Lexical processing of foreign-accented speech:

Rapid and flexible adaptation

Cover design: Phaidra Rullmann

ISBN: 9789076203447

Printed and bound by Ipskamp Drukkers b.v.

Lexical processing of foreign-accented speech: Rapid and flexible adaptation

Proefschrift

ter verkrijging van de graad van doctor

aan de Radboud Universiteit Nijmegen

op gezag van de rector magnificus prof. mr. S.C.J.J. Kortmann,

volgens besluit van het college van decanen

in het openbaar te verdedigen op vrijdag 22 februari 2013

om 10.30 uur precies

door

Marijt Johanne Witteman

geboren op 7 februari 1984

te Rotterdam

Promotoren: Prof. dr. J. M. McQueen
Prof. dr. A. Cutler

Copromotor : Prof. dr. A. Weber (Universität Tübingen)

Manuscriptcomissie: Prof. dr. T. Dijkstra
Prof. dr. M. Ernestus
Prof. dr. W. Duyck (Universiteit Gent)

The research reported in this thesis was supported by a grant from the Max-Planck-Gesellschaft zur Förderung der Wissenschaften, München, Germany.

Acknowledgements

Writing this dissertation often felt like a wave: there were a lot of highs as well as lows along the way. Though the process itself can be lonely at times, it would certainly not have been possible without the people I would like to thank here.

My first round of thanks goes out to Andrea Weber and James McQueen. Andrea, your door was always open for me and your cheerful attitude gave me new energy whenever I needed it. Your quick feedback on my work helped me to keep things going, especially near the end of my project. I really appreciated this. James, I couldn't have done it without your quick mind, attention to detail, and incredible grasp of theoretical concepts. Thank you for helping me to solve problems and providing sound advice in all kinds of situations.

Anne, as my third supervisor you created the environment in which it was possible for me to work on this dissertation. Thank you for doing that. There was always a possibility to present my data during MaRCS meetings. The feedback I got during those meetings proved to be very valuable for my thesis, so I would like to thank everyone involved.

I would especially like to thank (former) Adaptive Listening members: Andrea, Adriana, Eva, Frank, Laurence, Neil, and Odette, as well as the AdLis student-assistants: Anne Blankenhorn, Anne van Hoek, Héctor, and Felicia.

Not only were you there for research questions and help, I also look back fondly on our potluck dinners and the memorable conversations that took place there.

There were also a lot of people who helped me with practical matters during my PhD. Iris Mulders from Utrecht University allowed me to use her lab space, and that part of my data collection turned out to be incredibly fast and efficient because of her help. The Technical Group at the MPI was also always willing to lend a hand, and I would like to thank Alex, Ronald, Johan, and Ad in particular for their quick and helpful support.

Rian, ook jij mag hier natuurlijk niet ontbreken. Dankjewel voor al je hulp, zowel voor de kleine als grotere problemen! Rachel, thank you too, for listening to my rants and helping where needed. Phaidra, bedankt voor al je hulp met het ontwerpen van de omslag.

A thesis on understanding foreign-accented speech cannot be carried out without speakers willing to lend their voices for recordings. Jan, Carsten, Christian, Ruti, and Neil were my main 'victims' for these recordings. Thank you for that great contribution.

Many of my colleagues have become my friends during the past three years. I am extremely grateful that two of these friends agreed to be my paranymphs: Matthias and Neil. Matthias, met jou als kamergenoot was er altijd wat te lachen en dat heeft mij vaak door de dag heen gesleept. Ook aan onze buitenlandse conferentietripjes heb ik zeer goede herinneringen.

Neil, as my friend, colleague, and co-author you have fulfilled many different roles during my PhD. I don't think I have ever met anyone with whom I share an interest for so many random topics. You kept me up-to-date on many of these, listened to all of my updates, and were there to brainstorm about research whenever necessary. I am extremely glad to have two great friends by my side for my defense!

Laurence, wij hebben veel uren doorgebracht met praten over werkgerelateerde en wat minder relevante (werk)onderwerpen. Jouw kantoor bezoeken kwam altijd als een welkome afwisseling! Lies, ook jij was hierbij vaak aanwezig en je bent tijdens de 4Daagse ook nog een reddende engel geweest. Bedankt!

A special thanks goes out to all the MPI and IMPRS PhD-students, many of whom I also consider my friends now. I have so many good memories of drinks, dinner parties, or simple chats at the office. Thank you all for the gezelligheid!

Tyko, you and I started many projects together, and I am extremely pleased that we were able to finish some too! The rest will be left in the able hands of the next generation(s).

Mijn niet-MPI vrienden mogen hier natuurlijk ook niet ontbreken! Kors, bedankt voor de gezellige en culinair hoogstaande etentjes (als jij kookte dan...), dat er nog maar vele mogen volgen. Elisabeth en Kassandra: ik kijk altijd uit naar onze afspraken! Sybrine en Rosanne, ik vond het heerlijk om met jullie verhalen te vergelijken en bij te praten onder het genot van thee of lunch. And of course Greg! We chased that finish line together, and were able to share both frustrations and successes. Though we were usually far apart, it did not feel that way. And now we have both made it!

En dan de FnFers: Annika, Britt, Hebe, Hedwig, Joost, Joren, Katrine, Laura, Maaike, Merijn, Noortje, René, Richard, Sabrine, Sanne, Stefan en Tim, wij hebben zo veel avonturen meegemaakt samen dat alles opnoemen hier niet meer te doen is. Regelmatig prijs ik mijzelf gelukkig met zoveel vrienden waarmee ik kan lachen, maar waar ik ook terecht kan als het even minder gaat. Ik hoop dat er nog vele FnFs en Renesses zullen volgen. Katrine, jou wil ik nog even in het speciaal bedanken, want dankzij het samen schrijven in Leersum (tot diep in de nacht!) werden die gevreesde laatste maanden van mijn promotie niet alleen heel efficiënt, maar ook nog echt leuk!

Het laatste bedankje gaat uit naar mijn broer en mijn ouders. Jorrit, papa en mama, jullie waren er altijd: om leuke dingen te doen, maar ook op momenten dat het wat minder ging. Dankjewel voor de steun en dat jullie altijd in mij geloofd hebben.

Table of contents

Introduction

Chapter 1

Doing our best to understand foreign-accented speech has become a standard listening situation in our everyday lives. It is estimated that currently more than half of the world's population speaks at least two languages (Grosjean, 2010), with numbers steadily growing. This multilingualism not only entails that we regularly listen to speech that is not in our native language, but also that with increasing frequency we listen to our native language being spoken by non-native talkers. Pronunciation is one of the most difficult domains of second language (L2) acquisition to master (e.g., Flege, Munro, & MacKay, 1995; Munro, Flege, & MacKay, 1996), and the vast majority of non-native speakers will never achieve native-like pronunciation. That is, they maintain a foreign accent. In this thesis I will focus on how native listeners understand foreign-accented speech. Specifically, I will investigate the process of lexical adaptation: what circumstances aid and prohibit adaptation to foreign-accented speech (or, more specifically for this thesis: foreign-accented Dutch words)?

Foreign accents have been studied from different angles. Linguistic studies have mainly focused on intelligibility of foreign-accented speech (e.g., Bradlow & Bent, 2008; Munro & Derwing, 1995a; Derwing & Munro, 1997), and native listeners' ability to identify speakers by their accents (e.g., Flege, 1984). Studies in social psychology and sociolinguistics often focused on native attitudes towards foreign-accented speech and speakers (Niedzielski, 1999; Doeleman, 1998). Foreign accents also receive a lot of attention from foreign language institutes: many offer courses for L2 speakers to get rid of their foreign accents altogether. But is foreign-accented speech truly detrimental for native listening?

Native processing of foreign-accented speech has received relatively little attention in spoken language research. More knowledge about how foreign-accented speech is processed online is therefore greatly needed for the development of adequate theories of spoken-word recognition. Furthermore, studying foreign-accented speech will be informative about the speech perception system in general, because it can show how this system deals with variation in speech, and therefore provide insights about the flexibility of language processing.

Non-native speakers add all kinds of variation into a language. They might create different grammatical structures, speak more slowly, or pronounce a target word in a non-standard way. This thesis will focus on this last aspect: pronunciation. Importantly, the pronunciation variations of non-native speakers are not random. Because the features of foreign-accented speech arise primarily from an interaction of the phonological structures of the speaker’s native language and the target language, its phonetic characteristics are systematic and quite consistent across speakers (Flege, Schirru, & MacKay, 2003). Thus, although the characteristics of foreign-accented speech may differ between language combinations, the variation within a group of speakers of any given language combination is much smaller, because the native language is the driving force behind the variation. This leads to regular variations that are stereotypical for an accent. For example, while Dutch learners as a group sometimes say *indeet* instead of English *indeed*, Japanese learners of English as a group are more likely to say *indeedo* when trying to produce the same word (Weber, Broersma, & Aoyagi, 2011).

