinu]	T5^{F}	-	[luˈtinu]	T6^{F}	-	[ˈlutinu]	T7^{M}	2nd token
[ˈlutinu]	T5^{F}	-	[luˈtinu]	T6^{F}	-	[luˈtinu]	T7^{M}	1st token
[luˈtinu]	T5^{F}	-	[ˈlutinu]	T6^{F}	-	[ˈlutinu]	T7^{M}	1st token
[luˈtinu]	T5^{F}	-	[ˈlutinu]	T6^{F}	-	[luˈtinu]	T7^{M}	2nd token
[luˈtinu]	T5^{F}	-	[luˈtinu]	T6^{F}	-	[ˈlutinu]	T7^{M}	3rd token
[luˈtinu]	T5^{F}	-	[luˈtinu]	T6^{F}	-	[lutiˈnu]	T7^{M}	3rd token
[luˈtinu]	T5^{F}	-	[lutiˈnu]	T6^{F}	-	[luˈtinu]	T7^{M}	2nd token

[luˈtinu]	$T5^F$	-	[lutiˈnu]	$T6^F$	-	[lutiˈnu]	$T7^M$	1st token
[lutiˈnu]	$T5^F$	-	[luˈtinu]	$T6^F$	-	[luˈtinu]	$T7^M$	1st token
[lutiˈnu]	$T5^F$	-	[luˈtinu]	$T6^F$	-	[lutiˈnu]	$T7^M$	2nd token
[lutiˈnu]	$T5^F$	-	[lutiˈnu]	$T6^F$	-	[luˈtinu]	$T7^M$	3rd token
[ˈbikulu]	$T5^F$	-	[ˈbikulu]	$T6^F$	-	[biˈkulu]	$T7^M$	3rd token
[ˈbikulu]	$T5^F$	-	[biˈkulu]	$T6^F$	-	[ˈbikulu]	$T7^M$	2nd token
[ˈbikulu]	$T5^F$	-	[biˈkulu]	$T6^F$	-	[biˈkulu]	$T7^M$	1st token
[biˈkulu]	$T5^F$	-	[ˈbikulu]	$T6^F$	-	[ˈbikulu]	$T7^M$	1st token
[biˈkulu]	$T5^F$	-	[ˈbikulu]	$T6^F$	-	[biˈkulu]	$T7^M$	2nd token
[biˈkulu]	$T5^F$	-	[biˈkulu]	$T6^F$	-	[ˈbikulu]	$T7^M$	3rd token
[biˈkulu]	$T5^F$	-	[biˈkulu]	$T6^F$	-	[bikuˈlu]	$T7^M$	3rd token
[biˈkulu]	$T5^F$	-	[bikuˈlu]	$T6^F$	-	[biˈkulu]	$T7^M$	2nd token
[biˈkulu]	$T5^F$	-	[bikuˈlu]	$T6^F$	-	[bikuˈlu]	$T7^M$	1st token
[bikuˈlu]	$T5^F$	-	[biˈkulu]	$T6^F$	-	[biˈkulu]	$T7^M$	1st token
[bikuˈlu]	$T5^F$	-	[biˈkulu]	$T6^F$	-	[bikuˈlu]	$T7^M$	2nd token
[bikuˈlu]	$T5^F$	-	[bikuˈlu]	$T6^F$	-	[biˈkulu]	$T7^M$	3rd token
[ˈkitumi]	$T5^F$	-	[ˈkitumi]	$T6^F$	-	[kiˈtumi]	$T7^M$	3rd token
[ˈkitumi]	$T5^F$	-	[kiˈtumi]	$T6^F$	-	[ˈkitumi]	$T7^M$	2nd token
[ˈkitumi]	$T5^F$	-	[kiˈtumi]	$T6^F$	-	[kiˈtumi]	$T7^M$	1st token
[kiˈtumi]	$T5^F$	-	[ˈkitumi]	$T6^F$	-	[ˈkitumi]	$T7^M$	1st token
[kiˈtumi]	$T5^F$	-	[ˈkitumi]	$T6^F$	-	[kiˈtumi]	$T7^M$	2nd token
[kiˈtumi]	$T5^F$	-	[kiˈtumi]	$T6^F$	-	[ˈkitumi]	$T7^M$	3rd token
[kiˈtumi]	$T5^F$	-	[kiˈtumi]	$T6^F$	-	[krituˈmi]	$T7^M$	3rd token
[kiˈtumi]	$T5^F$	-	[kituˈmi]	$T6^F$	-	[kiˈtumi]	$T7^M$	2nd token
[kiˈtumi]	$T5^F$	-	[kituˈmi]	$T6^F$	-	[kituˈmi]	$T7^M$	1st token
[kituˈmi]	$T5^F$	-	[kiˈtumi]	$T6^F$	-	[kiˈtumi]	$T7^M$	1st token
[kituˈmi]	$T5^F$	-	[kiˈtumi]	$T6^F$	-	[kituˈmi]	$T7^M$	2nd token
[kituˈmi]	$T5^F$	-	[kituˈmi]	$T6^F$	-	[kiˈtumi]	$T7^M$	3rd token
[ˈmudini]	$T5^F$	-	[ˈmudini]	$T6^F$	-	[muˈdini]	$T7^M$	3rd token
[ˈmudini]	$T5^F$	-	[muˈdini]	$T6^F$	-	[ˈmudini]	$T7^M$	2nd token
[ˈmudini]	$T5^F$	-	[muˈdini]	$T6^F$	-	[muˈdini]	$T7^M$	1st token
[muˈdini]	$T5^F$	-	[ˈmudini]	$T6^F$	-	[ˈmudini]	$T7^M$	1st token
[muˈdini]	$T5^F$	-	[ˈmudini]	$T6^F$	-	[muˈdini]	$T7^M$	2nd token
[muˈdini]	$T5^F$	-	[muˈdini]	$T6^F$	-	[ˈmudini]	$T7^M$	3rd token
[muˈdini]	$T5^F$	-	[muˈdini]	$T6^F$	-	[mudiˈni]	$T7^M$	3rd token
[muˈdini]	$T5^F$	-	[mudiˈni]	$T6^F$	-	[muˈdini]	$T7^M$	2nd token
[muˈdini]	$T5^F$	-	[mudiˈni]	$T6^F$	-	[mudiˈni]	$T7^M$	1st token
[mudiˈni]	$T5^F$	-	[muˈdini]	$T6^F$	-	[muˈdini]	$T7^M$	1st token
[mudiˈni]	$T5^F$	-	[muˈdini]	$T6^F$	-	[mudiˈni]	$T7^M$	2nd token

[mudiˈni]	$T5^F$	-	[mudiˈni]	$T6^F$	-	[muˈdini]	$T7^M$	3rd token
[ˈlutinu]	$T5^F$	-	[ˈlutinu]	$T6^F$	-	[ˈlutinu]	$T7^M$	identical
[luˈtinu]	$T5^F$	-	[luˈtinu]	$T6^F$	-	[luˈtinu]	$T7^M$	identical
[lutiˈnu]	$T5^F$	-	[lutiˈnu]	$T6^F$	-	[lutiˈnu]	$T7^M$	identical
[ˈbikulu]	$T5^F$	-	[ˈbikulu]	$T6^F$	-	[ˈbikulu]	$T7^M$	identical
[biˈkulu]	$T5^F$	-	[biˈkulu]	$T6^F$	-	[biˈkulu]	$T7^M$	identical
[bikuˈlu]	$T5^F$	-	[bikuˈlu]	$T6^F$	-	[bikuˈlu]	$T7^M$	identical
[ˈkitumi]	$T5^F$	-	[ˈkitumi]	$T6^F$	-	[ˈkitumi]	$T7^M$	identical
[kiˈtumi]	$T5^F$	-	[kiˈtumi]	$T6^F$	-	[kiˈtumi]	$T7^M$	identical
[kituˈmi]	$T5^F$	-	[kituˈmi]	$T6^F$	-	[kituˈmi]	$T7^M$	identical
[ˈmudini]	$T5^F$	-	[ˈmudini]	$T6^F$	-	[ˈmudini]	$T7^M$	identical
[muˈdini]	$T5^F$	-	[muˈdini]	$T6^F$	-	[muˈdini]	$T7^M$	identical
[mudiˈni]	$T5^F$	-	[mudiˈni]	$T6^F$	-	[mudiˈni]	$T7^M$	identical

E. Trials for consonant discrimination (oddity task with catch trials)

1st token			2nd token			3rd token		answer
[buˈpili]	$T1^F$	-	[buˈpili]	$T2^F$	-	[buˈʃili]	$T4^M$	3rd token
[buˈpili]	$T1^F$	-	[buˈʃili]	$T2^F$	-	[buˈpili]	$T4^M$	2nd token
[buˈpili]	$T1^F$	-	[buˈʃili]	$T2^F$	-	[buˈʃili]	$T4^M$	1st token
[buˈʃili]	$T1^F$	-	[buˈpili]	$T2^F$	-	[buˈpili]	$T4^M$	1st token
[buˈʃili]	$T1^F$	-	[buˈpili]	$T2^F$	-	[buˈʃili]	$T4^M$	2nd token
[buˈʃili]	$T1^F$	-	[buˈʃili]	$T2^F$	-	[buˈpili]	$T4^M$	3rd token
[tiˈfubi]	$T1^F$	-	[tiˈfubi]	$T2^F$	-	[tiˈvubi]	$T4^M$	3rd token
[tiˈfubi]	$T1^F$	-	[tiˈvubi]	$T2^F$	-	[tiˈfubi]	$T4^M$	2nd token
[tiˈfubi]	$T1^F$	-	[tiˈvubi]	$T2^F$	-	[tiˈvubi]	$T4^M$	1st token
[tiˈvubi]	$T1^F$	-	[tiˈfubi]	$T2^F$	-	[tiˈfubi]	$T4^M$	1st token
[tiˈvubi]	$T1^F$	-	[tiˈfubi]	$T2^F$	-	[tiˈvubi]	$T4^M$	2nd token
[tiˈvubi]	$T1^F$	-	[tiˈvubi]	$T2^F$	-	[tiˈfubi]	$T4^M$	3rd token
[tiˈburi]	$T1^F$	-	[tiˈburi]	$T2^F$	-	[tiˈʃuri]	$T4^M$	3rd token
[tiˈburi]	$T1^F$	-	[tiˈʃuri]	$T2^F$	-	[tiˈburi]	$T4^M$	2nd token
[tiˈburi]	$T1^F$	-	[tiˈʃuri]	$T2^F$	-	[tiˈʃuri]	$T4^M$	1st token
[tiˈʃuri]	$T1^F$	-	[tiˈburi]	$T2^F$	-	[tiˈburi]	$T4^M$	1st token
[tiˈʃuri]	$T1^F$	-	[tiˈburi]	$T2^F$	-	[tiˈʃuri]	$T4^M$	2nd token
[tiˈʃuri]	$T1^F$	-	[tiˈʃuri]	$T2^F$	-	[tiˈburi]	$T4^M$	3rd token
[kuˈmitu]	$T1^F$	-	[kuˈmitu]	$T2^F$	-	[kuˈfitu]	$T4^M$	3rd token
[kuˈmitu]	$T1^F$	-	[kuˈfitu]	$T2^F$	-	[kuˈmitu]	$T4^M$	2nd token
[kuˈmitu]	$T1^F$	-	[kuˈfitu]	$T2^F$	-	[kuˈfitu]	$T4^M$	1st token

[kuˈfitu]	T1^F	-	[kuˈmitu]	T2^F	-	[kuˈmitu]	T4^M	1st token
[kuˈfitu]	T1^F	-	[kuˈmitu]	T2^F	-	[kuˈfitu]	T4^M	2nd token
[kuˈfitu]	T1^F	-	[kuˈfitu]	T2^F	-	[kuˈmitu]	T4^M	3rd token
[diˈpuki]	T1^F	-	[diˈpuki]	T2^F	-	[diˈbuki]	T4^M	3rd token
[diˈpuki]	T1^F	-	[diˈbuki]	T2^F	-	[diˈpuki]	T4^M	2nd token
[diˈpuki]	T1^F	-	[diˈbuki]	T2^F	-	[diˈbuki]	T4^M	1st token
[diˈbuki]	T1^F	-	[diˈpuki]	T2^F	-	[diˈpuki]	T4^M	1st token
[diˈbuki]	T1^F	-	[diˈpuki]	T2^F	-	[diˈbuki]	T4^M	2nd token
[diˈbuki]	T1^F	-	[diˈbuki]	T2^F	-	[diˈpuki]	T4^M	3rd token
[zuˈmitu]	T1^F	-	[zuˈmitu]	T2^F	-	[zuˈnitu]	T4^M	3rd token
[zuˈmitu]	T1^F	-	[zuˈnitu]	T2^F	-	[zuˈmitu]	T4^M	2nd token
[zuˈmitu]	T1^F	-	[zuˈnitu]	T2^F	-	[zuˈnitu]	T4^M	1st token
[zuˈnitu]	T1^F	-	[zuˈmitu]	T2^F	-	[zuˈmitu]	T4^M	1st token
[zuˈnitu]	T1^F	-	[zuˈmitu]	T2^F	-	[zuˈnitu]	T4^M	2nd token
[zuˈnitu]	T1^F	-	[zuˈnitu]	T2^F	-	[zuˈmitu]	T4^M	3rd token
[buˈpili]	T1^F	-	[buˈpili]	T2^F	-	[buˈpili]	T4^M	identical
[buˈʃili]	T1^F	-	[buˈʃili]	T2^F	-	[buˈʃili]	T4^M	identical
[tiˈfubi]	T1^F	-	[tiˈfubi]	T2^F	-	[tiˈfubi]	T4^M	identical
[tiˈvubi]	T1^F	-	[tiˈvubi]	T2^F	-	[tiˈvubi]	T4^M	identical
[tiˈburi]	T1^F	-	[tiˈburi]	T2^F	-	[tiˈburi]	T4^M	identical
[tiˈʃuri]	T1^F	-	[tiˈʃuri]	T2^F	-	[tiˈʃuri]	T4^M	identical
[kuˈmitu]	T1^F	-	[kuˈmitu]	T2^F	-	[kuˈmitu]	T4^M	identical
[kuˈfitu]	T1^F	-	[kuˈfitu]	T2^F	-	[kuˈfitu]	T4^M	identical
[diˈpuki]	T1^F	-	[diˈpuki]	T2^F	-	[diˈpuki]	T4^M	identical
[diˈbuki]	T1^F	-	[diˈbuki]	T2^F	-	[diˈbuki]	T4^M	identical
[zuˈmitu]	T1^F	-	[zuˈmitu]	T2^F	-	[zuˈmitu]	T4^M	identical
[zuˈnitu]	T1^F	-	[zuˈnitu]	T2^F	-	[zuˈnitu]	T4^M	identical

F. Trials for training of Group *Vowels* (AX and AXB tasks)

SESSIONS 1 TO 4	A			X		answer
Session 1: block 1	[zɛ]	T2^F	-	[zɨ]	T1^F	different
	[zɨ]	T2^F	-	[zɨ]	T1^F	identical
	[zɛ]	T2^F	-	[zɛ]	T1^F	identical
	[zɨ]	T2^F	-	[zɛ]	T1^F	different
	[zɨ]	T3^F	-	[zɨ]	T1^F	identical
	[zɛ]	T3^F	-	[zɛ]	T1^F	identical
	[zɛ]	T3^F	-	[zɨ]	T1^F	different
	[zɨ]	T3^F	-	[zɛ]	T1^F	different

	[zɨ]	$T1^F$	-	[zɨ]	$T2^F$	identical
	[zɛ]	$T1^F$	-	[zɛ]	$T2^F$	identical
	[zɛ]	$T1^F$	-	[zɨ]	$T2^F$	different
	[zɨ]	$T1^F$	-	[zɛ]	$T2^F$	different
	[zɛ]	$T3^F$	-	[zɨ]	$T2^F$	different
	[zɛ]	$T3^F$	-	[zɛ]	$T2^F$	identical
	[zɨ]	$T3^F$	-	[zɛ]	$T2^F$	different
	[zɨ]	$T3^F$	-	[zɨ]	$T2^F$	identical
	[zɛ]	$T1^F$	-	[zɨ]	$T3^F$	different
	[zɨ]	$T1^F$	-	[zɨ]	$T3^F$	identical
	[zɛ]	$T1^F$	-	[zɛ]	$T3^F$	identical
	[zɨ]	$T1^F$	-	[zɛ]	$T3^F$	different
	[zɨ]	$T2^F$	-	[zɨ]	$T3^F$	identical
	[zɛ]	$T2^F$	-	[zɛ]	$T3^F$	identical
	[zɨ]	$T2^F$	-	[zɛ]	$T3^F$	different
	[zɛ]	$T2^F$	-	[zɨ]	$T3^F$	different
Session 1: block 2	[zɨ]	$T2^F$	-	[zɨ]	$T1^F$	identical
	[zɨ]	$T2^F$	-	[ze]	$T1^F$	different
	[ze]	$T2^F$	-	[ze]	$T1^F$	identical
	[ze]	$T2^F$	-	[zɨ]	$T1^F$	different
	[ze]	$T3^F$	-	[zɨ]	$T1^F$	different
	[zɨ]	$T3^F$	-	[ze]	$T1^F$	different
	[zɨ]	$T3^F$	-	[zɨ]	$T1^F$	identical
	[ze]	$T3^F$	-	[ze]	$T1^F$	identical
	[ze]	$T1^F$	-	[zɨ]	$T2^F$	different
	[zɨ]	$T1^F$	-	[zɨ]	$T2^F$	identical
	[zɨ]	$T1^F$	-	[ze]	$T2^F$	different
	[ze]	$T1^F$	-	[ze]	$T2^F$	identical
	[zɨ]	$T3^F$	-	[zɨ]	$T2^F$	identical
	[zɨ]	$T3^F$	-	[ze]	$T2^F$	different
	[ze]	$T3^F$	-	[zɨ]	$T2^F$	different
	[ze]	$T3^F$	-	[ze]	$T2^F$	identical
	[zɨ]	$T1^F$	-	[ze]	$T3^F$	different
	[ze]	$T1^F$	-	[zɨ]	$T3^F$	different
	[zɨ]	$T1^F$	-	[zɨ]	$T3^F$	identical
	[ze]	$T1^F$	-	[ze]	$T3^F$	identical
	[zɨ]	$T2^F$	-	[ze]	$T3^F$	different
	[ze]	$T2^F$	-	[zɨ]	$T3^F$	different
	[ze]	$T2^F$	-	[ze]	$T3^F$	identical

