FACULTY OF ARTS
AND PHILOSOPHY

Perception of Non-Native Accented Speech:

An Experimental Study on Lexically-Guided Perceptual Learning of a Dutch Vowel Contrast

Gil Verbeke

Student ID: 01703389

Supervisor: Prof. Dr. Ellen Simon (Ghent University)
Co-Supervisors: Prof. Dr. Robert Hartsuiker (Ghent University)
Prof. Dr. Holger Mitterer (University of Malta)

A dissertation submitted to Ghent University in partial fulfilment of the requirements for the degree of Master of Arts in Linguistics and Literature, language combination: Dutch – English.

Academic year: 2020 – 2021

GHENT
UNIVERSITY

Abstract

Many studies have demonstrated that listeners can rapidly recalibrate their perceptual boundaries of specific L1 phoneme categories after exposure to a native speaker producing those phonemes noncanonically (Norris et al., 2003), or an L2 speaker producing those speech sounds with a non-native accent (Bradlow & Bent, 2008). The present study aims to gain further insight into how Dutch L1 listeners adapt to Italian accented front vowels, and how short-term experience with one L2 speaker's accent might help these listeners to interpret the speech of another L2 speaker.

The traditional experimental paradigm to study lexically-guided perceptual learning is used. In a lexical decision task, 100 native speakers of Belgian Dutch are exposed to 40 Dutch target words, which have either /ɪ/ or /i/ as the syllable nucleus, 60 fillers and 100 nonwords. All stimuli were produced by a female native speaker of Italian who is highly proficient in Dutch, but has a clearly noticeable Italian accent. Stimuli are presented in two conditions: participants either hear target words in which the /ɪ/-sound is replaced by an ambiguous sound in between [ɪ]-[i] and canonically produced /i/-words (/ɪ/-ambiguous), or the exact opposite pattern (/i/-ambiguous).

To assess if perceptual learning has developed, a phoneme categorization task is set up in which participants need to identify the front vowel in five Dutch /ɪ/-/i/ minimal pairs across two conditions: listeners either hear stimuli produced by the same female speaker or stimuli produced by a male-sounding speaker, whose voice was created from the female speaker's voice using the '*change gender*' function in *Praat*. This modification controls for comparable durational cues, but the change in pitch and formant frequencies leads listeners to perceive a male voice.

Neither for the female speaker nor for the male-sounding speaker did we observe auditory perceptual learning effects. That is, participants did not identify the ambiguous vowel in the minimal word pairs significantly differently depending on the exposure condition to which they had been assigned. As a corollary, we cannot ascertain whether perceptual learning of ambiguous vowel sounds could be transferred from one L2 speaker to another. A more in-depth analysis of the stimulus materials suggested that listeners may have been biased to hear the /ɪ/-word of some minimal pairs. A follow-up study with adjusted stimulus materials is therefore necessary to obtain a better understanding of how native speakers process L2 accented speech.

Keywords: speech perception, lexically-guided perceptual learning, L2 accented speech, cross-talker generalization, Dutch front vowels

Acknowledgements

"The only way to discover the limits of the possible is to go beyond them."
Arthur C. Clarke

"Don't you know that only fools are satisfied?"
Billy Joel, *Vienna*

What started out as a rudimentary idea at the end of August 2020 has eventually evolved into this final version of my master's dissertation. Using the determiner *my* is actually not fully justified, as I would not have been able to carry out this research project without the help of a rather unexpectedly large number of people. For that reason, I would like to express my deepest gratitude and appreciation to…

- ***Prof. Dr. Ellen Simon***, for your ongoing support and your insightful feedback on earlier drafts, but mostly for your eagerness to answer all of my thesis and non-thesis related questions;
- ***Prof. Dr. Robert Hartsuiker***, for helping me with the brainstorming process and data analysis, as well as sharing your expertise in experimental designs;
- ***Prof. Dr. Holger Mitterer***, for your willingness to co-supervise this project and helping me with the creation of the stimulus materials;
- ***Drs. Chi Zhang***, for providing me with useful code for my script, and ***Tristan Covemaeker***, for sacrificing (a lot of) your time to help me out with *JavaScript*, *jsPsych* and debugging, because it would not have been even remotely possible to run my experiment without you;
- ***Dr. Ludovic De Cuypere***, for helping me with the linear regression analyses and the interpretation of the statistical results;
- ***the Italian speaker***, for lending me your voice;
- ***all 126 participants***, for their willingness to participate in the online experiments;
- ***Thijs Gillioen*** and ***Chris Bulcaen***, for your assistance in helping me recruit participants;
- ***the writing groups***, for providing a great work environment while writing this dissertation;
- ***Merel Waeyaert*** and ***Thijs Gillioen***, for adding valuable comments to earlier drafts and proofreading this master's thesis;
- ***my parents***, ***Maaike*** and ***Maxime***, for supporting me during the writing process of this thesis;
- ***my friends***, for keeping me away from my computer when they felt that I needed to.

Table of Contents

Abstract .... **i**
Acknowledgements .... **ii**
List of Abbreviations .... **v**
List of Tables .... **vi**
List of Figures .... **vii**
1 Introduction **1**
2 Literature Review **5**
2.1 Perceptual Learning 5
2.2 Models of Speech Production and Perception 7
2.2.1 Speech Learning Model (SLM) 7
2.2.2 Perceptual Assimilation Model (PAM) 9
2.3 Lexically-Guided Perceptual Learning 11
2.4 Generalization of Perceptual Learning 14
2.4.1 Generalization Across Stimulus Materials 14
2.4.2 Generalization Across Speakers 16
3 The Present Study **19**
4 Pre-Test **23**
4.1 Method 23
4.1.1 Participants 23
4.1.2 Stimulus Materials 24
4.1.3 Procedure 29
4.1.4 Analysis 31
4.2 Results 32
4.2.1 Critical Words 33
4.2.2 Minimal Words 35
4.3 Discussion 37
5 Lexical Decision Task **39**
5.1 Method 39
5.1.1 Participants 39
5.1.2 Stimulus Materials 40
5.1.3 Procedure 42
5.1.4 Analysis 43
5.2 Results 45
5.3 Discussion 49

6 Phoneme Categorization Task **53**
6.1 Method 53
6.1.1 Participants 53
6.1.2 Stimulus Materials 54
6.1.3 Procedure 55
6.1.4 Analysis 56
6.2 Results 57
6.3 Discussion 61
7 General Discussion **65**
8 Conclusion **69**
9 References **71**
10 Appendices **87**
10.1 Participant Information 88
10.2 Participant Overview 92
10.3 Speaker Information 99
10.4 Stimuli Overview 111
10.5 Experiment Results 117
10.5.1 Pre-Test 117
10.5.2 Lexical Decision Task 124
10.5.3 Phoneme Categorization Task 131

Word count: 25,291 words *(in-text references, footnotes, quotes and tables included)*

List of Abbreviations

F0 = fundamental frequency

F1 = first formant

F2 = second formant

F3 = third formant

fpmw = frequency per million words

L1 = first or native language

L2 = second language

OLD20 = orthographic Levenshtein distance of a nonword's twenty closest lexical neighbors

PAM = Perceptual Assimilation Model

SDT = Signal Detection Theory

SLM = Speech Learning Model

List of Tables

Table 1: Overview of spectral and durational characteristics of the non-native speaker's vowel production by the two types of critical words. 26

Table 2: Fundamental frequency (F0) and formant frequencies of the minimal word pairs produced by the female speaker and the manipulated, male-sounding voice, averaged over the individual minimal words. 28

Table 3. Absolute and relative frequencies of the vowel categorization responses by continuum step across the four types of target words. 32

Table 4: Overview of the spectral and temporal characteristics of the non-native speaker's productions of the Dutch /ɑ/-/aː/ and /ɔ/-/oː/ contrasts. 40

Table 5: Absolute frequencies of the lexical decision responses by item types and exposure condition. 46

List of Figures

Figure 1: Perceptual assimilation patterns and discrimination performance for non-native speech segments and segmental contrasts based on Best (1995, pp. 194-195). Curly lines indicate that the assimilation type for a non-native speech contrast spans two different assimilation types for individual non-native sounds. 10

Figure 2: Schematic representation of a single experimental trial in the phoneme categorization task. 30

Figure 3: Relative frequencies of the categorization responses of the critical /ɪ/-words by continuum steps. 33

Figure 4: Relative frequencies of the categorization responses of the critical /i/-words by continuum steps. 34

Figure 5: Relative frequencies of the categorization responses of the minimal /ɪ/-/i/ word pairs with the female speaker's voice by continuum steps. 35

Figure 6: Relative frequencies of the categorization responses of the minimal /ɪ/-/i/ word pairs with the male speaker's voice by continuum steps. 36

Figure 7: Mean proportion of 'word'-responses by item type and exposure condition. 45

Figure 8: Mean response time (ms) and standard errors for /ɪ/-words and /i/-words by exposure condition. 47

Figure 9: Distribution of response latency (ms) across word types and exposure conditions. 48

Figure 10: Proportions of the phoneme categorization responses by exposure conditions and speaker conditions across the four continuum steps. 57

Figure 11: Probability of /i/-responses by (A) speaker condition, (B) exposure condition and (C) continuum step. 59

Figure 12: Caterpillar plots of the 95% confidence intervals for by-item and by-subject intercepts of the random effects. 60

1 Introduction

In our everyday communicative encounters, we are continuously confronted with the heterogeneous and versatile nature of spoken language. When multiple people are engaged in a conversation, for instance, speakers will not only differ in terms of which words and expressions they use to verbalize their thoughts, they will also differ substantially in how they physically produce those thoughts. That is, even when two speakers say the exact same word, the acoustic characteristics of one talker's production may vary seemingly infinitely from that of another talker, be it because they speak with another regional accent or because they have a different language background. Liberman et al. (1967) were among the first to investigate this 'lack of invariance' problem in speech, which implies that speech segments, such as vowels or consonants, can be realized by an extensive range of acoustically distinct sounds. Nevertheless, the human perceptual system has been shown to be sufficiently sophisticated to accommodate this large amount of acoustic variability, in turn enabling listeners to correctly associate the individual sounds in the continuous speech stream with the intended phonological categories and as such reconstruct the message that their interlocutor is trying to convey (for a review, see Weatherholtz & Jaeger, 2016). This complex interplay between production and perception lies at the core of the present paper, as we will examine how native speakers of a particular language process speech that deviates from the standard language pronunciation norm.

Fine-grained variation in how particular speech segments and words are produced has been found to originate from various factors. One factor is that individual speakers have an idiosyncratic speech articulation which considerably differs from that of others along several dimensions. Such interpersonal differences can mostly be attributed to physiological or anatomical variation, such as the size of the vocal tract (e.g., Sjerps et al., 2011), gender-related differences in pitch range (e.g., Hillenbrand et al., 1995; Jongman et al., 1998; Newman et al., 2001), or speech impairments (e.g., LeGendre et al., 2009, and Liss et al., 2002, on dysarthria). On top of these biologically defined criteria, it is even virtually impossible for the same speaker to produce the same utterance with identical acoustic features (McMurray & Jongman, 2011; Simpson, 2001). Besides, this type of intra-speaker variability may also be context-conditioned: people have been found to talk noticeably different than usual when they are, for example, intoxicated or speaking quickly (Liu & Jaeger, 2019), or when a specific speech situation demands a more formal register (e.g., Labov, 1972, and Babel & Munson, 2014, on speech styles).

Reaching beyond the characteristics of individual speakers, factors modulating speech production can also operate at the level of particular groups of speakers. Largely, we can divide group-level variability in spoken language production into two categories. On the one hand, the precise articulation of particular sounds may deviate structurally from the standard language variety within a single language community or even within a certain social group (i.e., *endogenous variation*). To give just one example, the concrete realization of the alveolar fricative /s/ has been found to vary across some regional and dialectal varieties of English. Speakers of New York/Long Island English, for instance, tend to produce this fricative in word-initial position more often as a postalveolar [ʃ], thus pronouncing *street* as [ʃtriːt] rather than [striːt] (Kraljic et al., 2008; Lawrence, 2000). On the other hand, acoustic departures from canonical speech can also manifest themselves in utterances produced by learners of a second language (hereafter L2) (i.e., *exogenous variation*). These L2 learners often speak with a noticeable non-native accent, which can in many cases be attributed to mismatches between the phonological inventory of the L2 speaker's first language (hereafter L1) and the phonemically distinct sounds in the target language (Flege, 1995; cf. Section 2.2.1). Nevertheless, after sufficient experience with the atypical pronunciations, language users have been shown to become increasingly adept at overcoming the initial interpretive difficulties posed by regionally and non-native accented speech (see Trude et al., 2013; Tzeng et al., 2021).

Both speaker-specific and group-level factors causing acoustic variability may thus force listeners to re-evaluate and update their long-term mental representations of particular speech sounds in light of what they perceive in the auditory input. Put differently, when listeners are repeatedly exposed to an acoustically distinct realization of a specific sound, this may drive them to temporarily change their phonological representation of that sound as tacitly stored in their minds. The process of adjusting one's representational space has frequently been referred to as the *adaptive plasticity hypothesis* in speech perception (e.g., Erb et al., 2013; Guediche et al., 2016). Seemingly oxymoronically, this implies that listeners must have highly stable linguistic representations of individual speech sounds so as to recognize them in the input, yet the boundaries of those speech categories are expected to be sufficiently flexible in order to allow listeners to effortlessly interpret peculiar pronunciations of a sound. This stability-flexibility trade-off has been of focal interest in many behavioral studies investigating how language users cope with the acoustic variability in accented speech, as this can shed light on how phonological representations of speech categories are stored in memory (see Idemaru & Holt, 2020). The large majority of these studies are concerned with different types of consonant contrasts (see Idemaru & Holt, 2020, on plosives; Chodroff & Wilson, 2020, and Kraljic & Samuel, 2005, on the sibilant fricatives /s/ and /ʃ/; Reinisch et al., 2014, on the nasal consonants /m/ and /n/; Scharenborg et al., 2011, on the liquids /r/ and /l/). How listeners can adjust their perceptual boundaries for vowels after exposure to L1 or L2 accented speech has, however, to date received significantly less attention (see, however, Faris et al., 2018; Liu & Holt, 2015; Llompart & Reinisch, 2019; Tyler et al., 2014; Weber et al., 2014).

According to the adaptive plasticity hypothesis, the human perceptual system can learn to fine-tune its performance in response to the unusual pronunciations of one particular speaker. What still needs to be addressed is how the adjustments made for one speaker can boost listeners' comprehension of a different speaker with a similar accent. Bradlow and Bent (2008) were among the first to show that perceptual adaptation, also called perceptual learning, induced by the acoustically distinct speech sounds of one non-native speaker can help listeners to understand the accented speech of another L2 speaker, whom listeners had not previously heard (see also Clark & Garrett, 2004; Xie et al., 2018). Notably, they concluded that generalization of perceptual learning across speakers only takes place if listeners have been acquainted with speech produced by multiple non-native speakers instead of a single L2 speaker. The findings of a recent large-scale replication did not align perfectly with those originally reported by Bradlow and Bent (2008). Specifically, Xie et al. (2021) provided evidence that exposure to multiple talkers with the same non-native accent does not facilitate cross-talker generalization of perceptual learning to a greater extent than would single-talker exposure. Although it is beyond question that many listeners benefit from high-variability input (e.g., Barriuso & Hayes-Harb, 2018; Bradlow et al., 1999; Sumner, 2011), Xie et al. suggest that successful generalization of perceptual adaptation is also strongly dependent on the acoustic proximity between the voices of the speakers (see also Alexander & Nygaard, 2019; Reinisch & Holt, 2014). That is, the relative benefit that listeners may derive from single- and multi-talker exposure is not only based on the quantity but also on the quality of the auditory input.

Building on speech perception research as discussed briefly above, the present experiment sets out to further flesh out how native listeners comprehend the accented speech of non-native speakers. Specifically, we will investigate whether, and by extension how, Dutch L1 listeners can adjust their perceptual system in response to the acoustic characteristics of a female Dutch L2 speaker's non-native accented vowel productions. Moreover, we also want to examine whether native speakers preserve knowledge about those atypically pronounced vowel categories, enabling them to interpret the accented speech of another, male-sounding L2 speaker, whose voice was digitally created from the female talker's voice.

To that end, the paper is organized as follows. A first section will zoom in on the current state of research in the field of language processing and perceptual learning, and provide the reader with two theoretical models of speech production and perception that will enable us to formulate our hypotheses (Section 2). These hypotheses and the two main research questions will be discussed in further detail in Section 3. The next chapters are subdivided with regard to the experimental paradigm that we adopted in the present study to answer the research questions, *viz.* a pre-test (Section 4), an auditory lexical decision task (Section 5) and a two-alternative phoneme categorization task (Section 6). Each of these last three macro-sections are arranged around a parallel micro-structure: (i) an overview of the experimental protocols and justification for particular methodological choices, (ii) a summary of the major findings, and (iii) a brief discussion of the results in view of the state of the art in perceptual learning research. All of these sections will be brought together in the general discussion, in which we will also address the limitations of the current experiment and suggestions for future research (Section 7). Key findings that have emerged throughout the paper will be recapitulated in the conclusion (Section 8).

2 Literature Review

This state-of-the-art section aims to provide the reader with a concise overview of how prior research on language processing and perceptual learning, both theoretically and empirically, has paved the way for the current experiment. In a first section, we will succinctly cover how the notion of perceptual learning has been described over time in order to arrive at our working definition (Section 2.1). We will then turn to two theoretical frameworks which attempt to capture the patterns underlying speech production and speech perception (Section 2.2). Next, we will shift our focus to one source of information that has been found to facilitate the identification and, by extension, the learning of atypically produced speech sounds, namely higher-level lexical knowledge (Section 2.3). To round off this literature review, we will consider how shifting the perceptual boundaries for particular speech sounds in response to a speaker's pronunciation variation could support future speech perception. Specifically, we will look at how listeners can avail themselves of their redeveloped categories to recognize novel words in the auditory input (Section 2.4.1), and also at how perceptual learning outcomes can be used to comprehend the accented speech of unfamiliar speakers (Section 2.4.2).

2.1 Perceptual Learning

From a very young age, the human perceptual system exhibits strong sensitivity to the acoustic-phonetic regularities in the incoming speech signal. During their first year of life, infants become increasingly attuned to the specific sound contrasts that appear in the spoken language of their parents or other caregivers, thereby gradually losing the ability to discriminate consonant and vowel pairs that are not distinctive in their L1 phonological inventory (Werker & Tees, 1984; see also Fort et al., 2017; Kuhl et al., 2006; Vihman, 2017; Werker, 2018). This type of phonetic learning is generally referred to as *perceptual narrowing* or *perceptual attunement*. One mechanism that has been shown to enable infants as young as one month of age to become used to the speech sounds of their first language is categorical speech perception (see Eimas et al., 1971). The phenomenon of categorical perception indicates that language users associate a wide range of phonetically and acoustically distinct sounds with only a relatively small, fixed set of language-specific speech categories (Bidelman et al., 2013; Kuhl et al., 1992). This does not imply, however, that when speakers have become accustomed to the phonemes of their first language, they will no longer be able to perceive differences between sounds falling outside those L1 speech categories. Although the /r/ and /l/ consonants are phonemically contrastive in English but not in Japanese, for instance, it has repeatedly been demonstrated that Japanese L1 speakers can become increasingly better at discriminating these two liquid consonants following a training session (see, among others, Aoyama et al., 2004; Bradlow et al., 1999; Iverson et al., 2016; Miyawaki et al., 1975). In other words, even when the process of perceptual attunement has been completed, language users will still be able to flexibly adapt to the acoustic regularities and deviations in the speech stream. In the literature, this process is generally referred to as *perceptual learning* (Norris et al., 2003) or *perceptual recalibration* (Bertelson et al., 2003).

At this point, it might be worthwhile to outline what is commonly understood by the terms perceptual learning or perceptual recalibration in the field of speech perception. Gibson and Gibson (1955) suggested in rather general terms that acoustic variability in the auditory speech input can lead listeners to grow "more sensitive to the variables of the stimulus array" (p. 40). Goldstone (1998) elaborated on this definition by arguing that perceptual learning should in fact be regarded as the formation of "relatively long-lasting changes to an organism's perceptual system that improve its ability to respond to its environment and are caused by this environment" (p. 586). Similar to infant language acquisition as noted above, Goldstone stresses the crucial role of the ambient language environment for successful speech perception. Gibson and Gibson's and Goldstone's perspectives can be integrated into the following working definition which will be used throughout this paper: perceptual learning can be defined as the temporary, or more sustainable, adjustments made by the human perceptual system to the perceptual boundaries between two or more language-specific speech sounds through repeated exposure to auditory stimuli that deviate noticeably from canonical pronunciations of those phonemes (see Guediche et al., 2014; Samuel & Kraljic, 2009; Scott, 2020; Weatherholtz & Jaeger, 2016). Evidently, perceptual learning needs to be regarded as a process, so the role of interpersonal differences in terms of how quickly and how efficiently listeners learn from the speech input may not be overlooked (see Golestani & Zatorre, 2009; Perrachione et al., 2011). Rather than further expanding on individual differences in perceptual learning (see, for instance, Banai & Lavie, 2021), we will direct our attention in the next paragraph to two highly influential conceptualizations of how language users produce and perceive speech.

2.2 *Models of Speech Production and Perception*

This subsection reviews two major theoretical frameworks on speech production and speech perception. First, we will address why particular consonants and vowels may sound slightly different when they are produced by non-native speakers compared to native speakers using Flege's (1995) *Speech Learning Model* (SLM). Secondly, the *Perceptual Assimilation Model* (PAM) as devised by Best (1995) will inform us on how listeners assimilate atypically pronounced speech segments to the sounds in the phonological system of their native language. By no means do we intend here to present an exhaustive overview of the different ways in which speech production and perception have been modeled. As mentioned in the introduction, these two models will ultimately enable us to formulate hypotheses on how native Dutch listeners understand the L2 accented speech of a non-native speaker (cf. Section 3).

2.2.1 Speech Learning Model (SLM)

In his Speech Learning Model (SLM), Flege (1995) seeks to conceptualize how second language learners' production of L2 consonants and vowels is mediated by their perception of those non-native speech sounds (for a recent review, see Flege & Bohn, 2021). Central to his theory is the assumption that learners of a second language will typically interpret L2 sounds through the 'grid' of their L1 phonological repertoire (Flege, 1995; see also Trubetzkoy, 1939). This principle of 'interlingual identification' suggests that there is a strong perceptual connection between L1 and L2 speech sounds. Flege claims that despite the phonetic dissimilarity between particular native and non-native sounds, some L2 learners will not discern a difference between such L1 and L2 phones, or pairs of sound contrasts. As a result of inaccurate perception, this group of language learners is expected to produce their L2 sounds with a detectable 'foreign' or non-native accent (Flege, 1995; cf. PAM in Section 2.2.2). It should be noted that some speakers may in fact be able to perceptually discriminate two non-native sounds, but they lack sufficient proficiency in that language to produce those speech categories with phonetically different properties. The inability to perceive differences between (pairs of) cross-linguistic sounds can, however, gradually dissolve as the learning process goes on, resulting in more adequate productions the non-native speech segments (see, among others, de Leeuw & Celata, 2019; Flege, 1995; Flege & Hammond, 1982; Zhang & Wang, 2007).

Fundamentally, the SLM thus predicts that the establishment of L2 speech categories and contrasts to a large extent depends on the interplay between the phonological system of the non-native speaker's first and second language. Flege (1988) posits a threefold distinction which tries to capture how L2 sounds may be perceived: (i) consonants and vowels in the L1 and L2 may be perceptually *identical*; (ii) they may differ slightly with regard to the specific acoustic properties of the phones, although they are sufficiently *similar* to be transcribed in with the same IPA symbol (e.g., acoustically, the fricative /s/ is not fully identical in Dutch and English, yet it has the same IPA symbol); (iii) there may also be particular L2 speech segments that cannot be correlated acoustically or articulatorily with any of the L1 sounds and are hence considered as *new* sounds, which are represented by different IPA symbols (Flege, 1988, pp. 227-228). By hypothesis, second language learners will usually acquire phonetic categories best when the acoustic-phonetic characteristics of the L2 sound are either indistinguishable from (i.e., *identical* categories), or differ noticeably from those of the closest L1 counterpart (i.e., *new* categories). The formation of L2 phonological representations may be inhibited by the perceptually induced mechanism of 'equivalence classification' (Flege, 1995, p. 239). This principle leads learners of a second language to perceive an L1 and L2 sound to be sufficiently *similar* to consider the two speech segments as instances of the same phonological category.

Note that we will use these types of non-native category acquisition and their relative chance of success to make predictions about the ability of the non-native speaker in the present study to produce phonetically distinct vowel sounds (cf. Section 3). Before we do so, we will briefly outline another model of speech perception in the next section, which can be used to predict how native Dutch listeners may perceive non-native accented sound contrasts.

2.2.2 Perceptual Assimilation Model (PAM)

The Perceptual Assimilation Model (PAM), as initially devised by Best (1995), is an attempt to formalize how language users systematically assimilate non-native speech segments to the phonological categories that are distinctive in the sound system of their native language. Although the original purpose of PAM was thus to examine how listeners map non-native speech segments onto the sound categories of their L1, we will exploit the assimilation patterns as described below to explore how native speakers of Dutch may perceive the non-native accented Dutch speech sounds of an L2 speaker. It should also be mentioned that listeners in PAM are regarded as "functional monolinguals, i.e., not actively learning or using an L2, and are linguistically naïve to the target language of the test stimuli" (Best & Tyler, 2007, p. 16). The specific emphasis on inexperienced listeners is of particular relevance for the present study, because participants in the current experiment will not have an advanced knowledge of the non-native speaker's L1.

In relation to the first cluster of assimilation patterns, L1 listeners are expected to deal with individual non-native speech sounds in three different ways (see Table 1, left-hand column). After exposure to a non-native phone, language users may be inclined to associate that speech segment with a closely related L1 phoneme. Specifically, such an L2 sound can be perceived as (i) "a good exemplar of that category", (ii) "an acceptable but not ideal exemplar of the category" or even (iii) "a notably deviant exemplar of the category" (Best, 1995, p. 194). A second possibility is that listeners identify the L2 segment as a speech sound that may be functional in some natural languages of the world, but they fail to assimilate the speech sound to a specific L1 category. Finally, the third assimilation pattern pertains to those sounds which are not considered to be part of the "universal phonetic domain" (Best, 1995, p. 186). These types of sounds are not regarded as belonging to human speech and can originate from different sources and actions, such as constrictions in the human vocal tract which are unnatural in phonological repertoires (e.g., coughing), humanly produced sounds (e.g., hand clapping) or non-human sounds stemming from human actions (e.g., slamming a door) (for other examples, see Best, 1995, p. 194).

Given that speech segments can share a place or manner of articulation, two non-native speech sounds can also be assimilated to L1 categories in a pairwise fashion due to their articulatory and phonetic similarity. Assimilation patterns of such contrasts depend by default on how the individual members of the contrast are assimilated (see Table 1, middle and right-hand column). Since participants will only have two vowel response options in the phoneme categorization task of the experiment (cf. Sections 3 and 6), we will limit ourselves here to the patterns of assimilation that are relevant to our experimental design (for further reading, see Best, 1995, pp. 193–199).[1] First, listeners may map each of the two non-native phones onto two contrastive L1 speech segments (*Two-Category Assimilation;* hereafter TC Type). Secondly, two perceptually less distinctive non-native segments can converge into a single L1 category, although listeners are still able to separate the standard or 'ideal' realizations of the phoneme from the more deviant renditions of the phone (*Category-Goodness Difference;* CG Type). A third possible assimilation pattern is that both non-native sounds are associated with the same native-language speech segment, while they are perceived as equally discrepant from canonical realizations of that particular L1 sound (*Single-Category Assimilation*; SC Type). If we analyze these three assimilatory processes in terms of how well listeners may detect a difference between L1 and L2 phonemes, it is expected that listeners will be able to discriminate the speech sounds in the TC Type, but discrimination performance may be somewhat less for the CG Type and even decline further for the SC Type. These possibilities will be further discussed in Section 3, where we will compare the vowel categories in the sound system of Dutch with those of the non-native speaker's L1.

Perceptual Assimilation Model (PAM)		
Non-native speech segments	**Non-native segmental contrasts**	
ASSIMILATION TYPE	ASSIMILATION TYPE	DISCRIMINATION PERFORMANCE
Assimilated to a native category	Two-category assimilation (TC Type)	excellent
	Category-Goodness Difference (CG Type)	moderate – very good
	Single-Category Assimilation (SC Type)	poor
	Uncategorized vs. Categorized (UC Type)	very good
Assimilated as uncategorizable speech sound	Both Uncategorizable (UU Type)	poor – very good
Not assimilated to speech (nonspeech sound)	Nonassimilable (NA Type)	good – very good

Figure 1: Perceptual assimilation patterns and discrimination performance for non-native speech segments and segmental contrasts based on Best (1995, pp. 194-195). Curly lines indicate that the assimilation type for a non-native speech contrast spans two different assimilation types for individual non-native sounds.

[1] For completeness, Figure 1 visualizes all six assimilation patterns for non-native speech contrasts.

2.3 Lexically-Guided Perceptual Learning

As emerged from the discussion of the speech production and perception models, how well listeners can discriminate two non-native speech sounds is strongly intertwined with the specific sound categories of their first language. Although discrimination might be poor to moderate at first, the human perceptual system can substantially fine-tune its performance after having accrued experience with the acoustic regularities in the auditory speech input. Clarke and Garrett (2004) demonstrated that as little as one minute of exposure to highly variable non-native accented speech suffices for listeners to become used to the acoustic features of that speaker's accent (see also Bradlow et al., 1999; Sumner, 2011). Perrachione et al. (2011) assert, however, that only listeners with a strong aptitude for perceptual discrimination will benefit from high-variability training, and that excessive exposure to irrelevant acoustic variation can even have a detrimental impact on listeners with weaker perceptual abilities (see also Antoniou & Wong, 2015, 2016; Fuhrmeister & Myers, 2017). While variability in the auditory speech input may thus be useful for many listeners, the fact that it does not facilitate speech perception for all listeners implies that there must be other perceptual mechanisms or 'tools' available (see Xie et al., 2021).

One source of information that listeners have repeatedly been found to leverage in order to disambiguate atypically produced speech sounds is prior lexical knowledge. [2] That is, language users can resolve the identity of a speech sound that has no straightforward phonological correspondence by drawing on their mental lexicon, hence the term *lexically-guided perceptual learning*. Norris et al. (2003) were the first to empirically investigate how higher-level lexical knowledge allows listeners to adjust their mental representations of individual speech sounds in response to perceptually ambiguous pronunciations in the speech input. In a lexical decision task, participants were presented with Dutch words ending in either /f/ (e.g., *aanhef* 'beginning') or /s/ (e.g., *hakmes* 'hatchet'), or Dutch nonwords with one of these fricatives in word-final position (e.g., *granklef* or *grankles*).[3] One group of listeners consistently heard ambiguous realizations of the fricative in /f/-final words, while the /s/-final words were presented in their canonical form (i.e., ambiguous [ˈaːn.hɛ?] but unambiguous [ˈhɑk.mɛs]), whereas the other group heard the exact opposite (i.e., unambiguous [ˈaːn.hɛf] but ambiguous [ˈhɑk.mɛ?]). In a subsequent phoneme categorization task, listeners were asked to identify which fricative they perceived on an [f]-[s] continuum. Participants

[2] Note that other contexts can also support perceptual learning, such as visual cues (e.g., Bertelson et al., 2003; McGurk & MacDonald, 1976; Reinisch & Mitterer, 2016; Vroomen et al., 2004), and other signal-based contexts, such as the covariation between specific acoustic cues (e.g., Clayards et al., 2008; Idemaru & Holt, 2014). Limitations of space preclude a more detailed discussion of these alternative contexts.

[3] Examples were taken from the original study (see Norris et al., 2003, p. 234).

in the /f/-ambiguous condition classified the ambiguous speech sounds on the continuum significantly more as /f/, whereas participants in the /s/-ambiguous condition gave more /s/-responses. Test performance of listeners who were presented with the set of nonwords with either of the two word-final ambiguous fricatives oscillated between that of the participants who heard ambiguous realizations of real /f/-final and /s/-final words. Based on these findings, Norris et al. (2003) conclude that the lexicality of the tokens was the driving force for listeners to modify their long-term phonological representation of the fricative consonant that sounded ambiguous in the lexical decision part of the experiment.

Robustness of exposure-driven perceptual learning as observed by Norris et al. (2003) has been confirmed in many replications and follow-up studies, also when the lexical items consisted of different ambiguous consonant and vowel contrasts (e.g., fricatives: Kraljic & Samuel, 2005; plosives: Ullas et al., 2020; vowels: Kim et al., 2020). Counter to McAuliffe & Babel's (2016) study, Drouin and Theodore (2018) provided evidence that the strength of perceptual learning does not depend on the nature of the short-term training that participants receive during the exposure phase. That is, ambiguous speech sounds in lexically disambiguating contexts were learned equally well in perception-focused tasks (e.g., counting the number of syllables per word, as in Samuel, 2016) as in comprehension-focused tasks (e.g., lexical or semantic decision tasks as in Norris et al., 2003, and Zhang & Samuel, 2014, respectively). Although Jesse and McQueen (2011) further confirmed that perceptual learning is robust, they added that there may be one constraint on this type of learning. Specifically, they found that the perceptual boundaries of an ambiguous sound are not adjusted when that speech sound occurs in word-initial position. Rather than acting as a constraint, this very finding, in fact, once again corroborates the idea that lexical information is needed to disambiguate the to-be-adapted speech sound.

It should also be noted that when the ambiguous speech segments are not embedded in a lexical context, exposure to such ambiguous sounds will not automatically lead a listener to adjust their mental representation of that phoneme. Eimas and Corbit (1973), for instance, demonstrated that when listeners were repeatedly exposed to one endpoint of a phonetic spectrum (e.g., /f/ in Norris et al., 2003), they are more likely to choose the other continuum endpoint in a two-alternative categorization task (i.e., /s/). This phenomenon is referred to as *selective adaptation* (see also Kraljic & Samuel, 2005; Samuel, 1986). Aftereffects of selective adaptation are diametrically opposed to those of perceptual learning. The former expects that continuously hearing an ambiguous production of one member of a contrast will maximize the likelihood that listeners will choose the opposite compound of the speech contrast. Eimas and Corbit (1973), and Samuel (1986) ascribe this negative aftereffect to the fatiguing of "linguistic feature detectors" in the brain by repetitively presenting the same speech category. On a lexically-guided perceptual learning account, conversely, listeners are expected to learn that the ambiguous sound represents an atypical production of a particular phonological category. As the surrounding lexical context could drive these listeners to reshape their mental representation of the intended speech category, it is expected that such retuning will cause them to identify the ambiguous sound more often as the intended phoneme.

The question then arises whether the findings reported by Norris et al. (2003) can be attributed to a fatiguing of "linguistic feature detectors", following selective speech adaptation, or whether there has been a true change in phonetic category representation. Clarke-Davidson et al. (2008) tried to answer this question by adopting the Signal Detection Theory (SDT) in combination with the traditional experimental paradigms for perceptual learning (i.e., a lexical decision task and a phoneme categorization task). Applied to the domain of speech perception, the SDT assumes that, under conditions of uncertainty, listeners' discrimination performance of perceptually ambiguous stimuli depends on two parameters (see Macmillan, 2002; Macmillan & Creelman, 2005). First, it is related to their individual ability to perceive a difference between the two speech sounds. Secondly, listeners can also develop a response bias over the course of the exposure phase. According to Clarke-Davidson and colleagues (2008), the results of their discrimination tasks and the signal detection analysis mutually point towards a true perceptual change of the ambiguous speech sound instead of a response bias. Corroborating evidence against the development of decision biases during exposure was also obtained from studies using eye-tracking (e.g., Mitterer & Reinisch, 2013) or examining neural activity with fMRI scans (e.g., Myers & Mesite, 2014). The effects of lexically-guided perceptual learning, as shown by Norris et al. (2003), can thus be regarded as genuinely perceptual and not simply a post-perceptual decision.

2.4 Generalization of Perceptual Learning

So far, this state-of-the-art section has largely addressed the idea that lexical contexts can robustly guide listeners to interpret atypical pronunciations of a particular speech sound. What still needs to be examined in more detail is how perceptual learning in general, and lexically-guided perceptual learning in particular, generalizes to novel situations. That is, once the phonological representations of speech categories have been updated for one speaker, how do listeners carry over knowledge about the acoustically ambiguous speech sounds to process lexical items and speakers that they have not encountered before. In what follows, we will discuss the generalization of perceptual learning across these two situations, *viz.* novel stimulus materials (Section 2.4.1) and novel speakers (Section 2.4.2). This overview is essential for understanding how the present study ties in with previous research on perceptual adaptation, as will be delineated in Section 3.

2.4.1 Generalization Across Stimulus Materials

As outlined in Section 2.3 above, Norris et al.'s (2003) seminal study showed that lexical contexts can help listeners understand atypical realizations of the /f/ and /s/ fricatives and that such lexically biasing contexts can also drive these listeners to change their perceptual boundaries of those speech categories. According to Norris et al. (2003), one of the main reasons why listeners retune their category boundaries is because it enables them to reconstruct which word the speaker may have intended to produce. In a follow-up study, McQueen et al. (2006) investigated whether perceptual learning takes place at a more abstract level than the lexicon itself. More precisely, they wanted to examine if listeners only adjust their phoneme boundaries in the specific lexical items that they perceive in the exposure phase, or whether perceptual adaptation also affects the prelexical representation of those speech sounds. The same familiarization task was used as in Norris et al. (2003), but rather than presenting participants only with sounds along an [f]-[s] continuum in the phoneme categorization task, the continuum steps were now embedded in minimal word pairs (e.g., [bri?], from *brief* 'letter' and *bries* 'breeze'). Similar to the original study, McQueen et al. (2006) found that listeners reported hearing either the /f/- or /s/-final word of minimal pair significantly more depending on which fricative was perceptually ambiguous in the test items during exposure.