Differences in pronunciation are very common. In fact, even within native speech (L1) there is a lot of variation in the way sounds are produced. This variation is caused by many different factors, such as speaker-specific characteristics, language context effects (such as coarticulation), and dialects (e.g., Peterson & Barney, 1952). This thesis focuses on foreign-accented speech rather than dialectal variations. Although foreign-accented speech and dialectal variation introduce variability to speech in similar ways (e.g., substitution of segments, change of phonological rules, allophonic variation), they also differ in several respects. Firstly, where variation within native language is often context-dependent, foreign-accented speech can be both dependent of and independent from context. This is because, as already noted, foreign-accented speech is dependent on the sounds of the speaker's native language (Flege, et al., 2003). When the target language includes speech sounds that are not found in the speaker's native language, foreign-accented speakers often replace target language sounds with native language sounds (e.g., *cattle* pronounced as *kettle* by Dutch speakers). Furthermore, while native speech varies within categories (i.e., an /i/, though differently pronounced by different dialectal speakers or in different segmental contexts, will typically remain within the /i/ category), a foreign-accented speaker might sometimes pronounce a sound that fall within the intended category, and at other times a sound that fall in a different category (e.g., /i/ pronounced as /ɪ/). Other characteristics of foreign-accented speech include the absence of phoneme contrasts or allophonic variations (e.g., *Italy* pronounced with a long /i/ as *Eataly* by Italian speakers) or differently-realized suprasegmental cues (e.g., *hyPOthesis*, pronounced as *hypoTHEsis* by Dutch and German speakers of English).

Because of these deviations from native speech, foreign-accented speech can sometimes lead to misidentification of words (Lane, 1963; Munro & Derwing, 1995a), or increased processing time compared to native speech (Munro & Derwing, 1995b). Although native listeners are slowed down initially when processing foreign-accented speech, this diminishes quickly, after as little as one minute, or two to four sentences of exposure (Clarke & Garrett, 2004). That process of adaptation to a speaker is the focus of this thesis: under what conditions do listeners adapt, and how do they do it?

Previous research on foreign-accented speech

Most research on foreign-accented speech has focused on intelligibility (e.g., Bradlow & Bent, 2008; Munro & Derwing, 1995a; Derwing & Munro, 1997). Munro (Munro & Derwing, 1995a, 1995b) defined intelligibility as "the extent to which the native speaker understands the intended message". Typically, native listeners in these studies are asked to transcribe (selected parts of) sentences spoken by foreign-accented talkers. Unsurprisingly, native speakers are usually better able to understand fellow native speakers than foreign-accented speakers, especially in noisy situations (Munro & Derwing, 1995a). Moreover, it has been found that intelligibility increases as the strength of accent of a speaker decreases (Bradlow & Bent, 2008), though an easily detectable foreign accent does not necessarily decrease intelligibility (Munro & Derwing, 1995a). Munro and Derwing (1997) therefore distinguish between intelligibility (measured by the number of correctly transcribed words) and comprehensibility, defined as listeners' perception of how intelligible an utterance is (assessed with rating scales). When native listeners rate the intelligibility, comprehensibility and accentedness (acoustic deviation from the target) of different speakers, intelligibility scores were rated higher than comprehensibility and accentedness scores, indicating that intelligibility does not necessarily suffer when a speaker has a noticeable accent.

For fellow non-native listeners, accentedness can sometimes even be beneficial. The interlanguage speech intelligibility benefit refers to the finding that non-native listeners may find it easier to understand a foreign-accented talker from the same language background (e.g., native Chinese participants listening to Chinese-accented English) than talkers of any other language background, and sometimes even a native speaker (Bent & Bradlow, 2003). There are, however, also studies showing no such benefit (e.g., Hayes-Harb, Smith, Bent, & Bradlow, 2008; Hongyan & van Heuven, 2007).

Exposure to multiple accented speakers has proven to be beneficial for understanding foreign-accented speech (Sidaras, Alexander, & Nygaard, 2009; Bradlow & Bent, 2008). Sidaras et al. found that native English listeners were able to attain higher transcription accuracy on novel words and utterances after a familiarization phase with multiple speakers compared to no exposure at all. Furthermore, participants who received accented familiarization were better at transcribing accented utterances than participants with only native training. Experience with one accent can sometimes even improve native listeners' transcription accuracy of speakers with a different non-native accent (Bradlow & Bent, 2008), though not all studies find such generalization effects (Jongman, Wade, & Sereno, 2003).

Previous work on foreign-accented speech has thus identified some of the reasons why foreign-accented speech is sometimes hard to understand, and indicated that additional exposure to an accent can aid listeners in adapting to the speaker. Importantly, it has been found that listeners are able to adapt to foreign speakers after a short period of listening to those speakers. In these studies, however, accentedness or intelligibility of a speaker was typically assessed on a general level (that is, not taking into account that not all words might be equally accented or intelligible). Since foreign accents are influenced by the speaker's native language phoneme repertoire, not every segment will be equally affected, as some segments will be shared with the target language and some will not. Thus, even within one language combination, there are words that can be strongly accented and words that are affected to a smaller extent. In this thesis, I will look at specific accent markers rather than overall intelligibility and contrast the recognition of more strongly accented words with that of less strongly accented words. I will examine the recognition of individual words that vary in perceived accentedness and acoustic deviation from the target. This perspective makes it possible to investigate in greater detail what makes foreign-accented speech hard to understand, and whether all types of variation lead to processing difficulty.

Processing variation in native speech

Like foreign-accented speech, native speech is rife with variation. Language is spoken by talkers with different characteristics (e.g. gender, speaking rate), and certain phonemes are pronounced differently depending on the surrounding context (e.g., Peterson & Barney, 1952). Research on variation in L1 can be divided up into studies on arbitrary variation (Connine, Blasko, & Titone, 1993; Marslen-Wilson, 1993; Marslen-Wilson & Zwitserlood, 1989) and studies on phonologically ruled variation (Gaskell & Marslen-Wilson, 1996, 1998; Gow, 2002; Pitt, 2009).

Research on arbitrary variation has shown that listeners are able to recognize mispronounced words under a variety of circumstances. But word recognition is very sensitive to phonetic mismatches, such that even a change in one phoneme is enough to prevent word recognition or even inhibit it (Marslen-Wilson, Nix, & Gaskell, 1995; Van Alphen & McQueen, 2006). Moreover, lexical access gets disrupted more strongly with every additional feature change (Connine, Blasko, & Titone, 1993).

Place assimilation is an example of phonologically ruled variation and refers to the phenomenon that a speech sound is influenced by surrounding speech sounds. For instance, the /t/ in *freight* may become a /p/ in the context of *freight bearer*, due to assimilation with the following stop consonant. Native English listeners were able to interpret both *freight bearer* and *freip bearer*, but when *freip* was presented in isolation, they did not accept this as an existing word form. Moreover, place assimilation was not accepted in unviable contexts (such as *freip carrier*; Gaskell & Marslen-Wilson, 1998). The context in which variation occurs thus plays an important role in whether listeners can accept these variant forms. Similar results have been found for studies on reduced word forms (e.g. *posman* for *postman*; Ernestus, Baayen, & Schreuder, 2002; Mitterer & Ernestus, 2006): listeners are able to recognize strongly reduced word forms in context, but not in isolation (Ernestus, et al., 2002). These findings thus suggest that listeners are able to reconstruct the underlying form from the surface variation, but only when the surface form occurs in a viable context.

Frequency effects

Some changes are more common in everyday speech than others. When adapting to variations in speech, listeners take information about how frequent certain language structures occur into account. That is, variant forms are more likely to be recognized correctly when they are presented in a phonological context in which they frequently occur, such as certain reductions including Dutch ['dam] for /'darɔm/ [therefore] (e.g., Ernestus et al., 2002; Mitterer & Ernestus, 2006), medial t-deletion, such as 'senner' or 'sennah' for English 'center'; (Pitt, 2009), and vowel raising (Dahan, Drucker, & Scarborough, 2008). For example, Pitt (2009) found that participants judged /t/-deleted variants as words only if the phonological environment in which the /t/-deletion occurred was common in production. A similar effect was found by Dahan et al. (2008), who manipulated English words ending in /g/ and /k/ such that they either contained a raised vowel (like the [ɛ] in 'bag') or an unraised vowel (similar to the [æ] in 'back'). Half of the words had contextually appropriate vowel raising, and the other half did not. Participants learned to understand the intended word, and even learned to expect the incorrect pronunciations when new items were presented. Moreover, adaptation proved to be extremely rapid: just a few trials were necessary to get a (limited) effect of generalization.

Experiments on the nasal flap in English (found in 'gentle' in casual American English) have also shown that lexical representations are stronger when people have more experience with this phenomenon (Ranbom & Connine, 2007). Studies on reduced speech have also looked at the importance of frequency of occurrence. Word-final /t/, for example, can be realized with an alveolar closure and an audible release, or as a glottal stop without a release. English listeners were able to recognize both forms equally quickly in a semantic priming task; however, when a was delay added to the task, participants did noticeably better with the typical variant (Sumner & Samuel, 2005).