	[zɨ]	$T2^{F}$	-	[zɨ]	$T3^{F}$	identical
Session 2: block 1	[za]	$T2^{F}$	-	[zɐ]	$T1^{F}$	different
	[zɐ]	$T2^{F}$	-	[za]	$T1^{F}$	different
	[zɐ]	$T2^{F}$	-	[zɐ]	$T1^{F}$	identical
	[za]	$T2^{F}$	-	[za]	$T1^{F}$	identical
	[za]	$T3^{F}$	-	[zɐ]	$T1^{F}$	different
	[za]	$T3^{F}$	-	[za]	$T1^{F}$	identical
	[zɐ]	$T3^{F}$	-	[zɐ]	$T1^{F}$	identical
	[zɐ]	$T3^{F}$	-	[za]	$T1^{F}$	different
	[zɐ]	$T1^{F}$	-	[za]	$T2^{F}$	different
	[zɐ]	$T1^{F}$	-	[zɐ]	$T2^{F}$	identical
	[za]	$T1^{F}$	-	[za]	$T2^{F}$	identical
	[za]	$T1^{F}$	-	[zɐ]	$T2^{F}$	different
	[zɐ]	$T3^{F}$	-	[zɐ]	$T2^{F}$	identical
	[zɐ]	$T3^{F}$	-	[za]	$T2^{F}$	different
	[za]	$T3^{F}$	-	[za]	$T2^{F}$	identical
	[za]	$T3^{F}$	-	[zɐ]	$T2^{F}$	different
	[za]	$T1^{F}$	-	[za]	$T3^{F}$	identical
	[zɐ]	$T1^{F}$	-	[zɐ]	$T3^{F}$	identical
	[za]	$T1^{F}$	-	[zɐ]	$T3^{F}$	different
	[zɐ]	$T1^{F}$	-	[za]	$T3^{F}$	different
	[za]	$T2^{F}$	-	[za]	$T3^{F}$	identical
	[zɐ]	$T2^{F}$	-	[za]	$T3^{F}$	different
	[za]	$T2^{F}$	-	[zɐ]	$T3^{F}$	different
	[zɐ]	$T2^{F}$	-	[zɐ]	$T3^{F}$	identical
Session 2: block 2	[zɔ]	$T2^{F}$	-	[zo]	$T1^{F}$	different
(filler)	[zo]	$T2^{F}$	-	[zɔ]	$T1^{F}$	different
	[zɔ]	$T2^{F}$	-	[zɔ]	$T1^{F}$	identical
	[zo]	$T2^{F}$	-	[zo]	$T1^{F}$	identical
	[zɔ]	$T3^{F}$	-	[zo]	$T1^{F}$	different
	[zo]	$T3^{F}$	-	[zɔ]	$T1^{F}$	different
	[zɔ]	$T3^{F}$	-	[zɔ]	$T1^{F}$	identical
	[zo]	$T3^{F}$	-	[zo]	$T1^{F}$	identical
	[zɔ]	$T1^{F}$	-	[zo]	$T2^{F}$	different
	[zɔ]	$T1^{F}$	-	[zɔ]	$T2^{F}$	identical
	[zo]	$T1^{F}$	-	[zɔ]	$T2^{F}$	different
	[zo]	$T1^{F}$	-	[zo]	$T2^{F}$	identical
	[zɔ]	$T3^{F}$	-	[zo]	$T2^{F}$	different
	[zɔ]	$T3^{F}$	-	[zɔ]	$T2^{F}$	identical

	[zo]	$T3^{F}$	-	[zɔ]	$T2^{F}$	different
	[zo]	$T3^{F}$	-	[zo]	$T2^{F}$	identical
	[zo]	$T1^{F}$	-	[zo]	$T3^{F}$	identical
	[zɔ]	$T1^{F}$	-	[zo]	$T3^{F}$	different
	[zɔ]	$T1^{F}$	-	[zɔ]	$T3^{F}$	identical
	[zo]	$T1^{F}$	-	[zɔ]	$T3^{F}$	different
	[zo]	$T2^{F}$	-	[zɔ]	$T3^{F}$	different
	[zɔ]	$T2^{F}$	-	[zɔ]	$T3^{F}$	identical
	[zo]	$T2^{F}$	-	[zo]	$T3^{F}$	identical
	[zɔ]	$T2^{F}$	-	[zo]	$T3^{F}$	different
Session 3: block 1	[zɨ]	$T2^{F}$	-	[zɨ]	$T1^{F}$	identical
	[zɨ]	$T2^{F}$	-	[zɐ]	$T1^{F}$	different
	[zɐ]	$T2^{F}$	-	[zɐ]	$T1^{F}$	identical
	[zɐ]	$T2^{F}$	-	[zɨ]	$T1^{F}$	different
	[zɨ]	$T3^{F}$	-	[zɐ]	$T1^{F}$	different
	[zɐ]	$T3^{F}$	-	[zɨ]	$T1^{F}$	different
	[zɨ]	$T3^{F}$	-	[zɨ]	$T1^{F}$	identical
	[zɐ]	$T3^{F}$	-	[zɐ]	$T1^{F}$	identical
	[zɨ]	$T1^{F}$	-	[zɐ]	$T2^{F}$	different
	[zɐ]	$T1^{F}$	-	[zɐ]	$T2^{F}$	identical
	[zɐ]	$T1^{F}$	-	[zɨ]	$T2^{F}$	different
	[zɨ]	$T1^{F}$	-	[zɨ]	$T2^{F}$	identical
	[zɐ]	$T3^{F}$	-	[zɐ]	$T2^{F}$	identical
	[zɐ]	$T3^{F}$	-	[zɨ]	$T2^{F}$	different
	[zɨ]	$T3^{F}$	-	[zɨ]	$T2^{F}$	identical
	[zɨ]	$T3^{F}$	-	[zɐ]	$T2^{F}$	different
	[zɐ]	$T1^{F}$	-	[zɨ]	$T3^{F}$	different
	[zɐ]	$T1^{F}$	-	[zɐ]	$T3^{F}$	identical
	[zɨ]	$T1^{F}$	-	[zɨ]	$T3^{F}$	identical
	[zɨ]	$T1^{F}$	-	[zɐ]	$T3^{F}$	different
	[zɨ]	$T2^{F}$	-	[zɨ]	$T3^{F}$	identical
	[zɨ]	$T2^{F}$	-	[zɐ]	$T3^{F}$	different
	[zɐ]	$T2^{F}$	-	[zɨ]	$T3^{F}$	different
	[zɐ]	$T2^{F}$	-	[zɐ]	$T3^{F}$	identical
Session 3: block 2	[zɐ]	$T2^{F}$	-	[zɐ]	$T1^{F}$	identical
	[zɛ]	$T2^{F}$	-	[zɐ]	$T1^{F}$	different
	[zɐ]	$T2^{F}$	-	[zɛ]	$T1^{F}$	different
	[zɛ]	$T2^{F}$	-	[zɛ]	$T1^{F}$	identical
	[zɐ]	$T3^{F}$	-	[zɐ]	$T1^{F}$	identical

	[zɐ]	$T3^F$	-	[zɛ]	$T1^F$	different
	[zɛ]	$T3^F$	-	[zɐ]	$T1^F$	different
	[zɛ]	$T3^F$	-	[zɛ]	$T1^F$	identical
	[zɐ]	$T1^F$	-	[zɐ]	$T2^F$	identical
	[zɛ]	$T1^F$	-	[zɛ]	$T2^F$	identical
	[zɛ]	$T1^F$	-	[zɐ]	$T2^F$	different
	[zɐ]	$T1^F$	-	[zɛ]	$T2^F$	different
	[zɛ]	$T3^F$	-	[zɛ]	$T2^F$	identical
	[zɐ]	$T3^F$	-	[zɛ]	$T2^F$	different
	[zɐ]	$T3^F$	-	[zɐ]	$T2^F$	identical
	[zɛ]	$T3^F$	-	[zɐ]	$T2^F$	different
	[zɐ]	$T1^F$	-	[zɛ]	$T3^F$	different
	[zɛ]	$T1^F$	-	[zɛ]	$T3^F$	identical
	[zɛ]	$T1^F$	-	[zɐ]	$T3^F$	different
	[zɐ]	$T1^F$	-	[zɐ]	$T3^F$	identical
	[zɛ]	$T2^F$	-	[zɐ]	$T3^F$	different
	[zɛ]	$T2^F$	-	[zɛ]	$T3^F$	identical
	[zɐ]	$T2^F$	-	[zɐ]	$T3^F$	identical
	[zɐ]	$T2^F$	-	[zɛ]	$T3^F$	different
Session 4: block 1	[ze]	$T2^F$	-	[ze]	$T1^F$	identical
	[zɛ]	$T2^F$	-	[zɛ]	$T1^F$	identical
	[zɛ]	$T2^F$	-	[ze]	$T1^F$	different
	[ze]	$T2^F$	-	[zɛ]	$T1^F$	different
	[zɛ]	$T3^F$	-	[ze]	$T1^F$	different
	[ze]	$T3^F$	-	[zɛ]	$T1^F$	different
	[zɛ]	$T3^F$	-	[zɛ]	$T1^F$	identical
	[ze]	$T3^F$	-	[ze]	$T1^F$	identical
	[ze]	$T1^F$	-	[ze]	$T2^F$	identical
	[zɛ]	$T1^F$	-	[zɛ]	$T2^F$	identical
	[ze]	$T1^F$	-	[zɛ]	$T2^F$	different
	[zɛ]	$T1^F$	-	[ze]	$T2^F$	different
	[ze]	$T3^F$	-	[zɛ]	$T2^F$	different
	[zɛ]	$T3^F$	-	[zɛ]	$T2^F$	identical
	[zɛ]	$T3^F$	-	[ze]	$T2^F$	different
	[ze]	$T3^F$	-	[ze]	$T2^F$	identical
	[ze]	$T1^F$	-	[ze]	$T3^F$	identical
	[zɛ]	$T1^F$	-	[ze]	$T3^F$	different
	[ze]	$T1^F$	-	[zɛ]	$T3^F$	different
	[zɛ]	$T1^F$	-	[zɛ]	$T3^F$	identical

	[zɛ]	$T2^{F}$	-	[ze]	$T3^{F}$	different
	[zɛ]	$T2^{F}$	-	[zɛ]	$T3^{F}$	identical
	[ze]	$T2^{F}$	-	[ze]	$T3^{F}$	identical
	[ze]	$T2^{F}$	-	[zɛ]	$T3^{F}$	different
Session 4: block 2	[ze]	$T2^{F}$	-	[ze]	$T1^{F}$	identical
	[zi]	$T2^{F}$	-	[zi]	$T1^{F}$	identical
	[ze]	$T2^{F}$	-	[zi]	$T1^{F}$	different
	[zi]	$T2^{F}$	-	[ze]	$T1^{F}$	different
	[ze]	$T3^{F}$	-	[ze]	$T1^{F}$	identical
	[ze]	$T3^{F}$	-	[zi]	$T1^{F}$	different
	[zi]	$T3^{F}$	-	[zi]	$T1^{F}$	identical
	[zi]	$T3^{F}$	-	[ze]	$T1^{F}$	different
	[ze]	$T1^{F}$	-	[zi]	$T2^{F}$	different
	[zi]	$T1^{F}$	-	[zi]	$T2^{F}$	identical
	[ze]	$T1^{F}$	-	[ze]	$T2^{F}$	identical
	[zi]	$T1^{F}$	-	[ze]	$T2^{F}$	different
	[zi]	$T3^{F}$	-	[ze]	$T2^{F}$	different
	[ze]	$T3^{F}$	-	[ze]	$T2^{F}$	identical
	[ze]	$T3^{F}$	-	[zi]	$T2^{F}$	different
	[zi]	$T3^{F}$	-	[zi]	$T2^{F}$	identical
	[zi]	$T1^{F}$	-	[zi]	$T3^{F}$	identical
	[ze]	$T1^{F}$	-	[zi]	$T3^{F}$	different
	[zi]	$T1^{F}$	-	[ze]	$T3^{F}$	different
	[ze]	$T1^{F}$	-	[ze]	$T3^{F}$	identical
	[zi]	$T2^{F}$	-	[zi]	$T3^{F}$	identical
	[zi]	$T2^{F}$	-	[ze]	$T3^{F}$	different
	[ze]	$T2^{F}$	-	[ze]	$T3^{F}$	identical
	[ze]	$T2^{F}$	-	[zi]	$T3^{F}$	different

SESSIONS 5 AND 6	**A**		**X**		**B**	**answer**
Block 1	[zɐ] $T1^{F}$	-	[zɐ] $T2^{F}$	-	[za] $T3^{F}$	A
	[za] $T1^{F}$	-	[za] $T2^{F}$	-	[zɐ] $T3^{F}$	A
	[za] $T1^{F}$	-	[zɐ] $T2^{F}$	-	[zɐ] $T3^{F}$	B
	[zɐ] $T1^{F}$	-	[za] $T2^{F}$	-	[za] $T3^{F}$	B
	[zɐ] $T1^{F}$	-	[za] $T2^{F}$	-	[za] $T3^{F}$	B
	[zɐ] $T1^{F}$	-	[zɐ] $T2^{F}$	-	[za] $T3^{F}$	A
	[za] $T1^{F}$	-	[za] $T2^{F}$	-	[zɐ] $T3^{F}$	A
	[za] $T1^{F}$	-	[zɐ] $T2^{F}$	-	[zɐ] $T3^{F}$	B
Block 2	[zɛ] $T1^{F}$	-	[ze] $T2^{F}$	-	[ze] $T3^{F}$	B

	[zɛ]	$T1^F$	-	[zɛ]	$T2^F$	-	[ze]	$T3^F$	A
	[zɛ]	$T1^F$	-	[zɛ]	$T2^F$	-	[ze]	$T3^F$	A
	[zɛ]	$T1^F$	-	[ze]	$T2^F$	-	[ze]	$T3^F$	B
	[ze]	$T1^F$	-	[zɛ]	$T2^F$	-	[zɛ]	$T3^F$	B
	[ze]	$T1^F$	-	[zɛ]	$T2^F$	-	[zɛ]	$T3^F$	B
	[ze]	$T1^F$	-	[ze]	$T2^F$	-	[zɛ]	$T3^F$	A
Block 3	[ze]	$T1^F$	-	[ze]	$T2^F$	-	[zɛ]	$T3^F$	A
	[zi]	$T1^F$	-	[zi]	$T2^F$	-	[ze]	$T3^F$	A
	[zi]	$T1^F$	-	[ze]	$T2^F$	-	[ze]	$T3^F$	B
	[ze]	$T1^F$	-	[zi]	$T2^F$	-	[zi]	$T3^F$	B
	[zi]	$T1^F$	-	[ze]	$T2^F$	-	[ze]	$T3^F$	B
	[zi]	$T1^F$	-	[zi]	$T2^F$	-	[ze]	$T3^F$	A
	[ze]	$T1^F$	-	[zi]	$T2^F$	-	[zi]	$T3^F$	B
	[ze]	$T1^F$	-	[ze]	$T2^F$	-	[zi]	$T3^F$	A
	[ze]	$T1^F$	-	[ze]	$T2^F$	-	[zi]	$T3^F$	A
Block 4	[zɛ]	$T1^F$	-	[zɛ]	$T2^F$	-	[zɨ]	$T3^F$	A
	[ze]	$T1^F$	-	[zɨ]	$T2^F$	-	[zɨ]	$T3^F$	B
	[zɛ]	$T1^F$	-	[zɨ]	$T2^F$	-	[zɨ]	$T3^F$	B
	[zɨ]	$T1^F$	-	[zɨ]	$T2^F$	-	[zɛ]	$T3^F$	A
Block 5	[zɨ]	$T1^F$	-	[zɛ]	$T2^F$	-	[zɛ]	$T3^F$	B
	[zɨ]	$T1^F$	-	[ze]	$T2^F$	-	[ze]	$T3^F$	B
	[zɨ]	$T1^F$	-	[zɨ]	$T2^F$	-	[ze]	$T3^F$	A
	[ze]	$T1^F$	-	[ze]	$T2^F$	-	[zɨ]	$T3^F$	A
Block 6	[zɐ]	$T1^F$	-	[zɐ]	$T2^F$	-	[zɨ]	$T3^F$	A
	[zɨ]	$T1^F$	-	[zɨ]	$T2^F$	-	[zɐ]	$T3^F$	A
	[zɨ]	$T1^F$	-	[zɨ]	$T2^F$	-	[zɐ]	$T3^F$	A
	[zɨ]	$T1^F$	-	[zɐ]	$T2^F$	-	[zɐ]	$T3^F$	B
	[zɐ]	$T1^F$	-	[zɐ]	$T2^F$	-	[zɨ]	$T3^F$	A
	[zɨ]	$T1^F$	-	[zɐ]	$T2^F$	-	[zɐ]	$T3^F$	B
	[zɐ]	$T1^F$	-	[zɨ]	$T2^F$	-	[zɨ]	$T3^F$	B
	[zɐ]	$T1^F$	-	[zɨ]	$T2^F$	-	[zɨ]	$T3^F$	B
Block 7	[zɐ]	$T1^F$	-	[zɛ]	$T2^F$	-	[zɛ]	$T3^F$	B
	[zɛ]	$T1^F$	-	[zɐ]	$T2^F$	-	[zɐ]	$T3^F$	B
	[zɛ]	$T1^F$	-	[zɛ]	$T2^F$	-	[zɐ]	$T3^F$	A
	[zɛ]	$T1^F$	-	[zɛ]	$T2^F$	-	[zɐ]	$T3^F$	A
	[zɐ]	$T1^F$	-	[zɛ]	$T2^F$	-	[zɛ]	$T3^F$	B
	[zɐ]	$T1^F$	-	[zɐ]	$T2^F$	-	[zɛ]	$T3^F$	A
	[zɐ]	$T1^F$	-	[zɐ]	$T2^F$	-	[zɛ]	$T3^F$	A
	[zɛ]	$T1^F$	-	[zɐ]	$T2^F$	-	[zɐ]	$T3^F$	B

G. Trials for training of Group *Stress* (AX and AXB tasks)

SESSIONS 1 TO 4	A			X		answer
Session 1: block 1	[ˈbifuʃi]	$T4^{M}$	-	[ˈbifuʃi]	$T1^{F}$	identical
	[ˈbifuʃi]	$T4^{M}$	-	[biˈfuʃi]	$T1^{F}$	different
	[biˈfuʃi]	$T4^{M}$	-	[ˈbifuʃi]	$T1^{F}$	different
	[biˈfuʃi]	$T4^{M}$	-	[biˈfuʃi]	$T1^{F}$	identical
	[ˈbifuʃi]	$T2^{F}$	-	[ˈbifuʃi]	$T4^{M}$	identical
	[ˈbifuʃi]	$T2^{F}$	-	[biˈfuʃi]	$T4^{M}$	different
	[biˈfuʃi]	$T2^{F}$	-	[ˈbifuʃi]	$T4^{M}$	different
	[biˈfuʃi]	$T2^{F}$	-	[biˈfuʃi]	$T4^{M}$	identical
	[biˈfuʃi]	$T1^{F}$	-	[biˈfuʃi]	$T2^{F}$	identical
	[ˈbifuʃi]	$T1^{F}$	-	[biˈfuʃi]	$T2^{F}$	different
	[biˈfuʃi]	$T1^{F}$	-	[ˈbifuʃi]	$T2^{F}$	different
	[ˈbifuʃi]	$T1^{F}$	-	[ˈbifuʃi]	$T2^{F}$	identical
Session 1: block 2	[ˈvumipi]	$T4^{M}$	-	[vuˈmipi]	$T1^{F}$	different
	[ˈvumipi]	$T4^{M}$	-	[ˈvumipi]	$T1^{F}$	identical
	[vuˈmipi]	$T4^{M}$	-	[ˈvumipi]	$T1^{F}$	different
	[vuˈmipi]	$T4^{M}$	-	[vuˈmipi]	$T1^{F}$	identical
	[ˈvumipi]	$T2^{F}$	-	[vuˈmipi]	$T4^{M}$	different
	[ˈvumipi]	$T2^{F}$	-	[ˈvumipi]	$T4^{M}$	identical
	[vuˈmipi]	$T2^{F}$	-	[vuˈmipi]	$T4^{M}$	identical
	[vuˈmipi]	$T2^{F}$	-	[ˈvumipi]	$T4^{M}$	different
	[vuˈmipi]	$T1^{F}$	-	[ˈvumipi]	$T2^{F}$	different
	[ˈvumipi]	$T1^{F}$	-	[ˈvumipi]	$T2^{F}$	identical
	[ˈvumipi]	$T1^{F}$	-	[vuˈmipi]	$T2^{F}$	different
	[vuˈmipi]	$T1^{F}$	-	[vuˈmipi]	$T2^{F}$	identical
Session 1: block 3	[ˈmupiʃu]	$T4^{M}$	-	[ˈmupiʃu]	$T1^{F}$	identical
	[muˈpiʃu]	$T4^{M}$	-	[muˈpiʃu]	$T1^{F}$	identical
	[muˈpiʃu]	$T4^{M}$	-	[ˈmupiʃu]	$T1^{F}$	different
	[ˈmupiʃu]	$T4^{M}$	-	[muˈpiʃu]	$T1^{F}$	different
	[ˈmupiʃu]	$T2^{F}$	-	[ˈmupiʃu]	$T4^{M}$	identical
	[muˈpiʃu]	$T2^{F}$	-	[ˈmupiʃu]	$T4^{M}$	different
	[ˈmupiʃu]	$T2^{F}$	-	[muˈpiʃu]	$T4^{M}$	different
	[muˈpiʃu]	$T2^{F}$	-	[muˈpiʃu]	$T4^{M}$	identical
	[muˈpiʃu]	$T1^{F}$	-	[muˈpiʃu]	$T2^{F}$	identical
	[muˈpiʃu]	$T1^{F}$	-	[ˈmupiʃu]	$T2^{F}$	different
	[ˈmupiʃu]	$T1^{F}$	-	[muˈpiʃu]	$T2^{F}$	different
	[ˈmupiʃu]	$T1^{F}$	-	[ˈmupiʃu]	$T2^{F}$	identical
Session 1: block 4	[fiˈluvi]	$T2^{F}$	-	[ˈfiluvi]	$T4^{M}$	different
	[ˈfiluvi]	$T2^{F}$	-	[ˈfiluvi]	$T4^{M}$	identical