McQueen et al. (2006) regard their findings as evidence that perceptual learning is more likely to change category boundaries at a more abstract prelexical or sublexical level rather than leading to adaptations to the post-lexical representations of speech sounds (for a similar view, see Hervais-Adelman et al., 2008; Mitterer et al., 2011; Nelson & Durvasula, 2021; Scharenborg et al., 2005). Note that since there is still no consensus on which specific linguistic units listeners use to recognize words in spoken language, there is also no agreement at which linguistic level perceptual adaptation occurs. Kraljic and Samuel (2006) claim that phonetic features (e.g., presence or absence of voice for consonants) pave the way for successful language processing, while Bowers et al. (2016) contend that phonemes are the cornerstones of speech perception. Yet another level is suggested by Mitterer et al. (2013), who posit that listeners are more likely to use context-dependent allophones than context-independent phonemes to identify words in the speech signal. This assertion is based on their finding that participants who heard an ambiguous realization of one position-conditioned allophone failed to generalize the perceptual learning effect to another allophone of the same phoneme. That is, when listeners were exposed to an ambiguous phone in between the word-final dark [ɫ] and the word-final alveolar approximant [ɹ] in Dutch words ending in either /l/ or /r/ (e.g., *nagel* 'nail' and *emmer* 'bucket'), they did not transfer the adapted speech sounds to novel test items in which /l/ or /r/ consonants occurred in word-initial position (e.g., the clear [l] and trill [r] in *lengte* 'length' and *reistas* 'traveling bag') (see also Mitterer et al., 2018).[4] As it falls beyond the scope of this study to determine which specific linguistic units are perceptually adjusted, and because some scholars have argued that the perceptual system will "latch onto virtually any systematic pattern in the input" (Samuel, 2020, p. 11), we will not adopt a particular position in this debate.

While the ambiguous speech segments in the vast majority of the lexically-guided perceptual learning studies above were consonants, research on how listeners adjust their vowel categories is relatively scant. Maye et al. (2008), however, examined to what extent listeners can learn atypical realizations of the English front vowels. Specifically, participants listened to a story in which front vowels were systematically shifted to represent an unfamiliar regionally accented variety of English that deviates from the standard English pronunciation (i.e., the high front vowel /i/ was lowered to [ɪ], /ɪ/ to [ɛ], /ɛ/ to [æ], and /æ/ to [a], so that for example *wicked* /ˈwɪk.ɪd/ sounded like *wecked* [ˈwɛk.ɪd]). Phonologically, words with the shifted vowel sounds do not correspond to real English words. Nevertheless, participants classified the accented words significantly more as real words compared to listeners who did not hear the vowel shifts during storytelling. Even when participants were presented with accented words that they had not encountered during storytelling, the word endorsement rates of the group of listeners who heard the vowel shifts during exposure were substantially higher than those of the participants who heard a speaker with a standard-sounding American accent. Again, this indicates that listeners used the lexically disambiguating contexts in the exposure phase to adjust their phonological boundaries to comprehend the speaker perceived during storytelling, which eventually also enabled them to comprehend novel words.

[4] Examples were taken from Mitterer et al. (2018, p. 91).

2.4.2 Generalization Across Speakers

Aside from approaching perceptual learning from the angle of generalization to novel lexical contexts, there are also several studies that have investigated whether the adjusted phoneme boundaries learned from the accented speech input of a single L1 speaker remain talker-specific, or whether the outcomes of perceptual learning can be transferred to understand the accent of other native speakers. Elaborating on Maye et al.'s (2008) study, Weatherholtz (2015) examined if exposure-driven perceptual learning of unfamiliar vowel shifts can also generalize to novel talkers. Contrary to the set of front vowels in Maye et al. (2008), participants in his study listened to a story in which all back vowels were lowered in a chain shift (i.e., the /u/ in *bouquet* was shifted to sound like [bʊˈkeɪ], /ʊ/ to [o] and /o/ to [ɑ]). Following the storytelling, participants performed an auditory lexical decision task as well as a word identification task. Converging evidence from both tasks showed that familiarization to a single talker was sufficient for listeners to generalize information about that speaker's idiosyncratic pronunciations not only to new words but also to novel talkers. Specifically, there was no difference in test performance for the novel lexical items when those tokens were produced by an unfamiliar female or male speaker's talker instead of the speaker to which they had been exposed during storytelling. Weatherholtz (2015) takes these findings as evidence that listeners can adjust their vowel spaces after a short exposure phase, which allows them to interpret accented speech independently of the talker they perceive.

Some other studies, however, have run up against a few constraints when it comes to cross-talker generalization. In Eisner & McQueen's (2005) extension of Norris et al. (2003), lexically-guided perceptual learning of the ambiguous fricative in between /f/ and /s/ was not successful when the test stimuli – steps along an [ɛf]-[ɛs] continuum – presented in the phoneme categorization task were produced by novel speakers. Only in two conditions were listeners found to effectively modify their perceptual boundaries of the fricatives: when the vowel sound in the [ɛf] or [ɛs] syllable was produced by the novel test talker and the ambiguous fricative by the exposure talker, or when the novel talker's realization of the critical consonant contrast had been excised from the test items of the categorization task and spliced into the exposure materials produced by the exposure talker (Eisner & McQueen, 2005). The absence of cross-talker generalization might in some conditions be explained by the lack of acoustic similarity between the different speakers. Reinisch and Holt (2014) established that learning effects only emerge when the phonetic-acoustic properties of the exposure talker's and the novel talker's phonological categories occupy a similar perceptual space (see also Xie et al., 2021; Xie & Myers, 2017). Moreover, the spectral characteristics of the phonemes under consideration have also been found to influence, at least partially, whether or not cross-talker generalization of the retuned speech sounds is blocked (Kraljic & Samuel, 2006; Reinisch & Holt, 2014). Kraljic and Samuel (2007), for instance, confirmed that perceptual learning of ambiguous stop consonants (e.g., /t/ vs. /d/) is speaker-independent whereas it remains speaker-specific for some ambiguous fricative contrasts (e.g., /s/ vs. /ʃ/). They suggest that the spectral cues of the consonants in the second contrast, which only differ in terms of spectral qualities and place of articulation, convey more information about the speaker producing them than would the temporal cues of stop consonants, which only differ in terms of duration and voicing (Kraljic & Samuel, 2007). Since fricative contrasts provide listeners with both segmental cues and speaker information, perceptual adaptation of these sounds is not expected to transfer to unfamiliar talkers.

Whereas the studies mentioned above have demonstrated that listeners can recalibrate their category boundaries after exposure to the pronunciation variation of a speaker with whom listeners share a first language, generalization to novel talkers has also been attested when listeners hear the non-native accented speech of an L2 speaker. Specifically, these atypically pronounced speech sounds could be considered an aspect of the non-native speaker's interlanguage, in which patterns and rules of the L1 are transferred into the linguistic system of the target language during the process of second language acquisition (Selinker, 1972; Selinker et al., 1975; for a set of definitions, see Tarone, 2018, pp. 2-3). A seminal study on perceptual learning of such exogenous variation was carried out by Bradlow and Bent (2008). They showed that generalization to a novel test talker only occurred when participants had been familiarized with multiple non-native speakers during the exposure phase as opposed to a single speaker (see also Sidaras et al., 2009). Likewise, Baese-Berk et al. (2013) found that short-term exposure to speakers who are from a variety of different language backgrounds will eventually result in accent-independent learning, which means that adaptation to the acoustic characteristics of one particular accent can help listeners interpret the speech of another L2 speaker who has a similar non-native accent. These findings on the effectiveness of input variability run to some extent counter to what Xie et al. (2021) observed in their recently conducted large-scale replication of Bradlow and Bent's (2008) original study. Crucially, they did not find substantially different learning outcomes between participants who were familiarized with the L2 accented speech of one talker compared to that of multiple speakers (cf. Huyck et al., 2017). This suggests that successful cross-talker generalization is not solely premised on how variable the acoustic properties are that listeners perceive during the familiarization phase of the experiment. In line with Reinisch and Holt (2014), Xie et al. (2021) argue that the degree of perceptual similarity between speakers' voices is also fundamental for the generalization of retuning across speakers.

3 The Present Study

In the present study, we will further explore the complex dynamics of how native listeners interpret non-native accented speech. Specifically, we will investigate whether Dutch L1 listeners adjust their mental representations of Dutch vowels after short-term experience with what could be considered non-native accented pronunciations of these vowel categories. Retuning their perceptual boundaries would then enable them to comprehend future utterances and novel words with vowel sounds that are produced in a similar way. Additionally, we also want to examine whether effectively adapting to one L2 speaker's non-native accented vowel productions can guide listeners to interpret acoustically similar vowels when they are realized by another, unfamiliar L2 speaker. Against this background, we can formulate the following two research questions (RQs):

RQ1: To what extent do native listeners of Dutch accommodate L2 accented pronunciations of Dutch vowels when they are produced by a non-native speaker, assuming that knowledge about this speaker's atypical pronunciation could be leveraged in future encounters with that speaker to interpret such atypically produced vowel sounds in novel words?

RQ2: When exposure to perceptually ambiguous vowels has indeed been found to facilitate word recognition, would listeners, by extension, generalize knowledge about one non-native speaker's accented vowels to comprehend the speech of another, unfamiliar L2 speaker, as these pronunciations could be typical of the Dutch interlanguage of those L2 speakers?

In an attempt to answer these two research questions, we familiarized Dutch L1 speakers with the non-native accented Dutch of a female native speaker of Italian. The speech contrast used in the current experiment is that between the Dutch front vowels /ɪ/ and /i/, which are characterized by different spectral qualities, but do not differ substantially in terms of duration (Adank et al., 2004; Simon et al., 2015). An Italian speaker of Dutch was selected because the phonological system of Italian lacks a similar vowel contrast; it only has the long front vowel /i/ (Krämer, 2009; Rogers & d'Arcangeli, 2004). While native speakers of Italian may perceive acoustic differences between the two Dutch front vowels, the absence of such a vowel contrast in their L1 may initially impose a burden on these speakers to produce the two vowels with spectrally distinct features (see Duguid, 2001). Empirical studies have indeed demonstrated that Italian L1 speakers encounter difficulties in discriminating the two members of the English /ɪ/-/i/ contrast, which is relatively similar to the Dutch one, and as a result produce both front vowels with more /i/-like properties (e.g., Flege et al., 1999; Flege & MacKay, 2004). According to the Speech Learning Model (Flege, 1995, 1988), initial difficulties and inaccuracies in vowel production will be remedied the more proficient the L2 learner becomes, especially with regard to L2 sounds for which there is no close equivalent in the L1. Since the Italian speaker in the present study is highly proficient in Dutch, we assume that she will be able to produce the /ɪ/- and /i/-vowels with spectrally different qualities (cf. Section 4.1.2).

To determine whether perceptual learning develops, and by extension generalizes to a novel L2 speaker, the traditional experimental design for assessing lexically-guided perceptual learning will be used, *viz.* (i) an auditory lexical decision task followed by (ii) a phoneme categorization task.[5] In the first task, participants will be familiarized with forty target words with either canonical productions of the /ɪ/-vowel but ambiguous realizations of the /i/-vowel, or vice versa. That is, *vlinder* ('butterfly') will be realized as [ˈvlɪn.dər], whereas the front vowel in *diefstal* ('theft') will be ambiguous (i.e., [ˈd?f.stɑl]). Ambiguous vowel sounds were created and acoustically manipulated by interpolating between prototypical /ɪ/- and /i/-productions (cf. Section 4.1.2), which listeners may perceive as non-native accented pronunciations that are typical of Italian accented Dutch. The second task will be used to verify if the Dutch L1 listeners interpreted the ambiguous sounds as atypical pronunciations of the Dutch front vowels. If so, this would drive them to identify the ambiguous vowels in minimal /ɪ/-/i/ words predominantly as either /ɪ/ or /i/, depending on whether /ɪ/-words or /i/-words consisted of the ambiguous vowels in the first task. If generalization of perceptual adaptation can occur across speakers, the ambiguous vowels will be identified in a similar way in the minimal words produced by a different L2 speaker. Based on this brief project outline and the literature review in the previous section, we can put forward the following hypotheses (Hs):

H1: Whether perceptual learning may develop over the course of the exposure phase will strongly depend on the Dutch L1 listeners' ability to discriminate the canonical and ambiguous front vowel. On a Perceptual Assimilation Model account (Best, 1995), exposure to the two types of front vowels is expected to trigger one of the following two assimilation patterns. That is, listeners may map the canonical and the ambiguous vowels onto two separate L1 categories (i.e., /ɪ/ and /i/). Alternatively, if listeners do not perceive the acoustic differences between the two types of sounds to be sufficiently large, or if they consider the canonical and ambiguous vowels as allophonic variants of the same phoneme instead of

[5] Note that some studies have also used transcription tasks to test perceptual learning (e.g., Bradlow & Bent, 2008; Sidaras et al., 2009; Xie et al., 2021). We preferred a lexical decision task in combination with a phoneme categorization task because this facilitates comparing the results obtained in the present study with those reported in the majority of previous studies on lexically-guided perceptual learning (e.g., Norris et al., 2003).

representing two individual phonemes, vowel productions may be clustered into a single L1 category (Best, 1995; cf. Section 2.2.2).[6] We hypothesize that participants in the current experiment will be able to distinguish the two types of vowel sounds (i.e., *Two-Category Assimilation*), because listeners are susceptible to subtle differences in spectral properties, which might lead them to perceive two sounds as belonging to different L1 categories (e.g., Boersma & Chládková, 2011; Escudero et al., 2009; Weatherholtz, 2015). Moreover, both vowel productions will be embedded in real Dutch words, which can help listeners disambiguate the acoustically ambiguous sounds. Since access to supplementary sources of information has repeatedly been shown to drive listeners to update their perceptual representations of speech sounds (cf. Section 2.3), we expect perceptual learning to take place, inducing listeners to identify the ambiguous vowel in novel lexical contexts in accordance with which vowel of the contrast was ambiguous during exposure.

H2: When perceptual learning of the female speaker's ambiguous vowels does indeed develop over the course of the lexical decision task, we can advance the following set of assumptions with regard to the generalization of learning across speakers. If participants perform the phoneme categorization task in a similar way when the minimal words are presented with the voice of the female exposure speaker or with that of an unfamiliar, male-sounding speaker, this could be taken as evidence that the Dutch L1 listeners successfully generalized knowledge about the female speaker's accent to interpret the non-native accented speech of another talker. At a more abstract level, generalization of learning would suggest that listeners adjusted their category boundaries independently of the exposure talker and thus may be perceived the atypical pronunciations as a typical feature of the Dutch interlanguage of Italian speakers. In the absence of similar categorization responses across the two speaker conditions (i.e., female voice vs. male-sounding voice), this would indicate that perceptual learning remained specific to the talker with whom the participants were familiarized during the lexical decision task and for whom they had adjusted their perceptual system in the first place. We assume that cross-talker generalization may occur, as previous studies have already shown that listeners can transfer abstract, context-independent learning to other speakers with a similar accent (see Bradlow & Bent, 2008; Clarke & Garrett, 2004; Xie et al.,

[6] Since we did not ask participants to motivate their responses in the phoneme categorization task, we will be unable to ascertain the reasoning behind mapping the canonical and ambiguous vowel onto the same L1 category. That is, categorization responses will not allow us to determine whether both vowel types were perceived as equally discrepant from the L1 sound (i.e., *Single-Category Assimilation*) or whether one vowel sound was perceived as a more acceptable variant of the L1 category than the other (i.e., *Category-Goodness Difference*).

2018). Note that a novel speaker was simulated by digitally generating a male-sounding voice from the acoustic properties of the female speaker's voice (cf. Section 4.1.2). This manipulation controlled for identical durational cues of the stimuli, but the spectral differences and change in pitch lead listeners to perceive a different, male-like voice. Importantly, this modification also ensures that both voices are sampled across a relatively similar perceptual space, which has been shown a crucial determinant for successful cross-talker generalization of perceptual learning (e.g., Reinisch & Holt, 2014; Xie & Myers, 2017).

Examining whether adapting to non-native accented speech improves native speakers' comprehension of novel lexical items and novel speakers is not only of theoretical interest but also of social relevance. On the one hand, analyzing whether listeners can reshape their perceptual boundaries of individual speech categories through exposure to non-native accented speech could further expand our understanding of how phonological representations are entrenched in memory. As such, this could provide us with a more in-depth insight into how listeners can seemingly effortlessly cope with the extreme acoustic variability of spoken language. On the other hand, research on language users' perception of non-native accented speech is also of social significance. Recall that mismatches between the phonological systems of a speaker's native and non-native language can cause these speakers to talk with a noticeable accent in their second language (cf. Section 2.2.1). Although people's language background is generally acknowledged as being part of their social identity (Jaspal, 2009; Norton, 2010; Tabouret-Keller, 1997), native speakers can sometimes hold negative attitudes towards speakers with a strong non-native accent. This has often been found to result in lower rates of employability success (e.g., Carlson & McHenry, 2006; Deprez-Sims & Morris, 2010) or, in extreme cases, lead to acts of discrimination (e.g., Fuertes et al., 2012; Gluszek & Dovidio, 2010). Understanding how native speakers process and comprehend L2 accented speech is therefore an important first step in discovering how negative attitudes towards non-native accents may develop.

In what follows, all of the assumptions discussed above will be tested in order to formulate an answer to the research questions of this study. We will first sketch how the ambiguous vowel sounds were created and how we ensured that those sounds were truly perceived as ambiguous by native speakers of Dutch (Section 4). Next, we will delineate the experimental design and report the results obtained in the lexical decision task (Section 5) and the phoneme categorization task (Section 6). The findings of the different components in the experiment will eventually be discussed in light of the hypotheses and the state of the art in the general discussion section (Section 7).

4 Pre-Test

In this section, we will discuss how lexical items with an ambiguous front vowel in the stressed syllable were created to use as stimuli for the auditory lexical decision task (Section 5) and the phoneme categorization task (Section 6) in the main experiment. Specifically, we wanted to determine when Dutch L1 listeners experience difficulties to classify an allegedly L2 accented vowel sound as either /ɪ/ or /i/. To that end, we constructed a vowel continuum that stretches between two prototypical realizations of the critical front vowels. As the perceptual boundaries of language-specific phonemes can vary between listeners and simply selecting the midway step on each [ɪ]-[i] continuum would fail to take into account potential individual differences (see Samuel, 2020), we administered a pre-test to mitigate potential selection biases. The results of this test will ultimately inform us on where the most ambiguous regions on the generated vowel continua are located.

4.1 Method

4.1.1 Participants

Seven female and three male native speakers of Belgian Dutch (mean age = 21.3 years, *SD* = 0.7 years) with normal hearing participated in the pre-test. Participants were recruited from the student population in Ghent, Belgium, through social media posts and in-class announcements. Nine of them were born and raised in the province of West-Flanders, whereas East-Flanders was the province of birth and residence of the remaining one participant. Before the start of the experiment, participants were informed about the purposes of the study and written informed consent was obtained from all ten students (cf. Appendix 1). Although participants did not receive course credit and were not compensated for their time, they were invited to register for a giveaway of five gift certificates worth twenty euros each. For a participant overview, see Appendix 2.

These ten participants were selected as such that the variation in linguistic and regional background was limited to ensure that their test performance could be reliably compared and analyzed. First, participants were not allowed to have an advanced knowledge of Italian or Spanish, or be enrolled in a program that specializes or offers elective courses in either or both of these two languages, in order to avoid influence from the Italian or Spanish vowel systems (cf. Section 3). Although French is also a Romance language sharing some phonological features with Italian and Spanish, it is impossible to rule out that listeners have a basic to more advanced command of French, given that French is one of the official languages in Belgium and it is instructed as a second language from primary school onwards. Secondly, only participants from the provinces of West-Flanders or

East-Flanders were eligible for participation. This second criterion was put in place due to regional variation in the pronunciation of the critical vowel contrast in non-standard Dutch in Flanders. Production studies have demonstrated that speakers from the Brabantine area, which mainly encompasses the provinces of Antwerp and Flemish-Brabant, do not always produce the front vowels /ɪ/ and /i/ with sufficiently different formant frequencies so as to create a qualitative sound contrast (see Simon et al., 2015; Verhoeven & Van Bael, 2002). Since participants from this region might not discern the difference between the canonical and the ambiguous vowels as adequately as speakers from other regions and hence skew the overall data, only people from West-Flanders and East-Flanders were invited to participate, because speakers of these two regional varieties of Dutch do produce spectrally distinct /ɪ/- and /i/-vowels.

4.1.2 Stimulus Materials

The set of stimuli in the pre-test consisted of forty Dutch critical words and five minimal word pairs, which will also be used in the lexical decision task (Section 5) and the phoneme categorization task (Section 6) of the main experiment, respectively. There were two language-internal criteria for selecting test items. First, all target words were monosyllabic and disyllabic content words, of which the nucleus of the stressed syllable was either the near-close /ɪ/-vowel (e.g., *vlinder* 'butterfly', /ˈvlɪn.dər/) or the close /i/-vowel (e.g., *dienst* 'service', /dinst/). Importantly, the critical words were selected in such a way that changing the target vowel by the other member of the contrast (/ɪ/ or /i/) does not yield an existing Dutch word (e.g., *priester* 'priest', /ˈpris.tər/ vs. */ˈprɪs.tər/).[7] Secondly, a position-specific criterion was added: /i/-vowels could not be immediately followed by a syllable-final /r/ (e.g., *bier* 'beer', [biːr]). Phonetically, vowel sounds that precede an /r/ at the end of a syllable tend to be lengthened (hence [iː] rather than [i] in the phonetic transcription of *bier*). To prevent these words from being perceptually more salient than other target words and to keep the durational variability between the critical vowels as small as possible, no target words with such a vowel-consonant combination were included in the stimulus list.

[7] Asterisks (*) refer to words and other constellations that are ungrammatical according to the prescriptive grammar of the language at issue. Similarly, phonological transcriptions which are preceded by an asterisk refer to combinations of sounds that do not correspond to real Dutch words.

Thirdly, only frequent words were included in the current list, as the degree of phonetic detail that listeners retain for particular words has been found to correlate with how familiar they are with those words (see White et al., 2013). A common approach to estimate word familiarity is to determine the 'frequency per million words' (*fpmw*) of each test item based on extensive corpora. Frequencies for the critical words in the present study were obtained from the SUBTLEX-NL database (Keuleers et al., 2010), which is a digitalized repository containing about 44 million Dutch words borrowed from subtitles of television programs and films. However, we decided to calculate a standardized Zipf value based on the proportional *fpmw* of each item, as the accuracy of the interpretation of word-specific *fpmws* has been shown to be highly contingent on the size of the corpus (van Heuven et al., 2014). Fundamentally, the Zipf scale is a logarithmic scale with values ranging from 1, which are low-frequency words, over 6, which are high-frequency lexical words, to 7, which are highly frequent, but often semantically neutral function words such as pronouns and copulas. All target words in this study have an average word frequency in Zipf values between 3.5 and 5.5 (M = 4.32, SD = 0.49), which means that test items are moderate-frequency and high-frequency words. Similarly, only minimal /ɪ/-/i/ word pairs with relatively high Zipf values were sampled, which are the following: *vis-vies* ('fish'-'dirty', Zipf values 4.71 and 4.30, respectively); *bid-bied* ('pray'-'bid', 4.29 and 4.34); *wit-wiet* ('white'-'weed', 4.52 and 4.18); *kist-kiest* ('case/box', 'chooses', 4.44 and 4.23); and *lig-lieg* ('lie', 'lie/deceive', 4.39 and 4.49).

All stimuli were recorded by a 45-year-old female native speaker of Italian, who was born in Naples, Campania, and relocated permanently to East-Flanders in 2006. Her command of Dutch corresponds to a C1 level of language proficiency as determined by the Common European Framework of Reference for Languages, of which she obtained a certificate after formal instruction at the Language Centre of Ghent University. Despite her advanced proficiency level in Dutch and daily interactions with native and non-native Dutch speakers, she has a clearly noticeable Italian accent when speaking Dutch (e.g., *gezegd* 'said' was frequently realized as [ge.ˈzext] instead of [ɣe.ˈzext], as the voiced velar fricative /ɣ/ is absent in Italian). She gave written informed consent to participate in this study, after which she was provided with written and oral instructions about the recording procedure (see Appendix 3). Stimulus materials were recorded twice with a neutral intonation contour using a *Marantz* audio recorder and took place in a room with as few external noise factors as possible. Any remaining noise was filtered out using *Audacity* (Audacity Team, 2018). The overall root-mean-square intensity of the stimuli recordings was scaled to 70 dB. Afterwards, the best of the two recordings of each target word was selected.

Acoustical analyses were performed using *Praat* software (Boersma & Weeninck, 2019) to assess how the critical vowels were produced by the Italian speaker. A number of production studies focusing on the vowel system of Southern Standard Dutch have recursively observed that /ɪ/ and /i/ only differ substantially in terms of spectral qualities, but not in terms of vowel duration (e.g., Adank et al., 2004; Simon et al., 2015). However, we expected that this observation might not hold completely for the L2 speaker's vowel productions, as there is no /ɪ/-/i/ contrast in Italian (Flege et al., 1999; Krämer, 2009; Rogers & d'Arcangeli, 2004). Since all critical words were recorded with both the canonical and the opposite vowel of the contrast to facilitate the creation of the ambiguous stimuli (e.g. [ɣɪts] and [ɣits] for *gids* 'guide'; see below), we can examine the formant frequencies and the duration of the Dutch L2 speaker's vowel productions, which are summarized in Table 1. In /ɪ/-words, the canonical vowel was on average half the duration of the non-canonical front vowel (83 ms vs. 150 ms), whereas in /i/-words the difference in canonical (114ms) and non-canonical vowel duration (75 ms) was approximately 40 ms. Moreover, the values of the first three formants at the vowel midpoint were similar across the two conditions, *viz.* F1 was consistently higher for [ɪ]-vowels as opposed to [i]-vowels, whereas F2 as well as F3 frequencies were on aggregate lower for [ɪ] compared to [i] (see Table 1 for all mean values and corresponding statistical analyses of the female Italian speaker's vowel productions). In short, these analyses demonstrate that the Dutch L2 speaker produced both front vowels with significantly different spectral properties, but contrary to standard Dutch, these vowels were also produced with a clear difference in duration.

Table 1: Overview of spectral and durational characteristics of the non-native speaker's vowel production by the two types of critical words.

		Canonical		Non-canonical		Paired Samples Test [8]
		M	*SD*	*M*	*SD*	(***: $p < 0.001$)
/ɪ/-words	F1 *(Hz)*	417	*37*	278	*40*	$t(19) = 10.299$ ***
	F2 *(Hz)*	1980	*187*	2318	*82*	$V = 2, n = 20$ ***
	F3 *(Hz)*	2559	*183*	3282	*145*	$V = 0, n = 20$ ***
	Duration *(ms)*	83	*22*	150	*45*	$t(19) = -6.191$ ***
/i/-words	F1 *(Hz)*	303	*53*	461	*38*	$V = 1, n = 20$ ***
	F2 *(Hz)*	2363	*71*	1848	*114*	$t(19) = 25.754$ ***
	F3 *(Hz)*	3238	*175*	2569	*101*	$V = 210, n = 20$ ***
	Duration *(ms)*	114	*48*	75	*18*	$t(19) = 3.930$ ***

[8] Differences in formant frequencies and vowel durations of the canonical and non-canonical vowels were compared using a *two-sample t-test*. Although vowel-specific formant frequencies could generally be assumed to follow a Gaussian distribution, normality of the data distribution could not be guaranteed in all cases due to the small number of tokens. To compare those values, the non-parametric *Wilcoxon signed-rank test* was used instead.

In order to discover what listeners generally experience as the perceptual boundaries of the /ɪ/- and /i/-vowel, and by extension the transition from one front vowel to the other, we constructed an 11-step continuum for all forty critical words and the five minimal word pairs. This was done through auditory morphing of the two natural recordings of the target words with /ɪ/ and /i/. That is, the canonical and non-canonical realization of the word *gids* 'guide' ([ɣɪts] and [ɣits], respectively) were morphed along 11 steps, of which the first step is 0% of the recording with the /i/-vowel and the eleventh step is 100% with the /i/-vowel (i.e., in incremental steps of 10%). Continua were generated using the STRAIGHT algorithm (Kawahara et al., 1999). This algorithm is based on the source-filter model, as it decomposes the speech input into (i) a voice source, (ii) a noise source and (iii) a pitch-adaptive spectral filter with Gaussian time windows of 10 ms. Since we worked with the time-aligned version of STRAIGHT to obtain a more natural resynthesized sound output, anchor points to indicate the boundaries between phonemes were manually annotated in TextGrids in *Praat*. Temporal anchors were at all times placed at zero-crossings in the waveform. In the case of *knie* ('knee'), for instance, anchors were set at the offset of the release phase of the voiceless stop /k/, and at the onset of voicing of the vowel (i.e, [k]|[n]|[i]). By doing so, we ensured that only speech segments with corresponding acoustic and spectral characteristics were morphed (i.e., the stop /k/ and the nasal /n/ in *knie* [kni] were morphed with the stop and nasal consonant of the non-canonical [knɪ]). For particular target words, only the portions around the critical vowel were spliced out for morphing, because some of the surrounding sounds tended to be acoustically problematic.[9] The acoustically manipulated vowel sounds were then reinserted into the original items afterwards. Besides, the STRAIGHT-algorithm also takes into account differences in segmental duration. For example, the nasal consonant in [kni] was about 85 ms long whereas it was 105 ms in the non-canonical form, so the consonant will be approximately 95 ms at the fifth step on the continuum (i.e., the 50% morph). Finally, the remaining noise that originated from combining the two recordings were manually excised from the resynthesized speech signals.

Since we want to examine whether listeners can generalize knowledge about ambiguous, or atypical, speech sounds from one speaker to another, we acoustically modified each of the five minimal word pairs produced by the female speaker using the '*change gender*' function in *Praat* to achieve a male-sounding voice (cf. RQ2 in Section 3). This gender manipulation lowers the formant frequencies and the fundamental frequency (F0) of the female speaker's voice to generate a perceivable male voice, and thus a new speaker, while maintaining the durational values of original

[9] This procedure was followed for *brief* 'letter', for example, as the fricative consonant in both the recording with the canonical and non-canonical production sounded unnatural because of the large amount of friction noise.

input (see Liu & Holt, 2015, for a similar approach). Specifically, formant frequencies of the female speaker's minimal word pairs were multiplied by a factor of 0.82, whereas a perceivably lower pitch was obtained by multiplying the F0 contour by a factor of 0.66 (cf. Mitterer et al., 2020). In addition, the median pitch of the generated words was shifted to 95 Hz. Generalized over the individual minimal pairs, an overview of the spectral and pitch properties of the /ɪ/-words and /i/-words produced by the female speaker and the generated male voice can be found in Table 2. Similar to the stimulus materials with the female speaker's voice, an 11-step continuum between the two members of each minimal pair needed to be created for the generated male voice to determine which vowel sounds participants cannot unambiguously identify as either the /ɪ/- or /i/-vowel. This was done by applying the *'change gender'* function to all 11 steps of the original continua produced by the female Dutch L2 speaker.[10]

Table 2: Fundamental frequency (F0) and formant frequencies of the minimal word pairs produced by the female speaker and the manipulated, male-sounding voice, averaged over the individual minimal words.

	/ɪ/-words				/i/-words			
	Female		Male		Female		Male	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
F0 (Hz)	178	*7*	96	*2*	183	*2*	95	*2*
F1 (Hz)	418	*29*	359	*28*	258	*15*	222	*9*
F2 (Hz)	1864	*145*	1596	*107*	2366	*54*	2951	*56*
F3 (Hz)	2567	*58*	2221	*31*	3245	*203*	2891	*59*

[10] This method for creating a continuum was preferred to changing the perceivable gender of the speaker first and then applying the STRAIGHT-algorithm to the modified /ɪ/-words and /i/-words, as the output of this second method did not yield natural-sounding stimuli due to the large presence of creaky voice.

4.1.3 Procedure

Participants in the pre-test were provided with a general outline of the study and after having agreed to the terms and conditions as stipulated in the informed consent sheet, they were sent an experiment procedure document and a web link that redirected them to the online experimental environment (see Appendix 1). Each participant was also assigned a unique four-digit code (2 letters and 2 numbers, e.g. AB12) to enter the experimental webpage.[11] The experiment was developed in a code editor using the *jsPsych* library (de Leeuw, 2015), which is an open-source JavaScript library specifically intended for web-based behavioral experiments.[12] Before the start of the experiment, participants were asked to fill out a small survey that was aimed at gathering general background information (cf. Section 4.1.1; see Appendix 2 for a detailed overview). As the phoneme categorization task was administered over headphones, participants were asked to specify the brand (and if possible also the type) of the headphones they would use at test and adjust their computer volume to a comfortable listening level. In an audio test, participants could play, and if necessary replay, a pure tone (sampling frequency at 44100.0 Hz and tone frequency at 440.0 Hz) that was sampled at the same frequency as the stimuli presented during the experiment.

Participants were given detailed on-screen instructions about the procedure of the two-alternative categorization test. Specifically, participants were asked to decide by key press whether they heard the /ɪ/-vowel, as in *kind* ('child'), or the /i/-vowel, as in *machine* ('machine'): for the first option, they needed to press the F-key on their keyboard, for the latter the J-key. Care was taken to minimize the influence of orthography on vowel perception by exemplifying the /i/-sound with words like *machine* or *piano*, in which the critical vowel sound is not linguistically encoded by the graphemes <ie> as in *koffie* ('coffee') or *diep* ('deep'). Moreover, they were explicitly asked to respond with the vowel sound they thought they heard, regardless of whether that word exists in Dutch. Participants were also encouraged to respond as quickly as possible, without sacrificing response accuracy. To reduce the time needed to complete the experiment, we selected seven continuum steps (i.e., the 10%, 30%, 40%, 50%, 60%, 70% and 90% morphs) for each of the critical /ɪ/- and /i/-words and the minimal word pairs with both the female and male-sounding voice rather than presenting the full continuum. The first and final steps on the continua were not included, as the 0% and 100% morphs simply represent the unambiguous endpoints and thus do not inform us on the location of the ambiguous region on the continuum. This adjustment resulted in a total of

[11] These randomly generated codes were used instead of names to secure the anonymity of the participants.
[12] Data sets and scripts used for the task in the pre-test are available online at https://github.com/TristanCovemaeker/GilVerbeke.

350 trials (7 steps × 40 critical words + 7 steps × 5 minimal pairs × 2 voices). Trials were randomized across all ten participants, with the constraint that no more than two different continuum steps of the same target word were presented consecutively. Before the onset of each trial and after an inter-trial interval of 500 ms, a fixation cross appeared in the center of the screen for 500 ms to indicate that the next trial would start soon (see Figure 2 for a visual representation of a single trial). Participants could only press a key after the audio had finished playing, and were reminded of the response options and the corresponding key. Every 50 trials, they were allowed to take a self-paced break and could continue the next series of trials by pressing the space bar. It took half an hour, on average, to complete the categorization task.

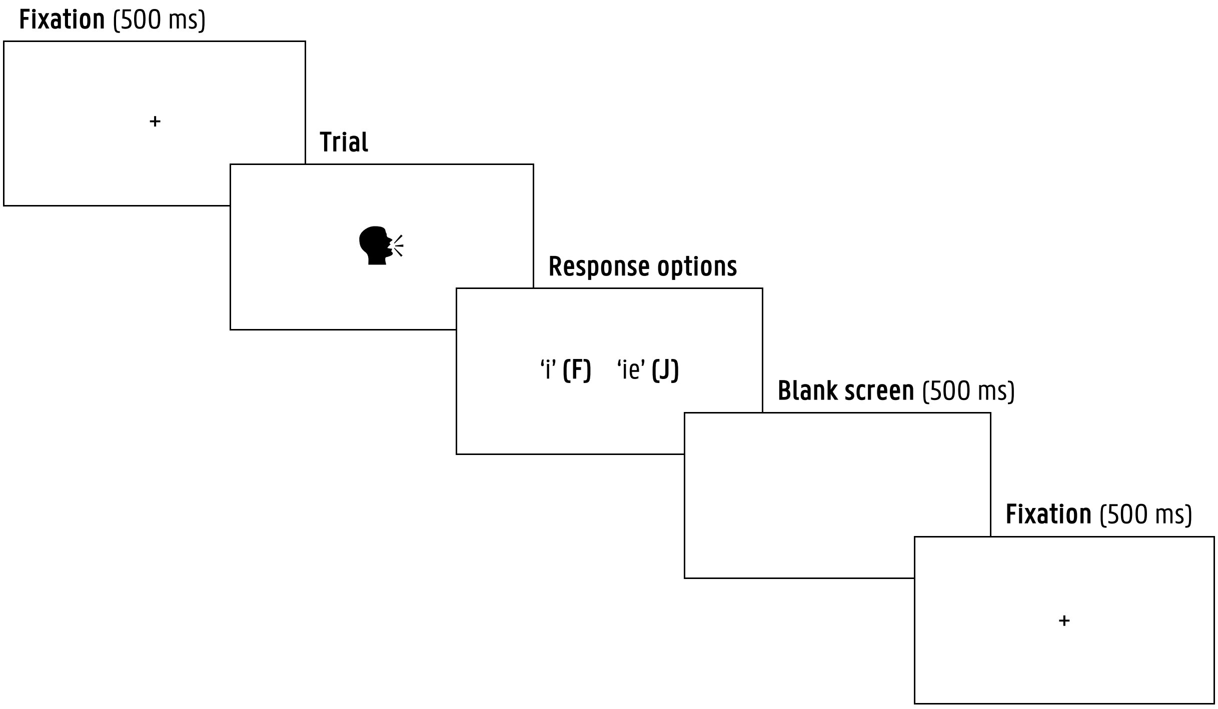

Figure 2: Schematic representation of a single experimental trial in the phoneme categorization task.