These studies thus show that native listeners can learn to processing variants after long-term exposure gained in everyday life. Another line of research has focused on short-term adaptation to native variations learned in a laboratory setting. These results are described below.

Perceptual learning in L1 speech

Listeners can use their phonological knowledge to shift their phoneme boundaries when they encounter a speaker talking in an idiosyncratic manner, with the goal of understanding that speaker. This phenomenon is called perceptual learning. One example is listeners' adaptation to artificially created sounds such as [?], midway between [s] and [f]. Listeners are able to use their lexical knowledge when adapting to these sounds in an exposure phase, that is, interpreting [?] as either [s] or [f] depending on the word context. This perceptual learning is known to generalize to new words in a test phase (e.g., McQueen, Cutler, & Norris, 2006) and was found using different tasks for exposure and test (e.g., McQueen, Norris, & Cutler, 2006).

Perceptual learning can occur in several situations: when a sound in a word is replaced by an ambiguous native sound (e.g., Norris, McQueen, & Cutler, 2003; McQueen, Norris, & Cutler, 2006), by another native sound (e.g., Maye, Aslin, & Tanenhaus, 2008), or even by a non-native sound

(Sjerps & McQueen, 2010). Adaptation is shown by listeners' adjustments of their category boundaries in categorization tasks, or their correct interpretation of words with mispronunciations in lexical decision tasks. This process takes place extremely quickly (e.g., Kraljic & Samuel, 2007) and is thought to be automatic (McQueen, Norris, & Cutler, 2006). That is, attention to the accent, in the form of conscious decisions on the mispronunciations, is not a requirement for perceptual learning effects to arise. In fact, even when listeners are instructed just to listen to a story (Eisner & McQueen, 2006) or asked simply to count the number of trials, perceptual learning effects are observed.

What is less clear, however, is whether foreign-accented speech can count on the same flexibility. Foreign-accented speech brings a different type of variation to the speech signal, namely variation that is driven by the native language of the speaker. Moreover, foreign-accented speech does not create variation on just one phoneme, but rather affects many segments, and to different degrees.

Models of spoken-word recognition

So how exactly are listeners able to recognize words that are pronounced differently? There are several theories on how listeners resolve variation in speech. Though the experiments presented in this thesis were not specifically designed to provide evidence for or refute certain claims, the results will have implications for models of spoken-word recognition.

Somehow listeners must be able to map the phonetic signal onto representations of known words in the mental lexicon. There are two main theories (each with several variations): representation-based and processing-based accounts. Representational models assume that the lexicon has entries for every single word, as well as for every variation on these words, including acoustic properties of spoken words that are not part of their phonology, such as speaker-specific indexical characteristics (e.g., Goldinger, 1998). Whether all these entries are separate entries, or one 'major' representation alongside with several 'minor' ones, however, remains an open question. Listeners are indeed able to retain detailed perceptual information about tokens of isolated words (Goldinger, 1998). Episodic representational models also state that all variation is encoded in the lexicon in fine-grained phonetic detail (e.g., Johnson, 2006) or in the form of a 'grainy spectrogram' (Pierrehumbert, 2001). These episodic traces are stored alongside the canonical variants of words. Finally, it has been proposed that the lexicon contains multiple abstract representations for variant forms (Ranbom & Connine, 2007). Although this view holds that both canonical and variant forms are stored, it does not include storage of indexical properties of spoken words, such as speech rate, pitch, voice quality and so on. What all representational models have in common is that they propose that disambiguation of accented (variant) forms takes place after lexical access. They are based on the findings that after experience with certain accented forms, listeners are able to recognize accented words better upon repeated exposure, which could be explained by listeners' ability to access the variant forms again in their mental lexicons.

Processing-based accounts, however, operate on the assumption that variation is resolved pre-lexically, between the speech signal and the lexicon. According to these views, the mental lexicon consists of canonical forms, and word recognition is achieved without explicit storage of pronunciation variants in the lexicon. Listeners learn from exposure to an accent how variations should be mapped on the (stored) canonical form (Lotto & Holt, 2006; Mitterer, Csépe, Honbolygo, & Blomert, 2006; Gaskell & Marslen-Wilson, 1998). Different mechanisms of how this pre-lexical process should take place have been proposed, ranging from general auditory mechanisms (Lotto & Holt, 2006; Mitterer, et al., 2006) to more abstract pattern recognition mechanisms (Gaskell & Marslen-Wilson, 1998). The discussion on the nature of these pre-lexical mechanisms has not been settled. But because these models localize adaptation processes to the pre-lexical level, they are able to explain why listeners can generalize from one mispronounced word to a similarly mispronounced word (e.g., McQueen, Cutler, & Norris, 2006; Sjerps & McQueen, 2010).

Cross-modal priming

Throughout this thesis, I use a cross-modal priming task and take facilitatory priming as a measure of word recognition (adaptation). Cross-modal priming is a task that is commonly used to tap into online language comprehension. In this task, listeners first hear a prime (word or nonword), followed by a visual target (word or nonword). Their task is to decide whether the visually presented target is a word or not by pressing a button. Participants' reaction times (RTs) are measured from visual target onset. All experimental trials have a word as their target, and fillers are added to the experiment to make the ratio between target words and target nonwords 1:1. When listeners do not make any errors, they thus have an equal number of 'word' and 'nonword' responses.

L1 research has shown that listeners are faster to respond if the prime and target are identical (e.g., in Dutch, [bœyk] - BUIK) compared to unrelated pairs (e.g., [dif] - BUIK), because the auditory prime facilitates the recogntion of the visual target (see e.g., Clarke & Garrett, 2004; McQueen, Cutler, & Norris, 2006; Marslen-Wilson, Nix, & Gaskell, 1995). Phonologically similar but non-identical auditory primes generally do not produce significant facilitatory priming, and sometimes even inhibitory priming (e.g., Van Alphen & McQueen, 2006). Therefore, cross-modal priming can be used to see whether listeners recognize the accented word (the prime) as the intended canonical Dutch word (the target). In this thesis, statistically significant facilitatory priming will be taken as evidence for successful online word recognition (Marslen-Wilson & Zwitserlood, 1989; Marslen-Wilson, Moss, & van Halen, 1996), and hence successful adaptation to the accented words.

Overview of the thesis

The major objective of this thesis is to discover how native listeners adapt to foreign-accented words: how can this adaptation be aided and what factors constrain it? Using cross-modal priming as a method, I will look at the processing difficulty caused by foreign-accented speech. Different types of experience will be investigated, as well as perceptual learning effects for foreign-accented speech. These results will then be related to models of spoken-word recognition.

Chapter 2 reports three cross-modal priming studies on long-term and short-term experience with foreign-accented speech. Experiment 2.1 investigates long-term experience with a specific accent, namely German-accented Dutch. Previous research has indicated that listeners who are exposed to dialectal variation (in this case: New York City English) on a daily basis accept both dialectal and non-dialectal (General American English) forms as words, whereas listeners who did not have (passive or active) experience with an accent do not accept the dialectal forms (Sumner & Samuel, 2009). Thus, listeners who are familiar with a dialect are apparently more flexible in form processing than inexperienced listeners are.

Experiment 2.1 tests whether this same flexibility holds in foreign-accented speech by comparing two groups of listeners, one with extensive prior experience with German-accented Dutch, and another with limited prior experience. Experiments 2.2 and 2.3 asks whether listeners with limited experience are able to adapt to a speaker after a short exposure phase

(Experiment 2.2), and whether we see speaker-generalization effects by exposing listeners to one speaker and testing them on another speaker of the same accent (Experiment 2.3). This chapter concentrates on what type of experience is required for native listeners to be able to process foreign-accented speech correctly.

The aim of Chapter 3 is to investigate whether acoustic similarity and comprehension go hand in hand, or whether the amount of acoustic deviation is not necessarily predictive of listening performance. This is based on the notion that although acoustic similarity indeed often predicts how listeners categorize non-native sounds, acoustic analyses cannot always explain cross-language perceptual difficulties (Flege, 1995). I thus explore the relationship between acoustic similarity (assessed with acoustic analyses), perceived accentedness (assessed with a rating study), and the degree of processing difficulty (assessed with a cross-modal priming task).

Chapter 4 explores the nature of adaptation to foreign-accented speech by investigating how automatic this adaptation is, as well as looking at how long-lasting it is. Perceptual learning studies have indicated that adaptation to native variation is automatic (McQueen, Norris, & Cutler, 2006) and that this process remains stable for at least 12 hours (Eisner & McQueen, 2006). In three cross-modal priming experiments I investigate adaptation to an accent unfamiliar to the listeners (Hebrew-accented Dutch). Experiment 3.1 serves as a baseline, while in Experiments 3.2 and 3.3 a phoneme monitoring exposure task was added while the test phase was delayed by one day and one week, respectively. The purpose of these experiments was thus to examine whether adaptation to foreign-accented speech works the same way, with respect to automaticity and stability as the perceptual learning mechanisms that underlie native speech processing.