	[fiˈluvi]	$T2^F$	-	[fiˈluvi]	$T4^M$	identical
	[ˈfiluvi]	$T2^F$	-	[fiˈluvi]	$T4^M$	different
	[fiˈluvi]	$T1^F$	-	[fiˈluvi]	$T2^F$	identical
	[ˈfiluvi]	$T1^F$	-	[ˈfiluvi]	$T2^F$	identical
	[fiˈluvi]	$T1^F$	-	[ˈfiluvi]	$T2^F$	different
	[ˈfiluvi]	$T1^F$	-	[fiˈluvi]	$T2^F$	different
	[fiˈluvi]	$T4^M$	-	[ˈfiluvi]	$T1^F$	different
	[ˈfiluvi]	$T4^M$	-	[fiˈluvi]	$T1^F$	different
	[fiˈluvi]	$T4^M$	-	[fiˈluvi]	$T1^F$	identical
	[ˈfiluvi]	$T4^M$	-	[ˈfiluvi]	$T1^F$	identical
Session 2: block 1	[bifuˈʃi]	$T4^M$	-	[biˈfuʃi]	$T1^F$	different
	[bifuˈʃi]	$T4^M$	-	[bifuˈʃi]	$T1^F$	identical
	[bifuˈʃi]	$T2^F$	-	[biˈfuʃi]	$T4^M$	different
	[bifuˈʃi]	$T2^F$	-	[bifuˈʃi]	$T4^M$	identical
	[bifuˈʃi]	$T1^F$	-	[bifuˈʃi]	$T2^F$	identical
	[bifuˈʃi]	$T1^F$	-	[biˈfuʃi]	$T2^F$	different
	[biˈfuʃi]	$T4^M$	-	[bifuˈʃi]	$T1^F$	different
	[biˈfuʃi]	$T4^M$	-	[biˈfuʃi]	$T1^F$	identical
	[biˈfuʃi]	$T2^F$	-	[biˈfuʃi]	$T4^M$	identical
	[biˈfuʃi]	$T2^F$	-	[bifuˈʃi]	$T4^M$	different
	[biˈfuʃi]	$T1^F$	-	[bifuˈʃi]	$T2^F$	different
	[biˈfuʃi]	$T1^F$	-	[biˈfuʃi]	$T2^F$	identical
Session 2: block 2	[vumiˈpi]	$T4^M$	-	[vumiˈpi]	$T1^F$	identical
	[vumiˈpi]	$T4^M$	-	[vuˈmipi]	$T1^F$	different
	[vumiˈpi]	$T2^F$	-	[vumiˈpi]	$T4^M$	identical
	[vumiˈpi]	$T2^F$	-	[vuˈmipi]	$T4^M$	different
	[vumiˈpi]	$T1^F$	-	[vumiˈpi]	$T2^F$	identical
	[vumiˈpi]	$T1^F$	-	[vuˈmipi]	$T2^F$	different
	[vuˈmipi]	$T4^M$	-	[vuˈmipi]	$T1^F$	identical
	[vuˈmipi]	$T4^M$	-	[vumiˈpi]	$T1^F$	different
	[vuˈmipi]	$T2^F$	-	[vumiˈpi]	$T4^M$	different
	[vuˈmipi]	$T2^F$	-	[vuˈmipi]	$T4^M$	identical
	[vuˈmipi]	$T1^F$	-	[vumiˈpi]	$T2^F$	different
	[vuˈmipi]	$T1^F$	-	[vuˈmipi]	$T2^F$	identical
Session 2: block 3	[filuˈvi]	$T4^M$	-	[filuˈvi]	$T1^F$	identical
	[filuˈvi]	$T4^M$	-	[fiˈluvi]	$T1^F$	different
	[filuˈvi]	$T2^F$	-	[fiˈluvi]	$T4^M$	different
	[filuˈvi]	$T2^F$	-	[filuˈvi]	$T4^M$	identical
	[filuˈvi]	$T1^F$	-	[filuˈvi]	$T2^F$	identical
	[filuˈvi]	$T1^F$	-	[fiˈluvi]	$T2^F$	different
	[fiˈluvi]	$T4^M$	-	[fiˈluvi]	$T1^F$	identical

	[fiˈluvi]	$T4^M$	-	[filuˈvi]	$T1^F$	different
	[fiˈluvi]	$T2^F$	-	[fiˈluvi]	$T4^M$	identical
	[fiˈluvi]	$T2^F$	-	[filuˈvi]	$T4^M$	different
	[fiˈluvi]	$T1^F$	-	[filuˈvi]	$T2^F$	different
	[fiˈluvi]	$T1^F$	-	[fiˈluvi]	$T2^F$	identical
Session 2: block 4	[mupiˈʃu]	$T4^M$	-	[muˈpiʃu]	$T1^F$	different
	[mupiˈʃu]	$T4^M$	-	[mupiˈʃu]	$T1^F$	identical
	[mupiˈʃu]	$T2^F$	-	[mupiˈʃu]	$T4^M$	identical
	[mupiˈʃu]	$T2^F$	-	[muˈpiʃu]	$T4^M$	different
	[mupiˈʃu]	$T1^F$	-	[muˈpiʃu]	$T2^F$	different
	[mupiˈʃu]	$T1^F$	-	[mupiˈʃu]	$T2^F$	identical
	[muˈpiʃu]	$T4^M$	-	[mupiˈʃu]	$T1^F$	different
	[muˈpiʃu]	$T4^M$	-	[muˈpiʃu]	$T1^F$	identical
	[muˈpiʃu]	$T2^F$	-	[mupiˈʃu]	$T4^M$	different
	[muˈpiʃu]	$T2^F$	-	[muˈpiʃu]	$T4^M$	identical
	[muˈpiʃu]	$T1^F$	-	[mupiˈʃu]	$T2^F$	different
	[muˈpiʃu]	$T1^F$	-	[muˈpiʃu]	$T2^F$	identical
Session 3: block 1	[vumiˈpi]	$T4^M$	-	[vumiˈpi]	$T1^F$	identical
	[vuˈmipi]	$T4^M$	-	[vuˈmipi]	$T1^F$	identical
	[vuˈmipi]	$T4^M$	-	[vuˈmipi]	$T1^F$	identical
	[ˈvumipi]	$T4^M$	-	[ˈvumipi]	$T1^F$	identical
	[vuˈmipi]	$T4^M$	-	[ˈvumipi]	$T1^F$	different
	[vuˈmipi]	$T4^M$	-	[vumiˈpi]	$T1^F$	different
	[vumiˈpi]	$T4^M$	-	[vuˈmipi]	$T1^F$	different
	[ˈvumipi]	$T4^M$	-	[vuˈmipi]	$T1^F$	different
	[vuˈmipi]	$T2^F$	-	[ˈvumipi]	$T4^M$	different
	[vumiˈpi]	$T2^F$	-	[vumiˈpi]	$T4^M$	identical
	[vumiˈpi]	$T2^F$	-	[vuˈmipi]	$T4^M$	different
	[vuˈmipi]	$T2^F$	-	[vuˈmipi]	$T4^M$	identical
	[ˈvumipi]	$T2^F$	-	[ˈvumipi]	$T4^M$	identical
	[vuˈmipi]	$T2^F$	-	[vumiˈpi]	$T4^M$	different
	[vuˈmipi]	$T2^F$	-	[vuˈmipi]	$T4^M$	identical
	[ˈvumipi]	$T2^F$	-	[vuˈmipi]	$T4^M$	different
	[vumiˈpi]	$T1^F$	-	[vuˈmipi]	$T2^F$	different
	[vuˈmipi]	$T1^F$	-	[vumiˈpi]	$T2^F$	different
	[vuˈmipi]	$T1^F$	-	[vuˈmipi]	$T2^F$	identical
	[vuˈmipi]	$T1^F$	-	[ˈvumipi]	$T2^F$	different
	[vumiˈpi]	$T1^F$	-	[vumiˈpi]	$T2^F$	identical
	[ˈvumipi]	$T1^F$	-	[ˈvumipi]	$T2^F$	identical
	[vuˈmipi]	$T1^F$	-	[vuˈmipi]	$T2^F$	identical
	[ˈvumipi]	$T1^F$	-	[vuˈmipi]	$T2^F$	different

Session 3: block 2	[mupiˈʃu]	T4^M	-	[muˈpiʃu]	T1^F	different
	[muˈpiʃu]	T4^M	-	[muˈpiʃu]	T1^F	identical
	[muˈpiʃu]	T4^M	-	[muˈpiʃu]	T1^F	identical
	[muˈpiʃu]	T4^M	-	[mupiˈʃu]	T1^F	different
	[ˈmupiʃu]	T4^M	-	[ˈmupiʃu]	T1^F	identical
	[mupiˈʃu]	T4^M	-	[mupiˈʃu]	T1^F	identical
	[ˈmupiʃu]	T4^M	-	[muˈpiʃu]	T1^F	different
	[muˈpiʃu]	T4^M	-	[ˈmupiʃu]	T1^F	different
	[muˈpiʃu]	T2^F	-	[muˈpiʃu]	T4^M	identical
	[mupiˈʃu]	T2^F	-	[muˈpiʃu]	T4^M	different
	[ˈmupiʃu]	T2^F	-	[ˈmupiʃu]	T4^M	identical
	[mupiˈʃu]	T2^F	-	[mupiˈʃu]	T4^M	identical
	[ˈmupiʃu]	T2^F	-	[muˈpiʃu]	T4^M	different
	[muˈpiʃu]	T2^F	-	[mupiˈʃu]	T4^M	different
	[muˈpiʃu]	T2^F	-	[ˈmupiʃu]	T4^M	different
	[muˈpiʃu]	T2^F	-	[muˈpiʃu]	T4^M	identical
	[mupiˈʃu]	T1^F	-	[muˈpiʃu]	T2^F	different
	[muˈpiʃu]	T1^F	-	[ˈmupiʃu]	T2^F	different
	[muˈpiʃu]	T1^F	-	[muˈpiʃu]	T2^F	identical
	[mupiˈʃu]	T1^F	-	[mupiˈʃu]	T2^F	identical
	[muˈpiʃu]	T1^F	-	[mupiˈʃu]	T2^F	different
	[ˈmupiʃu]	T1^F	-	[muˈpiʃu]	T2^F	different
	[ˈmupiʃu]	T1^F	-	[ˈmupiʃu]	T2^F	identical
	[muˈpiʃu]	T1^F	-	[muˈpiʃu]	T2^F	identical
Session 4: block 1	[ˈbifuʃi]	T4^M	-	[biˈfuʃi]	T1^F	different
	[biˈfuʃi]	T4^M	-	[bifuˈʃi]	T1^F	different
	[biˈfuʃi]	T4^M	-	[biˈfuʃi]	T1^F	identical
	[ˈbifuʃi]	T4^M	-	[ˈbifuʃi]	T1^F	identical
	[bifuˈʃi]	T4^M	-	[bifuˈʃi]	T1^F	identical
	[biˈfuʃi]	T4^M	-	[biˈfuʃi]	T1^F	identical
	[biˈfuʃi]	T4^M	-	[ˈbifuʃi]	T1^F	different
	[bifuˈʃi]	T4^M	-	[biˈfuʃi]	T1^F	different
	[biˈfuʃi]	T2^F	-	[biˈfuʃi]	T4^M	identical
	[ˈbifuʃi]	T2^F	-	[ˈbifuʃi]	T4^M	identical
	[biˈfuʃi]	T2^F	-	[biˈfuʃi]	T4^M	identical
	[bifuˈʃi]	T2^F	-	[biˈfuʃi]	T4^M	different
	[biˈfuʃi]	T2^F	-	[bifuˈʃi]	T4^M	different
	[biˈfuʃi]	T2^F	-	[ˈbifuʃi]	T4^M	different
	[ˈbifuʃi]	T2^F	-	[biˈfuʃi]	T4^M	different
	[bifuˈʃi]	T2^F	-	[bifuˈʃi]	T4^M	identical
	[bifuˈʃi]	T1^F	-	[biˈfuʃi]	T2^F	different

	[biˈfuʃi]	T1^F	-	[biˈfuʃi]	T2^F	identical
	[bifuˈʃi]	T1^F	-	[bifuˈʃi]	T2^F	identical
	[biˈfuʃi]	T1^F	-	[biˈfuʃi]	T2^F	identical
	[ˈbifuʃi]	T1^F	-	[ˈbifuʃi]	T2^F	identical
	[ˈbifuʃi]	T1^F	-	[biˈfuʃi]	T2^F	different
	[biˈfuʃi]	T1^F	-	[bifuˈʃi]	T2^F	different
	[biˈfuʃi]	T1^F	-	[ˈbifuʃi]	T2^F	different
Session 4: block 2	[fiˈluvi]	T4^M	-	[fiˈluvi]	T1^F	identical
	[filuˈvi]	T4^M	-	[fiˈluvi]	T1^F	different
	[ˈfiluvi]	T4^M	-	[ˈfiluvi]	T1^F	identical
	[ˈfiluvi]	T4^M	-	[fiˈluvi]	T1^F	different
	[fiˈluvi]	T4^M	-	[ˈfiluvi]	T1^F	different
	[fiˈluvi]	T4^M	-	[fiˈluvi]	T1^F	identical
	[fiˈluvi]	T4^M	-	[filuˈvi]	T1^F	different
	[filuˈvi]	T4^M	-	[filuˈvi]	T1^F	identical
	[fiˈluvi]	T2^F	-	[fiˈluvi]	T4^M	identical
	[fiˈluvi]	T2^F	-	[fiˈluvi]	T4^M	identical
	[fiˈluvi]	T2^F	-	[ˈfiluvi]	T4^M	different
	[ˈfiluvi]	T2^F	-	[ˈfiluvi]	T4^M	identical
	[fiˈluvi]	T2^F	-	[filuˈvi]	T4^M	different
	[ˈfiluvi]	T2^F	-	[fiˈluvi]	T4^M	different
	[filuˈvi]	T2^F	-	[fiˈluvi]	T4^M	different
	[filuˈvi]	T2^F	-	[filuˈvi]	T4^M	identical
	[fiˈluvi]	T1^F	-	[ˈfiluvi]	T2^F	different
	[filuˈvi]	T1^F	-	[filuˈvi]	T2^F	identical
	[fiˈluvi]	T1^F	-	[fiˈluvi]	T2^F	identical
	[ˈfiluvi]	T1^F	-	[fiˈluvi]	T2^F	different
	[ˈfiluvi]	T1^F	-	[ˈfiluvi]	T2^F	identical
	[filuˈvi]	T1^F	-	[fiˈluvi]	T2^F	different
	[fiˈluvi]	T1^F	-	[fiˈluvi]	T2^F	identical
	[fiˈluvi]	T1^F	-	[filuˈvi]	T2^F	different

SESSIONS 5 AND 6	**A**			**X**			**B**		**answer**
Block 1	[biˈfuʃi]	T1^F	-	[bifuˈʃi]	T2^F	-	[bifuˈʃi]	T4^M	B
	[biˈfuʃi]	T1^F	-	[biˈfuʃi]	T2^F	-	[ˈbifuʃi]	T4^M	A
	[bifuˈʃi]	T1^F	-	[biˈfuʃi]	T2^F	-	[biˈfuʃi]	T4^M	B
	[biˈfuʃi]	T1^F	-	[biˈfuʃi]	T2^F	-	[bifuˈʃi]	T4^M	A
	[ˈbifuʃi]	T1^F	-	[ˈbifuʃi]	T2^F	-	[biˈfuʃi]	T4^M	A
	[biˈfuʃi]	T1^F	-	[ˈbifuʃi]	T2^F	-	[ˈbifuʃi]	T4^M	B
	[bifuˈʃi]	T1^F	-	[bifuˈʃi]	T2^F	-	[biˈfuʃi]	T4^M	A
	[ˈbifuʃi]	T1^F	-	[biˈfuʃi]	T2^F	-	[biˈfuʃi]	T4^M	B

Block 2	[vuˈmipi]	$T1^F$	-	[vumiˈpi]	$T2^F$	-	[vumiˈpi]	$T4^M$	B
	[vuˈmipi]	$T1^F$	-	[ˈvumipi]	$T2^F$	-	[ˈvumipi]	$T4^M$	B
	[vuˈmipi]	$T1^F$	-	[vuˈmipi]	$T2^F$	-	[vumiˈpi]	$T4^M$	A
	[vumiˈpi]	$T1^F$	-	[vumiˈpi]	$T2^F$	-	[vuˈmipi]	$T4^M$	A
	[vuˈmipi]	$T1^F$	-	[vuˈmipi]	$T2^F$	-	[ˈvumipi]	$T4^M$	A
	[ˈvumipi]	$T1^F$	-	[vuˈmipi]	$T2^F$	-	[vuˈmipi]	$T4^M$	B
	[vumiˈpi]	$T1^F$	-	[vuˈmipi]	$T2^F$	-	[vuˈmipi]	$T4^M$	B
	[ˈvumipi]	$T1^F$	-	[ˈvumipi]	$T2^F$	-	[vuˈmipi]	$T4^M$	A
Block 3	[fiˈluvi]	$T1^F$	-	[ˈfiluvi]	$T2^F$	-	[ˈfiluvi]	$T4^M$	B
	[filuˈvi]	$T1^F$	-	[filuˈvi]	$T2^F$	-	[fiˈluvi]	$T4^M$	A
	[fiˈluvi]	$T1^F$	-	[filuˈvi]	$T2^F$	-	[filuˈvi]	$T4^M$	B
	[filuˈvi]	$T1^F$	-	[fiˈluvi]	$T2^F$	-	[fiˈluvi]	$T4^M$	B
	[ˈfiluvi]	$T1^F$	-	[fiˈluvi]	$T2^F$	-	[fiˈluvi]	$T4^M$	B
	[fiˈluvi]	$T1^F$	-	[fiˈluvi]	$T2^F$	-	[filuˈvi]	$T4^M$	A
	[ˈfiluvi]	$T1^F$	-	[ˈfiluvi]	$T2^F$	-	[fiˈluvi]	$T4^M$	A
	[fiˈluvi]	$T1^F$	-	[fiˈluvi]	$T2^F$	-	[ˈfiluvi]	$T4^M$	A
Block 4	[ˈmupiʃu]	$T1^F$	-	[ˈmupiʃu]	$T2^F$	-	[muˈpiʃu]	$T4^M$	A
	[mupiˈʃu]	$T1^F$	-	[mupiˈʃu]	$T2^F$	-	[muˈpiʃu]	$T4^M$	A
	[mupiˈʃu]	$T1^F$	-	[muˈpiʃu]	$T2^F$	-	[muˈpiʃu]	$T4^M$	B
	[muˈpiʃu]	$T1^F$	-	[mupiˈʃu]	$T2^F$	-	[mupiˈʃu]	$T4^M$	B
	[muˈpiʃu]	$T1^F$	-	[muˈpiʃu]	$T2^F$	-	[ˈmupiʃu]	$T4^M$	A
	[muˈpiʃu]	$T1^F$	-	[ˈmupiʃu]	$T2^F$	-	[ˈmupiʃu]	$T4^M$	B
	[ˈmupiʃu]	$T1^F$	-	[muˈpiʃu]	$T2^F$	-	[muˈpiʃu]	$T4^M$	B
	[muˈpiʃu]	$T1^F$	-	[muˈpiʃu]	$T2^F$	-	[mupiˈʃu]	$T4^M$	A
Block 5	[ʒuˈrigu]	$T1^F$	-	[ʒuˈrigu]	$T2^F$	-	[ʒuriˈgu]	$T4^M$	A
	[ˈʒurigu]	$T1^F$	-	[ʒuˈrigu]	$T2^F$	-	[ʒuˈrigu]	$T4^M$	B
	[ʒuˈrigu]	$T1^F$	-	[ʒuriˈgu]	$T2^F$	-	[ʒuriˈgu]	$T4^M$	B
	[ʒuˈrigu]	$T1^F$	-	[ˈʒurigu]	$T2^F$	-	[ˈʒurigu]	$T4^M$	B
	[ʒuriˈgu]	$T1^F$	-	[ʒuriˈgu]	$T2^F$	-	[ʒuˈrigu]	$T4^M$	A
	[ʒuriˈgu]	$T1^F$	-	[ʒuˈrigu]	$T2^F$	-	[ʒuˈrigu]	$T4^M$	B
	[ʒuˈrigu]	$T1^F$	-	[ʒuˈrigu]	$T2^F$	-	[ˈʒurigu]	$T4^M$	A
	[ˈʒurigu]	$T1^F$	-	[ˈʒurigu]	$T2^F$	-	[ʒuˈrigu]	$T4^M$	A
Block 6	[zitulii]	$T1^F$	-	[ziˈtuli]	$T2^F$	-	[ziˈtuli]	$T4^M$	B
	[ziˈtuli]	$T1^F$	-	[zitulii]	$T2^F$	-	[zitulii]	$T4^M$	B
	[ˈzituli]	$T1^F$	-	[ˈzituli]	$T2^F$	-	[ziˈtuli]	$T4^M$	A
	[ziˈtuli]	$T1^F$	-	[ziˈtuli]	$T2^F$	-	[ˈzituli]	$T4^M$	A
	[ziˈtuli]	$T1^F$	-	[ˈzituli]	$T2^F$	-	[ˈzituli]	$T4^M$	B
	[ˈzituli]	$T1^F$	-	[ziˈtuli]	$T2^F$	-	[ziˈtuli]	$T4^M$	B
	[ziˈtuli]	$T1^F$	-	[ziˈtuli]	$T2^F$	-	[zitulii]	$T4^M$	A
	[zitulii]	$T1^F$	-	[zitulii]	$T2^F$	-	[ziˈtuli]	$T4^M$	A