4.1.4 Analysis

This categorization task was administered to ascertain which step, or series of steps, on the [ɪ]-[i] continuum participants perceived as the most ambiguous vowel sounds. We initially considered continuum steps to be perceptually ambiguous when half of the participants classified the vowel as /ɪ/ and the other half as /i/ (i.e., 50%-50%). After completion of the test, however, some of the participants reported during a short informal debriefing that they had been guided by the canonical production of the target words as stored in their mental lexicon rather than that they had chosen the concrete vowel sound they perceived in the test items.[13] As lexical knowledge has indeed been found to bias listeners to perceive an ambiguous speech segment as a true phoneme when that particular sound yields a real word rather than a nonword (Ganong, 1980), we slightly raised the threshold value for the response proportions of the forty critical words (see Reinisch & Holt, 2014; Reinisch et al., 2013, for similar methodological approaches). Specifically, continuum steps of items which canonically have /ɪ/ as syllable nucleus were regarded as ambiguous when about 70% of the participants categorized the vowel sound as /i/. Phrased differently, seven out of ten participants needed to indicate hearing, for instance, [ˈsxil.dər] rather than [ˈsxɪl.dər] for *schilder* 'painter', although only the former option adheres to the Dutch pronunciation standard. Continuum steps of /i/-words were considered ambiguous when trials only received about 30% /i/-responses. Since the lexical context could not help listeners disambiguate the ambiguous vowel sound in the minimal word pairs, we selected continuum steps that had almost as many /ɪ/-responses as /i/-responses in trials with either the female or male-sounding voice. Moreover, instead of a single step, four consecutive steps on the minimal word pair continua were selected for the phoneme categorization task in the main experiment (cf. Section 6). Selecting multiple steps allows us to see whether listeners adjusted their phonological boundaries of the critical vowels after exposure to the ambiguous or atypical pronunciations of the vowels in the exposure phase.

[13] One participant reported the following: "Because I know that *fiets* ['bike'] is spelled with *ie*, I automatically hear more of an /i/-sound, regardless of whether the female speaker produces an [i]-like vowel." (personal communication with participant MM84; our translation).

4.2 Results

The results of the pre-test, pooled over participants and items, are summarized in Table 3. Data are further broken down by the seven continuum steps and the four types of words in which the vowel contrasts were presented, *viz.* (i) words with /ɪ/ as the nucleus of the stressed syllable, (ii) critical words with /i/, (iii) minimal /ɪ/-/i/ words with the female speaker's voice and (iv) with the male-sounding voice. Averaged over continuum steps, there seems to be an equilibrium in participants' responses, in that the difference in the number of /ɪ/-responses (n = 1748) and /i/-responses (n = 1752) across the four categories is almost negligible. When continuum steps are included in the analysis of participants' categorization performance, we can observe in Table 3 that listeners reported hearing either the /ɪ/- or /i/-vowel in the first half of the continuum steps, whereas they increasingly indicated hearing the opposite vowel of the contrast in the second half (cf. the S-shaped patterns in Figures 3 to 6 below). Although this pattern confirms that participants gradually mapped the vowel production onto the opposite vowel of the contrast, these aggregate response distributions do not inform us where exactly on the continuum the perceptual boundaries between the front vowels is located. In what follows, we will therefore look at each of the four categories in more detail to determine where the cut-off point between the two vowels lies in listeners' perception. We will address the results of the forty critical words in Section 4.2.1, and those of the minimal word pairs with the female speaker's voice and the male-sounding voice in Section 4.2.2.

Table 3. Absolute and relative frequencies of the vowel categorization responses by continuum step across the four types of target words.

	critical words				minimal words			
	/ɪ/-words (*n = 1400*)		/i/-words (*n = 1400*)		female voice (*n = 350*)		male voice (*n = 350*)	
Step	/ɪ/	/i/	/ɪ/	/i/	/ɪ/	/i/	/ɪ/	/i/
1	174 *(87)*	26 *(13)*	3 *(2)*	197 *(98)*	48 *(96)*	2 *(4)*	47 *(94)*	3 *(6)*
3	156 *(78)*	44 *(22)*	30 *(15)*	170 *(85)*	46 *(92)*	4 *(8)*	42 *(84)*	8 *(16)*
4	136 *(68)*	64 *(32)*	55 *(28)*	145 *(72)*	36 *(72)*	14 *(28)*	40 *(80)*	10 *(20)*
5	111 *(56)*	89 *(44)*	109 *(55)*	91 *(45)*	33 *(66)*	17 *(34)*	33 *(66)*	17 *(34)*
6	66 *(33)*	134 *(67)*	134 *(67)*	66 *(33)*	7 *(14)*	43 *(86)*	21 *(42)*	29 *(58)*
7	43 *(21)*	157 *(79)*	163 *(82)*	37 *(18)*	6 *(12)*	44 *(88)*	12 *(24)*	38 *(76)*
9	23 *(11)*	177 *(89)*	170 *(85)*	30 *(15)*	2 *(4)*	48 *(96)*	2 *(4)*	48 *(96)*
Total	709 *(51)*	691 *(49)*	664 *(47)*	736 *(53)*	178 *(51)*	172 *(49)*	197 *(56)*	153 *(44)*

4.2.1 Critical Words

With regard to the critical /ɪ/- and /i/-words, the relative distribution of the response options is shown to correlate with the particular steps on the generated continua, as is visualized in Figures 3 and 4 (cf. Table 3 for the corresponding values). That is, the further the continuum step is removed from the canonical realization of the vowel (i.e., step 0), the more the participant signals hearing the opposite vowel of the contrast. For the critical /ɪ/-words, this implies that listeners classified the ambiguous syllable nucleus incrementally more as /i/ than /ɪ/ towards the ninth continuum step, as is evidenced by the decreasing and increasing heights of the black and gray bars in Figure 3, respectively. For the critical /i/-words, the exact opposite bidirectional trend can be observed, as the number of /i/-responses gradually declines while the number of /ɪ/-responses increases. If we compare the gradient of the slopes between consecutive continuum steps in both figures, we can observe that the slope is the steepest between steps 5 and 6 for /ɪ/-words, whereas the largest transition in vowel response is most noticeable between steps 4 and 5 for the /i/-words. This leads us to conclude that the perceptual cut-off point, and hence the most ambiguous region on the continuum, is situated around step 5 for the critical words.

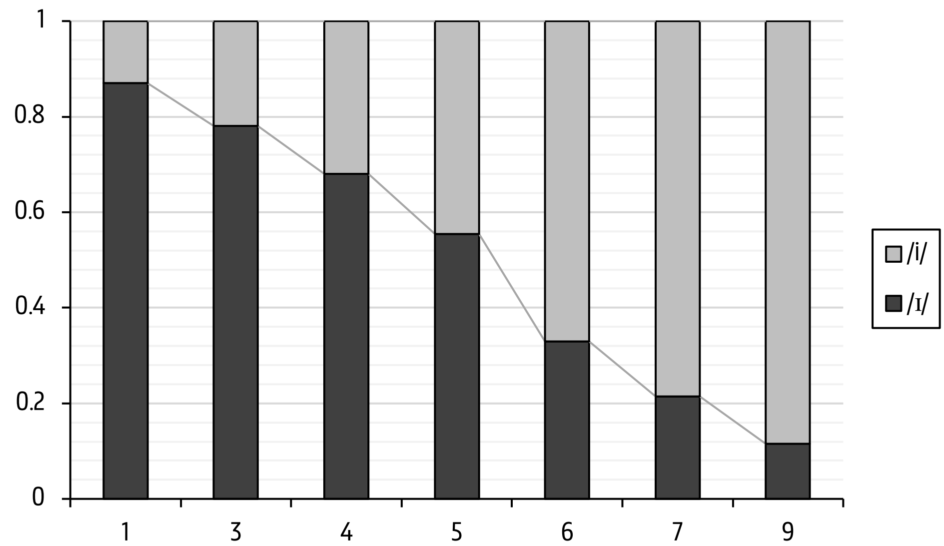

Figure 3: Relative frequencies of the categorization responses of the critical /ɪ/-words by continuum steps.

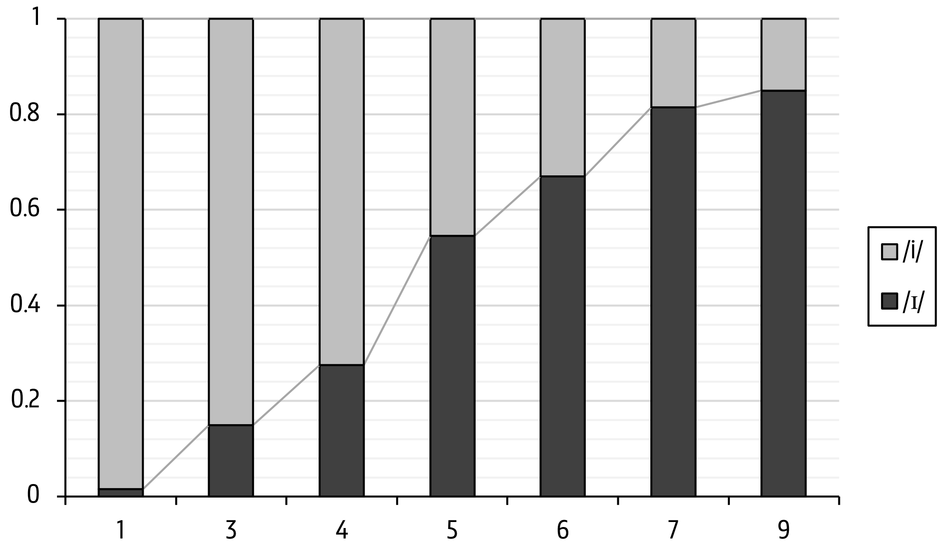

Figure 4: Relative frequencies of the categorization responses of the critical /i/-words by continuum steps.

On top of the analysis of the response distributions, we will also consider three additional parameters to determine which trials listeners perceived to be most ambiguous. First, since lexical knowledge may have influenced listeners' perception of the ambiguous vowel sounds (Ganong, 1980), step 6 for /ɪ/-words and step 4 for /i/-words were preferred over the midway step, because these steps received in nearly 70% of the cases responses of the non-canonical vowel of the contrast (cf. Section 4.1.4 on data analysis). Secondly, the response patterns of some individual participants did not always invariably correspond to the aggregate data. For example, the results obtained from participant MM84 did not display the bidirectional patterns as shown in Figures 3 and 4, in that even the ninth continuum step for /i/-words was identified in 70% of the trials as /i/ rather than /ɪ/. Additionally, it took remarkably longer for this participant to choose either of the two vowel options. On average, reaction times as measured from the offset of the trial were shorter, albeit minimal, towards the continuum endpoints (*M* = 976ms for steps 1 and 3, and 1089ms for steps 7 and 9) as opposed to the central, and hence perceptually more ambiguous steps (*M* = 1192ms for steps 4-6). However, this observation did not hold for MM84 across all four types of target words, while such a pattern could be established for the other participants in at least one of the four target word categories. The influence of the lexicon and individual differences in categorization performance thus called for a more rigorous approach. The ambiguous continuum step selected for the critical words was therefore in 40% of all lexical items (n = 16) different from the ambiguous steps of the aggregate data. For an overview of the selected steps for each target word, see Appendix 5.

4.2.2 Minimal Words

In addition to the forty critical words, the relative frequencies of /ɪ/- and /i/-responses for the minimal words with either a female or perceivable male speaker's voice are presented in Figures 5 and 6, respectively (see Table 3 above for the absolute values). Note that the threshold value did not need to be adjusted for these target words because the lexicon cannot help listeners to fill in the gap of the ambiguous vowel sound in the case of minimal word pairs. Moreover, instead of selecting just one step, we selected four consecutive continuum steps which were identified as perceptually most ambiguous by the listeners (i.e., continuum steps with about 50% /ɪ/- and 50% /i/-responses).

Contrary to the gradual shift in vowel discrimination for the two types of critical words as described above, the cut-off point between what participants categorized as either the /ɪ/- or the /i/-vowel is considerably sharper for the minimal words produced by the female L2 speaker. Looking at the distribution of responses between steps 5 and 6 on the continuum in Figure 5, we can see that the proportion of vowel responses drops from 66% to only 14% /ɪ/-responses. Accordingly, this involves that the proportion of /i/-responses remarkably increases between these steps (34% and 86%, respectively). Similar to the selection procedure for the critical /ɪ/- and /i/-words, extreme differences between the categorization responses of individual participants and the generalized data were also taken into account when selecting the appropriate continuum steps for each token. Specifically, for all but one of the minimal word pairs produced by the female Dutch L2 speaker, steps 3 to 6 were selected as the most ambiguous region based on pooled and individual categorization responses and response times. For the minimal pair *bid-bied*, however, steps 4 to 7 were selected.

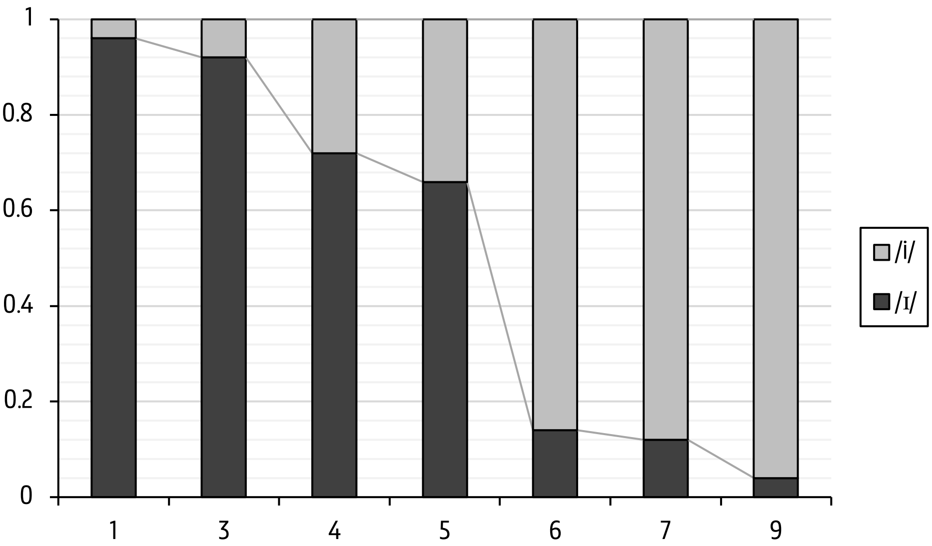

Figure 5: Relative frequencies of the categorization responses of the minimal /ɪ/-/i/ word pairs with the female speaker's voice by continuum steps.

For the minimal word pairs that were acoustically modified to be perceived as a male-sounding voice, the transition from mainly /ɪ/-responses to more /i/-responses is notably less pronounced than for the minimal words with the female voice. This is visualized in Figure 6 by the overall milder steepness of the slopes between two adjacent continuum steps. Nevertheless, the largest disagreement between listeners in terms of which of the two front vowel categories they perceived seems to be located between the fifth and sixth continuum step: on step 5, the vowel is identified in 66% of the trials as the /ɪ/-vowel, whereas it is classified as /ɪ/ in only 42% of the trials on the next continuum step. As outlined for the other three lexical types, we also performed an item-level analysis of the minimal word pairs to trace deviations from the general trend in terms of categorization responses. This analysis revealed that up to and including step 6 of the *kist-kiest* ‘chest/box-chooses’ and the *wit-wiet* ‘white-weed’ pair, at least eight out of ten participants still indicated hearing the /ɪ/-vowel. In other words, the cut-off point between the two vowels in these tokens only occurred between steps 6 and 7, or possibly even between step 7 and step 9. As a result, steps 5 to 9 were selected for *kist-kiest* as well as for *wit-wiet* as the perceptually most ambiguous steps. For the other three minimal word pairs with the male-sounding voice, steps 3 to 6 were chosen.

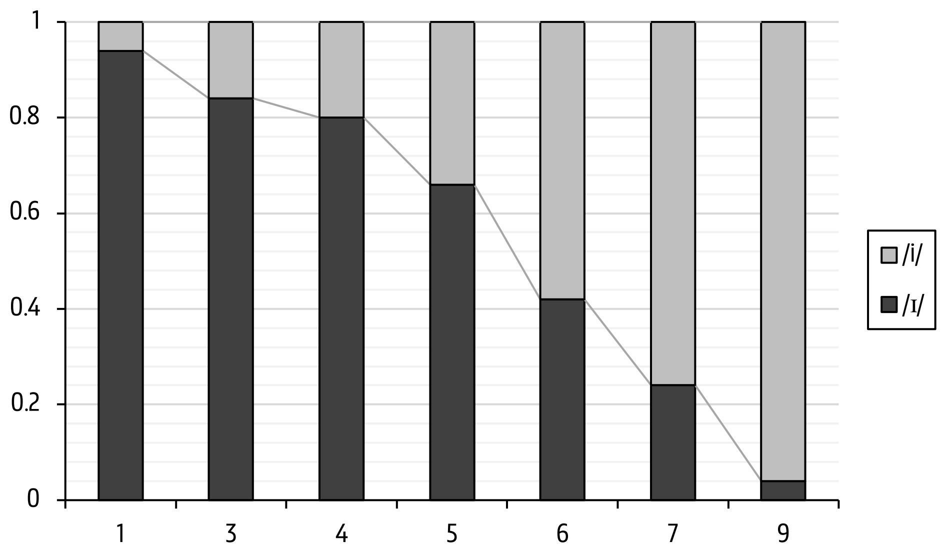

Figure 6: Relative frequencies of the categorization responses of the minimal /ɪ/-/i/ word pairs with the male speaker’s voice by continuum steps.

4.3 Discussion

The purpose of the pre-test was to determine where the ambiguous region on each item-specific continuum is situated, as this procedure has been deemed more efficient and reliable than simply choosing the midway step between the continuum endpoints (see Samuel, 2020). Ultimately, these ambiguous tokens will be used as stimulus materials for the lexical decision task (Section 5) and the phoneme categorization task (Section 6) in the main experiment. Remind that the most ambiguous region on the continua for the critical /ɪ/- and /i/-words was associated with those continuum steps that received approximately 70% responses of the opposite vowel in the contrast. For the minimal word pairs, the vowel sound in a continuum step was considered ambiguous when a step had as many /ɪ/-responses as /i/-responses. Moreover, item-specific analyses often revealed notable differences between participants in terms of how they classified the ambiguous vowels and how quickly they responded. Based on these parameters and observations, we selected in about 60% of the instances either step 4, 5 or 6 for the critical words as the continuum step containing the most ambiguous front vowel. Note that for the remaining 40% of the critical words the selected step was thus closer to either of the continuum endpoints. Finally, steps 3 to 6 were selected for four out of five minimal word pairs with a female voice and three out of five with a male voice.

Before turning to the main experiment, it might be of value to digress briefly on the possible causes for the often remarkable individual differences as well as on the implications of the selected continuum steps. The fact that response times of some participants were on average substantially longer than those of others may indicate that listeners were drawing on their lexical knowledge to fill in the gap of the ambiguous front vowel and did hence not intuitively indicate which vowel sound they perceived. To rule out excessive variability in response latency in the main experiment, we included a response time limit of 4000ms. With regard to the selected continuum steps for the minimal pairs, then, listeners might be inclined to categorize the ambiguous sound more often as the /ɪ/- than the /i/-vowel, because the steps selected for the majority of these target words lean somewhat more towards the /ɪ/-side of the generated continuum. The exact opposite may be true for some other minimal word pairs, in that steps 4 to 7 and 5 to 9 could in turn induce a bias towards /i/-responses. It should be noted, however, that participants in the pre-test heard both canonical and ambiguous realizations of each of the two front vowels (i.e., *within-subject design*), which has been found to impact on listeners' readiness to interpret ambiguous sounds as potentially atypical pronunciations of a particular speech category (see Norris et al., 2003). For that reason, listeners in the main experiment were assigned to separate conditions in which either all /ɪ/-words or all /i/-words consisted of an acoustically ambiguous vowel sound (i.e., *between-subject design*).

5 Lexical Decision Task

In this section, we will turn to the first major task of the experiment, which is the auditory lexical decision task. Native Dutch listeners were asked to classify a series of stimuli as either real Dutch words or words that are not part of the Dutch lexicon. Nevertheless, the main purpose was not to assess participants' ability to distinguish words and nonwords. Rather, this forced-choice task served as an exposure phase, during which listeners became familiar with the accented speech of the female non-native speaker of Dutch, and more importantly, with the lexical items in which the /ɪ/- or the /i/-vowel was acoustically ambiguous. Following previous studies on lexically-guided perceptual learning (cf. Section 2.3), the lexical context in which these vowels were embedded should enable listeners to identify which front vowel the L2 speaker may have intended to produce and as such encourage them to adjust their phonological representation of that vowel category accordingly. The results of the lexical decision task will inform us on how well native speakers of Dutch understand L2 accented speech, and whether they process real Dutch words differently than lexical items with atypical vowel pronunciations.

5.1 Method

5.1.1 Participants

116 native speakers of Belgian Dutch (mean age = 21.1 years, *SD* = 1.7) who did not participate in the pre-test, took part in the main experiment (for an overview, see Appendix 2). Female participants (n = 78) vastly outnumbered male participants (n= 36) and those who do not identify as either female or male (n = 2). 71% of the participants were born in West-Flanders (n = 82) whereas the remaining 29% was born in East-Flanders (n = 34). Except for two participants who were born in West-Flanders but grew up in East-Flanders, participants were raised in the same province of birth. Criteria for inclusion remained the same as those described in the method of the pre-test, namely self-reported normal hearing and no prior knowledge of Italian and Spanish (cf. Section 4.1.1). All participants were enrolled in a program offered at an institution of higher education in Ghent at the time of testing and were recruited through in-class announcements, university platforms and social media posts. Written informed consent was obtained from all students before the start of the experiment and they could once again register for a giveaway of five twenty euro gift cards as compensation for their time.

5.1.2 Stimulus Materials

The set of stimuli used in the lexical decision task consisted of 100 moderate- to high-frequency Dutch words and 100 phonotactically legal Dutch nonwords (see Appendix 4). Included in the former group were both the forty critical /ɪ/- and /i/-words as described in the methodology of the pre-test (Section 4.1.2), and sixty additional filler words, which were added to distract participants from focusing on the critical words. Fillers were subject to similar selection criteria as applied to the critical words, in that they were all real monosyllabic or disyllabic Dutch content words with a Zipf value ranging between 3.5 and 5.5 (M = 4.36; SD = 0.52). Contrary to the critical words, the nucleus of the stressed syllable was filled by the short /ɑ/ (e.g., *kapsel* 'haircut', /ˈkɑp.səl/) and long /aː/ vowel (e.g., *baas* 'boss', /baːs/), or by the short /ɔ/ (e.g., *mosterd* 'mustard', /ˈmɔs.tərt/) and long /oː/ vowel (e.g., *roos* 'rose', /roːs/). These or similar vowel contrasts are absent in the Italian phonological system. Although vowels can be lengthened in Italian depending on the prosodic structure of the utterance, there are no phonemic differences between vowels in terms of their length (Krämer, 2009; Rogers & d'Arcangeli, 2004). Despite the lack of /ɑ/-/aː/ and /ɔ/-/oː/ contrasts in Italian, the Dutch L2 speaker distinguished the short and the long variant in both contrasts through noticeably distinct temporal characteristics (i.e., the duration of the long vowel is approximately twice as long as that of the short vowel) and different spectral properties (see Table 4), although the differences in formant frequencies are slightly less pronounced than for the critical vowels in the /ɪ/-/i/ contrast (cf. Table 1 in Section 4.1.2).[14]

Table 4: Overview of the spectral and temporal characteristics of the non-native speaker's productions of the Dutch /ɑ/-/aː/ and /ɔ/-/oː/ contrasts.

	Vowel contrast 1				Vowel contrast 2			
	/ɑ/		/aː/		/ɔ/		/oː/	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
F1 (Hz)	762	*66*	821	*59*	517	*60*	412	*56*
F2 (Hz)	1301	*91*	1357	*64*	1001	*174*	916	*80*
F3 (Hz)	2436	*308*	2288	*266*	2710	*230*	2676	*304*
Duration (ms)	107	*33*	233	*61*	108	*32*	188	*64*

[14] No statistical analyses were conducted to verify if the spectral and temporal values differed significantly between the short and long variant of these contrasts. The rationale behind this is that (i) there is only a relatively small number of observations per category and (ii) that participants will not be explicitly asked to identify which member of these vowel contrasts they perceive in the phoneme categorization task.

In addition to the real Dutch words, 100 monosyllabic and disyllabic words which do not exist in Dutch and have a vowel other than /ɪ/ or /i/ in the stressed syllable were generated using *Wuggy* software (Keuleers & Brysbaert, 2010). *Wuggy* was preferred to other word generator programs, because it rules out an experimenter's preferences for particular combinations of letters. In order to generate nonwords, we made a random selection of highly frequent Dutch words to input in *Wuggy* (M_{Zipf} = 4.39, SD_{Zipf} = 0.81) (cf. Appendix 4). Since this program only outputs strings of letters that do not form real words in terms of orthography, the concrete phonetic realization of the candidate nonwords was also taken into account in the selection of the most suitable test items.[15] Moreover, we generated pseudowords rather than nonsense words (e.g., *krokkel* vs. *rkolkek*), as pseudowords adhere to the phonotactic patterns of Dutch and could therefore mitigate overtly noticeable differences between the real words and the nonwords at test. Finally, the value associated with one particular output option in *Wuggy* needed to be small, namely that for the orthographic Levenshtein distance of the generated nonword's twenty closest lexical neighbors (OLD20). If the OLD20 value is small, this means that the generated nonword can easily be transformed into a real word with only a few edit operations (e.g., substituting, inserting or omitting particular letters) (Keuleers & Brysbaert, 2010). OLD20 of the nonwords selected for the lexical decision task was never more than 2 (M = 1.44, SD = 0.28). In other words, maximally two orthographic edit operations are needed to turn the nonwords into a real Dutch word. These parameters should ensure that the differences between the real words (i.e., critical words and fillers) and the nonwords are not too conspicuous (see also Yarkoni et al., 2008, on the importance of OLD20 for response latency).

[15] Although a string of letters may not resemble a real word in terms of spelling, it can be phonetically identical to a real word. Such *pseudohomophones* were not included in the dataset (e.g., **blouw* [blɔu] vs. *blauw* 'blue' [blɔu]).

5.1.3 Procedure

The experiment procedure adopted for the lexical decision task largely corresponds to the one outlined for the pre-test (cf. Section 4.1.3). Participants were informed that the present study aimed to investigate how Dutch L1 speakers understand and accept speech produced by non-native speakers of Dutch. After giving written informed consent, participants were sent a procedure manual, including the link to the web-based experiment (see Appendix 1).[16] Before the start of the task, they were asked to fill in a small questionnaire on personal details (e.g., age, gender, province of birth and residence; cf. Section 5.1.1 above). Participants were instructed to complete the experiment over headphones in a quiet room. They were also encouraged to adjust their computer audio to a self-selected comfort level for listening. Written on-screen instructions to the lexical decision task explained that they would hear a semantically neutral carrier sentence in which the central word would be different across trials (i.e., *Ze heeft [X] gezegd* 'She said [X]').[17] Their task was to decide whether that central word was a real Dutch word (e.g., *appel* 'apple') or a nonword in Dutch (e.g., *krasp*) by pressing either the J- or F-key on their keyboard, respectively.[18]

Undisclosed to the participants was the fact that there were two conditions in the lexical decision task. Half of the participants were assigned to the /ɪ/-ambiguous condition and the other half to the /i/-ambiguous condition. In the first condition, the vowel sound in all test items that canonically have /ɪ/ as the nucleus of the stressed syllable (e.g., *kikker* 'frog', /ˈkɪ.kər/) were presented with an ambiguously sounding vowel, while the /i/-words were presented in their canonical realization (e.g., *fiets* 'bike', /fits/). Participants in the other condition heard exactly the opposite, in that all critical /ɪ/-words were produced naturally while the critical vowel in all /i/-words was acoustically ambiguous. Ambiguous vowel realizations, as represented by the steps along the [ɪ]-[i] continuum, were determined based on the results of the pre-test (cf. Section 4). A between-subject design was preferred here, as exposure to both ambiguous and unambiguous pronunciations of the same vowel has been found to negatively impact perceptual adaptation (see Kraljic et al., 2008).

In order to familiarize participants with the testing procedure, they first completed three practice trials with feedback about the correctness of their response. Shortly after the practice phase, the lexical decision task started in which feedback was no longer given. A total of 200 test items (100 nonwords + 60 fillers + 20 /ɪ/-words + 20 /i/-words) were randomized across participants, with

[16] Data sets and scripts used for the tasks in the main experiment are available online at https://github.com/TristanCovemaeker/GilVerbekeTest.
[17] Note that none of the vowels in the carrier sentence correspond to the ones in the critical words to neutralize unwanted priming effects.
[18] We deemed the term *nonword* slightly more straightforward than *pseudoword*, in particular because more than half of the participants were not enrolled in a language-oriented program of study.

the exception that no more than two trials of the same item type (i.e., nonwords, fillers, /ɪ/-word, /i/-word) could follow each other immediately. Participants were encouraged to respond as quickly as possible, without neglecting the accuracy of their response. If none of the two keys had been pressed after four seconds, a message appeared on the screen that no answer had been registered. After an inter-trial interval of 1000 ms, the next trial started automatically. Every forty trials, participants were allowed to take a self-paced break. It took approximately 15 minutes to complete this first task of the experiment.

5.1.4 Analysis

Data from some participants were excluded to ensure that all statistical analyses were conducted on a relatively homogeneous sample. In accordance with previous studies (e.g., Norris et al., 2003; Reinisch & Holt, 2014), responses of thirteen participants were excluded from further analysis, as they identified more than half of the critical /ɪ/- and /i/-words as nonwords, regardless of whether they were produced with a natural or an ambiguous vowel. Data of three additional participants were not analyzed either, because they did not complete the test over headphones or did not provide details of the headphones used at test. Finally, single lexical decision responses were not logged if participants exceeded the response time limit of 4000 ms. Due to this criterion, 73 trials (0.4%) falling outside this time window were excluded.

Data of the remaining 100 participants will be analyzed for lexical decision responses (word vs. nonword) and response latency (i.e., the time interval between the end of the trial and key-press). In an initial stage, we will examine the relative distribution of lexical decision responses across all four types of test items. This will inform us on how well the different types of target words were recognized in L2 accented speech. In a second phase, we will concentrate in more detail on the /ɪ/- and /i/-words, since this task predominantly functioned as an exposure phase to familiarize the participants with the atypical productions of either /ɪ/ and /i/. Moreover, we added reaction time as a second dependent variable, because we expected processing words with ambiguous speech sounds to require more cognitive effort than lexical items with naturally produced speech sounds, and will hence impact on how quickly participants make a categorization response. To that end, lexical decision performance for the forty critical words, as indexed by word endorsement and response latency, was analyzed with linear mixed-effects models (Baayen et al., 2008), using the *lme4* package (Bates et al., 2015) in *R* statistical software (R Core Team, 2020). Mixed-effect modeling was used here rather than repeated measures analyses of variance (ANOVA), because linear models have been found to be less susceptible to Type-I errors (i.e., falsely rejecting the null hypothesis) (see

Quené & van den Bergh, 2008). Mixed-effects models are also preferred with dichotomous categorical dependent variables, which is the case for word endorsement (i.e., word vs. nonword) (see Jaeger, 2008). Note that we built two separate models with either word endorsement (i.e., classifying the critical words as real Dutch words) or response latency (in ms) as the dependent variable. Both models included two fixed factors with two levels, namely *exposure condition* (/ɪ/-ambiguous vs. /i/-ambiguous) and *critical word type* (/ɪ/-word vs. /i/-word), of which the /ɪ/-words in the /ɪ/-ambiguous condition are be mapped onto the intercept. Variability between participants and the potential effect of specific test items were accounted for through the addition of by-subject and by-item intercepts in the regression models.

Although the *lme4* package has repeatedly been used in behavioral experiments to analyze perceptual learning, it does not output straightforward measures such as a p-value to determine whether a fixed factor significantly predicts the dependent variable.[19] In response to this gap, we also used the *lmerTest* (Kuznetsova et al., 2017) package to generate p-values for each of the fixed factors and their interactions. Note that all and only results from the linear regression analyses which have a p-value smaller than 0.05 will be reported as statistically significant (see Baayen, 2008). In addition to p-values, we will report three other terms per fixed factor. That is, (i) the estimated coefficients, or beta-values, of the fixed factors, which indicate the difference in predictive contribution on the dependent variable of one level of a fixed factor compared to the other level that is mapped onto the intercept; (ii) the standard error of the estimated coefficients; and finally (iii) Wald's z-score or t-value, depending which value is reported in the R output.

[19] The value that is generally used in mixed-effect modeling to determine statistical significance is Wald's *t*-value. If a factor has an absolute *t*-value larger than 1.96, this means that it has a significant impact on the dependent variable (see Hox, 2010). However, since p-values are assumed to be more accessible for other researchers to assess the effect of a particular factor, as Luke (2017) claims, we will also explicitly report probability values in addition to t-values.

5.2 *Results*

Averaged over the full group of participants, Figure 7 summarizes the word endorsement rate in the lexical decision task for each of the four item types, i.e. nonwords, fillers, /ɪ/-words and /i/-words. Proportions of 'word'-responses are also plotted by exposure condition, namely /ɪ/-ambiguous and /i/-ambiguous. Visual inspection of Figure 7 clearly shows that the vast majority of the target nonwords were accurately identified as not being part of the Dutch lexicon across both exposure conditions: participants who heard all /ɪ/-words with an ambiguous vowel correctly rejected 89% of the nonword items whereas this percentage was lower, albeit marginally, for the participants in the /i/-ambiguous condition (86%) (see Table 5 for the corresponding absolute values). Word endorsement accuracy for the fillers, then, was near-ceiling in both groups, in that 94% and 95% of the trials were correctly recognized as real words in the /ɪ/- and /i/-ambiguous conditions, respectively. These findings indicate that listeners managed to discriminate canonically produced real words and phonotactically legal nonwords equally well in both exposure conditions. Besides, the fact that approximately 10% of the nonwords received 'word'-responses suggests that the generated nonwords did at least to some extent resemble real Dutch words. Note that the relative frequencies of response options are preferred here over the absolute values, because the number of participants per condition is unequal (/ɪ/-ambiguous: n = 57 vs. /i/-ambiguous: n = 43). This asymmetry originated from the removal of data from thirteen participants as per exclusion criteria outlined in Section 5.1.4.

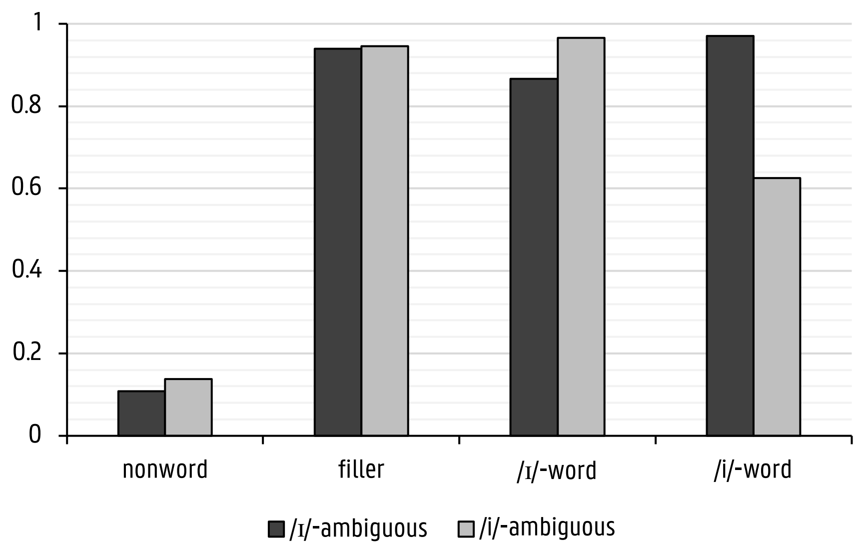

Figure 7: Mean proportion of 'word'-responses by item type and exposure condition.

Table 5: Absolute frequencies of the lexical decision responses by item types and exposure condition.

	/ɪ/-ambiguous			/i/-ambiguous			grand total
	word	nonword	*total*	word	nonword	*total*	
nonword	616	5055	*5671*	594	3695	*4289*	9960
filler	3197	208	*3405*	2435	138	*2573*	5978
/ɪ/-word	985	152	*1137*	829	30	*859*	1996
/i/-word	1104	34	*1138*	535	320	*855*	1993
grand total	5902	5449	*11351*	4393	4183	*8576*	19927

Compared to nonwords and fillers, the exposure condition appears to have had a more profound effect on participants' lexical decision performance for the critical words (see Table 5 and the right-hand side of Figure 7). Specifically, listeners in the /ɪ/-ambiguous condition identified /ɪ/-words more often as nonwords than /i/-words (i.e., 13% vs. 3%). The exact opposite is true for the participants assigned to the /i/-ambiguous condition: only 3% of the /ɪ/-words was classified as a nonword, while this percentage amounted to 37% for /i/-words. This response pattern indicates that, overall, canonical productions were endorsed more frequently as real words than ambiguously produced lexical items. It should be noted that the difference in relative frequencies of 'word'-responses between the two critical word types is noticeably larger in the /i/-ambiguous condition as opposed to the /ɪ/-ambiguous condition (10% vs. 34%, respectively).

To assess the influence of *exposure condition* and *critical word type* on lexical decision performance, we constructed a mixed-effects model, which also checks for by-subject and by-item variability. In advance, we conducted a logistic likelihood-ratio test, which indicated that a regression model with an interaction between the two fixed factors yields a better goodness of fit compared to one without interaction ($\chi^2(1) = 419.2$, $p < 0.001$). Statistical analyses of the model with an interaction show that, as expected, the exposure condition to which participants were assigned significantly predicted the proportion of 'word'-responses for /ɪ/-words ($b_{\text{Intercept}} = 2.88$, $SE = 0.40$, $z = 7.19$, $p < 0.001$, $b_{\text{Condition}} = 1.87$, $SE = 0.25$, $z = 7.37$, $p < 0.001$). Moreover, there was also an effect, albeit smaller, of critical word type ($b_{\text{WordType}} = 1.38$, $SE = 0.56$, $z = 2.45$, $p = 0.014$), suggesting that /i/-words were more likely to be identified as real words than /ɪ/-words in the /ɪ/-ambiguous condition. Finally, the interaction between exposure condition and item type was found to be highly significant ($b_{\text{WordType* Condition}} = -5.31$, $SE = 0.31$, $z = -17.22$, $p < .001$). The negative value of the estimate for the interaction suggests that /i/-words in the /i/-ambiguous condition were more likely to receive fewer 'word'-responses than /ɪ/-words in the alternative condition. Taken together, these findings show that word endorsement for the critical words is strongly modulated by exposure condition and critical word type, as well as their interaction.