In the last experimental chapter I examine the boundaries of adaptation to foreign-accented speech. Using cross-modal priming, listeners are exposed to either a speaker with a consistent accent, or to an inconsistently-accented speaker. Research on L1 speech has indicated that when listeners learn that an L1 speaker's mispronunciations are incidental (e.g., inconsistent), they do not show perceptual learning effects, whereas they do if this information is not provided (Kraljic, Samuel, & Brennan, 2008; Kraljic & Samuel, 2011). Foreign-accented speech is more variable than native speech (Hanulíková & Weber, 2012; Wade, Jongman, & Sereno, 2007). The critical question in Chapter 5 is therefore whether adaptation to foreign-accented speech depends on how consistent the accent is.

The final chapter (Chapter 6) summarizes the results and provides a discussion of the main findings of this thesis.

Foreign accent strength and listener familiarity with an accent co-determine speed of perceptual adaptation

Chapter 2

Witteman, M. J., Weber, A., and McQueen, J. M. (in press). Foreign accent strength and listener familiarity with an accent co-determine speed of perceptual adaptation. *Attention, Perception & Psychophysics*.

Abstract

We investigated how the strength of a foreign accent and varying types of experience with foreign-accented speech influence the recognition of accented words. In Experiment 1, native Dutch listeners with limited or extensive prior experience with German-accented Dutch completed a cross-modal priming experiment with strongly-, medium-, and weakly-accented words. Participants with limited experience were primed by the medium- and weakly-accented words, but not by the strongly-accented words. Participants with extensive experience were primed by all accent types. In Experiments 2 and 3, Dutch listeners with limited experience listened to a short story before doing the cross-modal priming task. In Experiment 2, the story was spoken by the priming-task speaker and either contained strongly-accented words or not. Strongly-accented exposure led to immediate priming by novel strongly-accented words, while exposure to the speaker without strongly-accented tokens led to priming only in the experiment's second half. In Experiment 3, listeners listened to the story with strongly-accented words spoken by a different German-accented speaker. Listeners were primed by the strongly-accented words, but again only in the experiment's second half. Together, these results show that adaptation to foreign-accented speech is rapid, but depends on accent strength, and on listener familiarity with those strongly-accented words.

Introduction

It is estimated that more than half of the world's population speak at least two languages (Grosjean, 2010), with numbers steadily growing. This multilingualism means that listening to foreign-accented speech has become a standard listening situation in metropolitan areas. In foreign-accented speech, native listeners are confronted with pronunciations that deviate from their language standards (e.g., Dutch speakers pronouncing *kettle* instead of *cattle*, or, for native Japanese speakers, something similar to *flied lice* instead of *fried rice*). How then do native listeners cope with foreign-accented speech? Is understanding of foreign-accented speech hindered only by large deviations from the intended pronunciation? Can inexperienced listeners adapt quickly to a new speaker? Furthermore, after adaptation to one non-native speaker, is it then possible to understand another speaker with the same accent? The current study addresses these questions.

Variation in native speech

Although foreign accents add variation to speech, native speech contains considerable variability too. Most research about variation in speech in fact stems from the first language (L1) domain. Research on L1 speech shows that even when one phoneme in a word is changed arbitrarily, the word can still be recognized (e.g., Connine, et al., 1993; Marslen-Wilson, 1993), at least as long as this change does not create a new word (Marslen-Wilson & Zwitserlood, 1989). However, the larger the phonemic deviation in a nonword is from the standard pronunciation of a word, the harder it is to recognize the word correctly (Connine, et al., 1993). When words are changed in ways typical for natural speech as in *greem bench,* where the word *green* is assimilated to what sounds like *greem* because the following context warrants nasal place assimilation, this does not prevent recognition of the intended word (e.g., Gaskell & Marslen-Wilson, 1996, 1998; Gow, 2002; Mitterer & Blomert, 2003). This effect does not occur, however, when the following context does not license place assimilation (Gaskell & Marslen-Wilson, 1996, 1998; Mitterer & Blomert, 2003). Reductions also form an interesting case for looking at deviation. Words that are frequently reduced in conversational speech can be recognized easily as their intended targets: hearing *posman,* for example, will facilitate recognition of unreduced *postman* (Ernestus, et al., 2002; Mitterer & Ernestus, 2006). Together these findings thus show that although the amount of deviation is important in recognizing word forms, the context in which the variation occurs has to be appropriate in order for successful recognition to occur. A foreign-accented speaker provides a natural context for deviant pronunciations that are consistent with that accent but typically not with the target language.

Frequency and familiarity effects

Frequency of occurrence is another important factor for recognition of deviant word forms. That is, variant forms are more likely to be recognized correctly when they are presented in a phonological context in which they frequently occur (such as certain reductions including Dutch ['dam] for

/'darɔm/ [therefore] (e.g., Ernestus, et al., 2002; Mitterer & Ernestus, 2006), medial t-deletion, such as 'senner' or 'sennah' for English 'center'; (Pitt, 2009), and vowel raising (Dahan, Drucker, & Scarborough, 2008). For example, Pitt (2009) found that participants judged /t/-deleted variants as words only if the phonological environment in which the /t/-deletion occurred was common in production. A similar effect was found by Dahan et al. (2008), who manipulated English words ending in /g/ and /k/ such that they would either contain a raised vowel (like the [ɛ] in *bag*) or an unraised vowel (similar to the [æ] in *back*). Half of the words had contextually appropriate vowel raising, and the other half did not. Participants learned to understand the intended word, and even learned to expect the incorrect pronunciations when new items were presented. Moreover, adaptation proved to be extremely rapid: just a few trials were necessary to get a (limited) effect of generalization.

Processing of words is also affected by the frequency of the variant. Words that are frequently pronounced without a schwa (e.g., *corporate* as *corp'rate*) were more likely to be judged as two-syllable words than words with a low schwa-deletion rate (Connine, Ranbom, & Patterson, 2008; Connine, et al., 1993). This influence of the distribution of variant representations has also been found in other languages, such as Dutch (Mitterer & Ernestus, 2006). Experiments on the nasal flap in English (found in *gentle* in American English) have also shown that lexical representations are stronger when people have more experience with this phenomenon (Ranbom & Connine, 2007). Reductions also form an interesting perspective on frequency of occurrence. Word-final /t/, for example, can be realized with an alveolar closure and an audible release, or as a glottal stop without a release. English listeners were able to recognize both forms equally quickly in a semantic priming task; however, when there was a delay added in the task, participants did noticeably better with the typical variant (Sumner & Samuel, 2005). Degree of familiarity with foreign-accented pronunciations should thus influence how listeners process them. In the present study we tested different types of familiarity effect: we first looked at the role of long-term experience with an accent on comprehension, and then investigated how rapidly comprehension of an unfamiliar accent could improve when participants were briefly exposed to that accent immediately before testing. We thus asked what type of experience is needed to process foreign-accented speech correctly.

Sumner and Samuel (2009) looked at the role of long-term experience by studying dialectal variation (New York City English vs. General American English) and found that speakers of a New York City dialect which drops word-final –r (turning bak[ɚ] ‘baker’ into bak[ə]) could instantly interpret these dialectal forms as the intended word, whereas non-dialectal speakers (speakers of General American) did not show such an effect. When tested again after a short (20-30 min) time lag, General American listeners had more trouble recognizing the dialectal forms compared to the standard forms, whereas for the NYC dialectal listeners there was no difference. Thus, listeners who are familiar (passively or actively) with a dialect are apparently more flexible in form processing than inexperienced listeners are:

experienced listeners can deal with more variation relative to the standard pronunciation when listening to dialectal speech. A similar result was found with British English listeners who had moved to the United States: they learned to interpret correctly the medial flap (ɾ) in "*todal*" (/toɾal/) as /t/ (thereby recognizing the intended word *total*; Scott & Cutler, 1984). These listeners seemed to have adapted their perceptual system to American English standards.

Effects of accent strength in foreign-accented speech?

The current study focuses on foreign-accented speech rather than dialectal variations. In foreign-accented speech, speakers often replace target language sounds with native language sounds when the target speech sounds are not found in the speaker's native language (e.g., *cattle* pronounced as *kettle* by Dutch speakers; Broersma & Cutler, 2011). When speech sounds are shared between the two languages, however, substitutions that involve a different category are usually not observed. The former types of alteration are often perceived as stronger accent markers than the latter. Thus, even within one language combination, there are words that can be strongly affected by foreign accents and words that are affected to a smaller extent. Here, we contrast the recognition of more strongly accented words with that of less strongly accented words.