H. Trials for training of Group *Vowels & Stress* (AX and AXB tasks)

SESSIONS 1 TO 4	A			X		answer
Session 1: block 1	[zi'za]	$T3^F$	-	[zi'za]	$T1^F$	identical
	['zizɐ]	$T2^F$	-	['zizɐ]	$T3^F$	identical
	[zi'za]	$T2^F$	-	['zizɐ]	$T3^F$	different
	['zizɐ]	$T3^F$	-	[zi'za]	$T1^F$	different
	[zi'za]	$T3^F$	-	['zizɐ]	$T1^F$	different
	['zizɐ]	$T2^F$	-	[zi'za]	$T3^F$	different
	[zi'za]	$T1^F$	-	[zi'za]	$T2^F$	identical
	[zi'za]	$T2^F$	-	[zi'za]	$T3^F$	identical
	['zizɐ]	$T1^F$	-	[zi'za]	$T2^F$	different
	['zizɐ]	$T1^F$	-	['zizɐ]	$T2^F$	identical
	['zizɐ]	$T3^F$	-	['zizɐ]	$T1^F$	identical
	[zi'za]	$T1^F$	-	['zizɐ]	$T2^F$	different
Session 1: block 2	[zi'zɛ]	$T3^F$	-	['zizɨ]	$T1^F$	different
	[zi'zɛ]	$T2^F$	-	[zi'zɛ]	$T3^F$	identical
	['zizɨ]	$T1^F$	-	['zizɨ]	$T2^F$	identical
	['zizɨ]	$T3^F$	-	['zizɨ]	$T1^F$	identical
	['zizɨ]	$T2^F$	-	[zi'zɛ]	$T3^F$	different
	[zi'zɛ]	$T2^F$	-	['zizɨ]	$T3^F$	different
	['zizɨ]	$T1^F$	-	[zi'zɛ]	$T2^F$	different
	['zizɨ]	$T3^F$	-	[zi'zɛ]	$T1^F$	different
	[zi'zɛ]	$T3^F$	-	[zi'zɛ]	$T1^F$	identical
	['zizɨ]	$T2^F$	-	['zizɨ]	$T3^F$	identical
	[zi'zɛ]	$T1^F$	-	[zi'zɛ]	$T2^F$	identical
	[zi'zɛ]	$T1^F$	-	['zizɨ]	$T2^F$	different
Session 1: block 3	['zizɨ]	$T3^F$	-	[zi'ze]	$T1^F$	different
	[zi'ze]	$T3^F$	-	[zi'ze]	$T1^F$	identical
	['zizɨ]	$T1^F$	-	[zi'ze]	$T2^F$	different
	[zi'ze]	$T2^F$	-	['zizɨ]	$T3^F$	different
	['zizɨ]	$T1^F$	-	['zizɨ]	$T2^F$	identical
	[zi'ze]	$T1^F$	-	['zizɨ]	$T2^F$	different
	[zi'ze]	$T2^F$	-	[zi'ze]	$T3^F$	identical
	['zizɨ]	$T2^F$	-	[zi'ze]	$T3^F$	different
	['zizɨ]	$T2^F$	-	['zizɨ]	$T3^F$	identical
	[zi'ze]	$T3^F$	-	['zizɨ]	$T1^F$	different
	['zizɨ]	$T3^F$	-	['zizɨ]	$T1^F$	identical
	[zi'ze]	$T1^F$	-	[zi'ze]	$T2^F$	identical

Session 1: block 4	[zi'zɔ]	T3^F	-	[zi'zɔ]	T1^F	identical
	[zi'zɔ]	T3^F	-	['zizu]	T1^F	different
	['zizu]	T3^F	-	['zizu]	T1^F	identical
	[zi'zɔ]	T2^F	-	['zizu]	T3^F	different
	[zi'zɔ]	T1^F	-	[zi'zɔ]	T2^F	identical
	['zizu]	T1^F	-	['zizu]	T2^F	identical
	['zizu]	T2^F	-	['zizu]	T3^F	identical
	['zizu]	T2^F	-	[zi'zɔ]	T3^F	different
	['zizu]	T3^F	-	[zi'zɔ]	T1^F	different
	[zi'zɔ]	T2^F	-	[zi'zɔ]	T3^F	identical
	['zizu]	T1^F	-	[zi'zɔ]	T2^F	different
	[zi'zɔ]	T1^F	-	['zizu]	T2^F	different
Session 2: block 1	['zazi]	T1^F	-	[zɐ'zi]	T2^F	different
	['zazi]	T2^F	-	[zɐ'zi]	T3^F	different
	[zɐ'zi]	T3^F	-	['zazi]	T1^F	different
	[zɐ'zi]	T3^F	-	[zɐ'zi]	T1^F	identical
	[zɐ'zi]	T2^F	-	[zɐ'zi]	T3^F	identical
	[zɐ'zi]	T2^F	-	['zazi]	T3^F	different
	['zazi]	T3^F	-	['zazi]	T1^F	identical
	[zɐ'zi]	T1^F	-	[zɐ'zi]	T2^F	identical
	[zɐ'zi]	T1^F	-	['zazi]	T2^F	different
	['zazi]	T3^F	-	[zɐ'zi]	T1^F	different
	['zazi]	T2^F	-	['zazi]	T3^F	identical
	['zazi]	T1^F	-	['zazi]	T2^F	identical
Session 2: block 2	['zɛzi]	T2^F	-	[zɨ'zi]	T3^F	different
	['zɛzi]	T3^F	-	['zɛzi]	T1^F	identical
	[zɨ'zi]	T1^F	-	['zɛzi]	T2^F	different
	[zɨ'zi]	T3^F	-	[zɨ'zi]	T1^F	identical
	[zɨ'zi]	T1^F	-	[zɨ'zi]	T2^F	identical
	['zɛzi]	T2^F	-	['zɛzi]	T3^F	identical
	['zɛzi]	T3^F	-	[zɨ'zi]	T1^F	different
	[zɨ'zi]	T2^F	-	['zɛzi]	T3^F	different
	['zɛzi]	T1^F	-	['zɛzi]	T2^F	identical
	[zɨ'zi]	T3^F	-	['zɛzi]	T1^F	different
	[zɨ'zi]	T2^F	-	[zɨ'zi]	T3^F	identical
	['zɛzi]	T1^F	-	[zɨ'zi]	T2^F	different
Session 2: block 3	['zezi]	T3^F	-	[zɨ'zi]	T1^F	different
	['zezi]	T3^F	-	['zezi]	T1^F	identical
	['zezi]	T2^F	-	['zezi]	T3^F	identical

	[zɨ'zi]	$T2^F$	-	[zɨ'zi]	$T3^F$	identical
	[zɨ'zi]	$T3^F$	-	['zezi]	$T1^F$	different
	[zɨ'zi]	$T1^F$	-	[zɨ'zi]	$T2^F$	identical
	[zɨ'zi]	$T3^F$	-	[zɨ'zi]	$T1^F$	identical
	['zezi]	$T1^F$	-	[zɨ'zi]	$T2^F$	different
	[zɨ'zi]	$T1^F$	-	['zezi]	$T2^F$	different
	[zɨ'zi]	$T2^F$	-	['zezi]	$T3^F$	different
	['zezi]	$T2^F$	-	[zɨ'zi]	$T3^F$	different
	['zezi]	$T1^F$	-	['zezi]	$T2^F$	identical
Session 2: block 4	['zɔzi]	$T2^F$	-	['zɔzi]	$T3^F$	identical
	[zu'zi]	$T1^F$	-	['zɔzi]	$T2^F$	different
	['zɔzi]	$T2^F$	-	[zu'zi]	$T3^F$	different
	['zɔzi]	$T3^F$	-	['zɔzi]	$T1^F$	identical
	['zɔzi]	$T1^F$	-	[zu'zi]	$T2^F$	different
	['zɔzi]	$T1^F$	-	['zɔzi]	$T2^F$	identical
	['zɔzi]	$T3^F$	-	[zu'zi]	$T1^F$	different
	[zu'zi]	$T3^F$	-	['zɔzi]	$T1^F$	different
	[zu'zi]	$T2^F$	-	[zu'zi]	$T3^F$	identical
	[zu'zi]	$T2^F$	-	['zɔzi]	$T3^F$	different
	[zu'zi]	$T1^F$	-	[zu'zi]	$T2^F$	identical
	[zu'zi]	$T3^F$	-	[zu'zi]	$T1^F$	identical
Session 3: block 1	[zɐ'za]	$T3^F$	-	[zɐ'za]	$T1^F$	identical
	[zɐ'za]	$T1^F$	-	[zɐ'za]	$T2^F$	identical
	['zazɐ]	$T1^F$	-	['zazɐ]	$T2^F$	identical
	['zazɐ]	$T3^F$	-	['zazɐ]	$T1^F$	identical
	['zazɐ]	$T3^F$	-	[zɐ'za]	$T1^F$	different
	['zazɐ]	$T2^F$	-	['zazɐ]	$T3^F$	identical
	['zazɐ]	$T2^F$	-	[zɐ'za]	$T3^F$	different
	['zazɐ]	$T1^F$	-	[zɐ'za]	$T2^F$	different
	[zɐ'za]	$T1^F$	-	['zazɐ]	$T2^F$	different
	[zɐ'za]	$T2^F$	-	[zɐ'za]	$T3^F$	identical
	[zɐ'za]	$T2^F$	-	['zazɐ]	$T3^F$	different
	[zɐ'za]	$T3^F$	-	['zazɐ]	$T1^F$	different
Session 3: block 2	['zazɨ]	$T1^F$	-	['zazɨ]	$T2^F$	identical
	[zɐ'zɛ]	$T1^F$	-	['zazɨ]	$T2^F$	different
	['zazɨ]	$T2^F$	-	['zazɨ]	$T3^F$	identical
	[zɐ'zɛ]	$T3^F$	-	['zazɨ]	$T1^F$	different
	['zazɨ]	$T2^F$	-	[zɐ'zɛ]	$T3^F$	different
	[zɐ'zɛ]	$T3^F$	-	[zɐ'zɛ]	$T1^F$	identical

	[zɐ'zɛ]	T2^F	-	[zɐ'zɛ]	T3^F	identical
	['zazɨ]	T3^F	-	[zɐ'zɛ]	T1^F	different
	[zɐ'zɛ]	T2^F	-	['zazɨ]	T3^F	different
	[zɐ'zɛ]	T1^F	-	[zɐ'zɛ]	T2^F	identical
	['zazɨ]	T3^F	-	['zazɨ]	T1^F	identical
	['zazɨ]	T1^F	-	[zɐ'zɛ]	T2^F	different
Session 3: block 3	['zazɨ]	T1^F	-	[zɐ'ze]	T2^F	different
	['zazɨ]	T1^F	-	['zazɨ]	T2^F	identical
	['zazɨ]	T2^F	-	[zɐ'ze]	T3^F	different
	['zazɨ]	T3^F	-	['zazɨ]	T1^F	identical
	[zɐ'ze]	T1^F	-	[zɐ'ze]	T2^F	identical
	[zɐ'ze]	T2^F	-	[zɐ'ze]	T3^F	identical
	[zɐ'ze]	T3^F	-	[zɐ'ze]	T1^F	identical
	[zɐ'ze]	T1^F	-	['zazɨ]	T2^F	different
	[zɐ'ze]	T3^F	-	['zazɨ]	T1^F	different
	['zazɨ]	T3^F	-	[zɐ'ze]	T1^F	different
	['zazɨ]	T2^F	-	['zazɨ]	T3^F	identical
	[zɐ'ze]	T2^F	-	['zazɨ]	T3^F	different
Session 3: block 4	[zɐ'zɔ]	T2^F	-	[zɐ'zɔ]	T3^F	identical
	[zɐ'zɔ]	T3^F	-	[zɐ'zɔ]	T1^F	identical
	['zazu]	T1^F	-	[zɐ'zɔ]	T2^F	different
	[zɐ'zɔ]	T3^F	-	['zazu]	T1^F	different
	['zazu]	T1^F	-	['zazu]	T2^F	identical
	['zazu]	T3^F	-	[zɐ'zɔ]	T1^F	different
	['zazu]	T3^F	-	['zazu]	T1^F	identical
	['zazu]	T2^F	-	[zɐ'zɔ]	T3^F	different
	[zɐ'zɔ]	T1^F	-	[zɐ'zɔ]	T2^F	identical
	[zɐ'zɔ]	T1^F	-	['zazu]	T2^F	different
	['zazu]	T2^F	-	['zazu]	T3^F	identical
	[zɐ'zɔ]	T2^F	-	['zazu]	T3^F	different
Session 4: block 1	[zɨ'za]	T1^F	-	[zɨ'za]	T2^F	identical
	[zɨ'za]	T2^F	-	['zɛzɐ]	T3^F	different
	['zɛzɐ]	T1^F	-	[zɨ'za]	T2^F	different
	[zɨ'za]	T3^F	-	[zɨ'za]	T1^F	identical
	['zɛzɐ]	T2^F	-	[zɨ'za]	T3^F	different
	[zɨ'za]	T2^F	-	[zɨ'za]	T3^F	identical
	[zɨ'za]	T3^F	-	['zɛzɐ]	T1^F	different
	['zɛzɐ]	T3^F	-	['zɛzɐ]	T1^F	identical
	['zɛzɐ]	T1^F	-	['zɛzɐ]	T2^F	identical

	[zɨ'za]	$T1^F$	-	['zɛzɐ]	$T2^F$	different
	['zɛzɐ]	$T2^F$	-	['zɛzɐ]	$T3^F$	identical
	['zɛzɐ]	$T3^F$	-	[zɨ'za]	$T1^F$	different
Session 4: block 2	['zɛzɨ]	$T2^F$	-	[zɨ'zɛ]	$T3^F$	different
	['zɛzɨ]	$T1^F$	-	[zɨ'zɛ]	$T2^F$	different
	[zɨ'zɛ]	$T3^F$	-	[zɨ'zɛ]	$T1^F$	identical
	[zɨ'zɛ]	$T2^F$	-	[zɨ'zɛ]	$T3^F$	identical
	[zɨ'zɛ]	$T2^F$	-	['zɛzɨ]	$T3^F$	different
	[zɨ'zɛ]	$T1^F$	-	[zɨ'zɛ]	$T2^F$	identical
	[zɨ'zɛ]	$T3^F$	-	['zɛzɨ]	$T1^F$	different
	['zɛzɨ]	$T2^F$	-	['zɛzɨ]	$T3^F$	identical
	['zɛzɨ]	$T1^F$	-	['zɛzɨ]	$T2^F$	identical
	[zɨ'zɛ]	$T1^F$	-	['zɛzɨ]	$T2^F$	different
	['zɛzɨ]	$T3^F$	-	[zɨ'zɛ]	$T1^F$	different
	['zɛzɨ]	$T3^F$	-	['zɛzɨ]	$T1^F$	identical
Session 4: block 3	['zɛzɨ]	$T3^F$	-	['zɛzɨ]	$T1^F$	identical
	['zɛzɨ]	$T1^F$	-	[zɨ'ze]	$T2^F$	different
	['zɛzɨ]	$T1^F$	-	['zɛzɨ]	$T2^F$	identical
	['zɛzɨ]	$T2^F$	-	[zɨ'ze]	$T3^F$	different
	[zɨ'ze]	$T3^F$	-	[zɨ'ze]	$T1^F$	identical
	[zɨ'ze]	$T1^F$	-	[zɨ'ze]	$T2^F$	identical
	[zɨ'ze]	$T2^F$	-	[zɨ'ze]	$T3^F$	identical
	[zɨ'ze]	$T2^F$	-	['zɛzɨ]	$T3^F$	different
	['zɛzɨ]	$T3^F$	-	[zɨ'ze]	$T1^F$	different
	[zɨ'ze]	$T3^F$	-	['zɛzɨ]	$T1^F$	different
	['zɛzɨ]	$T2^F$	-	['zɛzɨ]	$T3^F$	identical
	[zɨ'ze]	$T1^F$	-	['zɛzɨ]	$T2^F$	different
Session 4: block 4	[zɨ'zɔ]	$T3^F$	-	[zɨ'zɔ]	$T1^F$	identical
	['zɛzu]	$T2^F$	-	[zɨ'zɔ]	$T3^F$	different
	['zɛzu]	$T3^F$	-	[zɨ'zɔ]	$T1^F$	different
	['zɛzu]	$T3^F$	-	['zɛzu]	$T1^F$	identical
	[zɨ'zɔ]	$T1^F$	-	['zɛzu]	$T2^F$	different
	[zɨ'zɔ]	$T1^F$	-	[zɨ'zɔ]	$T2^F$	identical
	[zɨ'zɔ]	$T2^F$	-	['zɛzu]	$T3^F$	different
	['zɛzu]	$T2^F$	-	['zɛzu]	$T3^F$	identical
	['zɛzu]	$T1^F$	-	['zɛzu]	$T2^F$	identical
	[zɨ'zɔ]	$T2^F$	-	[zɨ'zɔ]	$T3^F$	identical
	[zɨ'zɔ]	$T3^F$	-	['zɛzu]	$T1^F$	different
	['zɛzu]	$T1^F$	-	[zɨ'zɔ]	$T2^F$	different