In addition to word endorsement, we also analyzed whether and, if so, to what extent the ambiguity of the critical words has an effect on how quickly participants respond by key-press. Seeing that words with ambiguous vowels are expected to require more cognitive effort than words with canonically produced vowels, it could be assumed that this difference in cognitive demand will be reflected in response latency. This assumption is confirmed by the data and visualized in Figure 8, in that the height of the black and gray bars, representing the mean response time in the /ɪ/-ambiguous and /i/-ambiguous condition respectively, for the /ɪ/-words seems to be in inverse proportion to the height of the bars for the /i/-words. As the bar plots in Figure 8 average over individual differences, the four boxplots in Figure 9 more clearly illustrate that the distribution of response time measures is right-skewed. This indicates that the response latency of some participants was considerably larger than the median value, hence the large number of outliers in Figure 9. To improve the distribution of the data and thus reduce between-trial variability, we applied a square-root transformation to the data, which means that the square root of each response time value was taken for further analysis (see Osborne, 2002).[20]

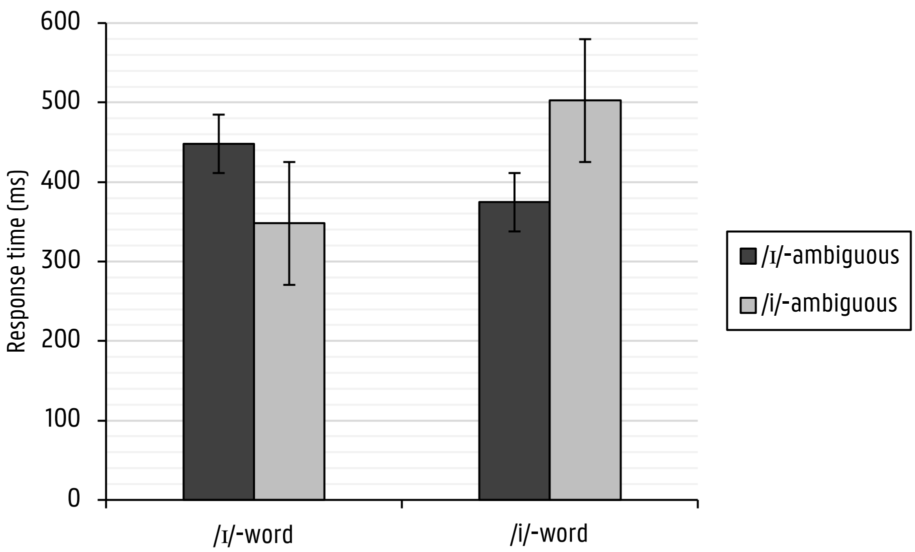

Figure 8: Mean response time (ms) and standard errors for /ɪ/-words and /i/-words by exposure condition.

[20] Although log-transformations are typically applied to improve the distribution of right-skewed data, such a transformation could not achieve a better spread of our data (see also Feng et al., 2014, on the implications and problems of log-transformation).

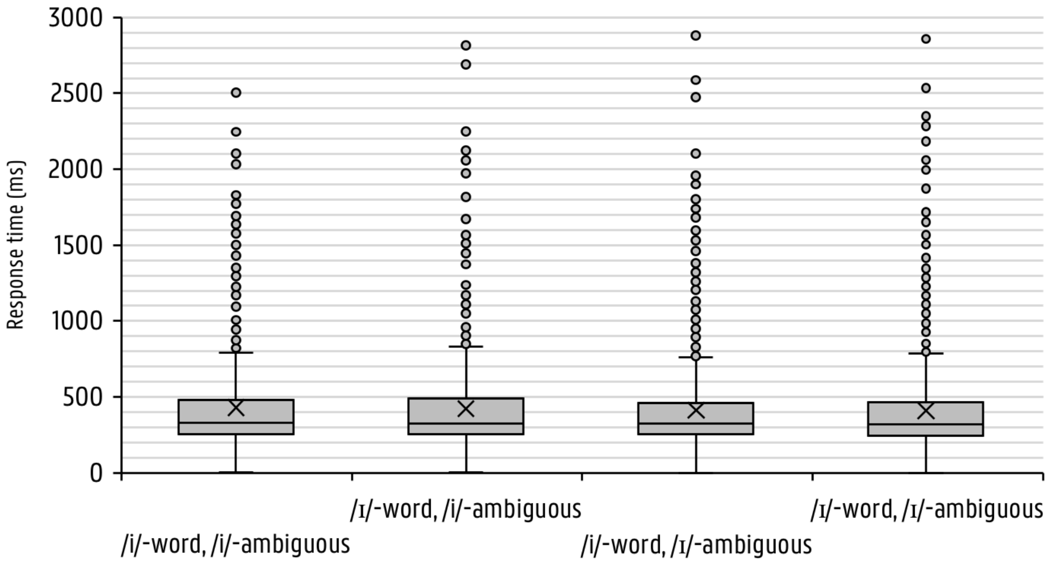

Figure 9: Distribution of response latency (ms) across word types and exposure conditions.

Similar to the analysis of word endorsement, we built a linear mixed-effects model for response time. Again, *exposure condition* (/ɪ/-ambiguous vs. /i/-ambiguous) and *critical word type* (/ɪ/-word vs. /i/-word) were entered as fixed factors and *participant* and *item* as random effects in the model. A likelihood ratio test showed that a linear model in which there is an interaction between the critical word type and exposure condition is a better predictor for response time than a model without such interaction between the fixed factors ($\chi^2(1) = 174.16$, $p < 0.001$). Specifically, when participants were assigned to the /ɪ/-ambiguous condition, response latency was significantly larger for /ɪ/-words compared to /i/-words ($b_{\text{Intercept}} = 19.96$, $SE = 0.48$, $t(111) = 41.46$, $p < 0.001$; $b_{\text{WordType}} = -1.72$, $SE = 0.50$, $t(48) = -3.43$, $p = 0.001$). Response time for the /ɪ/-words is expected to be significantly faster when listeners hear all /i/-words rather than /ɪ/-words with an ambiguous vowel in the lexical decision task ($b_{\text{Condition}} = -2.16$, $SE = 0.56$, $t(126) = -3.83$, $p < 0.001$). The interaction between the exposure conditions and the critical word types was significant too ($b_{\text{WordType*Condition}} = 5.15$, $SE = 0.39$, $t(3849) = 13.35$, $p < 0.001$). This implies that response latency could be reliably predicted based on which of the two front vowels sounded ambiguous at test in combination with the type of critical word on which participants were tested. In sum, participants assigned to the /ɪ/-ambiguous condition were thus expected to respond more quickly when hearing an /i/-word compared to an /ɪ/-word. The exact opposite pattern could be established for participants in the /i/-ambiguous condition, as is also reflected in the relative heights of the bars in Figure 8.

5.3 *Discussion*

The main purpose of the lexical decision task was to familiarize the Dutch L1 speakers with the L2 accented speech of the female speaker and the atypical pronunciations of the vowel sounds. To that end, participants listened to a series of semantically neutral sentences, of which the central word was changed across trials. Their task was to determine whether that central word was either a real Dutch word or a nonword by drawing on their mental lexicon. At test, participants were presented with 100 nonwords, 60 fillers and 40 critical /ɪ/- and /i/-words. One group of participants heard all /ɪ/-words with a syllable nucleus that could not unambiguously be specified as either the /ɪ/- or /i/-vowel, while the /i/-words were pronounced canonically. A second group heard an ambiguous front vowel in all /i/-words but not in /ɪ/-words. In order to disambiguate the acoustically ambiguous sound, listeners could avail themselves of higher-level lexical knowledge to reconstruct the word that the non-native speaker may have intended to produce (see, among others, Kraljic & Samuel, 2005; McQueen et al., 2006; Norris et al., 2003).

Disregarding the individual differences between listeners, we observed that participants succeeded in accurately distinguishing real words and nonwords. With regard to the generated nonword items, participants' response accuracy approximated 90%. The fact that no ceiling performance was achieved can be taken as evidence that opting for small values associated with the OLD20 output option in *Wuggy* must have been a good criterion for selecting the nonword stimuli (see also Yarkoni et al., 2008). For the fillers, then, nearly all trials were correctly identified as a real Dutch word. However, some filler tokens were almost unanimously considered to be nonwords. To give just one example: the Dutch word *aanval* 'attack' was categorized by 91 participants as a nonword. The main reason for this exception may be due to the non-native speaker's concrete realization of the fricative /v/, which was to some extent labialized in this word, thereby making the filler sound more like [ˈaːn.wɑl] as opposed to the standard Dutch pronunciation [ˈaːn.vɑl]. Thus, the speaker's L2 accent may have imposed a burden on listeners' intelligibility of a limited number of trials, preventing these listeners from recognizing the intended target word. Despite these minor response inaccuracies, we can conclude that participants were, overall, able to separate the nonwords from real Dutch words.

Although lexical decision performance for nonwords and fillers was highly comparable across exposure conditions, the proportion of 'word'-responses and response latency for /ɪ/- and /i/-words was found to vary significantly between the two groups. In terms of word endorsement, listeners tended to classify critical words with naturally produced vowels more often as real words than items with an ambiguous syllable nucleus. The effect of ambiguity was further corroborated by the asymmetrical response latencies across the word types per exposure condition. Specifically, it took longer for participants to decide whether a trial was a real word or a nonword when the item in that trial contained an ambiguously produced instead of a naturally produced vowel. Moreover, there was also a noteworthy difference between the two exposure conditions with regard to the proportion of 'word'-responses for items with ambiguous sounds. That is, participants in the /i/-ambiguous condition considered only about 60% of the /i/-words as real words in contrast to a word endorsement rate of 86% for ambiguous /ɪ/-words in the other condition. Put differently, listeners in the former condition were less prone to consider the ambiguous vowel as a non-native accented variant of a Dutch front vowel, which would in turn complete a real Dutch word (e.g., [ˈpr?s.tər] for *priester* 'priest', /ˈpris.tər/), than the participants in /ɪ/-ambiguous group. The fact that ambiguous /i/-words were less frequently perceived as acceptable instances of lexical items which canonically have /i/ is also reflected in the number of participants per condition that was excluded from further analysis: data from thirteen participants assigned to the /i/-ambiguous group were not analyzed, as they classified more than half of the critical words as nonwords (cf. Section 5.1.4).

The between-group difference in word endorsement of the ambiguous words suggests that participants were more readily inclined to accept [ˈv?ŋ.ər] for *vinger* ('finger', /ˈvɪŋ.ər/) than [ˈd?f.stɑl] for *diefstal* ('theft', /ˈdif.stɑl/). A plausible explanation for this trend could be listeners' experience with some varieties of Belgian Dutch. Recall that people from the Brabantine region in Flanders were not allowed to participate in the present study. Since speakers raised in this area frequently produce /ɪ/-vowels with formant frequencies which would – in Standard Dutch as well as in most other accents – be expected for /i/-vowels (Simon et al., 2015; Adank et al., 2004; cf. Section 4.1.1, on participant exclusion criteria), their ability to discriminate the front vowels was expected to be different than that of speakers who do produce these front vowels with spectrally distinct properties. Interestingly, it is precisely this regional variety of Belgian Dutch that currently functions as the supraregional language variety that is spoken in the vast majority of informal television programs and soap operas (e.g., De Caluwe, 2009; Van Hoof & Vandekerckhove, 2013). As a result, even people who were born outside this region will frequently be exposed to speakers from the Brabantine area, who produce the front vowel in both *vinger* and *diefstal* with more /i/-like spectral qualities.[21] Familiarity with endogenous variation can thus be considered one potential explanatory factor for the observed discrepancy in lexical decision performance.

[21] Note that despite the lack of substantial spectral differences, these front vowels are typically produced with different durational characteristics in the Brabantine dialect, in that the /i/-vowel is frequently realized as a long vowel, contrary to Standard Dutch and other regional dialects (see Verhoeven & Van Bael, 2002).

So far, we know that participants respond differently to /ɪ/-words and /i/-words in terms of word endorsement and response latency, depending on the exposure condition to which they had been allocated. What still needs to be determined is whether accumulated experience with the non-native speaker's accented speech during the lexical decision task induced listeners to learn that the ambiguous vowel sounds are typical of the Dutch interlanguage of the Italian L1 speaker. Remind that listeners could use the surrounding lexical context in which the ambiguous front vowels were presented to update or adjust the boundaries of their mental representation for the corresponding Dutch vowel category. To verify if lexically-guided perceptual learning occurred, listeners also had to perform a phoneme categorization task (Section 6). In this second task, the acoustically ambiguous vowels were embedded in lexical contexts that could no longer function as a diagnostic for listeners to determine whether the non-native speaker might have intended to produce either /ɪ/ or /i/. Such lexically ambiguous context will drive participants to draw on their phonological representations of the Dutch front vowels.

6 Phoneme Categorization Task

This section reports on the phoneme categorization task, which set out to verify whether lexically-guided perceptual learning took place over the course of the exposure phase (Section 5). More precisely, this second task was administered to test two hypotheses. First, we wanted to examine whether listeners had learned that the ambiguous sounds in either /ɪ/-words or /i/-words are non-native accented pronunciations of the Dutch L2 speaker. If prior lexical knowledge did indeed lead participants to modify their mental phonological representations of the /ɪ/- and /i/-vowel in response to the ambiguous vowels during exposure, we would expect that participants in the /ɪ/-ambiguous condition will identify the same, and acoustically similar, ambiguous vowels in novel lexical contexts more often as an /ɪ/-vowel than participants in the other condition. Importantly, to test whether listeners had truly remapped their vowel spaces, we integrated the ambiguous vowel sounds in minimal /ɪ/-/i/ word pairs, as these lexical contexts cannot help listeners to fill in the gap of the ambiguous syllable nucleus.

If participants are indeed found to successfully transfer knowledge about the ambiguous vowel realizations in the L2 speaker's speech to interpret other words, we also wanted to explore in a second stage whether participants would use that knowledge to interpret the non-native accented speech of another talker not previously encountered. Crucially, we digitally created a male-sounding voice from the female speaker's voice, as previous studies have demonstrated that the acoustic characteristics of an unfamiliar speaker's voice need to be amply commensurate with those of the exposure talker for successful generalization of perceptual learning to take place (see Reinisch & Holt, 2014; Xie & Myers, 2017). When participants are found to transfer learning outcomes from the female speaker's ambiguous vowels to comprehend the novel talker's speech, this could be regarded as further evidence that perceptual learning not only occurs at a sublexical or prelexical level (see McQueen et al., 2006; Mitterer et al., 2011; Sjerps & McQueen, 2010), but also independently of the specific speaker to whom participants are listening.

6.1 Method

6.1.1 Participants

Participants in the phoneme categorization task were the same 100 students recruited from the Ghent student population as in the lexical decision task. More details about the participants can be found in Section 5.1.1 and Appendix 2.

6.1.2 Stimulus Materials

Test items in the phoneme categorization task were five monosyllabic Dutch word pairs that only differ minimally in terms of their syllable nucleus: *bid-bied* ('pray'-'bid'), *kist-kiest* ('case/box'-'chooses'), *lig-lieg* ('lie', 'lie/deceive'), *vis-vies* ('fish'-'dirty') and *wit-wiet* ('white'-'weed'). Word frequency between the members of each minimal word pair was matched as closely as possible to avoid a decision bias in favor of more frequently occurring words (M_{Zipf} = 4.39, SD_{Zipf} = 0.16). As outlined in greater detail in Section 4.1.2, an 11-step continuum between the /ɪ/-word and the /i/-word of the minimal pair was generated to obtain stimuli with acoustically ambiguous vowels. In a second phase, the '*change gender*' function in *Praat* software was applied to all eleven steps on the minimal word continua, converting the female speaker's voice into a perceivable male speaker's voice. Since steps in the middle of the continuum are not necessarily perceived as the most ambiguous stimuli, a pre-test was administered to identify which steps participants could not unequivocally classify as either the /ɪ/- or /i/-word of the minimal pair (cf. Section 4). Rather than a single step, four consecutive continuum steps which had approximately 50% /ɪ/-responses and 50% /i/-responses were selected as stimulus materials for the current task. Note that the acoustic qualities of the first selected step will be slightly closer towards the /ɪ/-vowel, whereas the fourth step will have slightly more spectral properties in common with a canonical /i/-vowel.

Presenting participants with multiple steps should allow us to examine whether their categorization responses were affected by the exposure condition to which they had been assigned in the lexical decision task. Influence of the exposure condition would imply that, for instance, listeners in the /ɪ/-ambiguous condition classify the ambiguous vowels more often, and thus across more continuum steps, as the /ɪ/-vowel compared to participants in the /i/-ambiguous exposure condition. For both speaker voices, steps 3 to 6 were selected as the perceptually most ambiguous tokens for seven out of ten minimal word pairs based on the results of the pre-test. For the remaining pairs, the selected steps for the *bid-bied* pair with the female speaker's voice were shifted one step further towards the /i/-side of the continuum (i.e., steps 4 through 7), whereas for *kist-kiest* as well as for *wit-wiet*, steps 5, 6, 7 and 9 were selected when they will be presented with a male-sounding voice.[22]

[22] Since participants were only presented with seven of the eleven continuum steps (i.e., steps 1, 3, 4, 5, 6, 7, 9) in order to reduce the time needed to complete the pre-test, the ninth step was selected instead of the eighth step.

6.1.3 Procedure

Immediately after completing the lexical decision task, participants proceeded to the phoneme categorization task on the experimental webpage. They were instructed to identify the vowel sound they thought they heard in a series of Dutch words. Two response options were given: they could indicate hearing the /ɪ/-vowel as in *prins* ('prince') and *begin* ('beginning') or the /i/-vowel as in *diep* ('deep') and *machine* ('machine') by respectively pressing the F-key and J-key on their keyboard. As the larger part of the participants were not enrolled in a language-oriented program, graphemic representations (i.e., <i> for /ɪ/ and <ie> for /i/) were used instead of the IPA symbols. The use of graphemes was deemed less problematic in this task compared to the pre-test, because there is a one-to-one correspondence between the phoneme and the grapheme in all the stimuli.

In the interest of investigating cross-talker generalization, half of the participants heard the same female speaker as in the lexical decision task whereas the other half heard a male-sounding speaker. Note that participants were not informed about the possible change in speaker during the second part of the experiment. Intermixing these two speaker conditions with the two exposure conditions creates four different groups: (i) /ɪ/-ambiguous-female voice (n = 26); (ii) /ɪ/-ambiguous-male voice (n = 31); (iii) /i/-ambiguous-female voice (n = 23); and (iv) /i/-ambiguous-male voice (n = 20). The unequal distribution of the number of participants per condition can again be ascribed to the exclusion of participants based on their lexical decision performance (cf. Section 5.1.4). Nevertheless, this was not considered problematic, as a total of 20 participants per group is assumed to be sufficient for reliable statistical manipulations (see Simmons et al., 2011).

Each selected continuum step for all five minimal word pairs was presented eight times to ensure that none of the potential effects could be attributed to chance. This resulted in a total of 160 trials (5 word pairs × 4 continuum steps × 8 repetitions). Running order of the trials was pseudo-randomized, in that no different steps of the same minimal word pair could directly follow each other. Moreover, all four steps of each word pair needed to be presented first before the same 20 trials would be repeated in another pseudo-randomized order. Again, participants were asked to respond as quickly and as accurately as possible. If no response had been given after 4000ms, participants were notified that they had exceeded the response time limit. After an inter-trial interval of 1000ms, the next trial started automatically. Every 40 trials, participants were allowed to take a self-paced break. It took participants on average 12 minutes to complete the phoneme categorization task.

6.1.4 Analysis

Due to an unforeseen error in the script of the categorization task, 22 participants were not presented with the four selected continuum steps of the minimal word pair *bid-bied.* This means that in the current sample, the other four minimal pairs are represented slightly more often (i.e., 704 additional observations per pair). Nevertheless, this imperfect balance between word pairs is not expected to compromise the validity of the analysis due to the sufficiently large size of the sample (n = 15,296 observations). Conform the procedure used for the lexical decision task in this study, responses were not logged when the response time limit of 4000ms was exceeded, resulting in the removal of 32 (0.2%) additional trials.

Categorization responses across the two exposure and speaker conditions were estimated by a generalized linear mixed-effects model using the *lme4* package (Bates et al., 2015) in *R* (R Core Team, 2020). The main purpose of the analysis is to determine whether familiarization with the ambiguous sounds in lexically-biasing contexts may affect how the ambiguous vowels are categorized in four consecutive steps along an [ɪ]-[i] continuum. Moreover, we wanted to examine whether perceptual learning of the ambiguous vowels in the female L2 speaker's speech allowed listeners to carry over knowledge about the acoustic properties of those vowels to interpret the L2 accented speech of a novel male-sounding speaker. Confirmatory evidence for such a transfer would be that participants' phoneme categorization performance significantly differed between exposure conditions, but not between speaker conditions. With regard to these two pivotal research questions, *exposure condition* (/ɪ/-ambiguous vs. /i/-ambiguous), *speaker condition* (female voice vs. male voice) as well as *continuum step* (step 1 vs. 2 vs. 3 vs. 4) were entered as fixed factors in the regression model and random intercepts were included for *participant* and *items.* The reference level for the intercept of the model was set to the /ɪ/-ambiguous condition and the female speaker's voice. Similar to the lexical decision task, we will report the estimated coefficients for the variables in our model (beta-values), the standard errors (SE), z-values and p-values. Note that we re-coded the original levels of *continuum step* to facilitate comparisons between minimal word pairs (e.g., steps 3 to 6, and 4 to 7, were both re-labelled as steps ranging from 1 to 4). Moreover, since we want to determine to what extent the proportion of /ɪ/- and /i/-responses changes between two consecutive steps on the 4-step continuum, forward difference coding was assigned to the dependent variable *continuum step.* By applying this coding system, we can compare more efficiently how, for instance, the proportion of /i/-responses differs between step 1 and 2, step 2 and 3, and step 3 and 4, without comparing the increase in /i/-responses between two non-adjacent continuum steps (e.g., differences between step 1 and step 4).

6.2 *Results*

Generalized over participants and minimal word pairs, Figure 10 plots the proportion of /i/-responses per continuum step. It is further broken down by the two exposure conditions (/ɪ/-ambiguous vs. /i/-ambiguous) to which participants were assigned in the lexical decision task and the two speaker conditions (female voice vs. male voice), which represent the perceptual gender of the speaker in the phoneme categorization task. As anticipated, Figure 10 shows that the steps towards the right-hand side of the continuum were increasingly perceived as the /i/-vowel, which can be concluded from the upward slopes of the trend lines across all conditions. This pattern holds for virtually all steps, except for the transition between the first and the second step in the /i/-ambiguous-female voice condition. Although participants in this condition identified the ambiguous vowel slightly more as /i/ in the first step compared to the second step, the difference in the relative frequency of /i/-responses between these continuum steps seems small enough to be considered negligible (8.6% vs. 7%). More noteworthy is that the upper limit of the y-axis was adjusted to 50% for a better visibility of the proportions of /i/-responses across the four conditions. Despite the steep movement of the slopes for the male speaker between steps 3 and 4 (see below), the fact that trend lines never cross the 50% boundary suggests that /ɪ/-responses will remain dominant across all continuum steps.

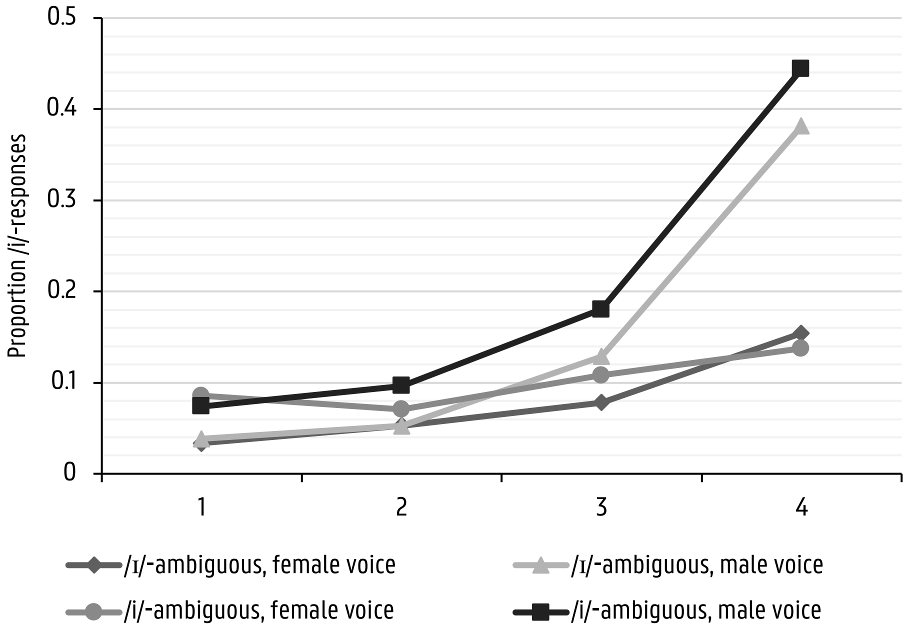

Figure 10: Proportions of the phoneme categorization responses by exposure conditions and speaker conditions across the four continuum steps.

To determine whether the increase in /i/-responses along the 4-step continuum can reliably be predicted based on the two exposure and the two speaker conditions, a linear regression model with mixed effects was fit. A log-likelihood ratio test demonstrated that an interaction between these two fixed factors did not significantly improve model goodness of fit compared to a model without interaction between exposure condition and speaker condition ($\chi^2(1) = 0.069$, $p = 0.79$). As such, participants' phoneme categorization will be analyzed using the reduced linear model (i.e., the one without an interacting term between the two conditions).

The relative predictive contribution of each of the fixed effects to the overall model yielded mixed results, as can be seen in Figure 11, which plots the increase in the probability of /i/-responses by all three factors separately. Recall that the intercept set in our model represents the categorization performance of the participants in the /ɪ/-ambiguous condition who listened to the female speaker. First, participants who heard the male-sounding voice in the categorization task were more likely to classify the ambiguous vowels in the minimal word pairs as /i/ than participants who heard the exposure speaker ($b_{\text{Intercept}} = -3.88$, $SE = 0.34$, $z = -11.35$, $p < 0.001$; $b_{\text{Speaker}} = 1.48$, $SE = 0.32$, $z = 4.67$, $p < 0.001$) (see panel A in Figure 11). Note that the positive effect for the male speaker's voice is also visually reflected in Figure 10 by the mild upward trend lines for the male speaker, especially towards the right side of the continuum. Secondly, no significant difference in effect was found between the two exposure conditions to which participants had been assigned in the lexical decision task ($b_{\text{Exposure}} = 0.45$, $SE = 0.32$, $z = 1.42$, $p = 0.15$). This means that, against our expectations, listeners did not interpret the ambiguous vowel more frequently as /i/ when they heard all /i/-words with an ambiguous vowel in the exposure phase compared to the group of listeners who heard all /ɪ/-words with an ambiguous vowel (see panel B in Figure 11). Finally, the results of the regression analysis show that each adjacent step was associated with a significantly higher chance of /i/-responses. That is, listeners were expected to indicate hearing /i/ more often than /ɪ/ between step 1 and step 2 ($b_{\text{Step1}} = -0.21$, $SE = 0.10$, $z = -2.04$, $p = 0.04$), step 2 and step 3 ($b_{\text{Step2}} = -0.78$, $SE = 0.09$, $z = -8.94$, $p < 0.001$) and step 3 and 4 ($b_{\text{Step3}} = -1.31$, $SE = 0.07$, $z = -18.92$, $p < 0.001$). Nevertheless, the likelihood of /i/-responses is at most about 20%, even for the final continuum step (see panel C in Figure 11).

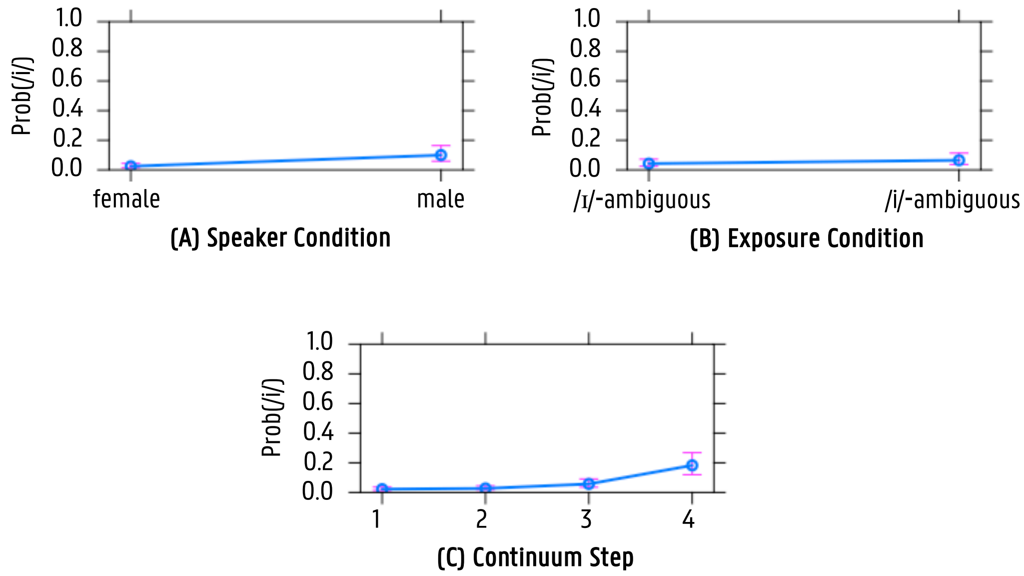

Figure 11: Probability of /i/-responses by (A) speaker condition, (B) exposure condition and (C) continuum step.

As the interaction between exposure and speaker conditions did not reach significance, we also examined whether the variability explained by the random effects could be used to clarify the patterns observed for participants' categorization responses. Participants and minimal word pairs were entered in the linear model, because their inclusion significantly improved model fit ($\chi^2(1) = 1491$, $p < 0.001$ and $\chi^2(1) = 210$, $p < 0.001$, respectively). Figure 12 plots the point estimates and the corresponding 95% confidence intervals for the random intercepts. Visual inspection of the left and right panel in Figure 12 suggests that variability caused by the idiosyncratic differences between the participants ($s^2 = 2.26$, $SD = 1.50$) was slightly larger than that caused by the differences between the five minimal word pair ($s^2 = 0.19$, $SD = 0.44$). For the by-subject intercepts, we can clearly see that the likelihood of participants perceiving the ambiguous vowel as the /i/-vowel is rather low. Notably, the majority of the intercepts are associated with a negative predicted value, which means that most participants were unlikely to identify the ambiguous sound as the /i/-vowel. What is most noticeable for the by-item intercepts, then, is that the minimal pairs *kist-kiest* and *wit-wiet* have a positive predicted value compared to zero values or a negative value for the other word pairs. This assumes that the former two minimal word pairs had a slightly higher probability of receiving /i/-responses in the categorization task. How the analysis of the random effects ties in with the complete regression analysis will be discussed in greater detail in the next section.

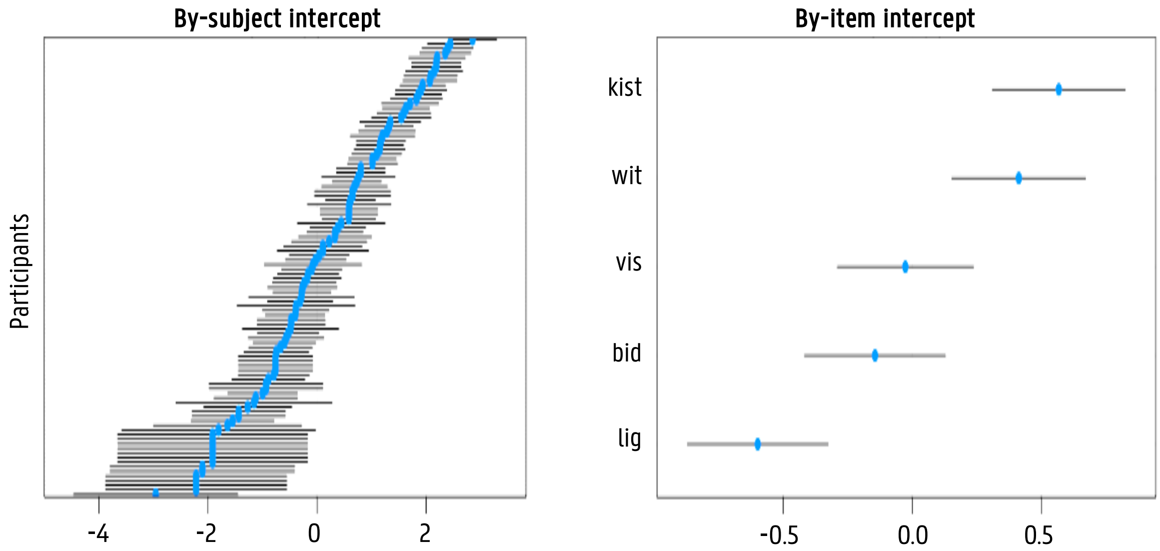

Figure 12: Caterpillar plots of the 95% confidence intervals for by-item and by-subject intercepts of the random effects.

6.3 Discussion

The central objective of this task was to explore whether short-term exposure to a non-canonical pronunciation variant of one member of the Dutch /ɪ/-/i/ contrast influences how listeners subsequently process novel words containing a similar ambiguous vowel when they are produced by either the same or an unfamiliar L2 speaker. Crucially, while atypical vowel sounds were integrated into lexically unambiguous contexts in the familiarization phase of the experiment (cf. Section 5), those vowels were embedded in lexically ambiguous contexts (i.e., minimal word pairs) in this second task to ensure that listeners would not be able to use lexical knowledge to determine which vowel canonically fills the syllable nucleus. Contrary to our hypotheses, the results of the categorization test showed that lexically-driven perceptual adaptation of the /ɪ/-/i/ category boundaries did not develop over the course of the lexical decision task. As no learning was observed for the female speaker, it is ruled out that listeners could generalize knowledge to interpret the non-native accented speech of a novel talker. This conclusion is primarily drawn on the following two findings. First, the exposure condition (/ɪ/-ambiguous vs. /i/-ambiguous) was not found to influence how participants identified the ambiguous vowels in the categorization task. Secondly, the absence of a significant interaction between exposure and speaker conditions could be regarded as further evidence that perceptual learning of the ambiguous vowels was impeded for both the exposure speaker and for the novel male-sounding speaker. Taken together, these outcomes suggest that exposure to lexical items with an acoustically ambiguous vowel did not lead listeners to adjust their stored linguistic representations of the front vowel.

The lack of generalization to novel situations prompts us to reflect on which factors might have prevented listeners from successful perceptual learning. Let us first consider the relation between exposure condition and categorization response. According to the perceptual learning paradigm, participants in /ɪ/-ambiguous condition are expected to identify the ambiguous speech segment in the categorization task more often as /ɪ/, because they adjusted their perceptual representation of the /ɪ/-category in response to the atypical pronunciations in the auditory input (see Norris et al., 2003; cf. Section 2.3). Participants in the /i/-ambiguous condition are in turn expected to give more /i/-responses. Although the categorization performance of the participants in the /ɪ/-ambiguous condition matches the expectation (i.e., exposure to ambiguous /ɪ/ results in more /ɪ/-responses), the trend observed for the categorization responses of the group of listeners in the /i/-ambiguous condition is diametrically opposed to this pattern. In fact, the response behavior of the participants in this second condition aligns more closely with the selective adaptation approach towards speech perception (see Eimas & Corbit, 1973; Kraljic & Samuel, 2005; cf. Section

2.3). This alternative paradigm assumes that repeated exposure to an ambiguous realization of one member of a speech contrast will increase the likelihood that listeners will give more responses of the opposite member of the contrast (i.e., exposure to ambiguous /i/ results in more /ɪ/-responses). Nevertheless, it seems highly unlikely that within the same experiment perceptual learning effects develop for one group of participants whereas selective adaptation outcomes are found for another group, as this has to our knowledge never been observed in previous studies on lexically-guided perceptual learning.[23]

Additionally, it might also be of particular interest to bring the between-group differences observed in the lexical decision task into the discussion of the phoneme categorization task. Recall that participants in the /i/-ambiguous condition were more reluctant to accept the ambiguous vowel in lexically disambiguating contexts as an atypical pronunciation of the /i/-vowel (cf. Section 5). On a *Perceptual Assimilation Model* account (Best, 1995), this signifies that, in contrast to the canonically produced front vowel in /ɪ/-words, listeners did not map the ambiguous sound onto the L1 category that would complete a real word in Dutch (cf. Section 2.2.2). That is, participants did not always consider the ambiguous vowel in tokens like [ˈspʔ.ɣəl] as an atypical realization of /i/, assuming that they would otherwise have endorsed it as a real word (i.e., *spiegel* 'mirror'). Listeners in the /ɪ/-ambiguous condition, conversely, assimilated the female speaker's canonical /i/-vowel as well as the ambiguous /ɪ/-vowel in more than 85% of the trials to the corresponding L1 phonological category. This disparity in assimilation could indicate that the ambiguous vowels in either the lexical decision task or the phoneme categorization task did not consist of sufficient /i/-like spectral qualities. Note that one spectral cue for distinguishing Dutch /ɪ/- and /i/-vowels is the frequencies of the first formant (F1), which is the acoustic correlate of vowel height (see Di Benedetto, 1989). As Boersma and Chládková (2011) showed, Dutch listeners' perceptual boundaries between two vowel categories are mainly based on differences in vowel height (i.e., horizontal boundaries). If we apply this observation to the front vowel contrast in our study, this may suggest that the F1 frequencies of the large majority of the ambiguous speech sounds were not sufficiently high to lead listeners to categorize the sound as the high vowel /i/. As a result, all tokens that did not reach the height target typically associated with /i/ may have been perceived as non-high, and hence not /i/. This may in turn have created a response bias in favor of the /ɪ/-variant of the minimal word pair.