Most research on foreign-accented speech has focused on its intelligibility (e.g., Bradlow & Bent, 2008; Munro & Derwing, 1995a; Derwing & Munro, 1997). In general, intelligibility increases as the strength of accent of a speaker decreases (Bent & Bradlow, 2003), and native listeners are known to benefit from more exposure to improve their understanding of low intelligibility speakers, whereas this is not necessary for highly intelligible speakers (Bradlow & Bent, 2008). Exposure from multiple speakers is also beneficial for understanding foreign-accented speech (Sidaras, Alexander, & Nygaard, 2009). Native English listeners were familiarized with multiple speakers of Spanish-accented English or with native English control speakers. Participants were able to attain higher transcription accuracy on novel words and utterances after a familiarization phase with multiple speakers compared to no exposure at all. Furthermore, participants who received accented familiarization were better at recognizing some of the accented vowels than participants with only native training. Which vowels participants were exposed to and how this affected learning of these vowels, however, was not manipulated systematically. Most of the studies on intelligibility used tasks in which participants were required to make judgments about sentences that were not controlled for specific accent markers. The present study will investigate specific accent markers and their effect on understanding accented speech. We examine the recognition of individual words that vary in perceived accentedness. The difference in perceived accentedness is mainly driven by vowels that deviate to a larger or smaller extent from the standard form. This perspective allows us to investigate in greater detail what makes foreign-accented speech hard to understand, and whether all types of variation lead to processing difficulty.

Familiarity to German-accented Dutch

As already noted, we were also interested in the effects of familiarity with a naturally occurring accent. In particular, we tested whether Dutch listeners who are either familiar or unfamiliar with German-accented Dutch are able to correctly interpret German-accented Dutch words. Research with second language (L2) listeners suggests that recognition of familiar variant forms is possible (Weber, et al., 2011). In an experiment in which Dutch and Japanese participants listened to either Dutch-accented English or Japanese-accented English, the L2 listeners could recognize accented words easily when they were produced in their own accent (e.g., Dutch listeners could recognize Dutch-accented English words easily and Japanese listeners Japanese-accented English words). But are participants also able to adapt to foreign-accented speech when the target language is their native language? In that case, listeners are constantly exposed to the native pronunciations from their fellow countrymen. Native Dutch speakers, for instance, will usually have far more experience with native Dutch than with German-accented Dutch. Can native listeners therefore easily understand only weakly-accented words, and do they, in order to understand strongly-accented words, need to attain a certain level of familiarity with the accent first?

Dutch listeners participated in three cross-modal priming experiments with Dutch as the target language, though spoken with a German accent. German-accented Dutch was chosen because native Dutch speakers are known to vary in how familiar they are with the accent. There are a substantial number of German students in the Netherlands, but they tend to study at Dutch universities close to the German border. Fewer German students are found in the center of the Netherlands. So it is possible to find Dutch listeners with either limited or extensive experience with German-accented Dutch.

A German accent in Dutch is particularly noticeable when it comes to vowels, and these vowels produce an excellent starting point for looking at effects of degree of accentedness. We therefore chose words with two particular Dutch diphthongs. Though both Dutch and German are Germanic languages, their vowel systems differ in a number of ways. Both languages have a large vowel inventory: Dutch has 13 monophthongs and 3 diphthongs (Gussenhoven, 1999); German has 12 monophthongs and 3 diphthongs (Kohler, 1999). While there is some overlap between the monophthongs, the diphthongs vary more across the two languages. Dutch has the three diphthongs /ɛɪ/, /œy/, and /ʌu/, while German has /aɪ/, /ɔɪ/, and /aʊ/. The two Dutch diphthongs that were the focus of this study, [œy] and [ɛɪ], are thus not part of the German vowel inventory. Both diphthongs are difficult for many learners of Dutch (Doeleman, 1998), but in particular /œy/ is a rare sound across languages and poses great difficulties for second-language learners. The Dutch [œy]-vowel was replaced with the German [ɔɪ] by the speaker of this experiment, and the Dutch [ɛɪ] was replaced with the German [aɪ]. Acoustically, the trajectories of [ɔɪ] and [œy] deviate more than those of [ɛɪ] and [aɪ]. This was confirmed for the speaker of our experiment with acoustical analyses (see Experiment 1 Method). German learners of Dutch are usually well aware of the large deviation between their pronunciation and the intended Dutch [œy]-vowel, while they are often oblivious to the smaller deviation between their [aɪ] pronunciation and the intended Dutch [ɛɪ]. In this study, Dutch words with [œy] were considered to be strongly accented and Dutch words with [ɛɪ] as medium accented. There was a third set of words without any segmental substitutions. These items contained only phonemes shared between Dutch and German but were nonetheless spoken by the same non-native speaker. These words were therefore considered to be weakly accented. The three different strengths of perceived accentedness were confirmed in a rating study (see Experiment 1 Method).

In summary, the current study had three goals. First, in Experiment 1 we investigated whether foreign-accented speech (in this study: German-accented Dutch) can be understood by native (Dutch) listeners with limited previous exposure, and contrasted their results with those of native listeners who were already highly familiar with the accent. Second, we asked whether effects of familiarity depend on how strongly accented the stimulus words were. Third, in two subsequent experiments we examined how short-term training on an accent influences word recognition, again as a function of strength of accent (Experiment 2), but also as a function of speaker (Experiment 3). The results of these experiments will be related to models of spoken-word recognition and the accounts they offer for how listeners cope with pronunciation variation – those based on representation (e.g. Goldinger, 1998; Johnson, 2006; Pierrehumbert, 2001; Ranbom & Connine, 2007) and those based on processing (e.g., Lotto & Holt, 2006; Mitterer, et al., 2006; Gaskell & Marslen-Wilson, 1998).

Experiment 1

Experiment 1 was designed to test whether familiarity with a foreign accent influences adaptation to that accent, as measured in terms of ease of word recognition. We compared Dutch listeners with limited experience with German-accented Dutch to listeners with extensive long-term experience with German-accented Dutch. Participants first listened to German-accented Dutch primes and then judged whether target words that appeared on a screen were Dutch words or nonwords. RTs to a target word are known to be faster when the auditory prime and the visually presented target word are identical than when the auditory prime is unrelated (see e.g., Clarke & Garrett, 2004; McQueen, Cutler, et al., 2006; Marslen-Wilson, Nix, & Gaskell, 1995). Phonologically similar but non-identical auditory primes generally do not produce significant facilitatory priming, and sometimes even inhibitory priming (e.g., Van Alphen & McQueen, 2006). We will therefore take statistically significant facilitatory priming as our measure of successful online word recognition (Marslen-Wilson & Zwitserlood, 1989; Marslen-Wilson, Moss, & van Halen, 1996), and hence of successful adaptation to the accent. Facilitatory priming of responses to visual target words after auditory primes produced with a foreign accent will thus be taken to suggest that listeners recognized the accented primes online as being the same Dutch words as the visual targets.

We expected that recognition of the variant forms would be more successful when the deviation from the standard was smaller (e.g., that we could observe significant priming for the weakly-accented words but not for the strongly-accented words), and that the more experience people have had with an accent, the easier it would be for them to adapt to it. Adaptation was measured separately for the first and second half of the experiment. It was possible that less-experienced listeners would not show priming in the first half, but would in the second half, after having had time to adapt to the accent. Given previous findings with native speech (e.g., Marslen-Wilson & Zwitserlood, 1989), it was predicted that Dutch listeners could interpret the weakly-accented words correctly, regardless of their previous experience with the German accent. Furthermore, we expected that participants with limited previous exposure to German-accented Dutch would have trouble understanding medium- and strongly-accented words, whereas participants with extensive experience with the accent would have no problems understanding such words.

Method

Participants

Two groups of participants were tested. The limited-experience group (n = 23, 19 females, *M* age 20.41 years), all native speakers of Dutch, was tested in Utrecht, a city in the middle of the Netherlands. These participants were recruited from the Utrecht University participant pool; the vast majority studied at Utrecht University. The extensive-experience group also consisted of 21 native speakers of Dutch (19 females, *M* age 22.97) and was tested in Nijmegen, a city in the east of the Netherlands, near the German border.

These participants were recruited from the MPI participant pool; the majority studied at the Radboud University Nijmegen.

There are many more German students enrolled at the Radboud University Nijmegen (approximately 8% of the students in 2011 were German) than at Utrecht University (approximately 1% German students in 2011). As such it could be expected that Dutch students in Nijmegen are in general more frequently exposed to German-accented Dutch than students in Utrecht are. But students in Utrecht can happen to have German friends or family with whom they communicate in Dutch, and students in Nijmegen can happen to major in a subject where only few German students are enrolled. In order to control more closely for the amount of prior experience with German-accented Dutch, a language history questionnaire was administered. One of the questions asked how often participants heard German-accented Dutch (possible answers: never, less than once a week, once a week, and multiple times a week). Another question asked from how many speakers participants heard German-accented Dutch (possible answers: 0-1, 2-5, 6-10, and more than 10). Only Utrecht-based students who reported hearing German-accented Dutch less than once a week from less than two speakers were included in the limited-experience group and only Nijmegen-based participants who reported hearing German-accented Dutch multiple times a week from more than two speakers were included in the extensive-experience group. Because the questionnaire was administered after the main experiment (to avoid a strong focus on German), a number of additional participants were tested but were not included in the analysis because they did not meet these criteria (14 participants in Utrecht, 10 participants in Nijmegen).