SESSIONS 5 AND 6	A			X			B		answer
Block 1	['zizɨ]	T1^{F}	-	['zizɨ]	T2^{F}	-	[zi'zɛ]	T3^{F}	A
	[zi'zɛ]	T1^{F}	-	['zizɨ]	T2^{F}	-	['zizɨ]	T3^{F}	B
	['zizɨ]	T1^{F}	-	[zi'zɛ]	T2^{F}	-	[zi'zɛ]	T3^{F}	B
	[zi'zɛ]	T1^{F}	-	[zi'zɛ]	T2^{F}	-	['zizɨ]	T3^{F}	A
Block 2	['zizɨ]	T1^{F}	-	['zizɨ]	T2^{F}	-	[zi'ze]	T3^{F}	A
	['zizɨ]	T1^{F}	-	[zi'ze]	T2^{F}	-	[zi'ze]	T3^{F}	B
	[zi'ze]	T1^{F}	-	['zizɨ]	T2^{F}	-	['zizɨ]	T3^{F}	B
	[zi'ze]	T1^{F}	-	[zi'ze]	T2^{F}	-	['zizɨ]	T3^{F}	A
Block 3	[zi'za]	T1^{F}	-	[zi'za]	T2^{F}	-	['zizɐ]	T3^{F}	A
	['zizɐ]	T1^{F}	-	[zi'za]	T2^{F}	-	[zi'za]	T3^{F}	B
	[zi'za]	T1^{F}	-	['zizɐ]	T2^{F}	-	['zizɐ]	T3^{F}	B
	['zizɐ]	T1^{F}	-	['zizɐ]	T2^{F}	-	[zi'za]	T3^{F}	A
Block 4	['zɛzi]	T1^{F}	-	['zɛzi]	T2^{F}	-	[zɨ'zi]	T3^{F}	A
	[zɨ'zi]	T1^{F}	-	[zɨ'zi]	T2^{F}	-	['zɛzi]	T3^{F}	A
	[zɨ'zi]	T1^{F}	-	['zɛzi]	T2^{F}	-	['zɛzi]	T3^{F}	B
	['zɛzi]	T1^{F}	-	[zɨ'zi]	T2^{F}	-	[zɨ'zi]	T3^{F}	B
Block 5	['zɛzɨ]	T1^{F}	-	['zɛzɨ]	T2^{F}	-	[zɨ'zɛ]	T3^{F}	A
	[zɨ'zɛ]	T1^{F}	-	[zɨ'zɛ]	T2^{F}	-	['zɛzɨ]	T3^{F}	A
	['zɛzɨ]	T1^{F}	-	[zɨ'zɛ]	T2^{F}	-	[zɨ'zɛ]	T3^{F}	B
	[zɨ'zɛ]	T1^{F}	-	['zɛzɨ]	T2^{F}	-	['zɛzɨ]	T3^{F}	B
Block 6	['zɛzɨ]	T1^{F}	-	['zɛzɨ]	T2^{F}	-	[zɨ'ze]	T3^{F}	A
	[zɨ'ze]	T1^{F}	-	['zɛzɨ]	T2^{F}	-	['zɛzɨ]	T3^{F}	B
	['zɛzɨ]	T1^{F}	-	[zɨ'ze]	T2^{F}	-	[zɨ'ze]	T3^{F}	B
	[zɨ'ze]	T1^{F}	-	[zɨ'ze]	T2^{F}	-	['zɛzɨ]	T3^{F}	A
Block 7	['zɛzɐ]	T1^{F}	-	['zɛzɐ]	T2^{F}	-	[zɨ'za]	T3^{F}	A
	[zɨ'za]	T1^{F}	-	['zɛzɐ]	T2^{F}	-	['zɛzɐ]	T3^{F}	B
	[zɨ'za]	T1^{F}	-	[zɨ'za]	T2^{F}	-	['zɛzɐ]	T3^{F}	A
	['zɛzɐ]	T1^{F}	-	[zɨ'za]	T2^{F}	-	[zɨ'za]	T3^{F}	B
Block 8	['zezi]	T1^{F}	-	['zezi]	T2^{F}	-	[zɨ'zi]	T3^{F}	A
	[zɨ'zi]	T1^{F}	-	['zezi]	T2^{F}	-	['zezi]	T3^{F}	B
	[zɨ'zi]	T1^{F}	-	[zɨ'zi]	T2^{F}	-	['zezi]	T3^{F}	A
	['zezi]	T1^{F}	-	[zɨ'zi]	T2^{F}	-	[zɨ'zi]	T3^{F}	B
Block 9	['zazi]	T1^{F}	-	['zazi]	T2^{F}	-	[zɐ'zi]	T3^{F}	A
	[zɐ'zi]	T1^{F}	-	[zɐ'zi]	T2^{F}	-	['zazi]	T3^{F}	A
	[zɐ'zi]	T1^{F}	-	['zazi]	T2^{F}	-	['zazi]	T3^{F}	B
	['zazi]	T1^{F}	-	[zɐ'zi]	T2^{F}	-	[zɐ'zi]	T3^{F}	B

Block 10	['zazɨ]	T1^F^	-	['zazɨ]	T2^F^	-	[zɐ'zɛ]	T3^F^	A
	[zɐ'zɛ]	T1^F^	-	[zɐ'zɛ]	T2^F^	-	['zazɨ]	T3^F^	A
	[zɐ'zɛ]	T1^F^	-	['zazɨ]	T2^F^	-	['zazɨ]	T3^F^	B
	['zazɨ]	T1^F^	-	[zɐ'zɛ]	T2^F^	-	[zɐ'zɛ]	T3^F^	B
Block 11	['zazɨ]	T1^F^	-	['zazɨ]	T2^F^	-	[zɐ'ze]	T3^F^	A
	[zɐ'ze]	T1^F^	-	[zɐ'ze]	T2^F^	-	['zazɨ]	T3^F^	A
	[zɐ'ze]	T1^F^	-	['zazɨ]	T2^F^	-	['zazɨ]	T3^F^	B
	['zazɨ]	T1^F^	-	[zɐ'ze]	T2^F^	-	[zɐ'ze]	T3^F^	B
Block 12	[zɐ'za]	T1^F^	-	[zɐ'za]	T2^F^	-	['zazɐ]	T3^F^	A
	['zazɐ]	T1^F^	-	[zɐ'za]	T2^F^	-	[zɐ'za]	T3^F^	B
	[zɐ'za]	T1^F^	-	['zazɐ]	T2^F^	-	['zazɐ]	T3^F^	B
	['zazɐ]	T1^F^	-	['zazɐ]	T2^F^	-	[zɐ'za]	T3^F^	A

Appendix VIII

Instructions presented to Hungarian participants in the beginning of the training phase of the perceptual training study (Chapter 5)

A. Instructions sent to Group *Vowels* in the beginning of the training (original version)

We are starting the training exercises!
You are getting now a new link from Gorilla site (your code is the same, always).
In this 2' video you have the general instructions (in English), or you can read them in Hungarian below.

***

Elkezdjük a gyakorló feladatokat 😊

Összesen három különböző típusú gyakorlorló feladatot készítettem, azonban a feladatok kiosztása közöttetek véletlenszerűen történik, nem az én döntésem, hogy melyik hallgató melyik feladatsorokat kapja.

A Te csoportod azokra a **portugál magánhangzók**ra összpontosít, melyek nehézséget okoznak a magyar anyanyelvűek számára. Ezek között a magánhangzók között vannak olyan beszédhangok, amik léteznek a magyar nyelvben, azonban attól árnyalatnyilag eltérően valósulnak meg; illetve vannak olyan beszédhangok is, amelyek egyáltalán nem jellemzőek a magyar nyelvre.

Ebben a gyakorlatban a magánhangzók kerülnek középpontba izolált (gyakran jelentéssel nem rendelkező) szótagokban, pl. Z+magánhangzó, "zi". Ebben a feladatban azért fókuszálunk szavak helyett egyszerű szótagokra, mert így nagyobb figyelem irányulhat a magánhangzó megvalósulására.

Ez a feladat sokban hasonlít a tesztfeladatokhoz, minden esetben két vagy három hangot fogsz hallani, és az lesz a feladatod, hogy eldöntsd, hogy az elhangzó felvételen ugyanazt a hangot hallod-e.

Összesen hat feladatot kell majd elvégezned, minden héten egyet. A feladatok elvégzése nagyjából 10 percet vesz igénybe, a feladatokat nehézség szempontjából állítottam sorba, ez azt jelenti, hogy az első héten kell elvégezni a legkönnyebb feladatsort, míg a hatadik héthez érve egyre nehezebbek következnek.

Kérlek, a kísérlet ideje alatt használj fülhallgatót és amennyire lehetséges, egy csöndes helyen tartózkodj!

Ha bármilyen technikai probémába ütközöl, vagy kérdésed merülne fel, kérlek írj nekemazonal!

Jó munkát!

B. Instructions sent to Group *Vowels* in the beginning of the training (English version)

We are starting the training exercises!

You are getting now a new link from Gorilla site (your code is the same, always).

In this 2' video you have the general instructions (in English), or you can read them in Hungarian below.

***

There are 3 groups of training, each group is practicing something different. The decision of which students practice what was random, not our decision.

You are in the group that will train the Portuguese vowels in which Hungarians show difficulties: some of these vowels exist in Hungarian but they are a little bit different, some don't exist in Hungarian.

The training has to focus on isolated vowels (inserted in a simple syllable: Z+vowel, zi for example), and not in words: vowels in words can change a lot, so first you have to practice each vowel separated, to be able to pay attention just to that vowel.

The training tasks are very similar to the pretest tasks: you will always hear a pair or threesome of sounds that you have to discriminate.

In total, there are 6 exercises, you will get one for week. The 6 exercises are organized from easier to more difficult, and each exercise takes only about 10 to 15 minutes.

Please, use headphones and try to be in a silent place, if possible.

If you experience any technical problem or have any question, please don't hesitate to write.

Good work!

C. Instructions sent to Group *Stress* in the beginning of the training (original version)

We are starting the training exercises!
You are getting now a new link from Gorilla site (your code is the same, always).
In this 2' video you have the general instructions (in English), or you can read them in Hungarian below.

***

Elkezdjük a gyakorló feladatokat 😊

Összesen három különböző típusú gyakorlorló feladatot készítettem, azonban a feladatok kiosztása közöttetek véletlenszerűen történik, nem az én döntésem, hogy melyik hallgató melyik feladatsorokat kapja.

A Te csoportod feladatsora a **szóhangsúly**ra fog összpontosítani. Habár a magyar nyelvben a szóhangsúly nem bír jelentésmegkülönböztető szereppel, hiszen mindig a szó első szótagján helyezkedik el, ezzel ellentétben a portugál nyelvben (az angolhoz hasonlóan) rendkívül fontos. Nézzünk rá egy példát:

Today is a PERfect day. (Ebben az esetben a perfect szó PER szótagja a hangsúlyos, ezért a szó melléknévi jelentésére utal.)

I need to perFECT my pronunciation. (Ebben az esetben azonban a FECT szótag a hangsúlyos, ezért a perfect szó igei jelentését értjük alatta.)

Hasonló példa a hangsúly jelentésmegkülönböztető szerepére a portugál nyelvben:

Tenho um bamBU no jardim. (*I have a bamboo in the garden.)* (A bamBU szó második szótagját ejtve hangsúlyosan főnévi jelentést kapunk)

Ele está a sentir-se BAMbo. (*He's feeling wobbly.*) (Míg az első szótagra helyezve a hangsúlyt melléknévi jelentést fogunk kapni.)

A gyakorlófeladatok erre az aspektusra fognak összpontosítani.

Ez a feladat sokban hasonlít a teszt feladatokhoz, minden esetben két vagy három hangot fogsz hallani, és az lesz a feladatod, hogy eldöntsd, hogy az elhangzó felvételen ugyanazt a hangot hallod-e.

Összesen hat feladatot kell majd elvégezned, minden héten egyet. A feladatok elvégzése nagyjából 10 percet vesz igénybe, a feladatokat nehézség szempontjából állítottam sorba, ez azt jelenti, hogy az első héten kell elvégezni a legkönnyebb feladatsort, míg az ötödik héthez érve egyre nehezebbek következnek.

Kérlek, a kísérlet ideje alatt használj fülhallgatót és amennyire lehetséges, egy csöndes helyen tartózkodj!

Ha bármilyen technikai probémába ütközöl, vagy kérdésed merülne fel, kérlek írj nekemazonal!

Jó munkát!

D. Instructions sent to Group *Stress* in the beginning of the training (English version)

We are starting the training exercises!

You are getting now a new link from Gorilla site (your code is the same, always).

In this 2' video you have the general instructions (in English), or you can read them in Hungarian below.

***

There are 3 groups of training, each group is practicing something different. The decision of which students practice what was random, not our decision.

You are in the group that will train the word stress. Although in Hungarian this is not a very important feature – the stress is almost always in the first syllable, in Portuguese is really important. Portuguese is like English, where word stress is fundamental. See for example:

Today is a PERfect day.

I need to perFECT my pronunciation.

In Portuguese the same occurs:

Tenho um bamBU no jardim. (I have a bamboo in the garden)

Ele está a sentir-se BAMbo. (He's feeling wobbly.)

Your training will focus exactly on this aspect: word stress.

The training tasks are very similar to the pretest tasks: you will always hear a pair or threesome of sounds that you have to discriminate.

In total, there are 6 exercises, you will get one for week. The 6 exercises are organized from easier to more difficult, and each exercise takes only about 10 to 15 minutes.

Please, use headphones and try to be in a silent place, if possible.

If you experience any technical problem or have any question, please don't hesitate to write.

Good work!

E. Instructions sent to Group *Vowels & Stress* in the beginning of the training (original version)

We are starting the training exercises!
You are getting now a new link from Gorilla site (your code is the same, always).
In this 2' video you have the general instructions (in English), or you can read them in Hungarian below.

***

Elkezdjük a gyakorló feladatokat 😊

Összesen három különböző típusú gyakorlorló feladatot készítettem, azonban a feladatok kiosztása közöttetek véletlenszerűen történik, nem az én döntésem, hogy melyik hallgató melyik feladatsorokat kapja.

A Te csoportod feladatai a **szóhangsúly**ra és a **magánhangzórövidülés** jelenségére fognak összpontosítani, amik alapvető különbségek a Magyar és portugál nyelvek között.

Habár a magyar nyelvben a szóhangsúly nem bír jelentésmegkülönböztető szereppel, hiszen mindig a szó első szótagján helyezkedik el, ezzel ellentétben a portugál nyelvben (az angolhoz hasonlóan) rendkívül fontos.

Today is a PERfect day. (Ebben az esetben a perfect szó PER szótagja a hangsúlyos, ezért a szó melléknévi jelentésére utal.)
I need to perFECT my pronunciation. (Ebben az esetben azonban a FECT szótag a hangsúlyos, ezért a perfect szó igei jelentését értjük alatta.)

Hasonló példa a hangsúly jelentésmegkülönböztető szerepére a portugál nyelvben:

Tenho um bamBU no jardim. (*I have a bamboo in the garden.)* (A bamBU szó második szótagját ejtve hangsúlyosan főnévi jelentést kapunk)
Ele está a sentir-se BAMbo. (*He's feeling wobbly.*) (Míg az első szótagra helyezve a hangsúlyt melléknévi jelentést fogunk kapni.)

Emellett a hangsúly helyzete befolyásolhatja a magánhangzók megvalósulását is adott szón belül, például a portugál nyelvben:

Ela é secreTÁria. (*She's a secretary.)*
Ela trabalha na secretaRIa. (*She works in the secretariat/office.)*

A gyakorlófeladatok erre az aspektusra fognak összpontosítani.

A feladatok sokban hasonlítanak a tesztfeladatokhoz, minden esetben két vagy három hangot fogsz hallani, és az lesz a feladatod, hogy eldöntsd, hogy az elhangzó felvételen ugyanazt a hangot hallod-e.

Összesen hat feladatot kell majd elvégezned, minden héten egyet. A feladatok elvégzése nagyjából 10 percet vesz igénybe, a feladatokat nehézség szempontjából állítottam sorba, ez azt jelenti, hogy az első héten kell elvégezni a legkönnyebb feladatsort, míg a hatadik héthez érve egyre nehezebbek következnek.

Kérlek, a kísérlet ideje alatt használj fülhallgatót és amennyire lehetséges, egy csöndes helyen tartózkodj!

Ha bármilyen technikai probémába ütközöl, vagy kérdésed merülne fel, kérlek írj nekem azonal!

Jó munkát!

F. Instructions sent to Group *Vowels & Stress* in the beginning of the training (English version)

We are starting the training exercises!

You are getting now a new link from Gorilla site (your code is the same, always).

In this 3' video you have the general instructions (in English), or you can read them in Hungarian below.

***

There are 3 groups of training, each group is practicing something different. The decision of which students practice what was random, not our decision.

You are in the group that will train the word stress and vowel reduction, a key difference between Hungarian and Portuguese languages.

Although in Hungarian this is not a very important feature – the stress is almost always in the first syllable, in Portuguese is really important. Portuguese is like English, where word stress is fundamental. See for example:

> Today is a PERfect day.
>
> I need to perFECT my pronunciation.

An example from Portuguese:

> Tenho um bamBU no jardim. (I have a bamboo in the garden)
>
> Ele está a sentir-se BAMbo. (He's feeling wobbly.)

Moreover, the location of the stress in a word can influence how the vowels are pronounced, take for example, in Portuguese:

> Ela é secreTÁria. (She's a secretary.)
>
> Ela trabalha na secretaRIa. (She works in the secretariat/office)

Your training will focus exactly on this aspect.

The training tasks are very similar to the pretest tasks: you will always hear a pair or threesome of sounds that you have to discriminate.

In total, there are 6 exercises, you will get one for week. The 6 exercises are organized from easier to more difficult, and each exercise takes only about 10 to 15 minutes.

Please, use headphones and try to be in a silent place, if possible.

If you experience any technical problem or have any question, please don't hesitate to write.

Good work!