[23] Vroomen et al. (2004), who examined visually-guided perceptual learning, did observe a change from perceptual learning to selective adaptation effects *within* the same condition, but not *between* conditions as in the present study.

The possibility that the stimulus materials were perceptually slightly closer to /ɪ/ than /i/ seems to be further supported by the difference in categorization responses across the speaker conditions and by the analysis of the by-item intercepts. When participants heard a male-sounding voice during the second part of the experiment, they categorized the ambiguous vowel more frequently as the /i/-vowel. Although this was found to be true for the steps towards the right endpoint of the 4-step continuum (see Figure 10), the proportion of /ɪ/-responses for the ambiguous vowel in the minimal word pairs remained almost the same across all steps when we pooled the categorization responses from participants in different exposure and speaker conditions (see panel C in Figure 11). What needs to be addressed here is that the original numbers of the continuum steps were re-coded as ranging from one to four to facilitate the comparison of the /ɪ/-/i/ proportions for the five minimal word pairs with either a female or male voice (cf. Section 6.1.4). Based on the results of the pre-test, steps 5 through 9 were selected for two of the five minimal word pairs presented in the speaker condition with the generated male voice (cf. Section 4). Seeing that step 9 is the final step before the continuum endpoint, and thus very close to natural productions of /i/, this may have driven listeners to hear more of an /i/-vowel in those steps in the male condition. This reasoning seems to be supported by the analysis of the random effects, in that these two minimal pairs were associated with a higher probability of /i/-responses than the three other word pairs, for which more steps towards the left-hand side of the original 11-step continuum were selected. These and other complications will be considered in the general discussion section below.

7 General Discussion

The present experiment was designed to explore whether native Dutch listeners adjust their perceptual boundaries between two Dutch front vowels after exposure to atypical pronunciations of one of these vowels in the non-native accented speech of an L2 speaker. Specifically, we wanted to gain insight into how these listeners overcome initial processing difficulties arising from a priori unexpected pronunciations that could be considered a typical feature of the Dutch interlanguage of non-native speakers. As cumulative experience with a particular speaker's accent has been shown to guide listeners to perceptually assimilate the atypically produced speech sounds to the intended speech segments (e.g., Drozdova et al., 2016; Norris et al., 2003; Reinisch & Holt, 2014; Trude et al., 2013; Tzeng et al., 2021), we familiarized participants in the present study with the accent of an Italian speaker of Dutch. Building on previous research on perceptual adaptation, we investigated how listeners can update their long-term representations of particular speech categories through the activation of higher-level lexical knowledge. Such perceptual adjustments could ultimately facilitate the recognition of other words with similar ambiguous sounds (see McQueen et al., 2006). Against this background, we aimed to answer the following two research questions. First, do listeners accommodate ambiguous pronunciations of the Dutch front vowels, allowing them to interpret novel words containing acoustically similar ambiguous vowel sounds? Secondly, if listeners are indeed found to have learned that the ambiguous productions may be typical of the L2 speaker's Dutch interlanguage, would these listeners also be able to generalize knowledge about the acoustic features of the non-native accented vowels, allowing them to interpret the accented speech of another L2 speaker with a similar accent?

Neither for the first nor the second research question formulated above did we obtain evidence that exposure-driven perceptual adaptation had occurred. That is, the ambiguous vowel sounds in the phoneme categorization task were identified in the vast majority of the trials as the /ɪ/-vowel, regardless of whether participants had been exposed to /ɪ/-words (e.g., *winkel* 'shop') or /i/-words (e.g., *fiets* 'bike') with an acoustically ambiguous syllable nucleus. The absence of perceptual learning was, in fact, already foreshadowed to some extent in the results of the lexical decision task. Specifically, ambiguous /i/-words were rejected noticeably more often than ambiguous /ɪ/-words (37% vs. 13%, respectively). This disparity shows that participants were less readily inclined to accept the ambiguous vowel in /i/-words as a non-native accented production of the /i/-vowel, based on the assumption that participants would otherwise have endorsed those tokens as real words. Labelling /i/-words with an ambiguous vowel as nonwords suggests that participants did not relax their perceptual boundaries of what falls within the scope of acceptable pronunciations of the /i/-vowel. With regard to the second research question, then, cross-talker generalization of the learning outcomes could not be reliably assessed in the current experiment, because generalization to novel talkers is premised on the idea that listeners first update their perceptual system in response to the L2 speaker they perceived in the exposure phase. Phrased differently, it is impossible to determine whether shifts in phonological representations remain speaker-specific, or if they are speaker-independent based on the present findings. In what follows, we will touch upon some factors which may have blocked learning and discuss how the limitations of this study could be addressed and overcome in a follow-up study.

As already briefly noted earlier, participants may have brought prior experience with how native and non-native speakers of Dutch vary in the concrete realization of the critical vowel contrast to the experiment. As specified in the recruitment protocol, participants were not allowed to be learners or fluent L2 speakers of Italian or Spanish (cf. Section 4.1.1). Although participants did thus not have prior knowledge about the sounds systems of these Romance languages, we cannot rule out that they were to some extent familiar with Italian- or Spanish-accented Dutch. When these Dutch L1 speakers had already previously communicated with Spanish or Italian learners of Dutch, they may have observed that these non-native speakers tend to produce the /ɪ/-vowel with /i/-like spectral qualities, because only the /i/-vowel is part of the Italian and Spanish phoneme inventory (Coe, 2001; Duguid, 2001; see also Weber et al., 2014, on English speakers' perception of Italian accented /ɪ/-/iː/ vowels). [24] On top of prior experience with exogenous linguistic variation, we also need to take into account the possible effect of endogenous variation (cf. Section 5.3). For the purposes of our experiment, no participants who were born or raised in the provinces of Antwerp or Flemish-Brabant were included, as the concrete realization of both front vowels in the regiolects of those provinces leans more towards what would in Standard Dutch and most other regional accents be regarded as the /i/-vowel (Adank et al., 2004; Simon et al., 2015). Recall that especially this L1 variety is the dominating supraregional colloquial language used in the present-day media landscape in Flanders (e.g., De Caluwe, 2009; Van Hoof & Vandekerckhove, 2013). Due to its ubiquitous presence, people born outside this region are thus frequently exposed to speakers saying [ˈvis] rather than the standard pronunciation [ˈvɪs] for *vis* ('fish'). At this point, we cannot fully ascertain to what extent these causal assumptions can account for the absence of perceptual learning. Adding an extra set of questions on participants' experience with both endogenous and exogenous variation at the end of the experiment could potentially shed further light on this.

[24] Note also the omnipresence of French-accented Dutch in the media in Flanders.

Another potential explanation for the results obtained in the current experiment might be related to the data collection method. While participants performed the experiment in a sound-damped booth in a laboratory setting in Norris et al.'s (2003) study and many follow-up studies, participants in the present study performed the lexical decision and the phoneme categorization tasks on an experimental web page without experimenter supervision. It is, however, very unlikely that this difference in test setting is responsible for the absence of perceptual adaption (see Germine et al., 2012). Earlier studies on speech perception in general (e.g., Burchill et al., 2018; Byun et al., 2015; Kunath & Weinberger, 2010) as well as studies on perceptual learning of particular speech segments (e.g., Kleinschmidt & Jaeger, 2015; Liu & Jaeger, 2018, 2019; Xie et al., 2021), which also used a web-based design, did succeed in reaching similar results as those reported in laboratory-run studies. Listeners in Liu and Jaeger's (2018) experiment, for instance, accomplished perceptual learning after exposure to atypical pronunciations of fricatives. As the spectral properties for distinguishing the individual members of a vowel contrast are considered to be perceptually more salient than those for a fricative sound contrast (see Escudero et al., 2009; Weatherholtz, 2015), listeners would be expected to perceive acoustically distinct vowel sounds in the auditory speech input, even if they did not listen to the audio in a sound-treated booth. We therefore conclude that the difference in test environment cannot be the explanation par excellence why no significant effect of exposure condition was found on participants' phoneme categorization performance.

What may have been slightly more problematic than running the experiment online is the set of stimulus materials selected to assess perceptual learning. Although a pre-test was administered to determine which tokens consisted of the perceptually most ambiguous /ɪ/- and /i/-vowels, there may have been an unforeseen /ɪ/-bias in the stimuli presented during the phoneme categorization task. That is, the ambiguous vowel sound may have had more /ɪ/-like acoustic properties in more than half of the steps on the four-step continua for minimal word pairs. This may in turn have impeded listeners in the /i/-ambiguous condition from retuning their mental representation of the /i/-vowel after repeated exposure to the ambiguous pronunciations (cf. Section 6.3). A follow-up experiment is therefore necessary to reveal if this presumed methodological issue did indeed block perceptual learning. In that study, the selected continuum steps for the second task could be replaced by steps that are located slightly more towards the right-hand side of the original 11-step continuum between the canonical [ɪ] and [i] endpoints. Ultimately, such an adjustment would then allow us to formulate more conclusive answers to our research questions.

8 Conclusion

The current experimental study set out to examine how native speakers of a particular language accomplish robust speech perception when they are confronted with the L2 accented speech of a non-native speaker. Additionally, we also investigated whether listeners might use knowledge about one speaker's pronunciation variation to comprehend the non-native accented utterances of another L2 speaker. This twofold goal was addressed by familiarizing Dutch L1 speakers with a series of non-native accented sentences, which were produced by a female native speaker of Italian. Crucially, the front vowel in lexical items which canonically have /ɪ/ (e.g., *winkel* 'shop') or /i/ (e.g., *liefde* 'love') as the nucleus of the stressed syllable was artificially manipulated, thereby resulting in a sound that could not unequivocally be identified as either /ɪ/ or /i/ as established by a pre-test (cf. Section 4). Note that half of the participants heard /ɪ/-words with an ambiguous vowel but naturally produced /i/-words, while the other half heard the exact opposite. These /ɪ/-words and /i/-words were presented along with nonwords and filler items in the lexical decision part of the experiment (cf. Section 5). In such a comprehension-oriented task, participants were expected to draw on prior lexical knowledge. Activating the mental lexicon has repeatedly been shown as an effective source of information to disambiguate acoustically ambiguous speech sounds (e.g., Maye et al., 2008; Norris et al., 2003; Reinisch & Holt, 2014). Immediately after the lexical decision task, participants also performed a phoneme categorization task, of which the results could show whether short-term experience with the acoustic variation in the non-native speaker's speech drove listeners to adjust the boundaries of their front vowel categories (cf. Section 6).

Based on participants' phoneme categorization performance, we concluded that lexically-guided perceptual learning of the ambiguous Dutch front vowels, and thus also the generalization of the learning outcomes to novel situations, did not occur. Specifically, participants were not found to interpret the ambiguous vowels in the minimal word pairs differently when they heard acoustically ambiguous variants of either /ɪ/-vowels or /i/-vowels in the exposure phase. This automatically ruled out that these listeners would be able to carry over perceptual learning effects to an unfamiliar, male-sounding non-native speaker. Taken together, the Dutch L1 listeners in the current experiment did not adjust their perceptual boundaries between the mental representations of the front vowels in response to the ambiguous vowel sounds in the Italian speaker's interlanguage. A brief reflection on the experimental design adopted in the present study suggested that, against our expectations, some of the ambiguous /i/-tokens may have been spectrally too close to the /ɪ/-vowel to be regarded by listeners as an atypical realization of the /i/-vowel. Moreover, the effect of prior experience with atypical pronunciations of the Dutch front vowels, be it in regionally accented or non-native accented speech, should also not be overlooked (cf. Section 7). These final notes emphasize that for future and follow-up research on speech perception in general and non-native accented speech perception in particular, it is of paramount importance to meticulously select the ambiguous stimulus materials, and that asking participants to elaborate slightly more on their linguistic background could potentially allow us to gain a better insight into how ambiguous speech sounds are truly perceived.

9 References

Adank, P., van Hout, R., & Smits, R. (2004). An acoustic description of the vowels of Northern and Southern Standard Dutch. *The Journal of the Acoustical Society of America*, *116*(3), 1729–1738. https://doi.org/10.1121/1.1779271

Alexander, J. D. E., & Nygaard, L. C. (2019). Specificity and generalization in perceptual adaptation to accented speech. *The Journal of the Acoustical Society of America, 145*(6), 3382–3398. https://doi.org/10.1121/1.5110302

Antoniou, M., & Wong, P. C. M. (2015). Poor phonetic perceivers are affected by cognitive load when resolving talker variability. *The Journal of the Acoustical Society of America*, *138*(2), 571–574. https://doi.org/10.1121/1.4923362

Antoniou, M., & Wong, P. C. M. (2016). Varying irrelevant phonetic features hinders learning of the feature being trained. *Journal of the Acoustical Society of America, 139*(1), 271–278. https://doi.org/10.1121/1.4939736

Aoyama, K., Flege, J. E., Guion, S. G., Akahane-Yamada, R., & Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics*, *32*(2), 233–250. https://doi.org/10.1016/S0095-4470(03)00036-6

Audacity Team (2018). *Audacity* (Version 2.4.2) [Computer software]. https://www.audacityteam.org

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R.* Cambridge University Press. https://doi.org/10.1017/CBO9780511801686

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Babel, M., & Munson, B. (2014). Producing socially meaningful linguistic variation. In M. Goldrick, V. Ferreira, & M. Miozzo (Eds.), *The Oxford handbook of language production* (pp. 308–325). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199735471.013.022

Baese-Berk, M. M., Bradlow, A. R., & Wright, B. A. (2013). Accent-independent adaptation to foreign accented speech. *The Journal of the Acoustical Society of America*, *133*(3), EL174–EL180. https://doi.org/10.1121/1.4789864

Banai, K., & Lavie, L. (2021). Rapid perceptual learning and individual differences in speech perception: The good, the bad, and the sad. *Auditory Perception & Cognition.* Advance online publication. https://doi.org/10.1080/25742442.2021.1909400

Barriuso, T. A., & Hayes-Harb, R. (2018). High variability phonetic training as a bridge from research to practice. *The CATESOL Journal, 30*(1), 177–194.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bertelson, P., Vroomen, J., & de Gelder, B. (2003). Visual recalibration of auditory speech identification: A McGurk aftereffect. *Psychological Science*, *14*(6), 592–597. https://doi.org/10.1046/j.0956-7976.2003.psci_1470.x

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 research* (pp. 171–204). York Press.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. 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.

Bidelman, G. M., Moreno, S., & Alain, C. (2013). Tracing the emergence of categorical speech perception in the human auditory system. *NeuroImage, 79*(1), 201–212. http://dx.doi.org/10.1016/j.neuroimage.2013.04.093

Boersma, P., & Chládková, K. (2011). Asymmetries between speech perception and production reveal phonological structure. *Proceedings of the International Congress of Phonetic Sciences XVII, Hong Kong,* 328–331.

Boersma, P., & Weeninck, D. (2019). *Praat: Doing phonetics by computer* (Version 6.1.10) [Computer software]. http://www.praat.org/

Bowers, J. S., Kazanina, N., & Andermane, N. (2016). Spoken word identification involves accessing position invariant phoneme representations. *Journal of Memory and Language*, *87*(1), 71–83. https://doi.org/10.1016/j.jml.2015.11.002

Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition*, *106*(2), 707–729. https://doi.org/10.1016/j.cognition.2007.04.005

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

Burchill, Z., Liu, L., & Jaeger, T. F. (2018). Maintaining information about speech input during accent adaptation. *PLoS ONE, 13*(8), e0199358. https://doi.org/10.1371/journal.pone.0199358

Byun, T. M., Halpin, P. F., & Szeredi, D. (2015). Online crowdsourcing for efficient rating of speech: A validation study. *Journal of Communication Disorders, 53*(1), 70–83. http://dx.doi.org/10.1016/j.jcomdis.2014.11.003

Carlson, H. K., & McHenry, M. A. (2006). Effect of accent and dialect on employability. *Journal of Employment Counseling, 43*(1), 70–83.

Chodroff, E., & Wilson, C. (2020). Acoustic–phonetic and auditory mechanisms of adaptation in the perception of sibilant fricatives. *Attention, Perception, and Psychophysics*, *82*(4), 2027–2048. https://doi.org/10.3758/s13414-019-01894-2

Clarke-Davidson, C. M., Luce, P. A., & Sawusch, J. R. (2008). Does perceptual learning in speech reflect changes in phonetic category representation or decision bias? *Perception and Psychophysics*, *70*(4), 604–618. https://doi.org/10.3758/PP.70.4.604

Clarke, C. M., & Garrett, M. F. (2004). Rapid adaptation to foreign-accented English. *The Journal of the Acoustical Society of America*, *116*(6), 3647–3658. https://doi.org/10.1121/1.1815131

Clayards, M., Tanenhaus, M. K., Aslin, R. N., & Jacobs, R. A. (2008). Perception of speech reflects optimal use of probabilistic speech cues. *Cognition*, *108*(3), 804–809. https://doi.org/10.1016/j.cognition.2008.04.004

Coe, N. (2001). Speakers of Spanish and Catalan. In M. Swan & B. Smith (Eds.), *Learner English: A teacher's guide to interference and other problems* (2nd ed., pp. 90–112). Cambridge University Press.

De Caluwe, J. (2009). Tussentaal wordt omgangstaal in Vlaanderen. *Nederlandse Taalkunde, 14*(1), 8–25.

de Leeuw, E., & Celata, C. (2019). Plasticity of native phonetic and phonological domains in the context of bilingualism. *Journal of Phonetics*, *75*(1), 88–93. https://doi.org/10.1016/j.wocn.2019.05.003

de Leeuw, J. R. (2015). jsPysch: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods, 47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

Deprez-Sims, A.-S., & Morris, S. B. (2010). Accents in the workplace: Their effects during a job interview. *International Journal of Psychology, 45*(6), 417–426. https://doi.org/10.1080/00207594.2010.499950

Di Benedetto, M.-G. (1989). Frequency and time variations of the first formant: Properties relevant to the perception of vowel height. *The Journal of the Acoustical Society of America, 86*(1), 67–77.

Drouin, J. R., & Theodore, R. M. (2018). Lexically guided perceptual learning is robust to task-based changes in listening strategy. *The Journal of the Acoustical Society of America*, *144*(2), 1089–1099. https://doi.org/10.1121/1.5047672

Drozdova, P., Van Hout, R., & Scharenborg, O. (2016). Lexically-guided perceptual learning in non-native listening. *Bilingualism: Language and Cognition, 19*(5), 914-920. http://dx.doi.org/10.1017/S136672891600002X

Duguid, A. (2001). Italian speakers. In M. Swan & B. Smith (Eds.), *Learner English: A teacher's guide to interference and other problems* (2nd ed., pp. 73–89). Cambridge University Press.

Eimas, P. D., & Corbit, J. D. (1973). Selective adaptation of linguistic feature detectors. *Cognitive Psychology, 4*(1), 99-109.

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, *171*(1), 303–306.

Eisner, F., & McQueen, J. M. (2005). The specificity of perceptual learning in speech processing. *Perception and Psychophysics*, *67*(2), 224–238. https://doi.org/10.3758/BF03206487

Erb, J., Henry, M. J., Eisner, F., & Obleser, J. (2013). The brain dynamics of rapid perceptual adaptation to adverse listening conditions. *Journal of Neuroscience*, *33*(26), 10688–10697. https://doi.org/10.1523/JNEUROSCI.4596-12.2013

Escudero, P., Benders, T., & Lipski, S. C. (2009). Native, non-native and L2 perceptual cue weighting for Dutch vowels: The case of Dutch, German, and Spanish listeners. *Journal of Phonetics, 37*(1), 452–465. https://doi.org/10.1016/j.wocn.2009.07.006

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), 1–19. https://doi.org/10.1016/j.wocn.2018.05.003

Feng, C., Wang, H., Lu, N., Chen, T., He, H., Lu, Y., & Tu, X. M. (2014). Log-transformation and its implications for data analysis. *Shanghai Archives of Psychiatry, 26*(2), 105–109. http://dx.doi.org/10.3969/j.issn.1002-0829.2014.02.009

Flege, J. E. (1988). The production and perception of foreign language speech sounds. In H. Winitz (Ed.), *Human communication and its disorders: A review 1988* (pp. 224–401). Ablex.

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. 233–277). York Press.

Flege, J. E., & Bohn, O.-S. (2021). The revised Speech Learning Model (SLM-r). In R. Wayland (Ed.), *Second language speech learning: Theoretical and empirical progress* (pp. 3–83). Cambridge University Press. https://doi.org/10.1017/9781108886901.003

Flege, J. E., & Hammond, R. M. (1982). Mimicry of non-distinctive phonetic differences between language varieties. *Studies in Second Language Acquisition*, *5*(1), 1–17. https://doi.org/10.1017/S0272263100004563

Flege, J. E., & MacKay, I. R. A. (2004). Perceiving vowels in a second language. *Studies in Second Language Acquisition*, *26*(1), 1–34.

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

Fort, M., Brusini, P., Carbajal, M. J., Sun, Y., & Peperkamp, S. (2017). A novel form of perceptual attunement: Context-dependent perception of a native contrast in 14-month-old infants. *Developmental Cognitive Neuroscience*, *26*(1), 45–51. https://doi.org/10.1016/j.dcn.2017.04.006

Fuertes, J. N., Gottdiener, W. H., Martin, H., Gilbert, T. C., & Giles, H. (2012). A meta-analysis of the effects of speakers' accents on interpersonal evaluations. *European Journal of Social Psychology, 42*(1), 120–133. https://doi.org/10.1002/ejsp.862

Fuhrmeister, P., & Myers, E. B. (2017). Non-native phonetic learning is destabilized by exposure to phonological variability before and after training. *The Journal of the Acoustical Society of America*, *142*(5), EL448–EL454. https://doi.org/10.1121/1.5009688

Ganong, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance, 6*(1), 110–125. https://doi.org/10.1037/0096-1523.6.1.110

Germine, L., Nakayama, K., Duchaine, B. C., Chabris, C. F., Chatterjee, G., & Wilmer, J. B. (2012). Is the Web as good as the lab? Comparable performance from Web and lab in cognitive/perceptual experiments. *Psychonomic Bulletin and Review, 19*(1), 847–857. https://doi.org/10.3758/s13423-012-0296-9

Gibson, J. J., & Gibson, E. J. (1955). Perceptual learning: Differentiation or enrichment? *Psychological Review*, *62*(1), 32–41.

Gluszek, A., & Dovidio, J. F. (2010). The way *they* speak: A social psychological perspective on the stigma of nonnative accents in communication. *Personality and Social Psychology Review, 14*(2), 214–237. https://doi.org/10.1177/1088868309359288

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology*, *105*(1), 585–612.

Golestani, N., & Zatorre, R. J. (2009). Individual differences in the acquisition of second language phonology. *Brain and Language*, *109*(2–3), 55–67. https://doi.org/10.1016/j.bandl.2008.01.005

Guediche, S., Blumstein, S. E., Fiez, J. A., & Holt, L. L. (2014). Speech perception under adverse conditions: Insights from behavioral, computational, and neuroscience research. *Frontiers in Systems Neuroscience*, *7*(1), 1–16. https://doi.org/10.3389/fnsys.2013.00126

Guediche, S., Fiez, J. A., & Holt, L. L. (2016). Adaptive plasticity in speech perception: Effects of external information and internal predictions. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(7), 1048–1059. https://doi.org/10.1037/xhp0000196

Hervais-Adelman, A., Davis, M. H., Johnsrude, I. S., & Carlyon, R. P. (2008). Perceptual learning of noise vocoded words: Effects of feedback and lexicality. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(2), 460–474. https://doi.org/10.1037/0096-1523.34.2.460

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of Acoustics Society of America*, *97*(5), 3099–3111.

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications* (2nd ed.). Routledge.

Huyck, J. J., Smith, R. H., Hawkins, S., & Johnsrude, I. S. (2017). Generalization of perceptual learning of degraded speech across talkers. *Journal of Speech, Language, and Hearing Research*, *60*(11), 3334–3341. https://doi.org/10.1044/2017_JSLHR-H-16-0300

Idemaru, K., & Holt, L. L. (2020). Generalization of dimension-based statistical learning. *Attention, Perception, and Psychophysics*, *82*(4), 1744–1762. https://doi.org/10.3758/s13414-019-01956-5

Idemaru, K., & Holt, L. L. (2020). Specificity of dimension-based statistical learning in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 40*(3), 1009–1021. https://doi.org/10.1037/a0035269

Iverson, P., Wagner, A., & Rosen, S. (2016). Effects of language experience on pre-categorical perception: Distinguishing general from specialized processes in speech perception. *The Journal of the Acoustical Society of America*, *139*(4), 1799–1809. https://doi.org/10.1121/1.4944755

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*(1), 434–446. https://doi.org/10.1016/j.jml.2007.11.007

Jaspal, R. (2009). Language and social identity: A psychosocial approach. *Psych-Talk, 64*(1)*,* 17–20.

Jesse, A., & McQueen, J. M. (2011). Positional effects in the lexical retuning of speech perception. *Psychonomic Bulletin and Review*, *18*(5), 943–950. https://doi.org/10.3758/s13423-011-0129-2

Jongman, A., Wayland, R., & Wong, S. (1998). Acoustic characteristics of English fricatives. *Working Papers of the Cornell Phonetics Laboratory*, *12*(3), 195–205. http://conf.ling.cornell.edu/plab/paper/wpcpl12-Jongman.pdf

Kawahara, H., Masuda-Katsuse, I., & de Cheveigné, A. (1999). Restructuring speech representations using a pitch-adaptive time-frequency smoothing and an instantaneous-frequency-based F0 extraction: Possible role of a repetitive structure in sounds. *Speech Communication*, *27*(3), 187–207. https://doi.org/10.1016/S0167-6393(98)00085-5

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods*, *42*(3), 627–633. https://doi.org/10.3758/BRM.42.3.627

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods*, *42*(3), 643-650. https://doi.org/10.3758/BRM.42.3.643

Kim, D., Clayards, M., & Kong, E. J. (2020). Individual differences in perceptual adaptation to unfamiliar phonetic categories. *Journal of Phonetics*, *81*(1), 100984. https://doi.org/10.1016/j.wocn.2020.100984

Kleinschmidt, D. F., & Jaeger, T. F. (2015). Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological Review*, *122*(2), 148–203. https://doi.org/10.1037/a0038695

Kraljic, T., & Samuel, A. G. (2005). Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology*, *51*(2), 141–178. https://doi.org/10.1016/j.cogpsych.2005.05.001

Kraljic, T., & Samuel, A. G. (2006). Generalization in perceptual learning for speech. *Psychonomic Bulletin and Review*, *13*(2), 262–268. https://doi.org/10.3758/BF03193841

Kraljic, T., & Samuel, A. G. (2007). Perceptual adjustments to multiple speakers. *Journal of Memory and Language*, *56*(1), 1–15. https://doi.org/10.1016/j.jml.2006.07.010

Kraljic, T., Brennan, S. E., & Samuel, A. G. (2008). Accommodating variation: Dialects, idiolects, and speech processing. *Cognition*, *107*(1), 54–81. https://doi.org/10.1016/j.cognition.2007.07.013

Krämer, M. (2009). *The phonology of Italian.* Oxford University Press.

Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*, *9*(2). https://doi.org/10.1111/j.1467-7687.2006.00468.x

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science, 255*(1), 606–608.

Kunath, S. A., & Weinberger, S. H. (2010). The wisdom of the crowd's ear: Speech accent rating and annotation with Amazon Mechanical Turk. *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk*, *Los Angeles, CA, USA,* 168–171.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). *lmerTest* package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Labov, W. (1972). *Sociolinguistic patterns*. University of Pennsylvania Press.

Lawrence, W. P. (2000). /str/ → /ʃtr/: Assimilation at a distance? *American Speech, 75*(1), 82–87.

LeGendre, S. J., Liss, J. M., & Lotto, A. (2009). Discriminating dysarthria type and predicting intelligibility from amplitude modulation spectra. *Journal of Acoustics Society of America*, *125*(4), 2530. https://doi.org/https://doi.org/10.1121/1.4783544

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

Liss, J. M., Spitzer, S. M., Caviness, J. N., & Adler, C. (2002). The effects of familiarization on intelligibility and lexical segmentation in hypokinetic and ataxic dysarthria. *The Journal of the Acoustical Society of America*, *112*(6), 3022–3030. https://doi.org/10.1121/1.1515793

Liu, L., & Jaeger, T. F. (2018). Inferring causes during speech perception. *Cognition*, *174*(1), 55–70. https://doi.org/10.1016/j.cognition.2018.01.003

Liu, L., & Jaeger, T. F. (2019). Talker-specific pronunciation or speech error? Discounting (or not) atypical pronunciations during speech perception. *Journal of Experimental Psychology:*

Human Perception and Performance, 45(12), 1562-1588. http://dx.doi.org/10.1037/xhp0000693

Liu, R., & Holt, L. L. (2015). Dimension-based statistical learning of vowels. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(6), 1783–1798. https://doi.org/https://doi.org/10.1037/xhp0000092

Llompart, M., & Reinisch, E. (2019). Robustness of phonolexical representations relates to phonetic flexibility for difficult second language sound contrasts. *Bilingualism*, *22*(5), 1085–1100. https://doi.org/10.1017/S1366728918000925

Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods, 49*(1), 1494–1502. https://dx.doi.org/10.3758/s13428-016-0809-y

Macmillan, N. A. (2002). Signal detection theory. In J. Wixted & H. Pashler (Eds.), *Stevens' handbook of experimental psychology, 3rd Edition, Vol. 4: Methodology in experimental psychology* (pp. 43–90). John Wiley & Sons.

Macmillan, N. A., & Creelman, D. C. (2005). *Detection theory: A user's guide* (2nd ed.). Lawrence Erlbaum Associates.

Maye, J., Aslin, R. N., & Tanenhaus, M. K. (2008). The weckud wetch of the wast: Lexical adaptation to a novel accent. *Cognitive Science*, *32*(3), 543–562. https://doi.org/10.1080/03640210802035357

McAuliffe, M., & Babel, M. (2016). Stimulus-directed attention attenuates lexically-guided perceptual learning. *The Journal of the Acoustical Society of America*, *140*(3), 1727. https://doi.org/https://doi.org/10.1121/1.4962529

McGurk, H., & Macdonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*(5588), 746–748. https://doi.org/10.1038/264746a0

McMurray, B., & Jongman, A. (2011). What information is necessary for speech categorization? Harnessing variability in the speech signal by integrating cues computed relative to expectations. *Psychological Review*, *118*(2), 219–246. https://doi.org/10.1037/a0022325

McQueen, J. M., Cutler, A., & Norris, D. (2006). Phonological abstraction in the mental lexicon. *Cognitive Science*, *30*(6), 1113–1126. https://doi.org/10.1207/s15516709cog0000_79

Mitterer, H., & Reinisch, E. (2013). No delays in application of perceptual learning in speech recognition: Evidence from eye tracking. *Journal of Memory and Language*, *69*(4), 527–545. https://doi.org/10.1016/j.jml.2013.07.002

Mitterer, H., Chen, Y., & Zhou, X. (2011). Phonological abstraction in processing lexical-tone variation: Evidence from a learning paradigm. *Cognitive Science*, *35*(1), 184–197. https://doi.org/10.1111/j.1551-6709.2010.01140.x

Mitterer, H., Eger, N. A., Reinisch, E. (2020). My English sounds better than yours: Second-language learners perceive their own accent as better than that of their peers. *PLoS ONE, 15*(2), e0227643. https://doi.org/10.1371/journal.pone.0227643

Mitterer, H., Reinisch, E., & McQueen, J. M. (2018). Allophones, not phonemes in spoken-word recognition. *Journal of Memory and Language*, *98*(1), 77–92. https://doi.org/10.1016/j.jml.2017.09.005

Mitterer, H., Scharenborg, O., & McQueen, J. M. (2013). Phonological abstraction without phonemes in speech perception. *Cognition*, *129*(2), 356–361. https://doi.org/10.1016/j.cognition.2013.07.011

Miyawaki, K., Jenkins, J. J., Strange, W., Liberman, A. M., Verbrugge, R., & Fujimura, O. (1975). An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception & Psychophysics*, *18*(5), 331–340. https://doi.org/10.3758/BF03211209

Myers, E. B., & Mesite, L. M. (2014). Neural systems underlying perceptual adjustment to non-standard speech tokens. *Journal of Memory and Language*, *76*(1), 80–93. https://doi.org/10.1016/j.jml.2014.06.007

Nelson, S., & Durvasula, K. (2021). Lexically-guided perceptual learning does generalize to new phonetic contexts. *Journal of Phonetics, 84*(1), 101019. https://doi.org/10.1016/j.wocn.2020.101019

Newman, R. S., Clouse, S. A., & Burnham, J. L. (2001). The perceptual consequences of within-talker variability in fricative production. *The Journal of the Acoustical Society of America*, *109*(1), 1181. https://doi.org/10.1121/1.1348009

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology*, *47*(2), 204–238. https://doi.org/10.1016/S0010-0285(03)00006-9

Norton, B. (2010). Language and identity. In N. H. Hornberger & S. L. McKay (Eds.), *Sociolinguistics and Language Education* (pp. 349–369). Multilingual Matters.

Osborne, J. W. (2002). Notes on the use of data transformations. *Practical Assessment, Research & Evaluation, 8*(6), 1–7. https://doi.org/10.7275/4vng-5608

Perrachione, T. K., Lee, J., Ha, L. Y. Y., & Wong, P. C. M. (2011). Learning a novel phonological contrast depends on interactions between individual differences and training paradigm design. *The Journal of the Acoustical Society of America*, *130*(1), 461–472. https://doi.org/10.1121/1.3593366

Quené, H., & van den Bergh, H. (2008). Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language, 59*(1), 413–425. https://doi.org/10.1016/j.jml.2008.02.002

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Reinisch, E., & Holt, L. L. (2014). Lexically-guided phonetic retuning of foreign-accented speech and its generalization. *Journal of Experimental Psychology: Human Perception and Performance, 40*(2), 539–555. https://doi.org/10.1037/a0034409

Reinisch, E., & Mitterer, H. (2016). Exposure modality, input variability and the categories of perceptual recalibration. *Journal of Phonetics*, *55*(1), 96–108. https://doi.org/10.1016/j.wocn.2015.12.004

Reinisch, E., Weber, A., & Mitterer, H. (2013). Listeners retune phoneme categories across languages. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(1), 75–86. https://doi.org/10.1037/a0027979

Reinisch, E., Wozny, D. R., Mitterer, H., & Holt, L. L. (2014). Phonetic category recalibration: What are the categories? *Journal of Phonetics*, *45*(1), 91–105. https://doi.org/10.1016/j.wocn.2014.04.002

Rogers, D., & d'Arcangeli, L. (2004). Italian. *Journal of the International Phonetic Association, 34*(1), 117–121. https://doi.org/10.1017/S0025100304001628

Samuel, A. G. (1986). Red herring detectors and speech perception: In defense of selective adaptation. *Cognitive Psychology, 18*(1)452–499.

Samuel, A. G. (2016). Lexical representations are malleable for about one second: Evidence for the non-automaticity of perceptual recalibration. *Cognitive Psychology*, *88*(1), 88–114. https://doi.org/10.1016/j.cogpsych.2016.06.007

Samuel, A. G. (2020). Psycholinguists should resist the allure of linguistic units as perceptual units. *Journal of Memory and Language*, *111*(1), 104070. https://doi.org/10.1016/j.jml.2019.104070

Samuel, A. G., & Kraljic, T. (2009). Perceptual learning for speech. *Attention, Perception & Psychophysics, 71*(6), 1207–1218. https://doi.org/10.3758/APP.71.6.1207

Scharenborg, O., Mitterer, H., & McQueen, J. M. (2011). Perceptual learning of liquids. *Proceedings of the 12th Annual Conference of the International Speech Communication Association (Interspeech 2011), Florence, Italy,* 149–152.

Scharenborg, O., Norris, D., ten Bosch, L., & McQueen, J. M. (2005). How should a speech recognizer work? *Cognitive Science*, *29*(6), 867–918. https://doi.org/10.1207/s15516709cog0000_37

Scott, M. (2020). Interaural recalibration of phonetic categories. *The Journal of the Acoustical Society of America, 147*(2), EL164–EL170. https://doi.org/10.1121/10.0000735

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics*, *10*(1), 209–232.

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics*, *10*(1), 209–232.

Selinker, L., Swain, M., & Dumas, G. (1975). The Interlanguage hypothesis extended to children. *Language Learning*, *25*(1), 139–152. https://doi.org/10.1111/j.1467-1770.1975.tb00114.x

Selinker, L., Swain, M., & Dumas, G. (1975). The Interlanguage hypothesis extended to children. *Language Learning*, *25*(1), 139–152. https://doi.org/10.1111/j.1467-1770.1975.tb00114.x

Sidaras, S. K., Alexander, J. E. D., & Nygaard, L. C. (2009). Perceptual learning of systematic variation in Spanish-accented speech. *The Journal of the Acoustical Society of America*, *125*(5), 3306. https://doi.org/10.1121/1.3101452

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant psychological. *Psychological Science, 22*(11), 1359–1366. https://doi.org/10.1177/0956797611417632

Simon, E., Debaene, M., & Van Herreweghe, M. (2015). The effect of L1 regional variation on the perception and production of standard L1 and L2 vowels. *Folia Linguistica, 49*(2), 521–553.

Simpson, A. P. (2001). Dynamic consequences of differences in male and female vocal tract dimensions. *The Journal of the Acoustical Society of America*, *109*(5), 2153–2164. https://doi.org/10.1121/1.1356020

Sjerps, M. J., Mitterer, H., & McQueen, J. M. (2011). Listening to different speakers: On the time-course of perceptual compensation for vocal-tract characteristics. *Neuropsychologia*, *49*(14), 3831–3846. https://doi.org/10.1016/j.neuropsychologia.2011.09.044

Sjerps, M., & McQueen, J. M. (2010). The bounds on flexibility in speech perception. *Journal of Experimental Psychology, 36*(1), 195–211. https://doi.org/10.1037/a0016803

Sumner, M. (2011). The role of variation in the perception of accented speech. *Cognition*, *119*(1), 131–136. https://doi.org/10.1016/j.cognition.2010.10.018

Tabouret-Keller, A. (1997). Language and identity. In. F. Coulmas (Ed.), *The handbook of sociolinguistics* (pp. 315–326). Blackwell Publishing.