In addition to their exposure to German-accented Dutch, we also asked participants about their knowledge of German. In The Netherlands, all students in upper educational levels have to follow German language courses for at least three years, and will thus have some knowledge of German. Since the educational programs for German are very similar across the country, the knowledge acquired in school should be comparable for our listener groups. Indeed, none of the participants in Experiment 1 studied German, none reported to be fluent in German, and German was always reported to be either their second or third non-native language.

All participants volunteered and were paid a small fee for participating. None reported a hearing disorder, and all had normal or corrected-to-normal vision. They all reported that English was the first non-native language they had learned, usually starting in school around the age of 10.

Materials

There were 48 critical items and 96 fillers. Each critical item was a combination of two auditory prime words and a visual target word. The targets were Dutch mono- or bisyllabic words (38 nouns, 7 adjectives, 2 adverbs, and 1 pronoun), and their corresponding primes were either the German-accented variants of the targets (identical primes) or phonologically and semantically unrelated Dutch words (unrelated primes). The 48 identical primes comprised 12 strongly-accented words, 12 medium-accented words, and 24 weakly-accented words. Strongly-accented primes were words with the Dutch vowel [œy] as in *huis*, 'house'. This diphthong is not part of the German phoneme inventory (see, e.g., Kohler, 1999), and German learners of Dutch mostly substitute it with [ɔɪ], a German diphthong that is perceptually and acoustically quite different from Dutch [œy] (Dutch [œy] starts front-central, half open and ends front-central, near close; German [ɔɪ] starts back-central, half closed and ends front-central, near open).

Medium-accented prime words contained the Dutch diphthong [ɛɪ] as in *lijst*, 'list'; this diphthong also does not exist in German (Kohler, 1999), and is usually substituted with German [aɪ] by German speakers of Dutch. The diphthong [aɪ] is not present in the Dutch phonemic inventory (Gussenhoven, 1999), but is phonetically relatively similar to the Dutch diphthong [ɛɪ]. German [aɪ] begins central, half open, and ends front, close-mid; Dutch [ɛɪ] begins front, open-mid and ends front, close-mid. Both vowels thus end in approximately the same place. We ensured that the remaining sounds (consonants and other vowels) of the strongly- and medium-accented words did not contain other obvious segmental substitutions by using only phonemes that were shared between the languages. Thus, except for the investigated diphthongs, all sounds were part of the Dutch and German phoneme inventories.

To ensure that the experimental words could not be interpreted as other existing Dutch words, we asked 10 additional native Dutch participants who did not hear German-accented Dutch multiple times a week to transcribe all 48 experimental words. Incorrect transcriptions were given, by as maximum of four participants, to only four words (two strongly-accented, one medium-accented and one weakly-accented). The incorrect transcriptions included one existing Dutch word (one word, and by only one participant), loan words (two words) and a name (one word). All other words were transcribed by all participants as the intended words.

The weakly-accented words contained only vowels and consonants that are present in both the Dutch and German phonemic inventory, such as [ɪ] and [ɛ] vowels and [m], [ŋ] and [b]. We thus minimized segmental variation and strength of perceived accentedness. The same subset of additional vowels and consonants that was used for the strongly- and medium-accented words was used for the weakly-accented words. To the extent that it can be expected that these sounds contribute to the perceived accentedness of words, they should do so to a comparable extent for all three accent types. The 48 unrelated primes matched overall in number of phonemes with their corresponding targets (e.g., 'ketting' [chain] and 'prikkel' [incentive]) and the overall lexical lemma frequency of unrelated primes was not different from the frequency of the targets (log frequency taken from the CELEX database (Baayen, Piepenbrock, & Van Rijn, 1993; $t(48) = .082$, $p = .935$). For a complete list of critical items, see Appendix A.

Of the 96 filler items, 24 had a Dutch word as the visual target. The remaining 72 fillers all had a nonword as their visual target. Eighteen of these items contained [ɔɪ] or [aɪ] in the prime, so that not every [ɔɪ] or [aɪ] prime would predict a yes-response. Therefore, the overall ratio of words and nonwords for the visual targets was 1:1, resulting in 50% "yes"-responses for errorless participants. Half of these 72 filler items were preceded by nonword primes, the others were preceded by existing Dutch word primes. Both the word and nonword auditory primes could contain the [ɔɪ] and [aɪ] vowel, again to ensure that participants could not form a response strategy based on the presence of these vowels in the items.

Speaker selection

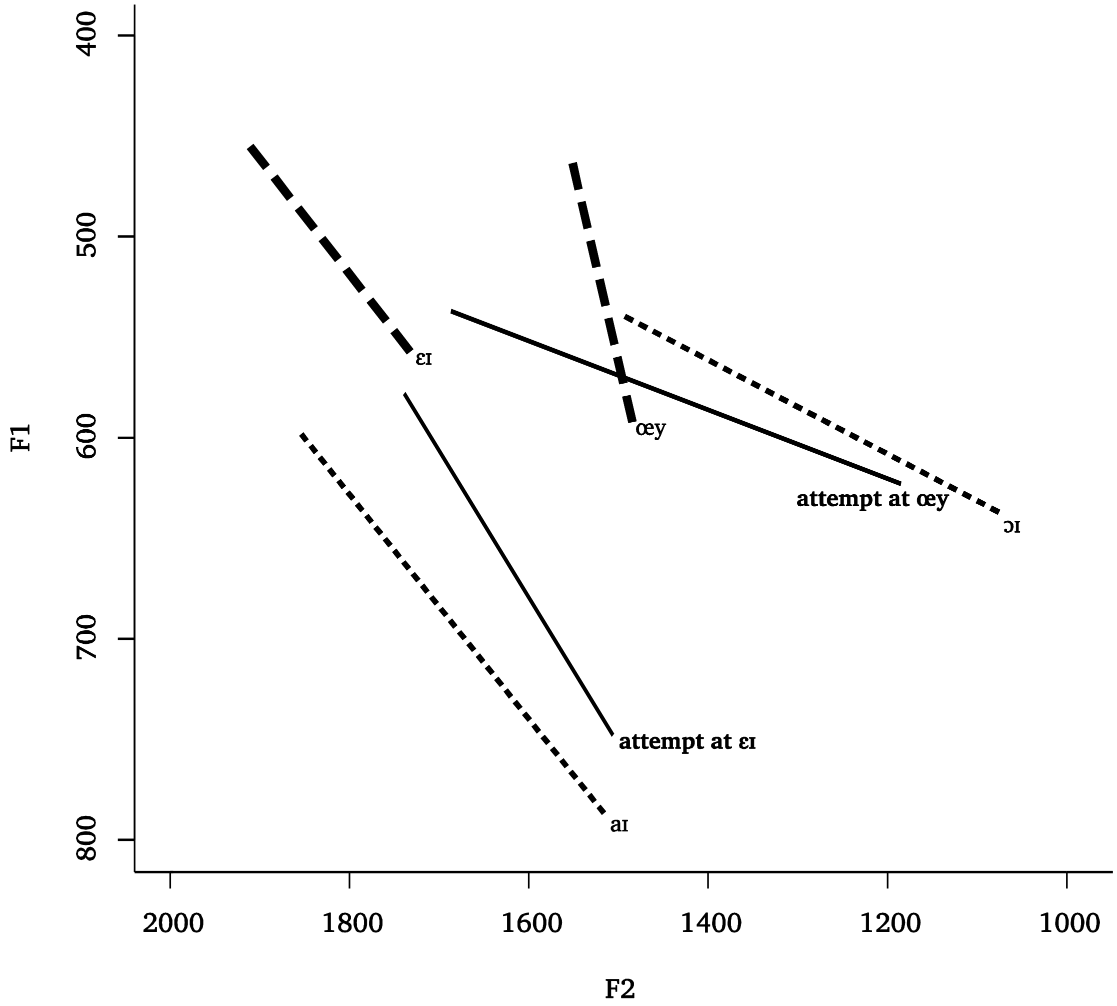
Seven native speakers of German were recorded while reading a short Dutch text. Speakers differed in their level of proficiency and time spent in the Netherlands. These recordings were made to find a learner of Dutch who would produce the requested mispronunciations spontaneously and consistently. The chosen speaker was a male native speaker of German, who grew up in Bavaria, in the south of Germany. At the time of recording, he had lived in the Netherlands for two years while studying in Dutch at the Radboud University Nijmegen. The speaker was quite fluent in Dutch, and knew the meaning of almost all Dutch words used in the experiment.

The Dutch word and nonword primes were recorded in pseudo-randomized order from a list of items given in their correct Dutch spelling. The speaker was not instructed to change his pronunciation, so all mispronunciations occurred naturally. The speaker produced the primes one by one, separated by a pause, in a clear citation style, recording each prime at least two times. The recordings were made in a sound-attenuated booth with a Sennheiser microphone and were stored directly onto a computer at a sample rate of 44 kHz. Primes were excised from the recording using the speech editor Praat (Boersma & Weenink, 2009), and the best tokens were selected by the first author, a native speaker of Dutch.