Appendix IX

Structure of the tests for stimuli validation, for the perceptual training study (Chapter 5)

A. Structure of the three sessions for test stimuli validation, conducted with the *Baseline* group

Session 1 96 trials	*vowel discrimination* [sɐ]-[sɛ], [sɛ]-s[e], [se]-[si], [sɐ]-[sɨ], [sɛ]-[sɨ], [se]-[sɨ], [sɐ]-[sɛ], [sɛ]-[se], [se]-[si], [sɐ]-[sɨ], [sɛ]-[sɨ], [se]-[sɨ] oddity task with catch trials
Session 2 135 trials	*stress discrimination* [ˈtudiki]-[tuˈdiki], [tuˈdiki]-[tudiˈki], [ˈkiɾuʃi]-[kiˈɾuʃi], [kiˈɾuʃi]-[kiɾuˈʃi], [ˈdutiku]-[duˈtiku], [duˈtiku]-[dutiˈku], [ˈtuduɾi]-[tuˈduɾi], [tuˈduɾi]-[tuduˈɾi], [ˈtikuɾu]-[tiˈkuɾu], [tiˈkuɾu]-[tikuˈɾu], [ˈkitumi]-[kiˈtumi], [kiˈtumi]-[kituˈmi], [ˈlutinu]-[luˈtinu], [luˈtinu]-[lutiˈnu], [ˈbikulu]-[biˈkulu], [biˈkulu]-[bikuˈlu], [ˈmudini]-[muˈdini], [muˈdini]-[mudiˈni] oddity task with catch trials
Session 3 48 trials	*consonant discrimination* [buˈʃili]-[buˈpili], [kuˈfitu]-[kuˈmitu], [diˈbuki]-[diˈpuki], [tiˈbuɾi]-[tiˈʃuɾi], [tiˈfubi]-[tiˈvubi], [zuˈmitu]-[zuˈnitu] oddity task with catch trials

B. Structure of the five sets of tests for training stimuli validation, conducted with the *Baseline* group

	SET 1
PART 1 48 trials	[zɛ]-[zɨ], [ze]-[zɨ] AX task, 24 trials for each contrast
PART 2 48 trials	[ˈzizɐ]-[ziˈza], [ˈzizɨ]-[ziˈzɛ], [ˈzizɨ]-[ziˈze], [ˈzizu]-[ziˈzɔ] AX task, 12 trials for each contrast
PART 3 24 trials	[ˈbifuʃi]-[biˈfuʃi], [vuˈmipi]-[vumiˈpi] AX task, 12 trials for each contrast

	SET 2
PART 1 48 trials	[za]-[zɐ], [zɔ]-[zo] AX task, 24 trials for each contrast
PART 2 48 trials	[ˈzazi]-[zɐˈzi], [ˈzɛzi]-[zɨˈzi], [ˈzezi]-[zɨˈzi], [ˈzɔzi]-[zuˈzi] AX task, 12 trials for each contrast
PART 3 24 trials	[ˈfiluvi]-[fiˈluvi], [muˈpiʃu]-[mupiˈʃu] AX task, 12 trials for each contrast

	SET 3
PART 1 48 trials	[zɐ]-[zɨ], [zɐ]-[zɛ] AX task, 24 trials for each contrast
PART 2 48 trials	[ˈzazɐ]-[zɐˈza], [ˈzazɨ]-[zɐ'zɛ], [ˈzazɨ]-[zɐˈze], [ˈzazu]-[zɐˈzɔ] AX task, 12 trials for each contrast
PART 3 24 trials	[biˈfuʃi]-[bifuˈʃi], [ˈmupiʃu]-[muˈpiʃu] AX task, 12 trials for each contrast

	SET 4
PART 1 48 trials	[zɛ]-[ze], [ze]-[zi] AX task, 24 trials for each contrast
PART 2 48 trials	[ˈzɛzɐ]-[zɨˈza], [ˈzɛzɨ]-[zɨˈzɛ], [ˈzɛzɨ]-[zɨˈze], [ˈzɛzu]-[zɨˈzɔ] AX task, 12 trials for each contrast
PART 3 24 trials	[fiˈluvi]-[filuˈvi], [ˈvumipi]-[vuˈmipi] AX task, 12 trials for each contrast

	SET 5
PART 1 24 trials	[zɐ]-[za], [zɛ]-[zɨ], [ze]-[zɨ], [zɐ]-[zɨ], [zɐ]-[zɛ], [zɛ]-[ze] AXB task, 4 trials for each contrast
PART 2 48 trials	[ˈzizɨ]-[ziˈzɛ], [ˈzizɨ]-[ziˈze], [ˈzizɐ]-[ziˈza], [ˈzɛzi]-[zɨˈzi], [ˈzɛzɨ]-[zɨˈzɛ], [ˈzɛzɨ]-[zɨˈze], ['zɛzɐ]-[zɨ'za], ['zezi]-[zɨ'zi], ['zazi]-[zɐ'zi], ['zazɨ]-[zɐ'zɛ], ['zazɨ]-[zɐ'ze], ['zazɐ]-[zɐ'za] AXB task, 4 trials for each contrast
PART 3 48 trials	[ˈbifuʃi]-[biˈfuʃi], [biˈfuʃi]-[bifuˈʃi], [ˈvumipi]-[vuˈmipi], [vuˈmipi]-[vumiˈpi], [ˈfiluvi]-[fiˈluvi], [fiˈluvi]-[filuˈvi], [ˈmupiʃu]-[muˈpiʃu], [muˈpiʃu]-[mupiˈʃu], [ˈʒuɾigu]-[ʒuˈɾigu], [ˈʒuɾigu]-[ʒuˈɾigu], [ˈzituli]-[ziˈtuli], [ziˈtuli]-[zituˈli] AXB task, 4 trials for each contrast

Appendix X

Results and statistical analysis for the main trials of the experimental tasks, for the perceptual training study (Chapter 5)

A. Results for consonant discrimination

group	N	ER	SD	SE	CI
Vowels	360	9.8	22.3	1.2	2.3
Stress	396	10.7	22.4	1.1	2.2
Vowels & Stress	432	11.7	25.1	1.2	2.4
Baseline	198	7.4	19.3	1.4	2.7

	contrast	N	ER	SD	SE	CI
Vowels	[b]-[ʃ]	60	6.7	16.5	2.1	4.3
	[f]-[v]	60	11.1	24.2	3.1	6.3
	[m]-[f]	60	7.8	16.2	2.1	4.2
	[n]-[m]	60	12.2	22.5	2.9	5.8
	[p]-[b]	60	11.9	19.4	2.5	5.0
	[p]-[ʃ]	60	8.9	18.3	2.4	4.7
Stress	[b]-[ʃ]	66	8.8	14.9	1.8	3.7
	[f]-[v]	66	7.3	18.3	2.3	4.5
	[m]-[f]	66	9.3	19.7	2.4	4.8
	[n]-[m]	66	14.1	26.2	3.2	6.4
	[p]-[b]	66	15.2	19.3	2.4	4.8
	[p]-[ʃ]	66	9.3	18.0	2.2	4.4
Vowels & Stress	[b]-[ʃ]	72	7.9	16.8	2.0	3.9
	[f]-[v]	72	12.7	27.4	3.2	6.4
	[m]-[f]	72	12.0	20.9	2.5	4.9
	[n]-[m]	72	21.8	28.7	3.4	6.7
	[p]-[b]	72	9.5	17.7	2.1	4.2
	[p]-[ʃ]	72	6.5	15.4	1.8	3.6
Baseline	[b]-[ʃ]	33	6.1	13.9	2.4	4.9
	[f]-[v]	33	7.1	17.7	3.1	6.3
	[m]-[f]	33	5.1	15.1	2.6	5.4
	[n]-[m]	33	9.1	20.6	3.6	7.3
	[p]-[b]	33	12.6	19.3	3.4	6.9
	[p]-[ʃ]	33	4.5	14.9	2.6	5.3

B. Results for vowel discrimination, for the catch trials

test	group	N	ER	SD	SE	CI
pretest	*Vowels*	120	17.5	44.5	4.1	8.0
	Stress	132	24.2	51.6	4.5	8.9
	Vowels & Stress	144	22.2	48.9	4.1	8.1
	Baseline	78	15.4	41.8	4.7	9.4
post-test[R]	*Vowels*	120	12.5	41.0	3.7	7.4
	Stress	132	15.9	43.4	3.8	7.5
	Vowels & Stress	144	12.5	39.4	3.3	6.5
	Baseline	78	15.4	41.8	4.7	9.4
post-test[N]	*Vowels*	120	19.2	46.5	4.2	8.4
	Stress	132	24.2	50.8	4.4	8.7
	Vowels & Stress	144	22.2	50.1	4.2	8.3
	Baseline	78	10.3	36.7	4.2	8.3

test	vowel	group	N	ER	SD	SE	CI
pretest	[a]	*Vowels*	20	0.0	9.4	2.1	4.4
		Stress	22	0.0	10.5	2.2	4.6
		Vowels & Stress	24	4.2	22.7	4.6	9.6
		Baseline	13	7.7	29.8	8.3	18.0
	[ɐ]	*Vowels*	20	20.0	37.8	8.4	17.7
		Stress	22	45.5	50.6	10.8	22.4
		Vowels & Stress	24	16.7	37.6	7.7	15.9
		Baseline	13	23.1	38.9	10.8	23.5
	[e]	*Vowels*	20	15.0	32.6	7.3	15.3
		Stress	22	18.2	38.1	8.1	16.9
		Vowels & Stress	24	12.5	34.9	7.1	14.8
		Baseline	13	23.1	43.7	12.1	26.4
	[ɛ]	*Vowels*	20	5.0	27.2	6.1	12.7
		Stress	22	9.1	30.0	6.4	13.3
		Vowels & Stress	24	12.5	31.9	6.5	13.5
		Baseline	13	7.7	26.3	7.3	15.9
	[ɨ]	*Vowels*	20	55.0	47.5	10.6	22.2
		Stress	22	68.2	44.8	9.6	19.9
		Vowels & Stress	24	66.7	45.6	9.3	19.3
		Baseline	13	7.7	28.1	7.8	17.0
	[i]	*Vowels*	20	10.0	32.0	7.2	15.0
		Stress	22	4.5	26.4	5.6	11.7
		Vowels & Stress	24	20.8	38.3	7.8	16.2
		Baseline	13	23.1	41.4	11.5	25.0

	[a]	Vowels	20	5.0	22.1	5.0	10.4
		Stress	22	0.0	10.5	2.2	4.6
		Vowels & Stress	24	0.0	10.1	2.1	4.3
		Baseline	13	7.7	29.8	8.3	18.0
	[ɐ]	Vowels	20	25.0	44.5	9.9	20.8
		Stress	22	31.8	46.7	10.0	20.7
		Vowels & Stress	24	12.5	31.9	6.5	13.5
		Baseline	13	23.1	38.9	10.8	23.5
	[e]	Vowels	20	15.0	38.7	8.7	18.1
		Stress	22	9.1	30.9	6.6	13.7
		Vowels & Stress	24	12.5	31.9	6.5	13.5
post-test[R]		Baseline	13	23.1	43.7	12.1	26.4
	[ɛ]	Vowels	20	0.0	9.4	2.1	4.4
		Stress	22	4.5	21.7	4.6	9.6
		Vowels & Stress	24	8.3	28.9	5.9	12.2
		Baseline	13	7.7	26.3	7.3	15.9
	[ɨ]	Vowels	20	30.0	48.5	10.8	22.7
		Stress	22	50.0	46.1	9.8	20.4
		Vowels & Stress	24	37.5	48.2	9.8	20.3
		Baseline	13	7.7	28.1	7.8	17.0
	[i]	Vowels	20	0.0	9.4	2.1	4.4
		Stress	22	0.0	10.5	2.2	4.6
		Vowels & Stress	24	4.2	23.8	4.9	10.0
		Baseline	13	23.1	41.4	11.5	25.0
	[a]	Vowels	20	0.0	9.4	2.1	4.4
		Stress	22	0.0	10.5	2.2	4.6
		Vowels & Stress	24	0.0	10.1	2.1	4.3
		Baseline	13	7.7	26.3	7.3	15.9
	[ɐ]	Vowels	20	5.0	24.8	5.5	11.6
		Stress	22	9.1	31.8	6.8	14.1
		Vowels & Stress	24	8.3	25.1	5.1	10.6
		Baseline	13	0.0	11.3	3.1	6.9
		Vowels	20	25.0	39.3	8.8	18.4
post-test[N]	[e]	Stress	22	68.2	46.7	9.9	20.7
		Vowels & Stress	24	62.5	50.0	10.2	21.1
		Baseline	13	15.4	38.9	10.8	23.5
	[ɛ]	Vowels	20	15.0	36.2	8.1	17.0
		Stress	22	27.3	42.6	9.1	18.9
		Vowels & Stress	24	12.5	33.4	6.8	14.1
		Baseline	13	23.1	44.8	12.4	27.1
	[ɨ]	Vowels	20	60.0	47.9	10.7	22.4
		Stress	22	31.8	46.7	10.0	20.7
		Vowels & Stress	24	50.0	50.1	10.2	21.1
		Baseline	13	0.0	11.3	3.1	6.9

[i]	Vowels	20	10.0	32.9	7.4	15.4
	Stress	22	9.1	26.0	5.5	11.5
	Vowels & Stress	24	0.0	10.1	2.1	4.3
	Baseline	13	15.4	33.5	9.3	20.3

C. Results for vowel discrimination, for the change trials

test	group	N	ER	SD	SE	CI
pretest	*Vowels*	140	32.7	34.8	2.9	5.8
	Stress	154	32.1	33.7	2.7	5.4
	Vowels & Stress	168	29.3	32.6	2.5	5.0
	Baseline	91	9.2	17.4	1.8	3.6
post-test[R]	*Vowels*	140	27.9	37.7	3.2	6.3
	Stress	154	25.1	31.6	2.5	5.0
	Vowels & Stress	168	27.9	35.3	2.7	5.4
	Baseline	91	9.2	17.4	1.8	3.6
post-test[N]	*Vowels*	140	38.3	37.0	3.1	6.2
	Stress	154	37.9	31.8	2.6	5.1
	Vowels & Stress	168	39.3	31.9	2.5	4.9
	Baseline	91	11.7	21.7	2.3	4.5

test	group	contrast type	N	ER	SD	SE	CI
pretest	*Vowels*	perceptual overlap	80	47.9	29.6	3.3	6.6
		stressed-unstressed	60	12.5	15.3	2.0	4.0
	Stress	perceptual overlap	88	47.5	27.1	2.9	5.8
		stressed-unstressed	66	11.6	16.0	2.0	3.9
	Vowels & Stress	perceptual overlap	96	40.5	30.1	3.1	6.1
		stressed-unstressed	72	14.4	17.9	2.1	4.2
	Baseline	perceptual overlap	52	10.3	17.9	2.5	5.0
		stressed-unstressed	39	7.7	11.6	1.9	3.8
post-test[R]	*Vowels*	perceptual overlap	80	43.1	35.2	3.9	7.8
		stressed-unstressed	60	7.5	11.9	1.5	3.1
	Stress	perceptual overlap	88	37.1	28.7	3.1	6.1
		stressed-unstressed	66	9.1	15.3	1.9	3.8
	Vowels & Stress	perceptual overlap	96	38.9	33.1	3.4	6.7
		stressed-unstressed	72	13.2	20.5	2.4	4.8
	Baseline	perceptual overlap	52	10.3	17.9	2.5	5.0
		stressed-unstressed	39	7.7	11.6	1.9	3.8

post-test^N	*Vowels*	perceptual overlap	80	47.7	34.7	3.9	7.7
		stressed-unstressed	60	25.8	25.1	3.2	6.5
	Stress	perceptual overlap	88	45.3	28.2	3.0	6.0
		stressed-unstressed	66	28.0	25.2	3.1	6.2
	Vowels & Stress	perceptual overlap	96	46.0	29.1	3.0	5.9
		stressed-unstressed	72	30.3	24.7	2.9	5.8
	Baseline	perceptual overlap	52	14.4	23.5	3.3	6.5
		stressed-unstressed	39	8.1	11.0	1.8	3.6

vowel contrast	test	group	N	ER	SD	SE	CI
[ɐ]-[ɛ]	pretest	*Vowels*	20	50.0	31.6	7.1	14.8
		Stress	22	56.1	23.0	4.9	10.2
		Vowels & Stress	24	39.6	25.9	5.3	10.9
		Baseline	13	3.8	8.1	2.2	4.9
	post-test^R	*Vowels*	20	41.7	36.6	8.2	17.1
		Stress	22	34.8	22.4	4.8	9.9
		Vowels & Stress	24	34.7	30.0	6.1	12.7
		Baseline	13	3.8	8.1	2.2	4.9
	post-test^N	*Vowels*	20	25.8	25.4	5.7	11.9
		Stress	22	26.5	18.9	4.0	8.4
		Vowels & Stress	24	20.1	18.3	3.7	7.7
		Baseline	13	5.1	11.9	3.3	7.2
[ɛ]-[e]	pretest	*Vowels*	20	64.2	30.6	6.8	14.3
		Stress	22	62.1	28.0	6.0	12.4
		Vowels & Stress	24	58.3	33.4	6.8	14.1
		Baseline	13	23.1	24.5	6.8	14.8
	post-test^R	*Vowels*	20	66.7	26.7	6.0	12.5
		Stress	22	52.3	33.7	7.2	14.9
		Vowels & Stress	24	66.7	32.2	6.6	13.6
		Baseline	13	23.1	24.5	6.8	14.8
	post-test^N	*Vowels*	20	63.3	24.0	5.4	11.2
		Stress	22	62.9	16.5	3.5	7.3
		Vowels & Stress	24	72.2	16.8	3.4	7.1
		Baseline	13	38.5	23.0	6.4	13.9

[e]-[i]	pretest	*Vowels*	20	38.3	22.5	5.0	10.5
		Stress	22	33.3	19.2	4.1	8.5
		Vowels & Stress	24	27.8	21.5	4.4	9.1
		Baseline	13	14.1	10.6	2.9	6.4
	post-test^R^	*Vowels*	20	40.0	33.9	7.6	15.9
		Stress	22	28.8	20.5	4.4	9.1
		Vowels & Stress	24	27.1	20.8	4.2	8.8
		Baseline	13	14.1	10.6	2.9	6.4
	post-test^N^	*Vowels*	20	70.8	30.9	6.9	14.4
		Stress	22	53.8	22.1	4.7	9.8
		Vowels & Stress	24	54.9	22.8	4.6	9.6
		Baseline	13	14.1	18.5	5.1	11.2
[ɐ]-[ɨ]	pretest	*Vowels*	20	39.2	16.3	3.6	7.6
		Stress	22	38.6	19.1	4.1	8.5
		Vowels & Stress	24	36.1	21.6	4.4	9.1
		Baseline	13	0.0	7.0	1.9	4.2
	post-test^R^	*Vowels*	20	24.2	16.7	3.7	7.8
		Stress	22	32.6	24.0	5.1	10.6
		Vowels & Stress	24	27.1	20.7	4.2	8.7
		Baseline	13	0.0	7.0	1.9	4.2
	post-test^N^	*Vowels*	20	30.8	22.6	5.0	10.6
		Stress	22	37.9	29.8	6.4	13.2
		Vowels & Stress	24	36.8	16.4	3.4	6.9
		Baseline	13	0.0	7.0	1.9	4.2
[a]-[ɐ]	pretest	*Vowels*	20	13.3	14.5	3.2	6.8
		Stress	22	14.4	18.4	3.9	8.2
		Vowels & Stress	24	17.4	20.4	4.2	8.6
		Baseline	13	7.7	10.5	2.9	6.4
	post-test^R^	*Vowels*	20	6.7	11.1	2.5	5.2
		Stress	22	12.1	20.0	4.3	8.9
		Vowels & Stress	24	13.9	26.3	5.4	11.1
		Baseline	13	7.7	10.5	2.9	6.4
	post-test^N^	*Vowels*	20	8.3	18.4	4.1	8.6
		Stress	22	28.0	32.2	6.9	14.3
		Vowels & Stress	24	25.7	30.3	6.2	12.8
		Baseline	13	9.0	11.0	3.1	6.7

[ɛ]-[ɨ]	pretest	*Vowels*	20	16.7	16.4	3.7	7.7
		Stress	22	10.6	13.9	3.0	6.1
		Vowels & Stress	24	14.6	15.5	3.2	6.5
		Baseline	13	2.6	5.1	1.4	3.1
	post-test^R	*Vowels*	20	11.7	12.0	2.7	5.6
		Stress	22	9.1	10.7	2.3	4.7
		Vowels & Stress	24	13.2	16.2	3.3	6.8
		Baseline	13	2.6	5.1	1.4	3.1
	post-test^N	*Vowels*	20	22.5	17.1	3.8	8.0
		Stress	22	20.5	16.7	3.6	7.4
		Vowels & Stress	24	24.3	18.7	3.8	7.9
		Baseline	13	7.7	7.4	2.1	4.5
[e]-[ɨ]	pretest	*Vowels*	20	7.5	10.5	2.4	4.9
		Stress	22	9.8	12.2	2.6	5.4
		Vowels & Stress	24	11.1	13.7	2.8	5.8
		Baseline	13	12.8	13.4	3.7	8.1
	post-test^R	*Vowels*	20	4.2	9.2	2.1	4.3
		Stress	22	6.1	10.0	2.1	4.4
		Vowels & Stress	24	12.5	13.7	2.8	5.8
		Baseline	13	12.8	13.4	3.7	8.1
	post-test^N	*Vowels*	20	46.7	15.7	3.5	7.3
		Stress	22	35.6	16.8	3.6	7.5
		Vowels & Stress	24	41.0	14.1	2.9	5.9
		Baseline	13	7.7	12.6	3.5	7.6