Tarone, E. (2018). Interlanguage. In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics* (pp. 1–7). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781405198431.wbeal0561.pub2

Tarone, E. (2018). Interlanguage. In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics* (pp. 1–7). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781405198431.wbeal0561.pub2

Trubetzkoy, N. (1939). *Principles of phonology*. University of California Press.

Trude, A. M., Tremblay, A., & Brown-Schmidt, S. (2013). Limitations on adaptation to foreign accents. *Journal of Memory and Language*, *69*(3), 349–367. https://doi.org/10.1016/j.jml.2013.05.002

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

Tzeng, C. Y., Nygaard, L. C., & Theodore, R. M. (2021). A second chance for a first impression: Sensitivity to cumulative input statistics for lexically guided perceptual learning. *Psychonomic Bulletin & Review*. Advance online publication. https://doi.org/10.3758/s13423-020-01840-6

Ullas, S., Formisano, E., Eisner, F., & Cutler, A. (2020). Interleaved lexical and audiovisual information can retune phoneme boundaries. *Attention, Perception, and Psychophysics*, *82*(4), 2018–2026. https://doi.org/10.3758/s13414-019-01961-8

van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology*, *67*(6), 1176-1190. http://dx.doi.org/10.1080/17470218.2013.850521

Van Hoof, S., & Vandekerckhove, B. (2013). Feiten en fictie: Taalvariatie in Vlaamse televisiereeksen vroeger en nu. *Nederlandse Taalkunde, 18*(1), 35–64.

Verhoeven, J., & Van Bael, C. (2002). Akoestische kenmerken van de Nederlandse klinkers in drie Vlaamse regio's. *Taal en Tongval, 54*(1). 1–23.

Vihman, M. M. (2017). Learning words and learning sounds: Advances in language development. *British Journal of Psychology*, *108*(1), 1–27. https://doi.org/10.1111/bjop.12207

Vroomen, J., van Linden, S., Keetels, M., de Gelder, B., & Bertelson, P. (2004). Selective adaptation and recalibration of auditory speech by lipread information: Dissipation. *Speech Communication*, *44*(1–4), 55–61. https://doi.org/10.1016/j.specom.2004.03.009

Weatherholtz, K. (2015). *Perceptual learning of systemic cross-category vowel variation* [Unpublished doctoral dissertation]. The Ohio State University.

Weatherholtz, K., & Jaeger, T. F. (2016). *Speech perception and generalization across talkers and accents.* Linguistics: Oxford Research Encyclopedias.

Weber, A., di Betta, A. M., & McQueen, J. M. (2014). Treack or trit: Adaptation to genuine and arbitrary foreign accents by monolingual and bilingual listeners. *Journal of Phonetics*, *46*(1), 34–51. https://doi.org/10.1016/j.wocn.2014.05.002

Werker, J. F. (2018). Perceptual beginnings to language acquisition. *Applied Psycholinguistics*, *39*(4), 703–728. https://doi.org/10.1017/S0142716418000152

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *7*(1), 49–63. https://doi.org/10.1016/S0163-6383(84)80022-3

White, K. S., Eiling, Y., Blumstein, S. E., & Morgan, J. L. (2013). Adults show less sensitivity to phonetic detail in unfamiliar words, too. *Journal of Memory and Language, 68*(1), 362–378. http://dx.doi.org/10.1016/j.jml.2013.01.003

Xie, X., & Myers, E. B. (2017). Learning a talker or learning an accent: Acoustic similarity constrains generalization of foreign accent adaptation to new talkers. *Journal of Memory and Language*, *97*(1), 30–46. https://doi.org/10.1016/j.jml.2017.07.005

Xie, X., Liu, L., & Jaeger, T. F. (2021). Cross-talker generalization in the perception of non-native speech: A large-scale replication. *Journal of Experimental Psychology General.* Advance online publication. https://doi.org/10.1037/xge0001039

Xie, X., Weatherholtz, K., Bainton, L., Rowe, E., Burchill, Z., Liu, L., & Jaeger, T. F. (2018). Rapid adaptation to foreign-accented speech and its transfer to an unfamiliar talker. *The Journal of the Acoustical Society of America*, *143*(4), 2013–2031. https://doi.org/10.1121/1.5027410

Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's N: A new measure of orthographic similarity. *Psychonomic Bulletin and Review*, *15*(5), 971–979. https://doi.org/10.3758/PBR.15.5.971

Zhang, X., & Samuel, A. G. (2014). Perceptual learning of speech under optimal and adverse conditions. *Journal of Experimental Psychology: Human Perception and Performance, 40*(1), 200–217. https://doi.org/10.1037/a0033182

Zhang, Y., & Wang, Y. (2007). Neural plasticity in speech acquisition and learning. *Bilingualism*, *10*(2), 147–160. https://doi.org/10.1017/S1366728907002908

10 Appendices

Overview of the Appendices

Appendix 1	Participant Recruitment	p. 88
Appendix 2	Participant Overview	p. 92
Appendix 3	Speaker Information	p. 99
Appendix 4	Stimuli Overview	p. 111
Appendix 5	Experiment Results	p. 117

10.1 Participant Information

Students interested in participating in the experiment were provided with a *general outline* of the study and an *informed consent sheet*. When participants had given written informed consent, the experimenter provided them with a concise *procedure manual* on how to perform the web-based experiment and the weblink to the experimental environment. Each participant was given a unique four-digit code, consisting of two letters and two numbers, to enter the experiment (cf. *ID* column in the tables in Section 10.2). All three document (in Dutch) can be consulted below.

FACULTEIT LETTEREN
EN WIJSBEGEERTE

Masterproef in de Taalkunde

Beste geïnteresseerde,

Mijn naam is Gil Verbeke en ik studeer Taal- en Letterkunde aan de Universiteit Gent. Voor mijn masterproef doe ik onderzoek naar de perceptie van moedertaalsprekers van het Nederlands op het taalgebruik van sprekers die het Nederlands als tweede taal (NT2) hebben. Ik analyseer met andere woorden de mate waarin Nederlandstaligen de spraak van een NT2-spreker begrijpen ('intelligibility') en aanvaarden ('acceptability').

Daarom ben ik op zoek naar een aantal moedertaalsprekers van het Nederlands die in **Gent studeren** en bereid zijn om deel te nemen aan een kort online luisterexperiment (± 30 minuten). Er zijn echter drie criteria waaraan u moet voldoen om te kunnen deelnemen: (i) u ondervindt **geen gehoorverlies**, (ii) u bent geboren én opgegroeid in de provincie **West-Vlaanderen** of **Oost-Vlaanderen** en (iii) u heeft **geen gevorderde kennis** van het **Italiaans** of **Spaans** en bent niet ingeschreven in een opleiding waar Italiaanse of Spaanse taalvakken als niet-keuzevakken worden aangeboden in het opleidingsprogramma.

Indien u geïnteresseerd bent om deel te nemen, dan zou u mij een scan of foto van een ondertekende *geïnformeerde toestemming* (zie p. 2) moeten bezorgen via e-mail (Gil.Verbeke@UGent.be) of MS Teams. In dat formulier kunt u zich ook registreren voor een verloting van **5 waardebonnen** van **Bol.com** t.w.v. **€20.** Na ontvangst van een ondertekend document bezorg ik u meer uitleg over en de link naar het online experiment.

Dit onderzoek wordt uitgevoerd onder supervisie van Prof. Dr. Ellen Simon *(UGent)*, Prof. Dr. Robert Hartsuiker *(UGent)* en Prof. Dr. Holger Mitterer *(L-Università ta' Malta)*. Mocht u nog bijkomende vragen hebben over de opzet van deze studie of wat er precies van u verwacht wordt, contacteer me dan gerust via e-mail of sociale media.

Alvast bedankt voor uw interesse!

Vriendelijke groeten,
Gil Verbeke

UNIVERSITEIT
GENT

FACULTEIT LETTEREN
EN WIJSBEGEERTE

Vakgroep Taalkunde
Blandijnberg 2, 9000 Gent, Tel.: 092 64 37 15
www.taalkunde.ugent.be

Geïnformeerde Toestemming

Ik, ondergetekende .... (voornaam + familienaam) bevestig hierbij dat ik als proefpersoon deelneem aan een onderzoek van de Vakgroep Taalkunde aan de Universiteit Gent en ten volle geïnformeerd ben over:

(1) De aard van het onderzoek: zijnde het doel van het onderzoek en de aard van de vragen, opdrachten die tijdens dit onderzoek gebruikt zullen worden en welke functie ze vervullen in het onderzoek;

(2) Het feit dat ik uit vrije wil deelneem aan het onderzoek;

(3) Het feit dat ik toestemming geef aan de proefleider om mijn resultaten te gebruiken en die toestemming elk moment kan intrekken, zonder opgave van reden, waardoor mijn resultaten niet langer deel zullen uitmaken van het onderzoek;

(4) Het feit dat niet deelnemen of mijn deelname aan het onderzoek stopzetten op geen enkele manier invloed heeft op mijn evaluatie en/of studiebegeleiding;

(5) Het feit dat alle resultaten van het onderzoek volledig anoniem zijn;

(6) Het feit dat ik op aanvraag steeds toegang heb tot een samenvatting van de onderzoeksbevindingen;

(7) Het feit dat ik ter compensatie voor mijn deelname aan het experiment kan deelnemen aan een verloting van **5 waardebonnen** van **Bol.com** t.w.v. **€20**. De proefleider kan me daarvoor op het volgende e-mailadres contacteren: ....

Gelezen en goedgekeurd op .... (datum)

Handtekening

De deelnemer

UNIVERSITEIT
GENT

FACULTEIT LETTEREN EN WIJSBEGEERTE

Procedure Luisterexperiment

Vooraleer u aan het experiment begint, is het belangrijk dat u tijdens de test aan een tafel zit in een **rustige omgeving** met zo weinig mogelijk externe geluidsfactoren (bv. straatlawaai, pratende voetgangers of andere vormen van ruis). Zorg ervoor dat:

- alle **tabbladen** op uw browser (bij voorkeur *Google Chrome* of *Safari*) afgesloten zijn;
- **meldingen** op uw computer tijdelijk uitgeschakeld worden;
- de **accu** van uw laptop voldoende opgeladen is;
- u een **hoofdtelefoon** bij de hand hebt om naar de geluidsfragmenten te luisteren.

Aan het begin van het experiment zal u gevraagd worden een **persoonlijke code** in te vullen. Uw unieke code is **XXXX** en dient louter als controlemiddel voor de proefleider indien er bepaalde problemen zouden optreden. Mochten er technische problemen zijn, probeer eerst de webpagina te refreshen. Als dit niet lukt, contacteer me dan zo snel mogelijk via e-mail (Gil.Verbeke@Ugent.be), via MS Teams of sociale media. Als er geen verdere vragen meer zijn, dan kunt u het experiment op onderstaande link starten:

KLIK HIER OM HET EXPERIMENT TE STARTEN

Hartelijk bedankt voor uw tijd en veel succes!

Gil Verbeke

UNIVERSITEIT GENT

1/1

10.2 Participant Overview

Pre-test [25]

ID	Age	Gender	Prov. (birth)	Prov. (raised)	Programme	Institution	Headphones
AH26	21	Male	East Flanders	East Flanders	Linguistics & Literature: Dutch-English	Ghent University	UR
DP37	21	Male	West Flanders	West Flanders	Master of Science in Rehabilitation Sciences and Physiotherapy	Ghent University	Turtle beach
HC63	21	Female	West Flanders	West Flanders	Linguistics & Literature	Ghent University	Apple (EarPods)
LA65	21	Female	West Flanders	West Flanders	Linguistics & Literature: Dutch-English	Ghent University	Bose
MM84	21	Female	West Flanders	West Flanders	Master of Science in Criminological Sciences	Ghent University	Marshall
MR76	21	Female	West Flanders	West Flanders	Linguistics & Literature: Dutch-English	Ghent University	Apple AirPods
NW28	21	Female	West Flanders	West Flanders	Linguistics & Literature: Comparative Modern Literature (MA)	Ghent University	Fresh n' Rebel
SG04	23	Female	West Flanders	West Flanders	Linguistics & Literature: English	Ghent University	JBL
SS33	22	Male	West Flanders	West Flanders	Marketing (BA)	Hogeschool Gent	Apple AirPods
VM40	21	Male	West Flanders	West Flanders	Bachelor of Science in Business Economics	Ghent University	Marshall

[25] *ID* = speaker identity as represented by a unique four-digit code consisting of two letters and two numbers; *Age* = participant age in years at the time of testing; *gender* = self-identified gender; *Prov. (birth)* = province in Flanders where participants were born; *Prov. (raised)* = province in Flanders where participants were raised; *Programme* = programme in which the participant was enrolled at the time of testing; *Institution* = institution of higher education in Ghent, Belgium, where the participant was enrolled; *Headphones* = brand (and in some cases also the type) of headphones participants used to listen to the audio.

Main Experiment

ID	Age	Gender	Prov. (birth)	Prov. (raised)	Programme	Institution	Headphones
BC71	22	Female	East Flanders	East Flanders	Linguistics & Literature: Dutch-English (Eduma)	Ghent University	Bose
BD00	19	Female	East Flanders	East Flanders	Linguistics & Literature: English-German	Ghent University	Sony WH-XB900N
BD54	22	Female	East Flanders	East Flanders	Linguistics & Literature	Ghent University	Steelseries Arctis 3
BF11	20	Female	West Flanders	West Flanders	Linguistics & Literature	Ghent University	Marshall Major III bluetooth
BH43	21	Female	East Flanders	East Flanders	Linguistics & Literature: English-Greek	Ghent University	Logitech
BJ47	21	X	East Flanders	East Flanders	Linguistics & Literature: English-Latin (Eduma)	Ghent University	Sony WH-1000
BP76	22	Female	West Flanders	West Flanders	Primary Education	Arteveldehogeschool	Sony WH-1000XM3
BT89	22	Female	West Flanders	West Flanders	Nursing (postgraduate after Midwifery)	Hogeschool Gent	Beats solo 2 Wireless
BU54	21	Male	East Flanders	East Flanders	Linguistics & Literature: Dutch-English	Ghent University	JVC HA-SR625
BU67	19	Female	East Flanders	East Flanders	Linguistics & Literature	Ghent University	Philips SHB7250
CF34	21	Male	West Flanders	West Flanders	Social Work	Arteveldehogeschool	JBL TUNE 750BTNC
CF47	21	Female	East Flanders	East Flanders	Linguistics & Literature	Ghent University	Urbanears, Plattan 2BT
CF70	22	Male	West Flanders	West Flanders	Accountancy	Hogeschool Gent	Marshall
CP39	21	Female	West Flanders	West Flanders	Communication Management	Arteveldehogeschool	Acer in-ear headphones
CU89	21	Male	West Flanders	West Flanders	Linguistics & Literature: Dutch-English (Eduma)	Ghent University	Sony MDR-ZX110
CX10	22	Female	West Flanders	West Flanders	Public Administration and Management	Ghent University	Marshall
CX41	21	Male	West Flanders	West Flanders	Psychology	Ghent University	TaoTronics
DE18	21	Female	East Flanders	East Flanders	Linguistics & Literature	Ghent University	In-ear headphones

DE53	19	Male	West Flanders	West Flanders	Linguistics & Literature: Dutch-English	Ghent University	JBL LIVE 500BT
DF25	20	Male	West Flanders	West Flanders	Marketing	Arteveldehogeschool	JBL
DT70	18	Female	East Flanders	East Flanders	Linguistics & Literature: Dutch-English	Ghent University	Apple Airpods 1
DV10	22	Female	West Flanders	West Flanders	Linking Programme: Life Sciences	Ghent University	Sony (MDRXB650BT)
DX27	23	Female	West Flanders	West Flanders	Psychology	Ghent University	AirPods
DX78	21	Female	West Flanders	West Flanders	Linguistics & Literature: Dutch-English	Ghent University	Apple
EE40	23	Male	West Flanders	West Flanders	Wood Technology	Ghent University	Beats studio 3
FD05	21	Male	East Flanders	East Flanders	Linguistics & Literature: English-Swedish	Ghent University	HyperX Cloud II
FE43	18	Female	East Flanders	East Flanders	Linguistics & Literature: Dutch-English	Ghent University	Blaupunkt 4633
FE46	21	Female	East Flanders	East Flanders	Linguistics & Literature	Ghent University	JBL Tune 500BT
FF20	18	Female	East Flanders	East Flanders	Linguistics & Literature: French-English	Ghent University	Sennheiser HD206
FF38	19	Female	West Flanders	West Flanders	Sociology	Ghent University	Marshall major 3
FG25	20	Male	East Flanders	East Flanders	Linguistics & Literature	Ghent University	Bose QC Earbuds
FI17	21	Female	West Flanders	East Flanders	Linguistics & Literature	Ghent University	Marshall
FI42	20	Female	East Flanders	East Flanders	Linguistics & Literature	Ghent University	No particular brand
FI03	20	Female	West Flanders	West Flanders	Communication Sciences	Ghent University	JBL
FJ55	23	Female	West Flanders	West Flanders	Educational Sciences	Ghent University	MM
FO47	20	Male	West Flanders	West Flanders	Marketing	Arteveldehogeschool	JBL TUNE500BT
FO97	22	Female	East Flanders	East Flanders	Linguistics & Literature: Dutch-English	Ghent University	Marshall
FU25	22	Female	West Flanders	West Flanders	Veterinary Medicine	Ghent University	Plantronics
FX75	21	Female	West Flanders	West Flanders	Veterinary Medicine	Ghent University	JBL
FX81	21	Female	West Flanders	West Flanders	Rehabilitation Sciences and Physiotherapy with Musculoskeletal Afflictions	Ghent University	In-ear headphones
GF66	24	Female	West Flanders	West Flanders	Communication Sciences	Ghent University	JBL T460BT
GJ73	22	Male	East Flanders	East Flanders	Linguistics & Literature: English-German (Eduma)	Ghent University	Urbanears Plattan II (no Bluetooth)
GR22	25	Female	East Flanders	East Flanders	Linguistics & Literature	Ghent University	Urbanears
GT29	21	Male	East Flanders	East Flanders	Linguistics & Literature: Dutch-English	Ghent University	JBL in-ears

GX22	19	Female	West Flanders	West Flanders	Communication Management	Arteveldehogeschool	JBL E65BTNC
GX23	20	Female	West Flanders	West Flanders	Bachelor of Science in Speech Language and Hearing Sciences (Audiology)	Ghent University	Samsung
HO46	21	Female	West Flanders	East Flanders	Linguistics & Literature: Dutch-English	Ghent University	Sony wh-1000xm3
IJ89	19	Female	East Flanders	East Flanders	Linguistics & Literature	Ghent University	JBL
IQ93	23	Male	East Flanders	East Flanders	Linguistics & Literature: English	Ghent University	JBL T210 in-ears
IU15	19	Female	East Flanders	East Flanders	Linguistics & Literature: Dutch-English	Ghent University	JBL tune 500
IU89	17	Female	East Flanders	East Flanders	Linguistics & Literature: Dutch-English	Ghent University	JBL E45BT
JD97	21	Female	East Flanders	East Flanders	Linguistics & Literature: Dutch-English	Ghent University	Flying Tiger
JU99	21	Female	West Flanders	West Flanders	Real Estate	Hogeschool Gent	Medion
KE71	22	Female	West Flanders	West Flanders	Master of Science in Educational Sciences	Ghent University	Marshall
KF51	21	Female	West Flanders	West Flanders	Educational Sciences	Ghent University	Beats Solo
KF72	22	Female	West Flanders	West Flanders	Linking Programme: Social Work	Ghent University	Apple in-ears
KF77	20	Male	East Flanders	East Flanders	Linguistics & Literature: French-English	Ghent University	Sony WH-1000XM3
KF87	22	Female	West Flanders	West Flanders	Linguistics & Literature: Latin-Greek (Eduma)	Ghent University	JBL tune 750BTNC
KF97	21	Female	West Flanders	West Flanders	Linguistics & Literature	Ghent University	Sony
LD19	19	Female	East Flanders	East Flanders	Linguistics & Literature: Dutch-English	Ghent University	Bose
LE65	23	Male	West Flanders	West Flanders	Linguistics & Literature: Dutch-English	Ghent University	Marshall
LF24	20	Female	East Flanders	East Flanders	Linguistics & Literature: Dutch-English	Ghent University	Sony in-ears
LI09	21	Male	West Flanders	West Flanders	Civil Engineering: Chemical Technology	Ghent University	Marshall
LM62	21	Female	West Flanders	West Flanders	Master of Science in Health Care Management and Policy	Ghent University	Sony WH-1000XM3
LO27	21	X	West Flanders	West Flanders	Eastern Languages and Cultures: Japan (MA1)	Ghent University	Apple (ear pods)
LX84	21	Male	West Flanders	West Flanders	Master in Pharmaceutical Sciences	Ghent University	Apple in-ears
MD88	21	Female	West Flanders	West Flanders	MSc in Biomedical Engineering	Ghent University	JBL
MF86	21	Female	West Flanders	West Flanders	Master of Arts in Multilingual Communication: Dutch-French-German	Ghent University	JBL
MF88	21	Female	West Flanders	West Flanders	Master of Science in Psychology (Personnel Management and Industrial Psychology)	Ghent University	Philips SHB3075

MF96	22	Female	West Flanders	West Flanders	Rehabilitation Sciences and Physiotherapy	Ghent University	JBL TUNE500BT
MH99	21	Female	West Flanders	West Flanders	Architecture	KU Leuven Campus Gent	CLAM ANC
MM50	20	Male	West Flanders	West Flanders	Linguistics & Literature	Ghent University	SONY MDR-XB550
MN15	23	Male	West Flanders	West Flanders	Master of Science in Engineering Technology	KU Leuven Campus Gent	Jabra Move v2.5.0
MN63	20	Female	West Flanders	West Flanders	Remedial Education	Hogeschool Gent	JBL
MO41	21	Female	West Flanders	West Flanders	Nursing	Arteveldehogeschool	Marshall
MR77	23	Male	West Flanders	West Flanders	Rehabilitation Sciences and Physiotherapy	Ghent University	Bose QC35ii
MT15	21	Female	West Flanders	West Flanders	Master of Science in Educational Sciences (Clinical Special Needs Education and Disability Studies)	Ghent University	JBL Tune 500 BT
MU11	20	Female	East Flanders	East Flanders	Linguistics & Literature	Ghent University	Huawei FreeBuds Pro
MU27	23	Female	West Flanders	West Flanders	Master of Science in Psychology (Personnel Management and Industrial Psychology)	Ghent University	Plantronics
MU75	21	Female	West Flanders	West Flanders	Pharmaceutical Sciences	Ghent University	SONY WH-XB900N
MX80	21	Male	West Flanders	West Flanders	Graphic & Digital Media: Audiovisual Design	Arteveldehogeschool	Beyerdynamic DT 770 PRO
MX89	19	Female	East Flanders	East Flanders	Linguistics & Literature: Latin-French	Ghent University	JBL Tune 500
ND78	20	Male	West Flanders	West Flanders	Marketing	Arteveldehogeschool	Creative
NF18	22	Female	West Flanders	West Flanders	Educational Sciences	Ghent University	Samsung
NF60	21	Female	West Flanders	West Flanders	Linking Programme: Master of Science in Health Care Management and Policy	Ghent University	/
NF64	21	Female	West Flanders	West Flanders	Eastern Languages & Cultures: China	Ghent University	Fresh n' Rebel
NF91	21	Male	West Flanders	West Flanders	Linguistics & Literature	Ghent University	JBL
NN33	21	Male	West Flanders	West Flanders	Bachelor of Education	Arteveldehogeschool	AKG
NN37	23	Female	West Flanders	West Flanders	Master of Arts in Applied Language Studies: Dutch-German (Eduma)	Ghent University	JBL 500 BT

NR34	18	Female	West Flanders	West Flanders	Bachelor in Sociology	Ghent University	Microsoft LifeChat LX-3000
NR99	19	Female	West Flanders	West Flanders	Bachelor of Laws	Ghent University	Sony
NU88	18	Female	East Flanders	East Flanders	Linguistics & Literature: English-Swedish	Ghent University	Sony - MDR-ZX110
NX04	21	Male	West Flanders	West Flanders	Bachelor Podiatry	Arteveldehogeschool	Jbl tune 500BT
NX62	19	Male	West Flanders	West Flanders	Bachelor in nutrition and dietetics	Odisee	No headphone
NX81	27	Female	West Flanders	West Flanders	Permanent Training Programme 'Couple, Family and Systemic Therapy'	Ghent University	Sony
NX92	22	Female	West Flanders	West Flanders	Linguistics & Literature	Ghent University	Sony WH 1000XM4
NY46	29	Female	East Flanders	East Flanders	Linguistics & Literature:	Ghent University	HEMA
OV64	18	Female	West Flanders	West Flanders	Linguistics & Literature: English-German	Ghent University	Skullcandy: Riff Wireless
PU79	22	Male	West Flanders	West Flanders	Master of Science in Psychology (Personnel Management and Industrial Psychology)	Ghent University	Sony
QX03	22	Female	West Flanders	West Flanders	Master of Science in Health Promotion	Ghent University	JBL
QX64	19	Female	West Flanders	West Flanders	Bachelor of Laws	Ghent University	Sony
QX97	23	Female	West Flanders	West Flanders	Communication Sciences	Ghent University	Skullcandy
SD99	21	Female	West Flanders	West Flanders	Bachelor Preschool Education	Hogeschool Gent	No
SF29	21	Female	West Flanders	West Flanders	Bachelor Early Childhood Education	Arteveldehogeschool	Over-ear headphones
SN78	20	Male	East Flanders	East Flanders	Linguistics & Literature: English-Swedish	Ghent University	Hifiman 400i 2020
SV59	23	Male	East Flanders	East Flanders	Linguistics & Literature	Ghent University	Sennheiser PXC 550
TF36	22	Male	West Flanders	West Flanders	Linguistics & Literature: Dutch-Swedish	Ghent University	Sennheiser HD 280 pro
TR42	22	Male	West Flanders	West Flanders	Master Veterinary Medicine	Ghent University	Sony
TX31	21	Female	West Flanders	West Flanders	Master of Medicine	Ghent University	Apple Earpods
TX60	20	Female	West Flanders	West Flanders	Business Engineering	Ghent University	JBL TUNE500BT
TY61	21	Female	West Flanders	West Flanders	Master of Science in Clinical Psychology	Ghent University	Beats

UI57	22	Female	West Flanders	West Flanders	Master of Science in Health Care Management and Policy	Ghent University	Marshall (MID A.N.C.
UN16	21	Male	West Flanders	West Flanders	Rehabilitation Sciences and Physiotherapy	Ghent University	JBL
UN26	22	Male	West Flanders	West Flanders	Wood Technology	Hogeschool Gent	Beats studio 3
UU48	22	Male	West Flanders	West Flanders	Master of Laws	Ghent University	Plantronics
XX28	22	Male	West Flanders	West Flanders	Rehabilitation Sciences and Physiotherapy	Ghent University	On-ear JBL

10.3 Speaker Information

After agreeing to participate in the present study, the female native speaker of Italian was provided with a *formalities and practicalities* document via email. This document consisted of:

1. a general outline of the project;
2. an informed consent sheet;
3. a short survey about the speaker's details and language background;
4. a step-by-step guideline for recording the stimulus materials.

In order to protect the anonymity of the speaker, a blank version of the document is added below.

FACULTEIT LETTEREN
EN WIJSBEGEERTE

Masterproef in de Taalkunde

Beste deelnemer,

Mijn naam is Gil Verbeke en ik studeer Taal- en Letterkunde (Nederlands-Engels) aan de Universiteit Gent. Voor mijn masterscriptie doe ik onderzoek naar de perceptie van moedertaalsprekers van het Nederlands op het taalgebruik van sprekers die het Nederlands als tweede/vreemde taal (NT2) hebben. Ik analyseer met andere woorden de mate waarin moedertaalsprekers de spraak van een NT2-spreker begrijpen ('intelligibility') en aanvaarden ('acceptability').

Daarom ben ik op zoek naar een NT2-spreker die voldoende vaardig is in het Nederlands om een lijst met stimuli in te spreken. De stimuli zijn semantisch neutrale zinnetjes waarin een reeks woorden wordt ingebed. De opname van de zinnetjes neemt normaal gezien niet meer dan een halfuur (±20-30 min) van uw tijd in beslag. In wat volgt kunt u nog drie andere documenten terugvinden: (i) een geïnformeerde toestemming, (ii) een korte vragenlijst, en (iii) een gedetailleerde handleiding voor de opname van de stimuli.

Dit onderzoeksproject wordt uitgevoerd onder supervisie van Prof. Dr. Ellen Simon (*Vakgroep Vertalen, Tolken en Communicatie*) en Prof. Dr. Robert J. Hartsuiker (*Vakgroep Experimentele Psychologie*). Mocht u nog bijkomende vragen hebben over het opzet van deze studie, contacteer me dan gerust via e-mail (Gil.Verbeke@Ugent.be) of telefonisch (0498/32.20.87).

Alvast bedankt voor uw deelname!

Met vriendelijke groeten,
Gil Verbeke

UNIVERSITEIT
GENT

1/11

FACULTEIT LETTEREN EN WIJSBEGEERTE

Vakgroep Taalkunde
Blandijnberg 2, 9000 Gent, Tel.: 092 64 37 15
www.taalkunde.ugent.be

Geïnformeerde Toestemming

Ik, ondergetekende .... (voornaam + familienaam) bevestig hierbij dat ik als proefpersoon deelneem aan een onderzoek van de Vakgroep Taalkunde aan de Universiteit Gent en ten volle geïnformeerd ben over:

(1) De aard van het onderzoek: zijnde het doel van het onderzoek en de aard van de vragen, opdrachten die tijdens dit onderzoek gebruikt zullen worden en welke functie ze vervullen in het onderzoek;
(2) Het feit dat ik uit vrije wil deelneem aan het onderzoek;
(3) Het feit dat ik toestemming geef aan de proefleider om mijn resultaten te gebruiken en die toestemming elk moment kan intrekken, zonder opgave van reden, waardoor mijn resultaten niet langer deel zullen uitmaken van het onderzoek;
(4) Het feit dat alle resultaten van het onderzoek volledig anoniem zijn;
(5) Het feit dat ik op aanvraag steeds toegang heb tot een samenvatting van de onderzoeksbevindingen.

Gelezen en goedgekeurd op .... (datum)

Handtekening

De deelnemer

UNIVERSITEIT GENT

Profiel van de NT2-Spreker

A. Personalia

Naam	
Geboortedatum	
Geboorteplaats	
Moedertaal (L1)	
Huidige woonplaats	
Aankomstjaar in Vlaanderen	

B. Taalgebruik

In welke context heeft u het Nederlands aangeleerd?

In welke mate beheerst u het Nederlands (ofwel geschat ofwel op basis van certificaat)? Omcirkel uw huidige taalniveau volgens het *Common European Framework of Reference for Languages* (CEFR).[1]					
Basisgebruiker		Onafhankelijke gebruiker		Vaardige gebruiker	
A1	*A2*	*B1*	*B2*	*C1*	*C2*
Eventuele toelichting:					

Hoe vaak spreekt u Nederlands met:						
	Dagelijks	Wekelijks	Maandelijks	Jaarlijks	Minder vaak dan jaarlijks	Nooit
a. gezinsleden						
b. familieleden						
c. buren						
d. collega's						
e. onbekenden						
f. andere:						

[1] Voor meer informatie over de precieze onderverdeling kunt u bij de volgende website van de Raad van Europa terecht: https://www.coe.int/en/web/common-european-framework-reference-languages/level-descriptions.

3/11

Procedure voor de Opname van de Stimuli

Dit document dient als handleiding tijdens de opname van de stimuli. Het is uiterst belangrijk dat de opname plaatsvindt in een ruimte waarin **externe geluidsfactoren**, zoals voorbijrijdende auto's, pratende voetgangers, drilboren of andere vormen van ruis, tot een **absoluut minimum** herleid worden om zo de kwaliteit van de opname te garanderen.

In wat volgt vindt u de stimuli die in dit onderzoek gebruikt zullen worden. Ze zijn onderverdeeld in twee groepen: (i) bestaande Nederlandse woorden en (ii) pseudo-woorden die fonotactisch gezien legaal zijn in het Nederlands. Aan het begin van elk deel wordt kort uitgelegd wat er precies van u verwacht wordt. Om te voorkomen dat bepaalde stimuli niet duidelijk hoorbaar zijn, om welke reden dan ook, zou ik u willen vragen om de stimuli in beide onderdelen **twee keer** in te spreken (±20-30 minuten in totaal). Het lijkt mij dan ook het beste dat u de recordings als **vier aparte audiobestanden** opslaat: twee voor het eerste deel en twee voor het tweede.

DEEL 1	NEDERLANDSE WOORDEN

In dit onderdeel vindt u een lijst van 110 hoogfrequente Nederlandse woorden die ingebed zijn in een standaardformulering. Het is belangrijk dat u deze zinnen als **één intonatie-eenheid** uitspreekt met een **neutrale dalende intonatie.** Probeer bovendien —in de mate van het mogelijke— telkens op **hetzelfde tempo**, en zo **duidelijk** én **natuurlijk** mogelijk te spreken. De woorden zijn per tien gegroepeerd om het overzichtelijk te houden. Het duurt ongeveer 30 seconden om één reeks in te spreken. Dit deel zal dus met andere woorden een elftal minuten (2 X 11 x ±30s = ±11 min) in beslag nemen. Gelieve de opnames voor dit onderdeel als **twee audiobestanden** op te slaan ('MAP_Verbeke_Deel1_1' en 'MAP_Verbeke_Deel1_2').

1	1. Ze heeft **last** gezegd. 2. Ze heeft **roman** gezegd. 3. Ze heeft **koorts** gezegd. 4. Ze heeft **vlinder** gezegd. 5. Ze heeft **bied** gezegd. 6. Ze heeft **knie** gezegd. 7. Ze heeft **dienst** gezegd. 8. Ze heeft **paard** gezegd. 9. Ze heeft **bitter** gezegd. 10. Ze heeft **winkel** gezegd.
2	1. Ze heeft **soldaat** gezegd. 2. Ze heeft **roos** gezegd. 3. Ze heeft **verstand** gezegd. 4. Ze heeft **ziek** gezegd. 5. Ze heeft **raak** gezegd. 6. Ze heeft **liever** gezegd. 7. Ze heeft **normaal** gezegd. 8. Ze heeft **ballon** gezegd. 9. Ze heeft **brief** gezegd. 10. Ze heeft **legaal** gezegd.
3	1. Ze heeft **fiets** gezegd. 2. Ze heeft **pot** gezegd. 3. Ze heeft **dokter** gezegd. 4. Ze heeft **verschil** gezegd. 5. Ze heeft **dorp** gezegd. 6. Ze heeft **droog** gezegd. 7. Ze heeft **liefde** gezegd. 8. Ze heeft **bravo** gezegd. 9. Ze heeft **bink** gezegd. 10. Ze heeft **water** gezegd.

4	1. Ze heeft **adel** gezegd. 2. Ze heeft **los** gezegd. 3. Ze heeft **kapsel** gezegd. 4. Ze heeft **priester** gezegd. 5. Ze heeft **zakdoek** gezegd. 6. Ze heeft **nobel** gezegd. 7. Ze heeft **lig** gezegd. 8. Ze heeft **vis** gezegd. 9. Ze heeft **advies** gezegd. 10. Ze heeft **zielig** gezegd.
5	1. Ze heeft **tafel** gezegd. 2. Ze heeft **lag** gezegd. 3. Ze heeft **titel** gezegd. 4. Ze heeft **vak** gezegd. 5. Ze heeft **yoga** gezegd. 6. Ze heeft **middag** gezegd. 7. Ze heeft **kist** gezegd. 8. Ze heeft **hal** gezegd. 9. Ze heeft **schilder** gezegd. 10. Ze heeft **foto** gezegd.
6	1. Ze heeft **vies** gezegd. 2. Ze heeft **donker** gezegd. 3. Ze heeft **ingang** gezegd. 4. Ze heeft **wit** gezegd. 5. Ze heeft **mosterd** gezegd. 6. Ze heeft **zon** gezegd. 7. Ze heeft **kikker** gezegd. 8. Ze heeft **gevaar** gezegd. 9. Ze heeft **doof** gezegd. 10. Ze heeft **bron** gezegd.
7	1. Ze heeft **uniek** gezegd. 2. Ze heeft **lieg** gezegd. 3. Ze heeft **marge** gezegd. 4. Ze heeft **baas** gezegd. 5. Ze heeft **afdruk** gezegd. 6. Ze heeft **klok** gezegd. 7. Ze heeft **spiegel** gezegd. 8. Ze heeft **schieten** gezegd. 9. Ze heeft **winter** gezegd. 10. Ze heeft **onrust** gezegd.

6/11

8	1. Ze heeft **kantoor** gezegd. 2. Ze heeft **boot** gezegd. 3. Ze heeft **genot** gezegd. 4. Ze heeft **zand** gezegd. 5. Ze heeft **dichter** gezegd. 6. Ze heeft **ton** gezegd. 7. Ze heeft **techniek** gezegd. 8. Ze heeft **vliegen** gezegd. 9. Ze heeft **visum** gezegd. 10. Ze heeft **offer** gezegd.
9	1. Ze heeft **diefstal** gezegd. 2. Ze heeft **ras** gezegd. 3. Ze heeft **single** gezegd. 4. Ze heeft **troon** gezegd. 5. Ze heeft **dikwijls** gezegd. 6. Ze heeft **staart** gezegd. 7. Ze heeft **bid** gezegd. 8. Ze heeft **actief** gezegd. 9. Ze heeft **aap** gezegd. 10. Ze heeft **frank** gezegd.
10	1. Ze heeft **mentaal** gezegd. 2. Ze heeft **winst** gezegd. 3. Ze heeft **wiet** gezegd. 4. Ze heeft **voordeel** gezegd. 5. Ze heeft **pakket** gezegd. 6. Ze heeft **hoofdstad** gezegd. 7. Ze heeft **plank** gezegd. 8. Ze heeft **kiest** gezegd. 9. Ze heeft **verlies** gezegd. 10. Ze heeft **ridder** gezegd.
11	1. Ze heeft **staal** gezegd. 2. Ze heeft **pittig** gezegd. 3. Ze heeft **vinger** gezegd. 4. Ze heeft **aanval** gezegd. 5. Ze heeft **gids** gezegd. 6. Ze heeft **vergif** gezegd. 7. Ze heeft **raadsel** gezegd. 8. Ze heeft **taal** gezegd. 9. Ze heeft **dinsdag** gezegd. 10. Ze heeft **baron** gezegd.