Acoustic measurements

We recorded the complete vowel space for our speaker both in Dutch and in German by having him pronounce all Dutch and German vowels separately in an hVba-context. This context was chosen because it minimizes influences of other segments on the vowels (Jenkins et al., 1997). All words were recorded at least twice per vowel in the speaker's natural accent, in clear citation, from correct Dutch or German spelling (e.g., the investigated Dutch vowels were written as *huiba* and *hijba*, the German vowels as *Heuba* and *Heiba*). These recordings were made in the same session as the recordings of the auditory primes. The best two tokens per vowel were selected by a native speaker of Dutch (the first author). For the critical German and Dutch diphthongs, we measured the first and second formants at the 25 and 75 percent points of the vowels. Figure 1 plots the average values for the first two formants for the Dutch diphthongs [œy] and [ɛɪ], as well as the German diphthongs [ɔɪ] and [aɪ] as produced by the speaker of the experiment, as well as the diphthongs [œy] and [ɛɪ] spoken by Dutch native speakers (data taken from Adank, Van Hout, & Smits, 2004).

Figure 1. F1 and F2 formant values of the two critical diphthongs, as pronounced by the experimental speaker (German-accented Dutch in solid lines, native German in thick dashed lines) compared to Dutch diphthongs from the Adank, et al., 2004, corpus (thin dashed lines).



When looking at the Dutch /ɛɪ/ in Figure 1 it can be seen that our speaker's F1 value is higher than the average Dutch speaker's F1, but the trajectories are quite similar. In comparison, our speaker's F2 is higher than the average Dutch speaker's F2 for Dutch /œy/, and the trajectories are furthermore quite different. As can also be seen in Figure 1, our speaker's Dutch /ɛɪ/ and his German /aɪ/ are quite similar, as well as his Dutch /œy/ and German /ɔɪ/. This supports the notion that our speaker substituted the Dutch diphthongs with existing German categories.

Rating study

To further ensure that the three types of identical primes were indeed perceived as varying in accent strength, twenty native Dutch speakers who did not participate in the priming study rated the items. We used all strongly-accented [œy]-primes, all medium-accented [ɛɪ]-primes, and half of the weakly-accented primes, so that there were 12 items of each type. We recorded two more sets of 36 items from two additional speakers: A native speaker of Dutch, and a native speaker of Italian with a very strong accent in Dutch. The reason for adding these two speakers was to add more variation to the materials, thereby avoiding artificial inflation of a perceived difference between prime types for the German speaker. During the rating study, participants heard one word at a time over closed headphones, immediately followed by a rating scale where they could indicate how accented the word was on a scale ranging from 1 (not accented) to 10 (very strongly accented).

The data were analyzed with paired-samples t-tests which indicated that indeed the strongly-accented [œy]-items (M = 7.98, SD = 1.18) were rated as more accented than the weakly-accented items (M = 4.73, SD = 1.67; $t(19) = 9.443$, $p < .001$) and the medium-accented [ɛɪ]-items (M = 7.01, SD = 1.22; $t(19) = 7.223$, $p < .001$). Furthermore, the medium-accented [ɛɪ]-items were rated as more accented than the weakly-accented items ($t(19) = 7.576$, $p < .001$). These results thus confirm the picture emerging from the acoustic measures: The strongly-accented words deviated most from the standard pronunciations and were indeed rated as more strongly accented than the medium-accented items.

Design and procedure of priming experiment

Participants were seated in a sound-attenuated booth and informed that they would first hear a Dutch word or nonword spoken by a German-accented speaker and then see a Dutch word or nonword on the screen. Two versions of the cross-modal priming experiment were created, so that every participant saw each visual target only once. To control for effects of presentation order, each participant received a different pseudo-randomized list. The first two items of the experiment were always fillers, and there were never two critical items in a row.

The participants' task was to decide as quickly and accurately as possible whether the word presented on the screen was an existing Dutch word or not. Participants responded by pushing one of two buttons on a button box in front of them. Yes-responses were always made with the dominant hand, and RTs were measured from visual target onset.

Auditory primes were presented binaurally over closed headphones at a comfortable listening level. Participants saw the visual targets on a computer screen situated about 50 cm in front of them. Visual targets were presented in white lowercase 24p Tahoma letters on a black background, 500 ms after the acoustic offset of the auditory primes. The visual targets stayed on the screen for 2000 ms, after which the next trial started. The experiment was created in Presentation (version 13, Neurobehavioural Systems Inc.) and controlled with NESU hardware (Nijmegen Experiment Set-Up). After the cross-modal priming experiment, participants were asked to fill out the language history questionnaire.

Results

Three items with weakly-accented primes were discarded from the analysis because they had high error rates (more than 10%). A possible reason for these error rates is that all three of these targets had a very low lexical frequency (for the exact items see Appendix A). These items were also excluded from analyses of all other experiments described here.

The remaining cross-modal priming data was analyzed with General Linear Model (GLM) Repeated Measures Analyses of Variance (ANOVAs) using a 3 (*accent type* - strong, medium, weak) x 2 (*priming* - identical vs. unrelated) x 2 (*half* – first and second half of the experiment) design. All of these were within–participant factors. We analyzed the results separately by participants (*F1*) and items (*F2*). In addition to these analyses, we conducted planned comparisons using paired sample t-tests to look at the priming effects. These were calculated separately by *half* and *accent type.*

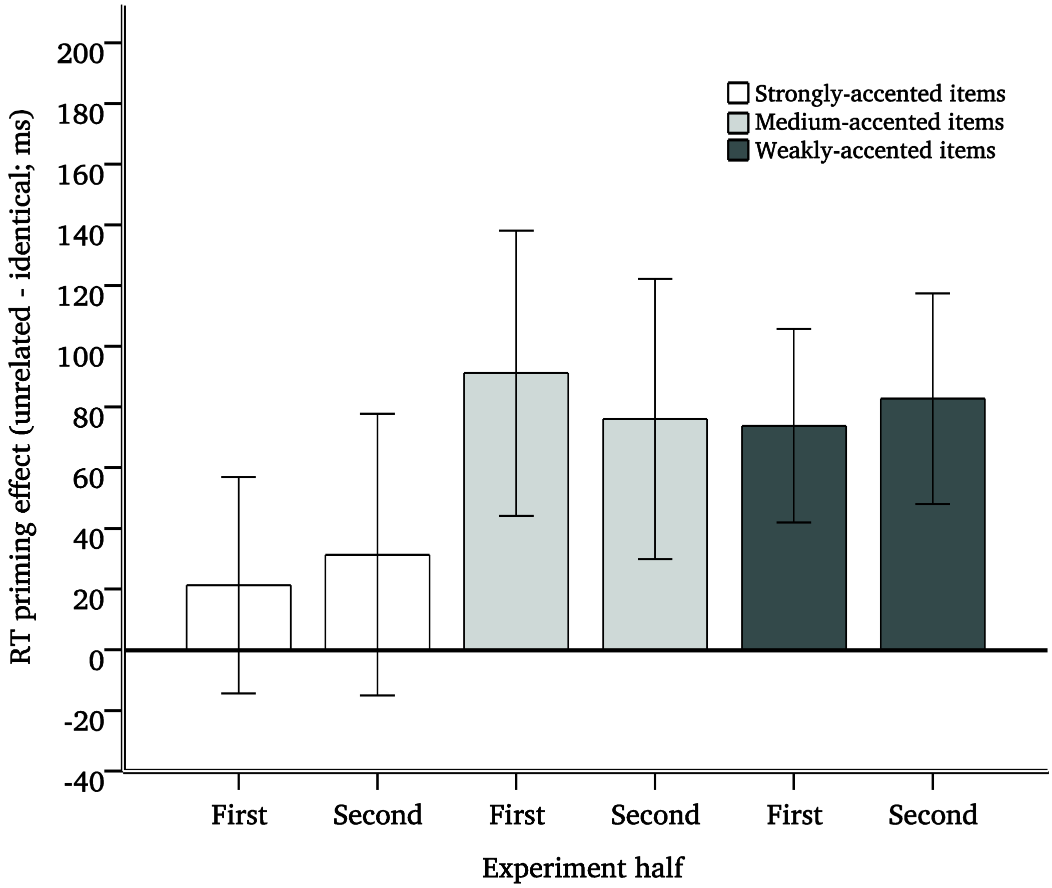
Limited Experience Group

In addition to the three weakly-accented items, a further 1.8% trials in which participants made errors, as well as trials with RTs that deviated more than 2.5 SD from the condition's overall mean, were discarded (together $< 5\%$ of all trials). Errors were distributed evenly across the conditions (see Appendix B), and due to their low overall occurrence, will not be analyzed statistically.

As shown in Figure 2 (calculated priming effects) and Appendix B (mean RTs), participants with limited experience were faster to respond to identical than to unrelated trials ($F1$ (1,22) = 49.028, $p < .001$; $F2$ (1,11) = 86.190, p

< .001). A main effect of *accent type* across participants furthermore indicated that participants reacted to target words with different speeds (*F1* (2,44) = 3.876, *p* = 0.028; *F2* (2,22) = 1.100, *p* = .350). Participants also got faster overall during the course of the experiment (*F1* (1,22) = 7.536, *p* = .019; *F2* (1,11) = 7.267, *p* = .013).