D. Results for stress discrimination, for the catch trials

test	group	N	ER	SD	SE	CI
pretest	*Vowels*	60	37.7	32.7	4.2	8.4
	Stress	66	32.1	27.7	3.4	6.8
	Vowels & Stress	72	36.4	29.9	3.5	7.0
	Baseline	36	22.8	19.1	3.2	6.5
post-test^R	*Vowels*	60	39.6	32.0	4.1	8.3
	Stress	66	34.8	36.4	4.5	8.9
	Vowels & Stress	72	33.0	35.1	4.1	8.2
	Baseline	36	21.5	21.2	3.5	7.2

	post-test^N	*Vowels*	60	28.3	25.1	3.2	6.5
		Stress	66	26.9	31.0	3.8	7.6
		Vowels & Stress	72	29.9	31.9	3.8	7.5
		Baseline	36	22.2	22.5	3.7	7.6

test	stress location	group	N	ER	SD	SE	CI
pretest	1st_syllable	*Vowels*	20	23.0	22.5	5.0	10.5
		Stress	22	21.8	26.0	5.5	11.5
		Vowels & Stress	24	24.2	19.6	4.0	8.3
		Baseline	12	25.0	16.7	4.8	10.6
	2nd_syllable	*Vowels*	20	58.0	22.9	5.1	10.7
		Stress	22	47.3	20.2	4.3	9.0
		Vowels & Stress	24	56.7	25.3	5.2	10.7
		Baseline	12	25.0	18.8	5.4	11.9
	3rd_syllable	*Vowels*	20	32.0	26.0	5.8	12.2
		Stress	22	27.3	16.4	3.5	7.3
		Vowels & Stress	24	28.3	17.6	3.6	7.4
		Baseline	12	18.3	14.4	4.1	9.1
post-test^R	1st_syllable	*Vowels*	20	30.0	22.3	5.0	10.4
		Stress	22	20.5	20.8	4.4	9.2
		Vowels & Stress	24	21.9	23.4	4.8	9.9
		Baseline	12	29.2	18.7	5.4	11.9
	2nd_syllable	*Vowels*	20	60.0	21.7	4.8	10.1
		Stress	22	63.6	27.7	5.9	12.3
		Vowels & Stress	24	53.1	34.3	7.0	14.5
		Baseline	12	22.9	17.8	5.2	11.3
	3rd_syllable	*Vowels*	20	28.8	26.0	5.8	12.2
		Stress	22	20.5	20.2	4.3	9.0
		Vowels & Stress	24	24.0	20.4	4.2	8.6
		Baseline	12	12.5	15.1	4.4	9.6
post-test^N	1st_syllable	*Vowels*	20	21.3	22.5	5.0	10.5
		Stress	22	17.0	19.4	4.1	8.6
		Vowels & Stress	24	22.9	29.2	6.0	12.3
		Baseline	12	20.8	23.0	6.6	14.6
	2nd_syllable	*Vowels*	20	32.5	19.7	4.4	9.2
		Stress	22	45.5	29.3	6.2	13.0
		Vowels & Stress	24	46.9	27.4	5.6	11.6
		Baseline	12	27.1	20.4	5.9	13.0
	3rd_syllable	*Vowels*	20	31.3	21.8	4.9	10.2
		Stress	22	18.2	19.8	4.2	8.8
		Vowels & Stress	24	19.8	15.5	3.2	6.5
		Baseline	12	18.8	14.5	4.2	9.2

E. Results for stress discrimination, for the change trials

test	group	N	ER	SD	SE	CI
pretest	*Vowels*	40	56.7	13.9	2.2	4.4
	Stress	44	60.3	18.0	2.7	5.5
	Vowels & Stress	48	55.8	14.3	2.1	4.2
	Baseline	24	32.4	10.3	2.1	4.3
post-test^R	*Vowels*	40	52.4	10.6	1.7	3.4
	Stress	44	40.8	17.6	2.7	5.4
	Vowels & Stress	48	47.7	13.0	1.9	3.8
	Baseline	24	29.9	9.9	2.0	4.2
post-test^N	*Vowels*	40	67.0	11.8	1.9	3.8
	Stress	44	57.4	14.0	2.1	4.3
	Vowels & Stress	48	64.5	14.1	2.0	4.1
	Baseline	24	35.6	11.9	2.4	5.0

test	group	stress location	N	ER	SD	SE	CI
pretest	*Vowels*	1st-2nd syllable	20	63.5	10.8	2.4	5.0
		2nd-3rd syllable	20	49.8	9.1	2.0	4.3
	Stress	1st-2nd syllable	22	66.2	12.5	2.7	5.6
		2nd-3rd syllable	22	54.4	16.9	3.6	7.5
	Vowels & Stress	1st-2nd syllable	24	60.4	8.4	1.7	3.5
		2nd-3rd syllable	24	51.1	14.6	3.0	6.2
	Baseline	1st-2nd syllable	12	38.9	6.5	1.9	4.1
		2nd-3rd syllable	12	25.8	4.8	1.4	3.1
post-test^R	*Vowels*	1st-2nd syllable	20	56.2	7.8	1.8	3.7
		2nd-3rd syllable	20	48.5	9.2	2.1	4.3
	Stress	1st-2nd syllable	22	50.2	12.4	2.6	5.5
		2nd-3rd syllable	22	31.4	11.6	2.5	5.1
	Vowels & Stress	1st-2nd syllable	24	52.4	10.1	2.1	4.3
		2nd-3rd syllable	24	42.9	10.8	2.2	4.6
	Baseline	1st-2nd syllable	12	36.1	5.8	1.7	3.7
		2nd-3rd syllable	12	23.6	5.2	1.5	3.3
post-test^N	*Vowels*	1st-2nd syllable	20	70.6	11.1	2.5	5.2
		2nd-3rd syllable	20	63.3	8.4	1.9	3.9
	Stress	1st-2nd syllable	22	62.7	11.4	2.4	5.1
		2nd-3rd syllable	22	52.1	11.0	2.3	4.9
	Vowels & Stress	1st-2nd syllable	24	68.6	10.6	2.2	4.5
		2nd-3rd syllable	24	60.4	13.1	2.7	5.5
	Baseline	1st-2nd syllable	12	41.3	8.9	2.6	5.6
		2nd-3rd syllable	12	29.9	8.6	2.5	5.5

F. Analysis of monosyllabic testing stimuli, with the *Baseline* Group

independent variables		levels
participant	13 levels:	PT003', 'PT005', 'PT006', etc.
stimuli set	2 levels:	'[gV]', '[sV]'
trial type	2 levels:	'catch', 'change'
vowel contrast	13 levels:	'[ɐ]-[ɛ]', '[ɛ]-[e]', '[e]-[i]', '[ɐ]-[ɨ]', '[a]-[ɐ]', '[ɛ]-[ɨ]', '[e]-[ɨ]', '[ɐ]', '[ɛ]', '[e]', '[ɨ]', '[i]'
		GLMER models (family = binominal)
null model		answer ~ 1 + (1|participant)
model 1		answer ~ stimuli + (1|participant)
model 2		answer ~ trial type + (1|participant)
model 3		answer ~ vowel contrast + (1|participant)
		ANOVA results
null model, model 1		χ^2(1, N = 3) = 0.8634, p = .3528
null model, model 2		χ^2(1, N = 3) = 0.7973, p = .3719
null model, model 3		χ^2(12, N = 14) = 107.99, p < 2.2e-16

G. Analysis of monosyllabic training stimuli, with the *Baseline* Group

independent variables		levels
participant	30 levels:	PT033', 'PT034', 'PT035', etc.
task type	2 levels:	'AX', 'AXB'
vowel contrast	7 levels:	'[ɐ]-[ɛ]', '[ɛ]-[e]', '[e]-[i]', '[ɐ]-[ɨ]', '[a]-[ɐ]', '[ɛ]-[ɨ]', '[e]-[ɨ]'
		GLMER models (family = binominal)
null model		answer ~ 1 + (1|participant)
model 1		answer ~ task + (1|participant)
model 2		answer ~ vowel contrast + (1|participant)
		ANOVA results
null model, model 1		χ^2(1, N = 3) = 0.278, p = .598
null model, model 2		χ^2(6, N = 8) = 71.734, p = 1.802e-13

H. Analysis of disyllabic training stimuli, with the *Baseline* Group

independent variables		**levels**
participant	30 levels:	PT033', 'PT034', 'PT035', etc.
task type	2 levels:	'AX', 'AXB'
vowel reduction	3 levels:	'reduction on the 1st syllable', 'reduction on the 2nd syllable' 'reduction on both syllables'
		GLMER models (family = binominal)
null model		answer ~ 1 + (1|participant)
model 1		answer ~ task + (1|participant)
model 2		answer ~ vowel reduction + (1|participant)
		ANOVA results
null model, model 1		χ^2(1, N = 4) = 3.7075, *p* = .05417
null model, model 2		χ^2(2, N = 5) = 1.6914, *p* = .4293

I. Analysis of trisyllabic testing stimuli, with the *Baseline* Group

independent variables		**levels**
participant	12 levels:	PT005', 'PT006', 'PT007', etc.
stimuli set	2 levels:	'pretest', 'novel'
trial type	2 levels:	'catch', 'change'
contrast location	5 levels:	'1st-2nd syllable', '2nd-3rd syllable', '1st syllable', '2nd syllable', '3rd syllable'
		GLMER models (family = binominal)
null model		answer ~ 1 + (1|participant)
model 1		answer ~ stimuli + (1|participant)
model 2		answer ~ trial type + (1|participant)
model 3		answer ~ contrast location + (1|participant)
		ANOVA results
null model, model 1		χ^2(1, N = 3) = 1.3449, *p* = .2462
null model, model 2		χ^2(1, N = 3) = 18.928, *p* = 1.357e-05
null model, model 3		χ^2(4, N = 6) = 6.741, *p* = 1.135e-09

J. Analysis of trisyllabic training stimuli, with the *Baseline* Group

independent variables		levels
participant	30 levels:	PT033', 'PT034', 'PT035', etc.
task type	2 levels:	'AX', 'AXB'
contrast location	2 levels:	'1st-2nd syllable', '2nd-3rd syllable'
		GLMER models (family = binominal)
null model		answer ~ 1 + (1|participant)
model 1		answer ~ task + (1|participant)
model 2		answer ~ contrast location + (1|participant)
		ANOVA results
null model, model 1		χ^2(1, N = 3) = 1.6001, *p* = .2059
null model, model 2		χ^2(1, N = 3) = 40.267, *p* = 2.215e-10

K. Analysis of stress vs. consonant contrast, for intervention groups

independent variables		levels
participant	66 levels:	'BGE03', 'ELTEBA12', 'CORV04', etc.
group	3 levels:	'Vowels', 'Stress', 'Vowels & Stress'
condition	3 levels:	'stress discrimination', 'consonant discrimination'
		GLMER models (family = binominal)
null model		answer ~ 1 + (1|participant)
model 1		answer ~ condition + (1|participant)
model 2		answer ~ condition*group + (1|participant)
		ANOVA results
null model, model 1		χ^2(1, N = 4) = 207.23, *p* < 2.2e-16
model 1, model 2		χ^2(4, N = 8) = 3.048, *p* = .5498

L. Analysis of catch trials for vowel discrimination (pretest), for intervention groups

independent variables		levels
participant	66 levels:	'BGE03', 'ELTEBA12', 'CORV04', etc.
group	3 levels:	'Vowels', 'Stress', 'Vowels & Stress'
vowel	6 levels:	'[a]', '[ɐ]', '[ɛ]', '[e]', '[ɨ]', '[i]'
		GLMER models (family = binominal)
null model		answer ~ 1 + (1|participant)
model 1		answer ~ vowel + (1|participant)
model 2		answer ~ vowel*group + (1|participant)

		ANOVA results
null model, model 1		χ^2(5, N = 7) = 92.562, *p* < 2.2e-16
model 1, model 2		χ^2(12, N = 19) = 12.48, *p* = .4079

M. Analysis of catch trials for stress discrimination (pretest), for intervention groups

independent variables		**levels**
participant	66 levels:	'BGE03', 'ELTEBA12', 'CORV04', etc.
group	3 levels:	'Vowels', 'Stress', 'Vowels & Stress'
stress location	6 levels:	'1st syllable', '2nd syllable', '3rd syllable'
		GLMER models (family = binominal)
null model		answer ~ 1 + (1|participant)
model 1		answer ~ vowel + (1|participant)
model 2		answer ~ vowel*group + (1|participant)
		ANOVA results
null model, model 1		χ^2(2, N = 4) = 84.012, *p* << 2.2e-16
model 1, model 2		χ^2(6, N = 10) = 2.507, *p* = .8677

N. Analysis of vowel discrimination, by test (including Baseline Group)

independent variables		**levels**
participant	79 levels:	'BGE03', 'ELTEBA12', 'PT014', etc.
group	4 levels:	'Vowels', 'Stress', 'Vowels & Stress', 'Baseline'
contrast type	2 levels:	'perceptual overlap', 'stressed-unstressed'
vowel contrast	7 levels:	'[ɐ]-[ɛ]', '[ɛ]-[e]', '[e]-[i]', '[ɐ]-[ɨ]', '[a]-[ɐ]', '[ɛ]-[ɨ]', '[e]-[ɨ]'
test	3 levels:	'pretest', 'post-test^{R}', 'post-test^{N}'
		GLMER models (family = binominal)
null model		answer ~ 1 + (1|participant)
model 1		answer ~ group + (1|participant)
model 2		answer ~ contrast type + (1|participant)
model 3		answer ~ contrast type*group + (1|participant)
model 4		answer ~ vowel contrast + (1|participant)
model 5		answer ~ vowel contrast*group + (1|participant)
		ANOVA results
		pretest trials
null model, model 1		χ^2(3, N = 5) = 84.204, *p* < 2.2e-16

null model, model 2	$\chi^2(1, N = 3) = 339.38$, $p < 2.2e-16$
model 2, model 3	$\chi^2(6, N = 9) = 104.99$, $p < 2.2e-16$
null model, model 4	$\chi^2(6, N = 8) = 427.52$, $p < 2.2e-16$
model 4, model 5	$\chi^2(21, N = 29) = 147.39$, $p < 2.2e-16$
	post-test^R trials
null model, model 1	$\chi^2(3, N = 5) = 45.175$, $p = 8.494e-10$
null model, model 2	$\chi^2(1, N = 3) = 319.57$, $p < 2.2e-16$
model 2, model 3	$\chi^2(6, N = 9) = 71.266$, $p = 2.248e-13$
null model, model 4	$\chi^2(6, N = 8) = 455.43$, $p < 2.2e-16$
model 4, model 5	$\chi^2(21, N = 29) = 118.56$, $p = 1.309e-15$
	post-test^N
null model, model 1	$\chi^2(3, N = 5) = 88.314$, $p < 2.2e-16$
null model, model 2	$\chi^2(1, N = 3) = 101.4$, $p < 2.2e-16$
model 2, model 3	$\chi^2(6, N = 9) = 90.303$, $p < 2.2e-16$
null model, model 4	$\chi^2(6, N = 8) = 379.88$, $p < 2.2e-16$
model 4, model 5	$\chi^2(21, N = 29) = 153.05$, $p < 2.2e-16$

O. Analysis of progress in vowel discrimination, from pre-test to post-test

independent variables		levels
participant	66 levels:	'BGE03', 'ELTEBA12', 'CORV01', etc.
group	3 levels:	'Vowels', 'Stress', 'Vowels & Stress'
contrast type	2 levels:	'perceptual overlap', 'stressed-unstressed'
vowel contrast	7 levels:	'[ɐ]-[ɛ]', '[ɛ]-[e]', '[e]-[i]', '[ɐ]-[ɨ]', '[a]-[ɐ]', '[ɛ]-[ɨ]', '[e]-[ɨ]'
test	3 levels:	'pretest', 'post-test^R', 'post-test^N'
		GLMER models (family = binominal)
null model		answer ~ 1 + (1|participant)
model 1		answer ~ test + (1|participant)
model 2		answer ~ test*group + (1|participant)
model 3		answer ~ contrast type*test + (1|participant)
model 4		answer ~ contrast type*test*group + (1|participant)
model 5		answer ~ vowel contrast*test + (1|participant)
model 6		answer ~ vowel contrast*test*group + (1|participant)
		ANOVA results
		from pretest to post-test^R
null model, model 1		$\chi^2(1, N = 3) = 12.782$, $p = .0003499$
model 1, model 2		$\chi^2(4, N = 7) = 4.6235$, $p = .3282$
model 1, model 3		$\chi^2(2, N = 5) = 688.99$, $p < 2.2e-16$

model 3, model 4	$\chi^2(8, N = 13) = 19.784$, $p = .01119$
model 1, model 5	$\chi^2(12, N = 15) = 895.95$, $p < 2.2e-16$
model 5, model 6	$\chi^2(28, N = 43) = 45.285$, $p = .02064$
	from pretest to post-test^N
null model, model 1	$\chi^2(1, N = 3) = 32.195$, $p = 1.395e-08$
model 1, model 2	$\chi^2(4, N = 7) = 3.3896$, $p = .4949$
model 1, model 3	$\chi^2(2, N = 5) = 447.9$, $p < 2.2e-16$
model 3, model 4	$\chi^2(8, N = 13) = 11.77$, $p = .1617$
model 1, model 5	$\chi^2(12, N = 15) = 774.5$, $p < 2.2e-16$
model 5, model 6	$\chi^2(28, N = 43) = 56.564$, $p = .001097$

P. Analysis of stress discrimination, by test (including Baseline Group)

independent variables		**levels**
participant	79 levels:	'BGE03', 'ELTEBA12', 'PT014', etc.
group	4 levels:	'Vowels', 'Stress', 'Vowels & Stress', 'Baseline'
contrast location	2 levels:	'1^{st}-2^{nd} syllable', '2^{nd}-3^{rd} syllable'
test	3 levels:	'pretest', 'post-test^R', 'post-test^N'

	GLMER models (family = binominal)
null model	answer ~ 1 + (1|participant)
model 1	answer ~ group + (1|participant)
model 2	answer ~ contrast location + (1|participant)
model 3	answer ~ contrast location*group + (1|participant)
	ANOVA results
	pretest trials
null model, model 1	$\chi^2(3, N = 5) = 19.679$, $p = .0001979$
null model, model 2	$\chi^2(1, N = 3) = 75.666$, $p = 2.2e-16$
model 2, model 3	$\chi^2(6, N = 9) = 22.39$, $p = .001029$
	post-test^R
null model, model 1	$\chi^2(3, N = 5) = 13.67$, $p = .003391$
null model, model 2	$\chi^2(1, N = 3) = 65.062$, $p = 7.257e-16$
model 2, model 3	$\chi^2(6, N = 9) = 24.104$, $p = .0004997$
	post-test^N
null model, model 1	$\chi^2(3, N = 5) = 24.39$, $p = 2.071e-05$
null model, model 2	$\chi^2(1, N = 3) = 37.821$, $p = 7.756e-10$
model 2, model 3	$\chi^2(6, N = 9) = 25.54$, $p = .0002711$

Q. Analysis of progress in stress discrimination, from pre-test to post-test

independent variables		levels
participant	66 levels:	'BGE03', 'ELTEBA12', 'CORV01', etc.
group	3 levels:	'Vowels', 'Stress', 'Vowels & Stress'
contrast location	2 levels:	'1st-2nd syllable', '2nd-3rd syllable'
test	3 levels:	'pretest', 'post-testR', 'post-testN'
		GLMER models (family = binominal)
null model		answer ~ 1 + (1|participant)
model 1		answer ~ test + (1|participant)
model 2		answer ~ test*group + (1|participant)
model 3		answer ~ contrast location*test + (1|participant)
model 4		answer ~ contrast location*test*group + (1|participant)
		ANOVA results
		from pretest to post-testR
null model, model 1		$\chi^2(1, N = 3) = 89.564$, $p < 2.2e-16$
model 1, model 2		$\chi^2(4, N = 7) = 32.066$, $p = 1.855e-06$
model 1, model 3		$\chi^2(2, N = 5) = 109.77$, $p < 2.2e-16$
model 3, model 4		$\chi^2(8, N = 13) = 42.798$, $p = 9.587e-07$
		from pretest to post-testN
null model, model 1		$\chi^2(1, N = 3) = 22.65$, $p = 1.944e-06$
model 1, model 2		$\chi^2(4, N = 7) = 28.408$, $p = 1.031e-05$
model 1, model 3		$\chi^2(2, N = 5) = 86.138$, $p < 2.2e-16$
model 3, model 4		$\chi^2(8, N = 13) = 30.693$, $p = .0001593$