DEEL 2	NEDERLANDSE PSEUDO-WOORDEN

In dit onderdeel zijn 140 pseudo-woorden (d.i., onbestaande woorden die aan de eisen van de Nederlandse fonotaxis beantwoorden) ingebed in een standaardformulering. Het is opnieuw van belang dat u deze zinnen als **één intonatie-eenheid** uitspreekt met een **neutrale dalende intonatie.** Probeer bovendien —in de mate van het mogelijke— telkens op **hetzelfde tempo**, en zo **duidelijk** én **natuurlijk** mogelijk te spreken. De woorden zijn per tien gegroepeerd om het overzichtelijk te houden. Bij sommige bisyllabische woorden is de **beklemtoonde syllabe onderstreept**. Het duurt opnieuw ongeveer 30 seconden om één reeks in te spreken, dus een vijftiental minuten (2 x 14 x ±30s = ±15min) in totaal. Gelieve de opnames voor dit onderdeel als **twee aparte audiobestanden** op te slaan ('MAP_Verbeke_Deel2_1' en 'MAP_Verbeke_Deel2_2').

1	1. Ze heeft **fraal** gezegd. 2. Ze heeft **ploog** gezegd. 3. Ze heeft **warken** gezegd. 4. Ze heeft **aven** gezegd. 5. Ze heeft **bool** gezegd. 6. Ze heeft **aanbol** gezegd. 7. Ze heeft **votten** gezegd. 8. Ze heeft **vogen** gezegd. 9. Ze heeft **kiekker** gezegd. 10. Ze heeft **tort** gezegd.
2	1. Ze heeft **haden** gezegd. 2. Ze heeft **unik** gezegd. 3. Ze heeft **sanen** gezegd. 4. Ze heeft **wopen** gezegd. 5. Ze heeft **vorn** gezegd. 6. Ze heeft **galzen** gezegd. 7. Ze heeft **famen** gezegd. 8. Ze heeft **krokkel** gezegd. 9. Ze heeft **geroom** gezegd. 10. Ze heeft **schaag** gezegd.

3	1. Ze heeft **zolden** gezegd. 2. Ze heeft **boors** gezegd. 3. Ze heeft **lamper** gezegd. 4. Ze heeft **vool** gezegd. 5. Ze heeft **grons** gezegd. 6. Ze heeft **toonzen** gezegd. 7. Ze heeft **oolde** gezegd. 8. Ze heeft **pansen** gezegd. 9. Ze heeft **diekwijls** gezegd. 10. Ze heeft **raps** gezegd.
4	1. Ze heeft **tattig** gezegd. 2. Ze heeft **vinger** gezegd. 3. Ze heeft **voolt** gezegd. 4. Ze heeft **mogens** gezegd. 5. Ze heeft **hank** gezegd. 6. Ze heeft **gebacht** gezegd. 7. Ze heeft **tittel** gezegd. 8. Ze heeft **schielder** gezegd. 9. Ze heeft **schaas** gezegd. 10. Ze heeft **mokker** gezegd.
5	1. Ze heeft **fits** gezegd. 2. Ze heeft **naafs** gezegd. 3. Ze heeft **vissum** gezegd. 4. Ze heeft **trak** gezegd. 5. Ze heeft **technik** gezegd. 6. Ze heeft **gieds** gezegd. 7. Ze heeft **smaren** gezegd. 8. Ze heeft **bopel** gezegd. 9. Ze heeft **advis** gezegd. 10. Ze heeft **bieter** gezegd.
6	1. Ze heeft **klamer** gezegd. 2. Ze heeft **vertaag** gezegd. 3. Ze heeft **spiggel** gezegd. 4. Ze heeft **lop** gezegd. 5. Ze heeft **bienk** gezegd. 6. Ze heeft **sotus** gezegd. 7. Ze heeft **dinst** gezegd. 8. Ze heeft **raafde** gezegd. 9. Ze heeft **spaag** gezegd. 10. Ze heeft **baarzel** gezegd.

7	1. Ze heeft **bromen** gezegd. 2. Ze heeft **bebot** gezegd. 3. Ze heeft **bloffen** gezegd. 4. Ze heeft **armel** gezegd. 5. Ze heeft **okte** gezegd. 6. Ze heeft **klort** gezegd. 7. Ze heeft **rons** gezegd. 8. Ze heeft **kang** gezegd. 9. Ze heeft **maags** gezegd. 10. Ze heeft **drokken** gezegd.
8	1. Ze heeft **vergief** gezegd. 2. Ze heeft **vaden** gezegd. 3. Ze heeft **fraat** gezegd. 4. Ze heeft **geband** gezegd. 5. Ze heeft **schoom** gezegd. 6. Ze heeft **kraden** gezegd. 7. Ze heeft **iengang** gezegd. 8. Ze heeft **baaks** gezegd. 9. Ze heeft **ootheid** gezegd. 10. Ze heeft **nocht** gezegd.
9	1. Ze heeft **lochten** gezegd. 2. Ze heeft **aanbag** gezegd. 3. Ze heeft **plap** gezegd. 4. Ze heeft **siengle** gezegd. 5. Ze heeft **balten** gezegd. 6. Ze heeft **rieder** gezegd. 7. Ze heeft **wienter** gezegd. 8. Ze heeft **mang** gezegd. 9. Ze heeft **mieddag** gezegd. 10. Ze heeft **ganger** gezegd.
10	1. Ze heeft **wienkel** gezegd. 2. Ze heeft **verlis** gezegd. 3. Ze heeft **vliender** gezegd. 4. Ze heeft **sponen** gezegd. 5. Ze heeft **zillig** gezegd. 6. Ze heeft **brif** gezegd. 7. Ze heeft **lanker** gezegd. 8. Ze heeft **vlaar** gezegd. 9. Ze heeft **actif** gezegd. 10. Ze heeft **zook** gezegd.

10/11

11	1. Ze heeft **broon** gezegd. 2. Ze heeft **morter** gezegd. 3. Ze heeft **gebocht** gezegd. 4. Ze heeft **afscheek** gezegd. 5. Ze heeft **haap** gezegd. 6. Ze heeft **slogen** gezegd. 7. Ze heeft **prister** gezegd. 8. Ze heeft **gonnen** gezegd. 9. Ze heeft **damen** gezegd. 10. Ze heeft **livver** gezegd.
12	1. Ze heeft **roog** gezegd. 2. Ze heeft **dodel** gezegd. 3. Ze heeft **braan** gezegd. 4. Ze heeft **kni** gezegd. 5. Ze heeft **gacht** gezegd. 6. Ze heeft **verbool** gezegd. 7. Ze heeft **waars** gezegd. 8. Ze heeft **difstal** gezegd. 9. Ze heeft **schitten** gezegd. 10. Ze heeft **zomel** gezegd.
13	1. Ze heeft **pietig** gezegd. 2. Ze heeft **dast** gezegd. 3. Ze heeft **baam** gezegd. 4. Ze heeft **lard** gezegd. 5. Ze heeft **krankt** gezegd. 6. Ze heeft **zik** gezegd. 7. Ze heeft **malken** gezegd. 8. Ze heeft **stolm** gezegd. 9. Ze heeft **zals** gezegd. 10. Ze heeft **zoren** gezegd.
14	1. Ze heeft **verschiel** gezegd. 2. Ze heeft **wost** gezegd. 3. Ze heeft **vliggen** gezegd. 4. Ze heeft **lifde** gezegd. 5. Ze heeft **diechter** gezegd. 6. Ze heeft **voormaal** gezegd. 7. Ze heeft **pops** gezegd. 8. Ze heeft **nonzig** gezegd. 9. Ze heeft **diensdag** gezegd. 10. Ze heeft **wienst** gezegd.

10.4 Stimuli Overview

ID	Item [26]	Type	Freq.Million	Zipf	ID	Item	Type	Freq.Million	Zipf
1	*dikwijls (often)*	/ɪ/-word	3.25	3.51	1	*visum (visa)*	/i/-word	3.7	3.57
2	*bink (hunk)*	/ɪ/-word	3.29	3.52	2	*actief (active)*	/i/-word	9.72	3.99
3	*pittig (spicy)*	/ɪ/-word	3.75	3.57	3	*knie (knee)*	/i/-word	10.24	4.01
4	*single (single)*	/ɪ/-word	4.07	3.61	4	*uniek (unique)*	/i/-word	12.74	4.11
5	*bitter (bitter)*	/ɪ/-word	4.32	3.64	5	*diefstal (theft)*	/i/-word	13.51	4.13
6	*vlinder (butterfly)*	/ɪ/-word	6.13	3.79	6	*titel (title)*	/i/-word	18.23	4.26
7	*kikker (frog)*	/ɪ/-word	8.23	3.92	7	*fiets (bike)*	/i/-word	21.75	4.34
8	*vergif (poison)*	/ɪ/-word	10.31	4.01	8	*zielig (pathetic)*	/i/-word	24.86	4.40
9	*schilder (painter)*	/ɪ/-word	10.5	4.02	9	*spiegel (mirror)*	/i/-word	27.44	4.44
10	*gids (guide)*	/ɪ/-word	11.39	4.06	10	*priester (priest)*	/i/-word	31.42	4.50
11	*ridder (knight)*	/ɪ/-word	13.58	4.13	11	*advies (advice)*	/i/-word	33.23	4.52
12	*ingang (entrance)*	/ɪ/-word	16.35	4.21	12	*techniek (technique)*	/i/-word	49.07	4.69
13	*winst (profit)*	/ɪ/-word	21.63	4.34	13	*verlies (loss)*	/i/-word	49.07	4.69
14	*dinsdag (Tuesday)*	/ɪ/-word	22.02	4.34	14	*brief (letter)*	/i/-word	73.84	4.87
15	*winter (winter)*	/ɪ/-word	22.36	4.35	15	*vliegen (fly)*	/i/-word	89.69	4.95
16	*middag (afternoon)*	/ɪ/-word	22.64	4.35	16	*dienst (service)*	/i/-word	92.73	4.97
17	*dichter (poet)*	/ɪ/-word	27.24	4.44	17	*ziek (ill)*	/i/-word	129.2	5.11
18	*vinger (finger)*	/ɪ/-word	28.91	4.46	18	*schieten (shoot)*	/i/-word	132.34	5.12
19	*verschil (difference)*	/ɪ/-word	54.22	4.73	19	*liever (rather)*	/i/-word	172.1	5.24
20	*winkel (shop)*	/ɪ/-word	65.13	4.81	20	*liefde (love)*	/i/-word	208.9	5.32
1a	*vis (fish)*	MWP-/ɪ/	50.8	4.71	1b	*vies (dirty)*	MWP-/i/	19.8	4.30
2a	*bid (pray)*	MWP-/ɪ/	19.46	4.29	2b	*bied (bid)*	MWP-/i/	21.86	4.34
3a	*wit (white)*	MWP-/ɪ/	33.48	4.52	3b	*wiet (weed)*	MWP-/i/	15.23	4.18
4a	*kist (case/chest)*	MWP-/ɪ/	27.67	4.44	4b	*kies (choose)*	MWP-/i/	16.85	4.23
5a	*lig (lie)*	MWP-/ɪ/	24.4	4.39	5b	*lieg (lie)*	MWP-/i/	30.8	4.49

[26] *ID* = identity of each item per type; *Item* = (pseudo)lexical item presented at test; *Type* = item type [/ɪ/-word = word with /ɪ/ as syllable nucleus; /i/-word = word with /i/ as syllable nucleus; MWP = minimal word pair]; *Freq.Million* = item frequency per million words based on the SUBTLEX-NL corpus; *Zipf* = Zipf value associated with each item.

ID	Item	Type	Freq.million	Zipf	ID	Item	Type	Freq.million	Zipf
1	*yoga (yoga)*	Filler	3.73	3.57	31	*zand (sand)*	Filler	19.99	4.30
2	*mental (mental)*	Filler	3.84	3.58	32	*hal (hallway)*	Filler	20.38	4.31
3	*marge (margin)*	Filler	3.98	3.60	33	*voordeel (advantage)*	Filler	21.04	4.32
4	*onrust (unrest)*	Filler	4.16	3.62	34	*vak (subject)*	Filler	22.02	4.34
5	*adel (nobility)*	Filler	4.28	3.63	35	*klok (clock)*	Filler	23.9	4.38
6	*pakket (parcel)*	Filler	4.85	3.69	36	*aap (monkey)*	Filler	28.56	4.46
7	*zakdoek (handkerchief)*	Filler	4.92	3.69	37	*bron (source)*	Filler	29.93	4.48
8	*mosterd (musterd)*	Filler	4.99	3.70	38	*pot (pot)*	Filler	30.62	4.49
9	*ballon (balloon)*	Filler	5.28	3.72	39	*taal (language)*	Filler	36.29	4.56
10	*afdruk (imprint)*	Filler	5.9	3.77	40	*verstand (sense)*	Filler	38.19	4.58
11	*nobel (noble)*	Filler	5.9	3.77	41	*aanval (attack)*	Filler	47.59	4.68
12	*hoofdstad (capital)*	Filler	6.43	3.81	42	*soldaat (soldier)*	Filler	53.03	4.72
13	*kapsel (hairdo)*	Filler	7.55	3.88	43	*last (load)*	Filler	53.49	4.73
14	*baron (baron)*	Filler	7.68	3.89	44	*dorp (village)*	Filler	53.99	4.73
15	*genot (enjoyment)*	Filler	8.12	3.91	45	*donker (dark)*	Filler	64.44	4.81
16	*raadsel (riddle)*	Filler	9.42	3.97	46	*zon (sun)*	Filler	68.67	4.84
17	*staal (steel)*	Filler	10.18	4.01	47	*tafel (table)*	Filler	83.4	4.92
18	*roman (novel)*	Filler	10.98	4.04	48	*paard (horse)*	Filler	83.63	4.92
19	*plank (plank)*	Filler	11.22	4.05	49	*raak (apt)*	Filler	83.67	4.92
20	*roos (rose)*	Filler	11.71	4.07	50	*lag (lay)*	Filler	87.04	4.94
21	*legal (legal)*	Filler	11.75	4.07	51	*gevaar (danger)*	Filler	91.54	4.96
22	*troon (throne)*	Filler	12.1	4.08	52	*boot (boat)*	Filler	95.35	4.98
23	*offer (offer)*	Filler	12.37	4.09	53	*frank (frank)*	Filler	104.69	5.02
24	*ton (barrel)*	Filler	13.72	4.14	54	*foto (picture)*	Filler	119.16	5.08
25	*koorts (fever)*	Filler	14.77	4.17	55	*normaal (normal)*	Filler	120.6	5.08
26	*ras (race)*	Filler	16.46	4.22	56	*kantoor (office)*	Filler	124.42	5.09
27	*staart (tail)*	Filler	17.95	4.25	57	*baas (boss)*	Filler	167.21	5.22
28	*doof (deaf)*	Filler	18.59	4.27	58	*los (loose)*	Filler	184.8	5.27
29	*bravo (bravo)*	Filler	19.32	4.29	59	*dokter (doctor)*	Filler	244.07	5.39
30	*droog (dry)*	Filler	19.64	4.29	60	*water (water)*	Filler	244.5	5.39

ID	Input	Zipf	Pseudoword	OLD20[27]	OLD20_Diff	Ned1	Ned1_Diff	Overlap Ratio	Maximum deviation	Summed deviation
1	*dapper*	4.33	lamper	1.45	0	11	2	2/3	179	552
2	*aanbod*	4.43	aanbag	1.8	0.1	4	0	2/3	176	293
3	*aanval*	2.50	aanbol	1.8	0.1	4	0	2/3	113	173
4	*afspraak*	4.99	afscheek	1.8	-0.1	4	3	2/3	52	112
5	*appel*	4.01	armel	1.65	0.05	7	1	2/3	-215	571
6	*ogen*	3.05	aven	1	0.05	21	1	2/3	835	1190
7	*baard*	2.92	baaks	1.55	0.4	9	-6	2/3	-4	6
8	*boom*	2.66	baam	1	0.05	21	2	2/3	-2	3
9	*vaarwel*	4.53	baarzel	1.9	-0.2	2	0	2/3	-86	129
10	*bouten*	3.19	balten	1.2	0.25	16	-4	2/3	207	282
11	*bezet*	4.24	bebot	1.7	0.25	6	-3	2/3	-257	550
12	*treffen*	4.11	bloffen	1.75	0	5	2	2/3	120	242
13	*boot*	4.98	bool	1.05	0.1	19	-12	2/3	-34	50
14	*boord*	4.77	boors	1.5	0.25	10	-3	2/3	11	16
15	*vogel*	4.51	bopel	1.85	0.15	3	-1	2/3	190	441
16	*traan*	3.42	braan	1.3	-0.1	14	4	2/3	14	17
17	*drogen*	3.54	bromen	1.45	0.05	11	1	2/3	65	200
18	*brein*	4.22	broon	1.5	0.15	10	-1	2/3	3	3
19	*haven*	4.23	damen	1	-0.05	21	4	2/3	-229	590
20	*vast*	5.82	dast	1	0.05	24	1	2/3	63	69
21	*motel*	4.20	dodel	1.55	0.15	9	-1	2/3	-239	542
22	*trekken*	4.91	drokken	1.55	0.05	9	1	2/3	-81	149
23	*jaren*	5.22	famen	1.2	-0.05	16	3	2/3	254	637
24	*sjaal*	3.72	fraal	1.65	-0.1	7	4	2/3	3	3
25	*fruit*	4.11	fraat	1.7	0.1	6	0	2/3	-4	5

[27] Variable descriptions are taken from Keuleers & Brysbaert (2010, pp. 632-633): *OLD20* = "[c]hecking this option computes the average orthographic Levenshtein distance between the generated candidate and its 20 most similar words in the lexicon"; *OLD20_Diff* = "difference in OLD20 between the generated nonword and the reference word is also shown"; *Ned1* = "number of orthographic neighbors at edit distance 1. This is the number of words that can be made from the candidate by substituting, deleting, or inserting a single letter"; *overlap ratio* = "the number of segments that overlap in the generated sequence and the reference sequence is shown as a fraction"; *maximum deviation* = "largest difference in transition frequencies between the subsyllabic segments in the generated sequence and those in the reference sequence"; *summed deviation* = "sum of all transition frequency deviations (absolute values) and a column showing where in the string the maximally deviating transition is situated".

ID	Input	Zipf	Pseudoword	OLD20	OLD20_Diff	Ned1	Ned1_Diff	Overlap ratio	Maximum deviation	Summed deviation
26	*vacht*	3.52	gacht	1.4	0.1	12	0	2/3	12	20
27	*golven*	3.96	galzen	1.6	0	8	2	2/3	-57	166
28	*vinger*	4.46	ganger	1.65	0.05	7	1	2/3	-452	689
29	*gewicht*	4.23	gebacht	1.65	0.1	7	0	2/3	295	480
30	*gezond*	4.50	geband	1.35	-0.1	13	4	2/3	355	652
31	*gezicht*	5.26	gebocht	1.65	0.15	7	-1	2/3	-177	363
32	*geloof*	5.59	geroom	1.7	0.1	6	0	2/3	-230	673
33	*gommen*	2.05	gonnen	1.55	0.3	9	-4	2/3	-223	478
34	*grens*	4.58	grons	1.55	0.1	9	0	2/3	8	13
35	*hoop*	5.56	haap	1.1	0.1	18	0	2/3	-8	11
36	*laten*	6.11	haden	1	0.05	27	-2	2/3	-239	482
37	*hand*	5.30	hank	1.1	0	18	2	2/3	18	26
38	*hang*	4.45	kang	1.2	0.25	16	-3	2/3	-8	12
39	*slager*	3.82	klamer	1.65	0.1	7	0	2/3	65	187
40	*sport*	4.30	klort	1.55	0.25	9	-3	2/3	2	3
41	*praten*	5.80	kraden	1.3	-0.1	14	4	2/3	-239	702
42	*kracht*	4.99	krankt	1.55	0.2	9	-2	2/3	-29	44
43	*spikkel*	2.13	krokkel	1.7	0.05	6	1	2/3	-45	69
44	*donker*	4.81	lanker	1.4	0.05	12	1	2/3	179	482
45	*lamp*	4.14	lard	1.35	0	13	2	2/3	-7	7
46	*wachten*	5.51	lochten	1.4	0.2	12	-2	2/3	96	163
47	*pop*	4.38	lop	1	0.05	34	3	2/3	44	48
48	*maagd*	4.35	maags	1.7	-0.1	6	4	2/3	7	10
49	*danken*	4.32	malken	1.2	-0.05	16	3	2/3	-248	746
50	*mand*	3.63	mang	1	-0.1	22	6	2/3	10	18

ID	Input	Zipf	Pseudoword	OLD20	OLD20_Diff	Ned1	Ned1_Diff	Overlap ratio	Maximum deviation	Summed deviation
51	*moment*	5.40	mogens	1.8	-0.1	4	4	2/3	-65	193
52	*lekker*	5.44	mokker	1.5	0.35	10	-5	2/3	-90	152
53	*dokter*	5.38	morter	1.65	0	7	2	2/3	89	239
54	*naald*	3.93	naafs	1.95	0.1	1	0	2/3	-11	11
55	*nacht*	5.31	nocht	1.45	0.1	11	0	2/3	-11	13
56	*zonnig*	3.47	nonzig	1.95	0.25	1	-3	2/3	-56	66
57	*orde*	4.70	okte	1.3	0.05	14	1	2/3	239	475
58	*einde*	5.05	oolde	1.2	0.25	16	-3	2/3	-213	303
59	*eenheid*	4.39	ootheid	1.85	-0.05	3	3	2/3	-44	108
60	*tanken*	3.70	pansen	1.3	0.15	14	-1	2/3	-241	540
61	*plan*	5.15	plap	1.35	0.2	13	-2	2/3	-12	20
62	*ploeg*	3.99	ploog	1.65	-0.1	7	4	2/3	2	3
63	*pond*	4.57	pops	1.1	-0.05	18	3	2/3	-10	16
64	*liefde*	5.32	raafde	1.6	0.1	8	0	2/3	33	53
65	*rand*	4.28	raps	1.3	0.35	14	-5	2/3	-10	15
66	*rots*	4.16	rons	1	0	20	2	2/3	-12	16
67	*loog*	4.40	roog	1	0.05	21	-1	2/3	-7	10
68	*jagen*	4.51	sanen	1.25	0.1	15	0	2/3	286	610
69	*schaal*	3.99	schaag	1.6	0.15	8	-1	2/3	-94	95
70	*schaar*	3.80	schaas	1.5	0.1	10	0	2/3	91	96
71	*schoon*	4.69	schoom	1.65	0.1	7	0	2/3	-46	47
72	*dromen*	4.89	slogen	1.4	0.05	12	1	2/3	112	271
73	*dwalen*	3.34	smaren	1.45	-0.1	11	4	2/3	-247	640
74	*forum*	3.12	sotus	1.9	0.15	2	-1	2/3	87	115
75	*graag*	5.76	spaag	1.6	0.35	8	-5	2/3	-9	14

ID	Input	Zipf	Pseudoword	OLD20	OLD20_Diff	Ned1	Ned1_Diff	Overlap ratio	Maximum deviation	Summed deviation
76	*storen*	4.34	sponen	1.4	0.4	12	-6	2/3	-624	1010
77	*storm*	4.47	stolm	1.6	0	8	2	2/3	-2	3
78	*pittig*	3.57	tattig	1.65	-0.05	7	3	2/3	67	126
79	*peinzen*	2.47	toonzen	1.9	0.25	2	-3	2/3	35	81
80	*kort*	4.67	tort	1	0.05	27	-1	2/3	-23	25
81	*zwak*	4.55	trak	1	-0.5	20	12	2/3	25	35
82	*goden*	4.47	vaden	1	-0.05	22	5	2/3	835	1063
83	*verhaal*	5.30	verbool	1.75	0.3	5	-4	2/3	-227	309
84	*verraad*	4.28	vertaag	1.6	-0.15	8	5	2/3	63	124
85	*zwaar*	4.90	vlaar	1.55	-0.1	9	4	2/3	-23	24
86	*boven*	5.39	vogen	1	0.05	20	-1	2/3	226	557
87	*vuil*	4.38	vool	1.5	0.35	10	-5	2/3	7	14
88	*voelt*	5.22	voolt	1.6	0.45	8	-7	2/3	-6	9
89	*voordeel*	4.32	voormaal	1.85	0.05	3	1	2/3	-252	481
90	*vorm*	4.59	vorn	1.5	-0.05	10	3	2/3	2	3
91	*vonden*	4.76	votten	1.15	0.2	17	-2	2/3	-242	615
92	*waard*	5.05	waars	1.25	-0.1	15	4	2/3	11	12
93	*wassen*	4.44	warken	1.25	0.15	15	-1	2/3	241	421
94	*roken*	4.64	wopen	1.2	0.25	16	-5	2/3	-396	1034
95	*west*	4.12	wost	1.1	0.15	18	-5	2/3	-14	16
96	*zand*	4.30	zals	1.45	0.25	11	-3	2/3	-15	20
97	*zenden*	3.79	zolden	1.05	0.1	19	-6	2/3	-213	341
98	*nobel*	3.77	zomel	1.7	-0.1	6	4	2/3	233	514
99	*zaak*	5.38	zook	1.15	0.2	17	-2	2/3	-11	14
100	*tonen*	4.57	zoren	1	0.05	23	-1	2/3	-475	815

10.5 Experiment Results

10.5.1 Pre-Test

Critical /ɪ/-words [28]					
Item	**Continuum Step**	**Response**		**Response Time (ms)**	
		/ɪ/	/i/	Mean	*SD*
bink	***1***	10	-	1126	*886*
	3	10	-	1051	*776*
	4	10	-	862	*732*
	5	10	-	1257	*770*
	6	5	5	1173	*1449*
	7	4	6	1076	*1107*
	9 (X)	2	8	645	*194*
bitter	***1***	9	1	912	*606*
	3	9	1	1713	*2326*
	4	9	1	1303	*879*
	5	4	6	2326	*3282*
	6 (X)	3	7	1006	*850*
	7	1	9	803	*757*
	9	1	9	829	*707*
dichter	***1***	6	4	1233	*1317*
	3	4	6	741	*603*
	4	-	10	1262	*1732*
	5	-	10	1247	*1286*
	6 (X)	2	8	564	*525*
	7	1	9	1296	*968*
	9	2	8	1191	*1056*
dikwijls	***1***	10	-	1386	*1233*
	3	10	-	1308	*1197*
	4	10	-	1327	*974*
	5	5	5	1005	*838*
	6	1	9	741	*522*
	7 (X)	2	8	1315	*1108*
	9	1	9	821	*890*
dinsdag	***1***	9	1	473	*452*
	3	8	2	922	*894*
	4	8	2	1115	*979*
	5	8	2	1818	*2674*
	6	5	5	1487	*1808*
	7	4	6	1034	*913*
	9 (X)	1	9	590	*490*
gids	***1***	10	-	569	*235*
	3	10	-	961	*752*
	4	8	2	808	*584*
	5	10	-	1350	*1451*
	6	7	3	1581	*1607*
	7 (X)	4	6	1458	*888*
	9	1	9	990	*979*

[28] **(X)** = step selected for the lexical decision task.

ingang	***1***	4	6	1546	*1093*
	3	3	7	1236	*1326*
	4	2	8	1780	*2208*
	5 (X)	3	7	1417	*1039*
	6	2	8	852	*1219*
	7	1	9	1923	*1920*
	9	2	8	665	*730*
kikker	***1***	9	1	819	*591*
	3	9	1	541	*227*
	4	10	-	742	*386*
	5	10	-	628	*285*
	6 (X)	3	7	1807	*3079*
	7	2	8	586	*398*
	9	1	9	716	*642*
middag	***1***	9	1	857	*855*
	3	7	3	1032	*696*
	4 (X)	3	7	1934	*2407*
	5	2	8	1570	*2374*
	6	1	9	1056	*811*
	7	2	8	572	*244*
	9	1	9	737	*631*
pittig	***1***	10	-	623	*229*
	3	9	1	1267	*1810*
	4	6	4	1514	*1404*
	5 (X)	6	4	2367	*2100*
	6	1	9	1339	*1704*
	7	1	9	934	*1005*
	9	-	10	949	*682*
ridder	***1***	10	-	785	*507*
	3	9	1	1279	*1110*
	4	9	1	1540	*1420*
	5	5	5	1412	*1076*
	6 (X)	3	7	1026	*807*
	7	-	10	982	*950*
	9	-	10	897	*704*
schilder	***1***	10	-	1385	*1211*
	3	8	2	1180	*921*
	4	7	3	889	*845*
	5	8	2	823	*472*
	6	4	6	1221	*1025*
	7 (X)	4	6	1316	*1339*
	9	-	10	1278	*1318*
single	***1***	6	4	1171	*1161*
	3	5	5	1588	*1660*
	4	4	6	1175	*1478*
	5	-	10	1525	*2226*
	6 (X)	3	7	1196	*1665*
	7	2	8	852	*795*
	9	1	9	994	*1530*

vergif	***1***	9	1	760	*660*
	3	10	-	1002	*997*
	4	9	1	974	*857*
	5	7	3	1490	*1485*
	6	4	6	1289	*1460*
	7 (X)	3	7	805	*919*
	9	2	8	795	*556*
verschil	***1***	8	2	1214	*708*
	3	7	3	1302	*927*
	4	7	3	1553	*2023*
	5	5	5	2051	*1978*
	6 (X)	3	7	1107	*843*
	7	1	9	1312	*1012*
	9	-	10	1456	*1986*
vinger	***1***	9	1	675	*415*
	3	9	1	641	*239*
	4	8	2	703	*494*
	5	10	-	780	*458*
	6	7	3	1223	*2104*
	7	5	5	1402	*1634*
	9 (X)	5	5	1106	*1457*
vlinder	***1***	10	-	535	*577*
	3	9	1	580	*587*
	4	10	-	1080	*1042*
	5	5	5	1423	*1625*
	6 (X)	3	7	929	*672*
	7	1	9	581	*360*
	9	1	9	990	*1200*
winkel	***1***	9	1	1136	*1073*
	3	5	5	1434	*2127*
	4	2	8	1147	*1384*
	5 (X)	3	7	1051	*1198*
	6	2	8	940	*826*
	7	2	8	1135	*1472*
	9	1	9	754	*618*
winst	***1***	7	3	1270	*846*
	3	7	3	1751	*1589*
	4	6	4	1464	*1473*
	5	5	5	1174	*961*
	6 (X)	4	6	2001	*1833*
	7	1	9	1275	*1483*
	9	-	10	895	*885*
winter	***1***	10	-	1001	*1276*
	3	8	2	1428	*1496*
	4	8	2	1156	*776*
	5	5	5	1286	*1087*
	6 (X)	3	7	630	*240*
	7	2	8	1112	*1554*
	9	1	9	1762	*2779*

Critical /i/-words					
Item	**Continuum Step**	**Response**		**Response Time (ms)**	
		/ɪ/	/i/	Mean	*SD*
actief	***1***	-	10	395	*190*
	3	-	10	1056	*1013*
	4	1	9	591	*387*
	5	4	6	1169	*1195*
	6	7	3	1109	*1017*
	7 (X)	7	3	858	*676*
	9	8	2	1311	*1502*
advies	***1***	-	10	650	*432*
	3	-	10	527	*381*
	4	1	9	555	*341*
	5	2	8	603	*442*
	6	4	6	887	*1318*
	7	5	5	597	*386*
	9 (X)	7	3	1739	*3421*
brief	***1***	-	10	1443	*2807*
	3	2	8	730	*466*
	4	5	5	1281	*1130*
	5	8	2	1306	*1219*
	6	8	2	1231	*1185*
	7 (X)	7	3	1290	*935*
	9	10	-	708	*718*
diefstal	***1***	-	10	544	*660*
	3	2	8	901	*804*
	4	1	9	996	*995*
	5	4	6	866	*812*
	6	3	7	863	*590*
	7 (X)	8	2	762	*1056*
	9	8	2	1609	*2559*
dienst	***1***	-	10	1021	*847*
	3	3	7	1226	*1499*
	4	4	6	2193	*2786*
	5 (X)	7	3	1105	*899*
	6	10	-	1545	*1761*
	7	10	-	1006	*888*
	9	10	-	1098	*934*
fiets	***1***	-	10	479	*261*
	3	-	10	697	*482*
	4	1	9	654	*506*
	5	3	7	1012	*1423*
	6	2	8	749	*678*
	7 (X)	7	3	859	*685*
	9	9	1	1016	*770*

knie	***1***	-	10	698	*482*
	3	2	8	882	*587*
	4	2	8	1094	*1053*
	5	2	8	1182	*953*
	6 (X)	5	5	1817	*2490*
	7	9	1	1508	*1521*
	9	8	2	2212	*3264*
liefde	***1***	-	10	556	*277*
	3	-	10	1003	*796*
	4	1	9	1066	*853*
	5	2	8	1107	*1251*
	6	5	5	1684	*1041*
	7	6	4	1826	*1682*
	9 (X)	8	2	1013	*1038*
liever	***1***	-	10	784	*670*
	3	-	10	968	*1242*
	4	3	7	1959	*2025*
	5 (X)	7	3	874	*466*
	6	9	1	1703	*1890*
	7	10	-	1076	*1028*
	9	9	1	1617	*1196*
priester	***1***	-	10	655	*499*
	3	-	10	385	*141*
	4	-	10	1703	*3805*
	5	3	7	1491	*2313*
	6	5	5	789	*527*
	7	5	5	1088	*1642*
	9 (X)	8	2	1368	*1602*
schieten	***1***	-	10	1018	*840*
	3	4	6	1311	*1382*
	4 (X)	8	2	1802	*1319*
	5	10	-	1041	*1307*
	6	10	-	822	*584*
	7	10	-	1500	*1707*
	9	9	1	1065	*898*
spiegel	***1***	-	10	822	*922*
	3	1	9	1078	*1084*
	4	3	7	1054	*696*
	5 (X)	8	2	1279	*1130*
	6	10	-	1149	*1196*
	7	10	-	1415	*1315*
	9	10	-	1741	*2039*
techniek	***1***	1	9	1508	*1459*
	3	2	8	1127	*827*
	4	2	8	1542	*1424*
	5	3	7	1508	*1370*
	6	2	8	1083	*735*
	7	3	7	929	*622*
	9 (X)	3	7	1243	*1035*

titel	***1***	1	9	681	*385*
	3	1	9	1363	*1176*
	4	3	7	1036	*903*
	5 (X)	8	2	1132	*919*
	6	9	1	895	*729*
	7	10	-	722	*537*
	9	9	1	818	*577*
uniek	***1***	-	10	790	*695*
	3	1	9	547	*344*
	4	3	7	597	*320*
	5	4	6	999	*742*
	6	4	6	746	*476*
	7 (X)	8	2	1167	*844*
	9	8	2	1066	*562*
verlies	***1***	-	10	709	*308*
	3	1	9	809	*906*
	4	1	9	881	*1113*
	5	3	7	844	*833*
	6 (X)	6	4	879	*607*
	7	10	-	1109	*951*
	9	10	-	929	*1155*
visum	***1***	1	9	994	*1147*
	3	4	6	1244	*1320*
	4	4	6	1000	*1277*
	5	8	2	1902	*2014*
	6 (X)	7	3	1363	*881*
	7	9	1	1859	*1641*
	9	8	2	1318	*1487*
vliegen	***1***	-	10	672	*623*
	3	2	8	1156	*1181*
	4 (X)	5	5	1341	*1001*
	5	10	-	2123	*2569*
	6	9	1	1472	*1061*
	7	10	-	1377	*793*
	9	10	-	1244	*1470*
ziek	***1***	-	10	438	*253*
	3	1	9	542	*244*
	4	1	9	544	*177*
	5 (X)	7	3	1267	*1223*
	6	10	-	961	*537*
	7	10	-	964	*746*
	9	10	-	1291	*1681*
zielig	***1***	-	10	812	*620*
	3	4	6	971	*733*
	4	6	4	1324	*878*
	5 (X)	6	4	922	*765*
	6	9	1	737	*586*
	7	9	1	889	*847*
	9	8	2	1303	*1532*

Minimal Word Pairs									
Item	**Cont. Step**	**Female voice**				**Male voice**			
		Response		Response Time (ms)		Response		Response Time (ms)	
		/ɪ/	/i/	*Mean*	SD	/ɪ/	/i/	*Mean*	SD
bid bied	***1***	9	1	1346	1470	9	1	1607	1713
	3	10	-	934	615	8	2	1305	1366
	4	9	1	1432	546	6	4	1869	1775
	5	5	5	1242	1156	6	4	1666	2139
	6	2	8	1126	1331	2	8	1027	983
	7	1	9	1026	937	1	9	1098	1060
	9	1	9	980	1226	-	10	925	632
kist kiest	***1***	10	-	646	449	10	-	838	920
	3	9	1	952	836	10	-	1324	1682
	4	9	1	1292	1075	10	-	652	295
	5	7	3	746	387	8	2	1535	1517
	6	2	8	1305	1057	8	2	1155	729
	7	2	8	1133	1173	6	4	959	757
	9	1	9	982	953	2	8	1051	1124
lig lieg	***1***	10	-	734	413	9	1	1133	1052
	3	9	1	1586	2404	8	2	601	326
	4	7	3	1044	782	7	3	766	360
	5	9	1	824	562	5	5	1369	1051
	6	1	9	1032	715	2	8	1208	803
	7	1	9	896	593	1	9	956	599
	9	-	10	709	525	-	10	755	742
vis vies	***1***	9	1	1284	1689	9	1	679	575
	3	10	-	1545	1847	7	3	1091	924
	4	6	4	900	516	7	3	1361	1189
	5	5	5	885	566	7	3	1611	1914
	6	1	9	1126	1224	3	7	1228	1262
	7	1	9	652	489	-	10	1029	980
	9	-	10	676	264	-	10	490	104
wit wiet	***1***	10	-	1183	1036	10	-	971	691
	3	8	2	981	972	9	1	994	721
	4	5	5	1076	961	10	-	701	397
	5	7	3	996	697	7	3	879	641
	6	1	9	1000	975	6	4	1067	866
	7	1	9	855	933	4	6	903	944
	9	-	10	745	485	-	10	586	202

10.5.2 Lexical Decision Task

Critical /ɪ/-words										
Item	**/ɪ/-ambiguous**					**/i/-ambiguous**				
	Response			Response time (ms)		Response			Response time (ms)	
	Word	*Non-word*	*X*[29]	*Mean*	*SD*	*Word*	*Non-word*	*X*	*Mean*	*SD*
	985	**152**	**3**	**448**	**354**	**829**	**30**	**1**	**348**	**249**
bink	27	30	-	599	509	21	22	-	541	334
bitter	43	13	1	440	314	43	-	-	311	112
dichter	56	1	-	397	305	43	-	-	308	146
dikwijls	44	13	-	438	227	43	-	-	315	129
dinsdag	54	2	1	505	505	43	-	-	331	202
gids	50	7	-	412	285	43	-	-	348	278
ingang	57	-	-	364	263	43	-	-	277	137
kikker	55	2	-	390	226	43	-	-	335	127
middag	57	-	-	367	234	43	-	-	305	196
pittig	28	28	1	595	437	43	-	-	308	130
ridder	25	32	-	596	414	43	-	-	367	307
schilder	56	1	-	359	190	43	-	-	298	150
single	53	4	-	645	457	39	3	1	583	564
vergif	46	11	-	550	444	43	-	-	288	223
verschil	57	-	-	372	261	43	-	-	385	247
vinger	53	4	-	397	259	42	1	-	383	254
vlinder	57	-	-	332	153	43	-	-	344	307
winkel	56	1	-	436	464	43	-	-	296	132
winst	54	3	-	361	270	42	1	-	331	188
winter	57	-	-	410	330	40	3	-	308	175

[29] *X* = response time limit of 4000ms was exceeded, hence no response was logged.