R. Analysis of progress in training, for Group *Vowels*

independent variables		levels
participant	20 levels:	'CORV05', 'ELTEBA01', 'MET06', etc.
session	6 levels:	'session 1', 'session 2', 'session 3', 'session 4' 'session 5', 'session 6'
		GLMER models (family = binominal)
null model		answer ~ 1 + (1|participant)
model 1		answer ~ session + (1|participant)
		ANOVA results
		for the subset dataframe with the contrast '[ɐ]'-'[ɛ]'
null model, model 1		$\chi^2(2, N = 4) = 2.0996$, $p = .35$

	for the subset dataframe with the contrast '[ɛ]'-'[e]'
null model, model 1	$\chi^2(2, N = 4) = 7.9296$, $p = .01897$
	for the subset dataframe with the contrast '[e]'-'[i]'
null model, model 1	$\chi^2(2, N = 4) = .9479$, $p = .6225$
	for the subset dataframe with the contrast '[ɐ]'-'[ɨ]'
null model, model 1	$\chi^2(2, N = 4) = 7.8992$, $p = .01926$
	for the subset dataframe with the contrast '[a]-[ɐ]'
null model, model 1	$\chi^2(2, N = 4) = 2.9518$, $p = .2286$
	for the subset dataframe with the contrast '[ɛ]'-'[ɨ]'
null model, model 1	$\chi^2(2, N = 4) = 3.6404$, $p = .162$
	for the subset dataframe with the contrast '[e]'-'[ɨ]'
null model, model 1	$\chi^2(2, N = 4) = 8.4947$, $p = .0143$

S. Analysis of progress in training, for Group *Stress*

independent variables		**levels**
participant	22 levels:	'BGE08', 'CORV01', 'ELTEBA04', etc.
session	6 levels:	'session 1', 'session 2', 'session 3', 'session 4' 'session 5', 'session 6'
		GLMER models (family = binominal)
null model		answer ~ 1 + (1|participant)
model 1		answer ~ session + (1|participant)
		ANOVA results
		for the subset dataframe with the contrast '1st-2nd syllable'
null model, model 1		$\chi^2(4, N = 6) = 11.1$, $p = .02546$
		for the subset dataframe with the contrast '2nd-3rd syllable'
null model, model 1		$\chi^2(4, N = 6) = 51.822$, $p = 1.503e-10$

T. Analysis of progress in training, for Group *Vowels & Stress*

independent variables		**levels**
participant	22 levels:	'CORV17', 'DEB12', 'KAROL08', etc.
session	6 levels:	'session 1', 'session 2', 'session 3', 'session 4' 'session 5', 'session 6'
		GLMER models (family = binominal)
null model		answer ~ 1 + (1|participant)
model 1		answer ~ session + (1|participant)

ANOVA results	
for the subset dataframe with the vowel reduction on the 1st syllable'	
null model, model 1	$\chi^2(2, N = 4) = 42.976$, $p = 4.655e-10$
for the subset dataframe with the vowel reduction on the 2nd syllable'	
null model, model 1	$\chi^2(2, N = 4) = 8.8723$, $p = .01184$
for the subset dataframe with the vowel reduction on both syllables'	
null model, model 1	$\chi^2(3, N = 5) = 53.782$, $p = 1.249e-11$

U. Analysis of extra-linguistic factors, in general improvement, by participant

independent variables		**levels**
participant	66 levels:	'BGE03', 'ELTEBA12', 'MET06', etc.
progress in training	3 levels:	'positive', 'none or negative'
affiliation	10 levels:	'BGE', 'CORV', 'ELETEBA', etc.
contact with EP	2 levels:	'none', 'normal'
weekly course load	3 levels:	'270', 180', '90'
		GLM models (family = binominal)
null model		improvement ~ 1
model 1		improvement ~ progress in training
model 2		improvement ~ affiliation
model 3		improvement ~ contact with EP
model 4		improvement ~ weekly course load
		ANOVA results
null model, model 1		$\chi^2(2) = 0.15213$, $p = .9268$
null model, model 2		$\chi^2(9) = 6.7699$, $p = .6611$
null model, model 3		$\chi^2(1) = 0.01466$, $p = .9036$
null model, model 4		$\chi^2(2) = 2.9343$, $p = .2306$

Appendix XI

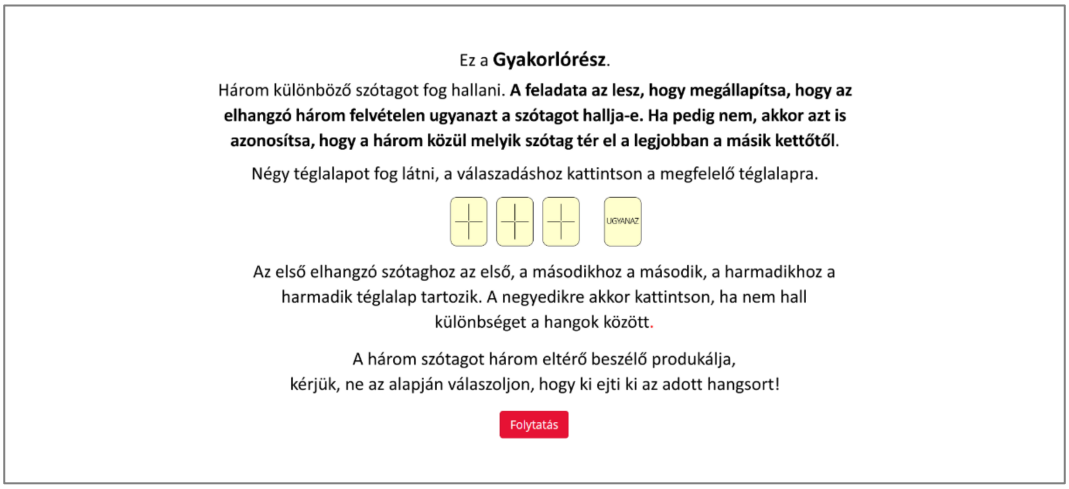
Screenshots of experimental tasks, for the perceptual training study

(Chapter 5)

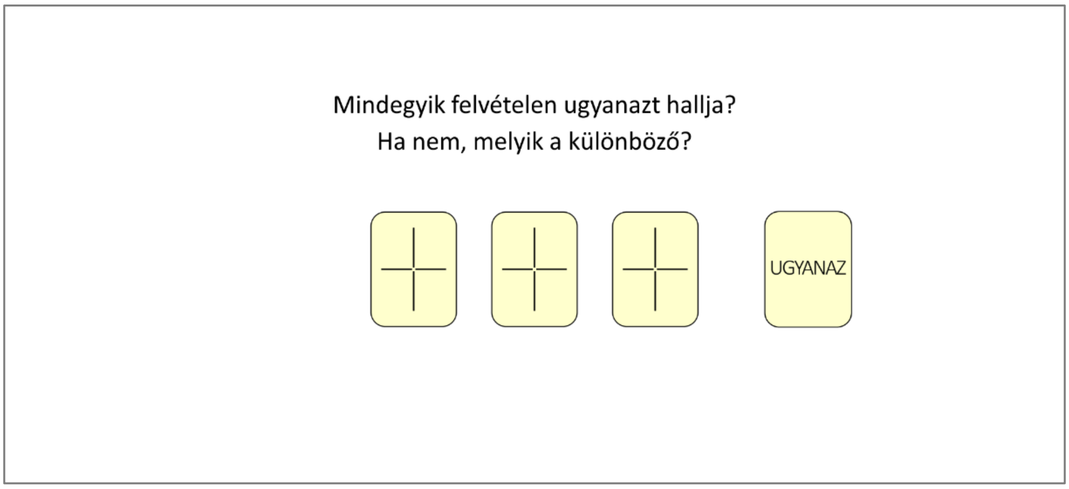
A. Screenshots from the general instructions (tests)

B. Screenshot of the volume adjustment section (tests)

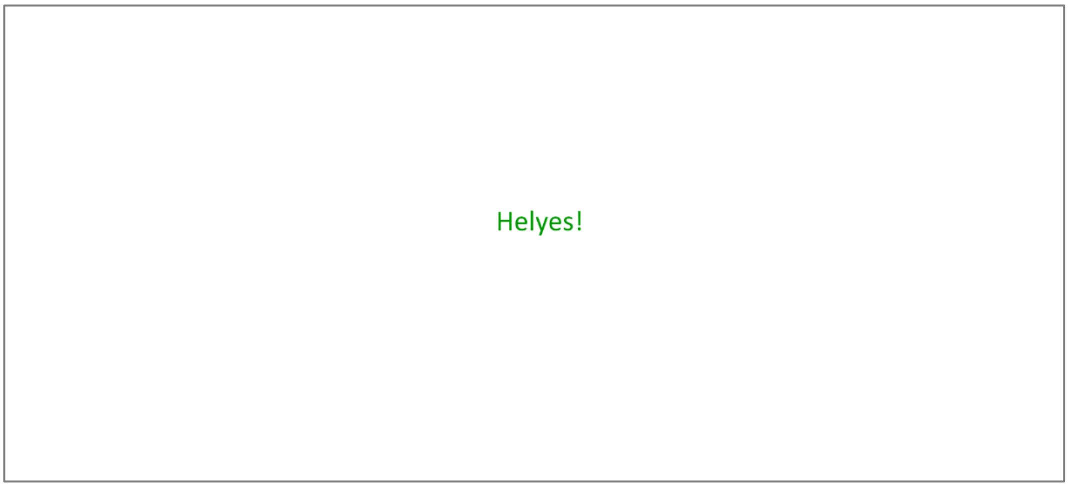
C. Screenshot of the instructions of the familiarization task (tests)



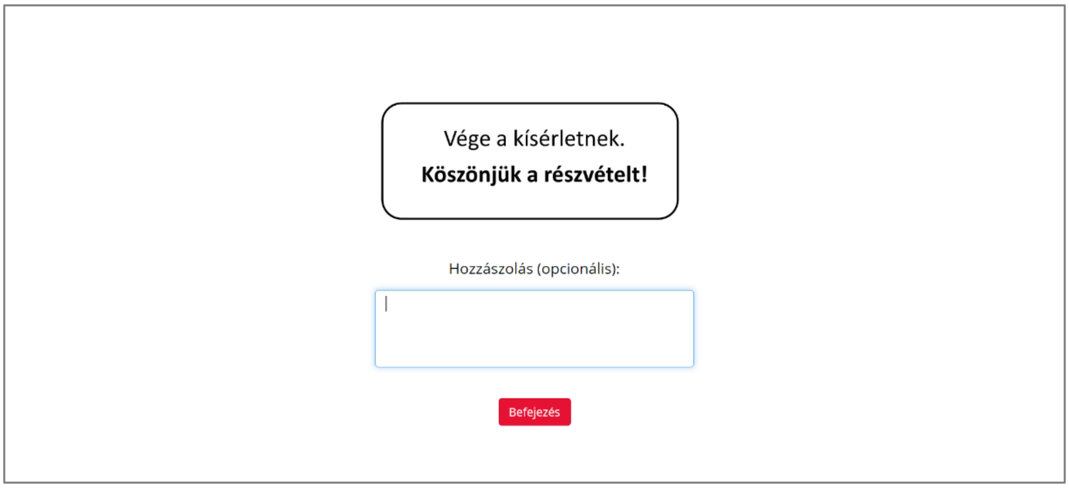
D. Screenshot of the oddity trials (tests)



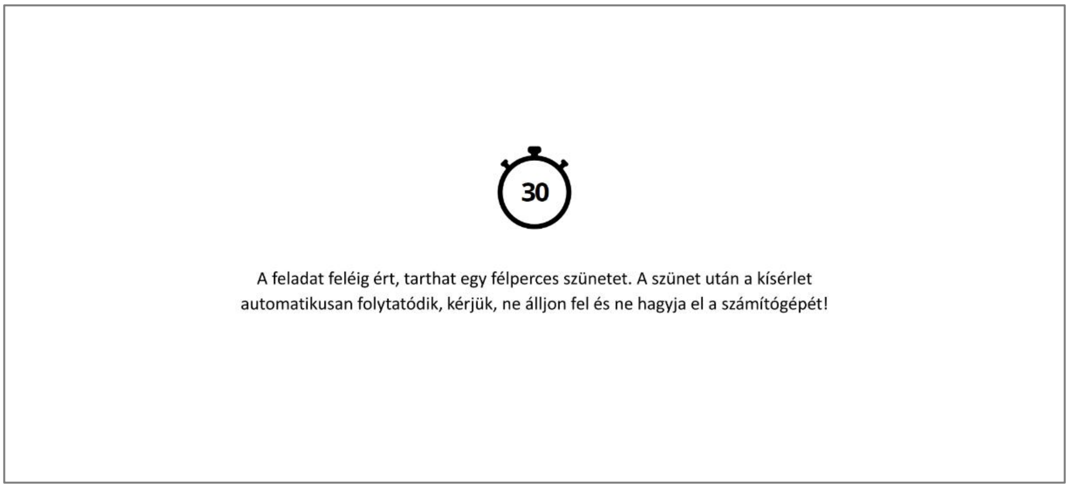
E. Screenshots of feedback in the familiarization trials (tests) and training trials



F. Screenshot of the end of a session (tests)

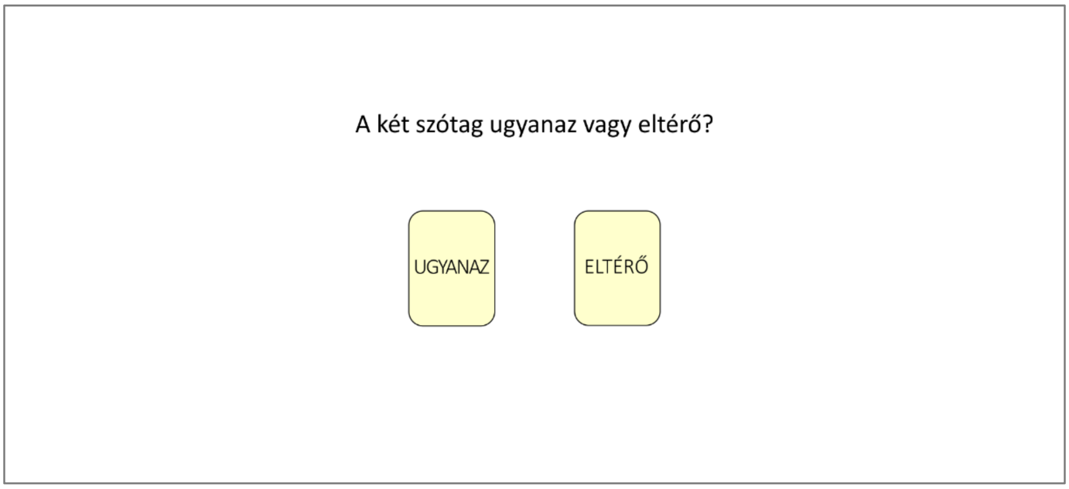


G. Screenshot of the pauses

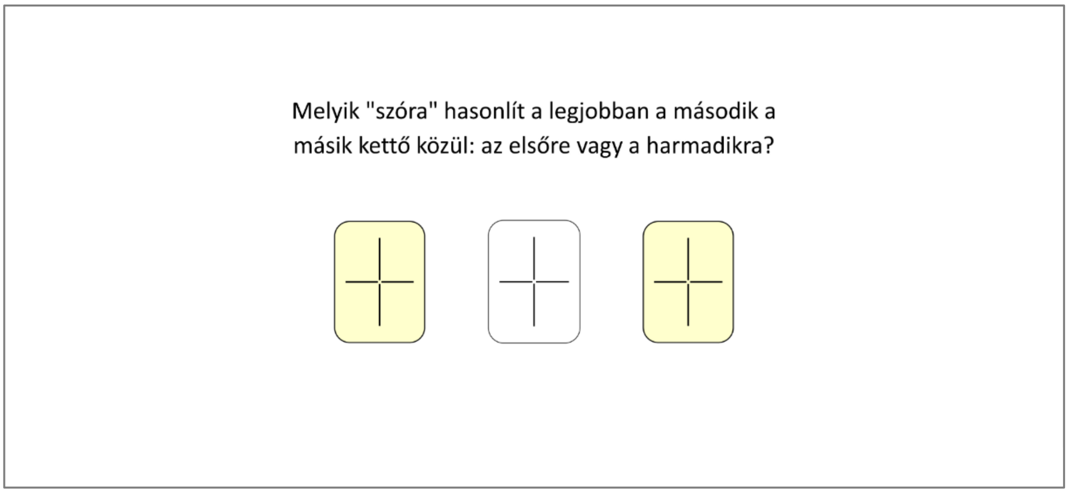


H. Screenshot of the first instruction in the training sessions

I. Screenshot of an AX trial (training)



J. Screenshot of an AXB trial (training)



Appendix XII

Description of the bonus tasks, included in the perceptual training study (Chapter 5)

A. Bonus task for the first training session

The first task topic was the expression “Vai passear macacos” (literaly, 'Go walk monkeys!'). First, students had to hear a recording of the sentence, after, they had to choose, from three possibilities, the literal meaning of the sentence (translated to Hungarian). The final task was to decide on the meaning with which EP speakers use it, also among three choices.

Figure XII.1. Computer screen of the first bonus-task

B. Bonus task for the second training session

This task was focused on the nicknames for the hand fingers, for children: “Pequenino, seu vizinho, pai de todos, fura bolos, mata-piolhos”. First, students had to hear a recording of the phrase, without any image or explanation, as many times as they wanted. After, they had to reconstruct the sequence of nicknames. In the fourth screen, an explanation was provided, and finally, the terms currently used (by adults) were displayed.

Figure XII.2. Computer screen of the second bonus task

C. Bonus task for the third training session

In the first screen of this task, six images, with written subtitles in Portuguese and recordings of each subtitle. The six words were retrieved from three popular expressions, each expression containing two of the target-words. Students were asked to memorize each word and its meaning. After, they were presented with an audio recording of the popular expression, with two of the target words, and they had to decide, from three possibilities, which was the expression. After choosing, they were presented with the literal and metaphorical meaning of the expression.

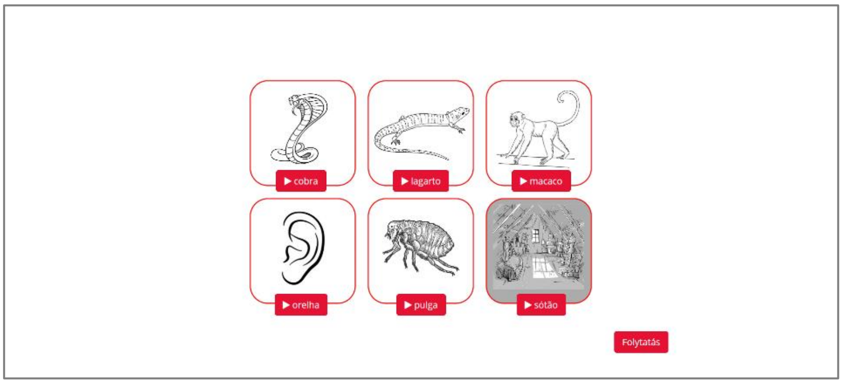


Figure XII.3. Computer screen of the third bonus task

D. Bonus task for the fourth training session

The topic of this session was a popular children's song. In the first step, the first quatrain was presented on the screen in writing, with 3 missing words. Students had to fill the gaps by listening to a recording of the song (the 1st quatrain) and with the help of three images. After this task was completed, they could hear the complete version of the song, with the original lyrics displayed.

Figure XII.4. Computer screen of the fourth bonus task

E. Bonus task for the fifth training session

This task was focused on a tongue twister with the consonant [ʀ]: "O rato roeu a rolha da garrafa de rum do Rei da Rússia". Participants were presented with six images, corresponding to each of the terms in the phrase. They were asked to think on a possible phrase with those terms. Finally, they were presented with the phrase, a recording, and the meaning.

Figure XII.5. Computer screen of the fifth bonus task

F. Bonus task for the sixth training session

The last task consisted on a series of body gestures that have a meaning (e.g., to shrug your shoulders, as 'not caring about something'). Six brief videos without sound were displayed, each with a gesture. For each gesture, participants had to decided, among five options, the meaning of the gesture, and after, a sentence in Portuguese with the same meaning was presented. In the end of the six videos, we presented a longer video with the author of the thesis and a translator, repeating and explaining – in Portuguese and Hungarian – the gestures.