Critical /i/-words										
Item	/ɪ/-ambiguous					/i/-ambiguous				
	Response			Response time (ms)		Response			Response time (ms)	
	Word	*Non-word*	*X*	*Mean*	*SD*	*Word*	*Non-word*	*X*	*Mean*	*SD*
	1104	**34**	**2**	**375**	**301**	**535**	**320**	**5**	**502**	**368**
actief	56	1	-	388	276	43	-	-	357	168
advies	57	-	-	409	403	36	7	-	503	394
brief	57	-	-	302	273	16	26	1	533	365
diefstal	57	-	-	347	249	41	2	-	362	244
dienst	56	1	-	379	245	10	33	-	557	320
fiets	56	1	-	315	199	32	10	1	438	419
knie	56	-	1	407	365	26	17	-	493	288
liefde	57	-	-	312	170	28	15	-	601	443
liever	51	6	-	470	535	26	16	1	594	379
priester	56	-	1	336	249	43	-	-	365	203
schieten	57	-	-	351	248	14	29	-	580	418
spiegel	56	1	-	393	357	38	5	-	377	184
techniek	41	16	-	527	409	27	15	1	726	548
titel	54	3	-	331	215	30	13	-	475	499
uniek	57	-	-	371	276	35	8	-	369	187
verlies	56	1	-	337	218	16	27	-	569	422
visum	54	3	-	466	287	24	19	-	568	333
vliegen	57	-	-	334	157	27	16	-	598	451
ziek	56	1	-	387	330	12	31	-	485	274
zielig	57	-	-	330	232	11	31	1	505	325

Fillers										
Item	**/ɪ/-ambiguous**					**/i/-ambiguous**				
	Response			Response time (ms)		Response			Response time (ms)	
	Word	*Non-word*	*X*	*Mean*	*SD*	*Word*	*Non-word*	*X*	*Mean*	*SD*
	3197	**208**	**15**	**410**	**334**	**2435**	**138**	**7**	**371**	**278**
aanval	6	50	1	544	441	2	41	-	483	300
aap	57	-	-	361	225	43	-	-	342	253
adel	54	1	2	386	257	43	-	-	343	170
afdruk	57	-	-	390	309	43	-	-	371	339
baas	57	-	-	368	349	42	-	1	316	183
ballon	56	1	-	428	342	43	-	-	361	354
baron	47	10	-	500	368	41	2	-	406	311
boot	57	-	-	379	296	42	1	-	383	325
bravo	57	-	-	396	280	42	1	-	361	218
bron	50	7	-	414	258	37	6	-	404	211
dokter	57	-	-	409	266	43	-	-	309	185
donker	57	-	-	355	247	43	-	-	299	157
doof	55	2	-	362	372	37	6	-	418	389
dorp	53	4	-	495	504	43	-	-	291	144
droog	55	2	-	396	349	43	-	-	328	197
foto	56	-	1	342	271	43	-	-	338	337
frank	50	7	-	705	506	38	5	-	722	701
genot	54	2	1	410	340	42	1	-	383	216
gevaar	54	2	1	418	362	39	4	-	406	356
hal	40	17	-	663	507	30	12	1	560	467
hoofdstad	51	4	2	471	385	43	-	-	311	132
kantoor	57	-	-	381	247	42	1	-	298	127
kapsel	56	1	-	369	311	42	1	-	366	286
klok	55	2	-	313	182	43	-	-	357	268
koorts	56	-	1	369	334	43	-	-	331	192
lag	54	3	-	406	281	43	-	-	352	240
last	55	2	-	380	335	42	-	1	306	163
legaal	57	-	-	394	353	42	1	-	368	208
los	56	1	-	389	330	41	2	-	325	201
marge	55	1	1	376	257	43	-	-	316	138
mentaal	56	1	-	324	198	43	-	-	360	211
mosterd	57	-	-	295	323	43	-	-	348	240
nobel	34	22	1	492	356	33	10	-	457	368
normaal	57	-	-	336	210	43	-	-	323	177
offer	57	-	-	336	208	43	-	-	277	136

onrust	54	3	-	453	367	40	3	-	397	241
paard	57	-	-	331	174	42	1	-	360	231
pakket	57	-	-	420	462	43	-	-	317	165
plank	57	-	-	316	139	43	-	-	355	221
pot	56	1	-	413	332	42	1	-	347	286
raadsel	52	5	-	536	424	34	8	1	436	263
raak	52	4	1	453	410	41	2	-	415	203
ras	50	7	-	490	325	37	6	-	385	224
roman	55	2	-	344	147	43	-	-	316	180
roos	56	1	-	418	314	43	-	-	415	322
soldaat	57	-	-	337	279	42	-	1	346	211
staal	55	1	1	427	322	41	1	1	391	277
staart	57	-	-	367	301	43	-	-	387	256
taal	55	-	2	330	220	43	-	-	347	311
tafel	56	1	-	385	398	43	-	-	344	181
ton	51	6	-	475	394	41	2	-	415	379
troon	52	5	-	459	327	42	1	-	302	152
vak	32	25	-	597	484	24	18	1	508	409
verstand	56	1	-	459	369	43	-	-	409	409
voordeel	57	-	-	356	262	43	-	-	365	382
water	56	1	-	332	234	43	-	-	333	155
yoga	56	1	-	340	143	43	-	-	321	214
zakdoek	57	-	-	413	252	42	1	-	410	266
zand	55	2	-	361	241	43	-	-	334	138
zon	57	-	-	438	437	43	-	-	367	248

Nonwords										
Item	**/ɪ/-ambiguous**					**/i/-ambiguous**				
	Response			Response time (ms)		Response			Response time (ms)	
	Word	*Non-word*	*X*	*Mean*	*SD*	*Word*	*Non-word*	*X*	*Mean*	*SD*
	616	**5055**	**29**	**527**	**416**	**594**	**3695**	**11**	**498**	**379**
aanbag	2	55	-	642	546	2	41	-	488	363
aanbol	3	54	-	571	378	8	35	-	619	432
afscheek	2	55	-	556	432	5	38	-	437	296
armel	1	56	-	471	418	3	40	-	499	403
aven	5	51	1	597	484	2	41	-	484	290
baaks	-	57	-	415	247	-	43	-	353	174
baam	16	40	1	514	445	13	30	-	461	273
baarzel	-	57	-	455	347	-	43	-	438	297
balten	9	47	1	450	274	5	38	-	426	323
bebot	4	53	-	534	468	4	39	-	556	357
bloffen	3	53	1	543	449	6	37	-	558	451
bool	7	50	-	638	466	5	38	-	553	341
boors	8	49	-	528	364	9	34	-	504	336
bopel	4	53	-	403	310	2	40	1	399	329
braan	3	54	-	539	443	4	38	1	564	489
bromen	16	41	-	504	367	10	33	-	501	368
broon	1	56	-	438	317	4	39	-	552	514
damen	13	43	1	874	688	12	31	-	774	618
dast	-	57	-	537	474	3	40	-	480	291
dodel	-	57	-	491	407	3	40	-	457	292
drokken	3	54	-	501	379	2	41	-	477	410
famen	-	57	-	495	447	1	42	-	408	367
fraal	4	52	1	618	488	4	38	1	600	425
fraat	5	52	-	529	410	4	39	-	602	432
gacht	41	16	-	547	450	33	10	-	376	190
galzen	1	55	1	406	317	1	42	-	452	466
ganger	7	50	-	759	588	7	36	-	801	689
gebacht	2	54	1	477	370	2	41	-	489	342
geband	3	54	-	520	313	4	39	-	514	358
gebocht	-	57	-	505	392	2	41	-	494	505
geroom	5	52	-	637	459	8	35	-	643	454
gonnen	37	20	-	684	558	31	12	-	410	246
grons	3	53	1	533	408	3	40	-	539	370
haap	7	49	1	630	488	13	30	-	645	477
haden	2	55	-	602	413	5	38	-	644	392

hank	34	23	-	655	537	27	15	1	590	573
kang	8	49	-	558	419	5	38	-	568	384
klamer	1	56	-	452	288	4	39	-	434	347
klort	-	55	2	386	398	-	43	-	351	202
kraden	2	55	-	498	463	-	43	-	352	196
krankt	-	56	1	521	383	3	40	-	505	286
krokkel	-	57	-	400	283	-	43	-	400	242
lamper	-	56	1	400	304	-	43	-	413	266
lanker	6	51	-	542	421	11	32	-	464	257
lard	-	57	-	330	135	1	42	-	494	414
lochten	2	55	-	533	406	1	42	-	384	246
lop	10	46	1	580	424	10	33	-	488	355
maags	1	56	-	500	462	1	42	-	387	193
malken	34	23	-	438	286	25	18	-	513	467
mang	16	40	1	566	370	15	28	-	622	508
mogens	6	49	2	670	499	8	35	-	652	341
mokker	7	50	-	575	537	3	39	1	515	337
morter	4	53	-	529	456	3	40	-	504	413
naafs	-	57	-	409	278	-	43	-	386	258
nocht	3	53	1	671	581	7	36	-	506	326
nonzig	-	57	-	395	231	-	43	-	410	252
okte	-	57	-	451	314	1	42	-	541	439
oolde	-	57	-	517	302	-	42	1	491	319
ootheid	9	48	-	622	390	19	24	-	662	561
pansen	-	57	-	413	260	-	43	-	422	227
plap	2	55	-	394	211	3	40	-	345	219
ploog	6	50	1	469	434	4	39	-	471	318
pops	1	55	1	678	505	2	41	-	528	400
raafde	4	53	-	567	381	4	39	-	533	478
raps	5	51	1	508	425	5	38	-	485	426
rons	1	56	-	578	456	2	41	-	419	290
roog	3	54	-	455	283	1	42	-	498	413
sanen	8	49	-	522	326	4	39	-	391	290
schaag	29	28	-	548	348	26	16	1	576	305
schaas	1	56	-	432	271	-	43	-	421	271
schoom	7	50	-	574	415	9	34	-	408	218
slogen	24	33	-	634	554	23	20	-	475	394
smaren	-	56	1	409	242	-	43	-	374	300
sotus	-	57	-	466	285	2	41	-	503	433
spaag	5	50	2	444	310	8	35	-	428	309
sponen	3	54	-	478	351	-	43	-	462	285

stolm	3	54	-	510	382	3	40	-	487	382
tattig	1	56	-	423	371	1	42	-	418	281
toonzen	1	56	-	609	479	3	40	-	561	397
tort	4	53	-	563	460	5	38	-	412	235
trak	4	53	-	484	331	2	40	1	545	374
vaden	4	53	-	481	402	5	37	1	531	328
verbool	4	53	-	562	333	5	38	-	578	443
vertaag	1	56	-	469	312	1	42	-	495	337
vlaar	6	51	-	516	331	10	33	-	568	357
vogen	13	43	1	572	493	6	37	-	528	377
vool	3	54	-	551	440	4	39	-	544	384
voolt	2	55	-	414	274	-	43	-	522	262
voormaal	22	35	-	750	524	14	29	-	668	625
vorn	46	11	-	433	388	31	12	-	490	465
votten	9	47	1	515	361	6	37	-	491	351
waars	2	55	-	591	417	1	42	-	466	243
warken	5	52	-	484	317	11	32	-	553	331
wopen	4	53	-	551	472	3	40	-	487	323
wost	2	55	-	572	413	4	38	1	501	312
zals	1	56	-	556	455	2	41	-	375	237
zolden	9	47	1	701	533	10	32	1	653	491
zomel	1	56	-	482	413	1	42	-	460	477
zook	7	49	1	472	439	6	37	-	382	310
zoren	3	54	-	573	507	3	40	-	497	421

10.5.3 Phoneme Categorization Task

bid – bied																					
Step	**It.**	**Female voice**										**Male voice**									
		/ɪ/-ambiguous					**/i/-ambiguous**					**/ɪ/-ambiguous**					**/i/-ambiguous**				
		Response			Response time (ms)		Response			Response time (ms)		Response			Response time (ms)		Response			Response time (ms)	
		/ɪ/	/i/	X	*M*	*SD*	/ɪ/	/i/	X	*M*	*SD*	/ɪ/	/i/	X	*M*	*SD*	/ɪ/	/i/	X	*M*	*SD*
1		**140**	**12**	-	**415**	**350**	**122**	**22**	-	**357**	**240**	**204**	**3**	**1**	**488**	**349**	**116**	**4**	-	**413**	**265**
	1	19	-	-	426	268	17	1	-	400	232	25	-	1	633	233	15	-	-	738	417
	2	18	1	-	390	279	16	2	-	390	302	26	-	-	446	191	13	2	-	479	221
	3	19	-	-	421	385	15	3	-	368	301	25	1	-	442	302	14	1	-	364	170
	4	16	3	-	346	217	16	2	-	326	165	26	-	-	426	248	15	-	-	336	153
	5	17	2	-	404	407	15	3	-	379	281	26	-	-	630	453	14	1	-	371	231
	6	16	3	-	466	480	14	4	-	333	231	26	-	-	371	227	15	-	-	382	273
	7	18	1	-	383	330	14	4	-	370	231	26	-	-	494	461	15	-	-	292	128
	8	17	2	-	486	406	15	3	-	290	155	24	2	-	472	480	15	-	-	340	162
2		**134**	**18**	-	**404**	**386**	**127**	**17**	-	**409**	**384**	**207**	**1**	-	**477**	**317**	**117**	**3**	-	**418**	**238**
	1	19	-	-	469	524	17	1	-	624	567	26	-	-	779	375	15	-	-	599	256
	2	17	2	-	334	266	17	1	-	374	359	26	-	-	469	191	14	1	-	437	194
	3	16	3	-	332	159	16	2	-	317	196	26	-	-	498	237	15	-	-	328	117
	4	16	3	-	341	318	16	2	-	430	287	25	1	-	416	286	14	1	-	409	218
	5	16	3	-	350	238	15	3	-	337	256	26	-	-	366	217	15	-	-	416	326
	6	17	2	-	511	598	16	2	-	499	565	26	-	-	421	256	15	-	-	347	230
	7	16	3	-	421	322	15	3	-	372	369	26	-	-	437	305	15	-	-	406	235
	8	17	2	-	476	466	15	3	-	320	253	26	-	-	428	439	14	1	-	403	229

		115	**37**	-	**424**	**314**	**121**	**23**	-	**389**	**312**	**191**	**17**	-	**522**	**357**	**111**	**9**	-	**426**	**302**
	1	19	-	-	474	389	18	-	-	448	208	25	1	-	739	432	13	2	-	804	453
	2	16	3	-	439	247	16	2	-	329	114	24	2	-	529	315	14	1	-	490	280
	3	16	3	-	394	283	15	3	-	285	177	25	1	-	532	389	13	2	-	379	175
3	4	14	5	-	425	411	15	3	-	374	288	25	1	-	468	310	14	1	-	447	373
	5	15	4	-	441	301	14	4	-	505	506	23	3	-	500	377	14	1	-	358	204
	6	11	8	-	444	281	14	4	-	329	244	23	3	-	488	411	14	1	-	342	177
	7	12	7	-	445	400	16	2	-	474	440	22	4	-	459	268	15	-	-	286	108
	8	12	7	-	330	154	13	5	-	365	306	24	2	-	457	282	14	1	-	298	165
	4	**96**	**56**	-	**447**	**361**	**113**	**31**	-	**401**	**287**	**184**	**23**	**1**	**532**	**372**	**116**	**4**	-	**495**	**313**
	1	14	5	-	647	340	16	2	-	588	334	24	2	-	672	287	15	-	-	817	426
	2	16	3	-	506	409	14	4	-	500	312	24	2	-	478	310	15	-	-	618	398
	3	15	4	-	386	229	15	3	-	318	180	22	4	-	500	342	15	-	-	429	207
4	4	11	8	-	478	571	14	4	-	387	200	22	4	-	467	277	14	1	-	459	300
	5	9	10	-	434	264	13	5	-	377	293	23	3	-	596	500	15	-	-	418	184
	6	11	8	-	368	170	13	5	-	285	231	23	2	1	434	287	15	-	-	436	272
	7	10	9	-	399	498	14	4	-	418	262	23	3	-	557	441	13	2	-	419	209
	8	10	9	-	355	129	14	4	-	338	363	23	3	-	546	446	14	1	-	360	199
Total		**485**	**123**	-	**422**	**353**	**483**	**93**	-	**389**	**310**	**786**	**44**	**2**	**505**	**349**	**460**	**20**	-	**438**	**282**

kist – kiest																					
Step	**It.**	**Female voice**										**Male voice**									
		/ɪ/-ambiguous					**/i/-ambiguous**					**/ɪ/-ambiguous**					**/i/-ambiguous**				
		Response			Response time (ms)		Response			Response time (ms)		Response			Response time (ms)		Response			Response time (ms)	
		/ɪ/	/i/	X	*M*	*SD*	/ɪ/	/i/	X	*M*	*SD*	/ɪ/	/i/	X	*M*	*SD*	/ɪ/	/i/	X	*M*	*SD*
1		**206**	**2**	-	**416**	**414**	**169**	**14**	**1**	**417**	**367**	**223**	**25**	-	**462**	**284**	**143**	**17**	-	**455**	**349**
	1	25	1	-	635	457	23	-	-	604	547	29	2	-	577	271	19	1	-	508	173
	2	25	1	-	367	359	22	1	-	459	317	27	4	-	559	335	19	1	-	612	558
	3	26	-	-	403	326	21	2	-	362	273	27	4	-	471	337	17	3	-	394	189
	4	26	-	-	440	555	19	4	-	373	249	28	3	-	418	235	18	2	-	416	266
	5	26	-	-	270	179	20	3	-	402	453	29	2	-	371	193	19	1	-	424	313
	6	26	-	-	331	223	21	2	-	427	367	27	4	-	444	315	19	1	-	385	198
	7	26	-	-	342	277	22	1	-	334	233	27	4	-	390	267	15	5	-	484	545
	8	26	-	-	540	637	21	1	1	372	369	29	2	-	469	250	17	3	-	421	305
2		**203**	**5**	-	**402**	**331**	**173**	**11**	-	**393**	**308**	**218**	**28**	**2**	**496**	**316**	**129**	**30**	**1**	**442**	**262**
	1	25	1	-	518	497	23	-	-	544	315	28	3	--	592	272	19	1	-	588	305
	2	25	1	-	356	312	21	2	-	402	331	27	3	1	521	296	17	3	-	516	391
	3	25	1	-	419	280	22	1	-	429	503	27	4	-	555	423	16	4	-	432	232
	4	26	-	-	402	313	21	2	-	354	234	27	4	-	550	422	15	5	-	383	218
	5	25	1	-	464	402	21	2	-	402	240	29	2	-	450	240	16	4	-	387	294
	6	26	-	-	340	192	21	2	-	300	180	28	2	1	501	328	16	3	1	404	157
	7	25	1	-	397	352	22	1	-	342	251	26	5	-	410	255	15	5	-	409	203
	8	26	-	-	322	186	22	1	-	368	285	26	5	-	386	183	15	5	-	415	186

3		**203**	**5**	-	**425**	**333**	**165**	**18**	**1**	**413**	**337**	**181**	**67**	-	**572**	**416**	**107**	**53**	-	**450**	**240**
	1	26	-	-	553	532	22	-	1	569	396	23	8	-	748	507	15	5	-	556	220
	2	25	1	-	411	342	21	2	-	405	259	25	6	-	492	320	15	5	-	413	147
	3	25	1	-	328	223	20	3	-	411	455	21	10	-	529	288	12	8	-	493	250
	4	26	-	-	399	309	19	4	-	403	298	23	8	-	591	404	15	5	-	415	240
	5	25	1	-	438	361	21	2	-	399	266	20	11	-	661	401	16	4	-	371	188
	6	25	1	-	498	308	20	3	-	349	209	21	10	-	462	362	11	9	-	442	248
	7	26	-	-	379	197	21	2	-	428	485	24	7	-	542	469	12	8	-	448	311
	8	25	1	-	392	273	21	2	-	346	211	24	7	-	549	492	11	9	-	462	273
4		**195**	**11**	**2**	**437**	**426**	**165**	**19**	-	**425**	**348**	**103**	**145**	-	**554**	**347**	**46**	**114**	-	**526**	**340**
	1	25	1	-	586	501	21	2	-	543	310	20	11	-	705	388	7	13	-	742	437
	2	24	2	-	467	480	20	3	-	385	247	15	16	-	504	200	6	14	-	618	421
	3	24	1	1	380	222	21	2	-	472	380	12	19	-	563	488	7	13	-	477	213
	4	24	1	1	400	419	20	3	-	385	228	10	21	-	598	383	6	14	-	507	408
	5	24	2	-	448	389	19	4	-	351	282	9	22	-	480	239	6	14	-	442	230
	6	25	1	-	463	598	21	2	-	462	517	12	19	-	497	286	5	15	-	501	345
	7	24	2	-	354	406	21	2	-	440	442	11	20	-	561	350	6	14	-	433	241
	8	25	1	-	396	283	22	1	-	360	282	14	17	-	525	343	3	17	-	486	285
Total		**807**	**23**	**2**	**420**	**378**	**672**	**62**	**2**	**412**	**340**	**725**	**265**	**2**	**521**	**347**	**425**	**214**	**1**	**468**	**303**

lig – lieg																					
Step	**It.**	**Female voice**										**Male voice**									
		/ɪ/-ambiguous					**/i/-ambiguous**					**/ɪ/-ambiguous**					**/i/-ambiguous**				
		Response			Response time (ms)		Response			Response time (ms)		Response			Response time (ms)		Response			Response time (ms)	
		/ɪ/	/i/	X	*M*	*SD*	/ɪ/	/i/	X	*M*	*SD*	/ɪ/	/i/	X	*M*	*SD*	/ɪ/	/i/	X	*M*	*SD*
1		**201**	**6**	**1**	**395**	**282**	**166**	**17**	**1**	**369**	**308**	**246**	**2**	-	**476**	**399**	**156**	**3**	**1**	**384**	**190**
	1	26	-	-	484	337	23	-	-	439	196	31	-	-	711	484	20	-	-	533	166
	2	25	1	-	385	293	23	-	-	326	203	31	-	-	527	379	19	1	-	440	208
	3	24	2	-	350	308	22	1	-	300	188	30	1	-	480	332	19	1	-	406	200
	4	26	-	-	389	255	21	2	-	341	225	31	-	-	509	542	20	-	-	356	193
	5	26	-	-	388	310	18	5	-	456	481	31	-	-	421	264	19	-	1	323	179
	6	24	1	1	330	164	20	3	-	424	466	31	-	-	322	165	20	-	-	320	145
	7	25	1	-	376	251	19	3	1	368	358	31	-	-	420	328	20	-	-	340	134
	8	25	1	-	455	306	20	3	-	299	137	30	1	-	420	482	19	1	-	349	209
2		**200**	**7**	**1**	**390**	**281**	**164**	**19**	**1**	**385**	**357**	**240**	**8**	-	**421**	**299**	**156**	**4**	-	**422**	**325**
	1	26	-	-	659	382	22	1	-	536	565	29	2	-	617	302	19	1	-	598	423
	2	26	-	-	325	207	21	2	-	404	355	30	1	-	479	246	20	-	-	501	463
	3	25	1	-	358	233	22	1	-	374	388	30	1	-	438	407	18	2	-	415	284
	4	25	1	-	313	187	20	3	-	348	269	30	1	-	394	298	20	-	-	351	203
	5	24	2	-	438	317	20	3	-	515	457	31	-	-	403	325	20	-	-	375	221
	6	25	1	-	303	272	20	3	-	347	244	30	1	-	319	166	20	-	-	378	284
	7	24	1	1	351	224	20	2	1	327	186	30	1	-	390	246	20	-	-	371	317
	8	25	1	-	369	228	19	4	-	227	110	30	1	-	330	264	19	1	-	390	294

3		**192**	**16**	-	**449**	**319**	**154**	**30**	-	**385**	**330**	**240**	**7**	**1**	**454**	**370**	**155**	**5**	-	**464**	**382**
	1	26	-	-	514	321	21	2	-	505	362	30	-	1	759	434	20	-	-	806	522
	2	25	1	-	421	313	21	2	-	308	162	29	2	-	408	231	18	2	-	499	431
	3	25	1	-	324	232	21	2	-	329	168	29	2	-	521	532	20	-	-	377	261
	4	23	3	-	496	332	20	3	-	386	285	31	-	-	423	374	20	-	-	408	234
	5	24	2	-	495	450	18	5	-	348	164	31	-	-	424	399	20	-	-	413	261
	6	22	4	-	477	343	19	4	-	353	400	31	-	-	363	258	19	1	-	497	568
	7	23	3	-	490	297	16	7	-	344	299	29	2	-	388	273	19	1	-	328	150
	8	24	2	-	378	192	18	5	-	509	564	30	1	-	353	203	19	1	-	382	250
4		**168**	**39**	**1**	**415**	**321**	**145**	**39**	-	**404**	**304**	**224**	**23**	**1**	**503**	**418**	**135**	**25**	-	**463**	**359**
	1	25	1	-	530	331	20	3	-	569	345	26	4	1	732	450	16	4	-	705	563
	2	24	2	-	417	315	21	2	-	413	280	29	2	-	491	359	17	3	-	514	381
	3	23	3	-	340	190	20	3	-	302	169	29	2	-	419	271	15	5	-	450	205
	4	23	3	-	366	240	18	5	-	396	297	28	3	-	571	588	16	4	-	404	212
	5	18	8	-	481	553	16	7	-	344	321	27	4	-	529	484	18	2	-	383	193
	6	18	8	-	440	328	17	6	-	460	386	28	3	-	471	372	19	1	-	408	314
	7	18	7	1	351	181	16	7	-	339	205	28	3	-	394	348	16	4	-	458	507
	8	19	7	-	393	268	17	6	-	411	333	29	2	-	424	332	18	2	-	386	231
Total		**761**	**68**	**3**	**412**	**302**	**629**	**105**	**2**	**386**	**325**	**950**	**40**	**2**	**464**	**375**	**602**	**37**	**1**	**433**	**324**

vis – vies																					
Step	It.	Female voice										Male voice									
		/ɪ/-ambiguous					/i/-ambiguous					/ɪ/-ambiguous					/i/-ambiguous				
		Response			Response time (ms)		Response			Response time (ms)		Response			Response time (ms)		Response			Response time (ms)	
		/ɪ/	/i/	X	*M*	*SD*	/ɪ/	/i/	X	*M*	*SD*	/ɪ/	/i/	X	*M*	*SD*	/ɪ/	/i/	X	*M*	*SD*
1	**1**	**194**	**12**	**2**	**506**	**507**	**164**	**19**	**1**	**488**	**411**	**245**	**2**	**1**	**514**	**399**	**135**	**25**	-	**468**	**364**
	1	24	1	1	860	669	22	1	-	702	442	31	-	-	877	636	19	1	-	867	639
	2	24	1	1	489	535	22	1	-	562	503	31	-	-	547	420	15	5	-	486	295
	3	26	-	-	491	621	17	4	1	465	370	30	1	-	532	379	17	3	-	513	336
	4	24	2	-	413	417	19	4	-	442	372	31	-	-	408	233	16	4	-	372	210
	5	24	2	-	462	322	22	2	-	351	202	31	-	-	426	244	18	2	-	361	209
	6	25	1	-	451	397	21	2	-	384	351	30	1	-	413	345	17	3	-	383	328
	7	22	4	-	481	571	21	2	-	457	258	30	-	1	482	342	17	3	-	345	181
	8	25	1	-	411	328	20	3	-	542	600	31	-		425	247	16	4	-	418	236
2	**2**	**188**	**20**	-	**566**	**550**	**168**	**15**	**1**	**436**	**399**	**236**	**11**	**1**	**514**	**353**	**132**	**27**	**1**	**479**	**397**
	1	26	-	-	872	762	23	-	-	720	664	30	-	1	688	453	17	2	1	867	593
	2	26	-	-	450	344	20	2	1	431	356	31	-	-	520	359	16	4	-	580	363
	3	24	2	-	635	695	21	2	-	422	264	29	2	-	563	382	17	3	-	480	429
	4	21	5	-	520	502	20	3	-	537	436	25	6	-	598	414	15	5	-	346	171
	5	22	4	-	500	458	21	2	-	377	240	31	-	-	491	391	19	1	-	424	297
	6	22	4	-	561	514	21	2	-	278	165	30	1	-	371	187	16	4	-	441	493
	7	24	2	-	439	460	20	3	-	453	478	30	1	-	423	222	16	4	-	329	192
	8	23	3	-	548	493	22	1	-	272	165	30	1	-	462	254	16	4	-	386	208

3	**3**	**194**	**14**	-	**457**	**417**	**164**	**20**	-	**484**	**481**	**229**	**18**	**1**	**561**	**433**	**128**	**32**	-	**455**	**353**
	1	25	1	-	570	391	22	1	-	904	813	30	1	-	988	623	15	5	-	799	639
	2	26	-	-	481	485	21	2	-	575	389	29	1	1	552	393	17	3	-	525	345
	3	24	2	-	397	255	20	3	-	403	374	26	5	-	546	354	15	5	-	383	218
	4	24	2	-	326	216	19	4	-	479	425	29	2	-	563	359	15	5	-	422	234
	5	23	3	-	500	581	19	4	-	437	488	31	-	-	486	408	16	4	-	469	290
	6	23	3	-	605	614	21	2	-	383	351	28	3	-	482	364	17	3	-	348	211
	7	25	1	-	364	259	21	2	-	333	166	28	3	-	502	348	19	1	-	312	177
	8	24	2	-	412	289	21	2	-	356	382	28	3	-	370	288	14	6	-	381	275
4	**4**	**165**	**42**	**1**	**538**	**503**	**154**	**30**	-	**518**	**449**	**199**	**47**	**2**	**609**	**446**	**112**	**48**	-	**490**	**392**
	1	21	4	1	779	590	19	4	-	759	716	24	5	2	953	603	17	3	-	860	603
	2	21	5	-	532	497	22	1	-	547	496	25	6	-	588	349	14	6	-	538	330
	3	22	4	-	588	599	18	5	-	499	432	23	8	-	614	475	13	7	-	375	191
	4	19	7	-	513	272	21	2	-	535	430	27	4	-	551	355	13	7	-	386	227
	5	19	7	-	504	539	19	4	-	494	261	25	6	-	543	384	14	6	-	417	196
	6	22	4	-	379	266	19	4	-	545	502	23	8	-	551	387	13	7	-	690	572
	7	21	5	-	616	723	18	5	-	317	175	27	4	-	527	438	14	6	-	392	240
	8	20	6	-	406	272	18	5	-	445	285	25	6	-	564	434	14	6	-	265	132
Total		**741**	**88**	**3**	**517**	**497**	**650**	**84**	**2**	**481**	**436**	**909**	**78**	**5**	**549**	**411**	**507**	**132**	**1**	**473**	**376**

wit – wiet																					
Step	**It.**	**Female voice**										**Male voice**									
		/ɪ/-ambiguous					**/i/-ambiguous**					**/ɪ/-ambiguous**					**/i/-ambiguous**				
		Response			Response time (ms)		Response			Response time (ms)		Response			Response time (ms)		Response			Response time (ms)	
		/ɪ/	/i/	X	*M*	*SD*	/ɪ/	/i/	X	*M*	*SD*	/ɪ/	/i/	X	*M*	*SD*	/ɪ/	/i/	X	*M*	*SD*
1		**207**	**1**	-	**391**	**267**	**181**	**3**	-	**344**	**226**	**233**	**14**	**1**	**495**	**346**	**153**	**7**	-	**422**	**220**
	1	26	-	-	536	318	23	-	-	469	283	30	1	-	643	299	20	-	-	540	226
	2	26	-	-	361	205	23	-	-	427	297	30	1	-	490	292	19	1	-	446	195
	3	25	1	-	388	281	23	-	-	300	144	29	2	-	519	318	20	-	-	461	268
	4	26	-	-	370	240	22	1	-	311	202	27	3	1	481	397	19	1	-	393	204
	5	26	-	-	403	296	22	1	-	369	202	29	2	-	434	310	18	2	-	405	220
	6	26	-	-	330	201	22	1	-	297	239	29	2	-	449	314	19	1	-	381	217
	7	26	-	-	323	152	23	-	-	306	147	29	2	-	484	405	19	1	-	320	146
	8	26	-	-	418	354	23	-	-	276	201	30	1	-	456	404	19	1	-	427	232
2		**206**	**2**	-	**426**	**357**	**184**	-	-	**378**	**311**	**233**	**15**	-	**480**	**337**	**151**	**9**	-	**416**	**269**
	1	26	-	-	431	257	23	-	-	565	341	30	1	-	706	368	18	2	-	686	367
	2	26	-	-	504	447	23	-	-	336	200	29	2	-	482	291	19	1	-	436	222
	3	25	1	-	379	242	23	-	-	304	311	28	3	-	486	319	18	2	-	361	305
	4	26	-	-	454	511	23	-	-	353	313	27	4	-	481	434	20	-	-	375	162
	5	25	1	-	411	307	23	-	-	384	329	29	2	-	487	379	19	1	-	383	223
	6	26	-	-	369	282	23	-	-	357	212	30	1	-	416	252	19	1	-	424	260
	7	26	-	-	395	278	23	-	-	413	468	30	1	-	377	196	19	1	-	341	183
	8	26	-	-	465	454	23	-	-	314	175	30	1	-	407	320	19	1	-	324	236

3		**203**	**5**	-	**408**	**358**	**180**	**4**	-	**380**	**329**	**203**	**45**	-	**566**	**478**	**122**	**38**	-	**411**	**238**
	1	26	-	-	497	352	23	-	-	462	276	26	5	-	1015	723	15	5	-	661	350
	2	25	1	-	362	260	23	-	-	335	273	25	6	-	494	233	18	2	-	445	165
	3	25	1	-	374	279	23	-	-	338	188	25	6	-	466	344	17	3	-	450	280
	4	25	1	-	332	193	22	1	-	420	239	26	5	-	547	443	14	6	-	349	164
	5	25	1	-	429	435	23	-	-	363	383	24	7	-	480	321	15	5	-	362	203
	6	26	-	-	452	444	22	1	-	268	157	27	4	-	464	326	14	6	-	354	146
	7	25	1	-	432	446	22	1	-	381	282	25	6	-	527	510	14	6	-	325	133
	8	26	-	-	386	384	22	1	-	470	611	25	6	-	536	527	15	5	-	344	216
4		**204**	**3**	**1**	**383**	**276**	**182**	**2**	-	**362**	**270**	**29**	**218**	**1**	**503**	**283**	**13**	**146**	**1**	**428**	**245**
	1	26	-	-	464	203	23	-	-	438	293	8	22	1	767	407	3	17	-	634	313
	2	26	-	-	350	164	23	-	-	400	311	3	28	-	546	265	1	19	-	447	304
	3	26	-	-	361	319	23	-	-	391	306	3	28	-	470	243	2	18	-	375	170
	4	25	1	-	396	415	23	-	-	365	240	3	28	-	532	240	2	18	-	359	165
	5	26	-	-	362	301	23	-	-	369	358	6	25	-	452	249	2	18	-	405	271
	6	24	1	1	369	263	22	1	-	355	281	2	29	-	378	151	2	18	-	404	198
	7	25	1	-	373	248	23	-	-	299	160	2	29	-	433	199	1	19	-	399	232
	8	26	-	-	387	243	22	1	-	277	132	2	29	-	456	291	-	19	1	402	175
Total		**820**	**11**	**1**	**402**	**317**	**727**	**9**	-	**366**	**287**	**698**	**292**	**2**	**511**	**369**	**439**	**200**	**1**	**419**	**243**