Figure 2. Experiment 1: Priming effects and confidence intervals (CIs) for participants with limited experience by experiment halves and accent type.



The F1 analysis showed a significant interaction across participants between *accent type* and *priming* (*F1* (2,44) = 4.580, *p* = .016; *F2* (2,22) = 2.541, *p* = .102), reflecting that participants did not show equal priming effects for all accent types. This was investigated further using planned pair-wise comparisons (see Table 1). The interaction between *half* and *priming* was not significant (*F1* < 1, *F2* < 1), neither was the three-way interaction (*F1* < 1, *F2* < 1). The pair-wise comparisons show that participants had no problems with adapting to the weakly- and medium-accented items, but could not adapt to the strongly-accented items during the experiment.

Table 1. Planned Pairwise Comparisons of Priming Effects for all Accent Types Across Participants and Items for Participants with Limited Experience with German-accented Dutch.

		Participant analysis			Item analysis		
Condition	Half	*df*	*t*1	*p*	*df*	*t*2	*p*
Strongly-accented items	First	1,22	1.233	.230	1,11	.997	.340
Strongly-accented items	Second	1,22	1.399	.176	1,11	1.781	.103
Medium-accented items	First	1,22	4.026	.001	1,11	3.348	.007
Medium-accented items	Second	1,22	3.414	.002	1,11	3.546	.005
Weakly-accented items	First	1,22	4.803	.000	1,20	3.599	.002
Weakly-accented items	Second	1,22	4.947	.000	1,20	2.301	.032

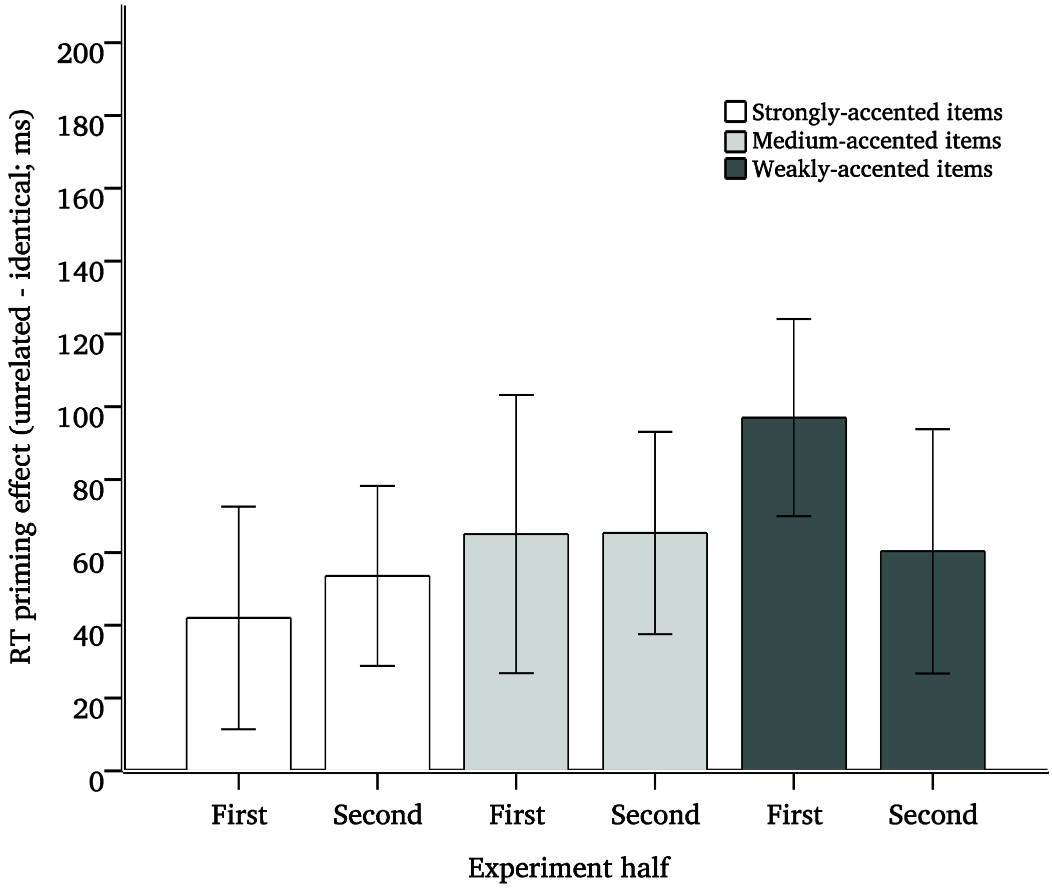
Extensive experience group

Trials on which participants made errors (2.0%) were excluded from the analyses. In addition, we excluded trials on which the RTs deviated more than 2.5 SD from the condition's overall mean (together < 5% of all trials). Inspection of the errors showed no specific patterns.

As shown in Figure 3 (calculated priming effects) and Appendix B (mean RTs), participants with extensive experience with German-accented Dutch were faster to respond to identical than to unrelated items (*F1* (1,20) = 81.727, $p < .001$; *F2* (1,11) = 29.828, $p < .001$), and responded similarly to the three different accent types (*F1* < 1; *F2* < 1). Participants were faster in the second half of the experiment compared to the first (*F1* (1,20) =

5.301, $p = .033$; *F2* (1,11) = 8.957, $p = .014$). Across participants there was a marginally significant interaction between *accent type* and *priming* (*F1* (2,40) = 3.073, $p = .058$; *F2* < 1), reflecting that priming effects differed in size for the accent types. The remaining interactions did not reach significance.

Figure 3: Experiment 1: Priming effects and CIs for participants with extensive experience by experiment halves and accent Type



Planned pair-wise comparisons were used to look at the priming effects for each accent type separately. Table 2 displays the statistical analyses of the priming effects across participants and items. Participants with extensive experience showed significant priming from the start of the experiment for all accent types, thereby showing that had previously adapted to the accent, and were able to apply this knowledge rapidly to the experimental speaker.

Table 2. Planned Pair-wise Comparisons of Priming Effects for all Accent Types Across Participants and Items for Participants with Extensive Experience with German-accented Dutch.

	Participant analysis				Item analysis		
Condition	Half	*df*	*t*1	*p*	*df*	*t*2	*p*
Strongly-accented items	First	1,20	3.009	.007	1,11	1.302	.220
Strongly-accented items	Second	1,20	4.855	.000	1,11	1.871	.099
Medium-accented items	First	1,20	3.562	.002	1,11	2.079	.032
Medium-accented items	Second	1,20	5.172	.000	1,11	3.642	.004
Weakly-accented items	First	1,20	6.012	.000	1,20	2.155	.016
Weakly-accented items	Second	1,20	3.263	.004	1,20	2.156	.050

Discussion

Taken together, the results of Experiment 1 thus show that listeners with limited experience with German-accented Dutch and those with extensive experience can immediately interpret the medium- and weakly-accented items correctly, but only the listeners with extensive experience show facilitatory priming for strongly-accented items. We also analyzed both listener groups (limited experience and extensive experience) in one overall Repeated Measures Analysis with group as an additional between-participants factor. These analyses showed an effect of group marginally significant across participants ($F1$ (1,43) = 3.737, p = .060, $F2$ (1,20) = 18.223, p < .001), with the experienced listeners responding faster overall than the inexperienced listeners, but there were no significant interactions between any of the factors. This could imply that the difference between the groups with respect to processing strongly-accented words was possibly less pronounced than the separate analyses suggest or that this effect depends on the strength of elements of the accent. Moreover, it is possible that even Dutch listeners with extensive experience with German-accented Dutch may still encounter some difficulties understanding strongly-accented words.

Experiment 1 allowed us to study the role of long-term exposure to German-accented Dutch. It showed not only that the degree of experience influences how easily a listener can adapt to foreign-accented speech, but also that this effect depends on the strength of the accent. In Experiment 2, we wanted to shed more light on the role of short-term experience on word recognition and in particular on the effect of short-term exposure on adaptation to strongly-accented words. Research on perceptual learning has shown that the word recognition process is not only sensitive to long-term listening experience, but can also adapt after a short amount of exposure. There is some evidence for rapid adaptation to foreign-accented speech (see e.g., Clarke & Garrett, 2004; Bradlow & Bent, 2008). This is in line with perceptual learning research with artificial accents or speech sounds constructed to be ambiguous showing flexibility in online word-recognition processes (e.g., Maye, Aslin, & Tanenhaus, 2008; McQueen, Cutler, et al., 2006) as well as in the perception of phoneme categories (e.g., Norris, McQueen, & Cutler, 2003; Eisner & McQueen, 2006).

Experiment 2

In Experiment 2, we asked for the first time whether there is short-term adaptation to individual vowels in a real foreign accent (as opposed to an artificial accent). We again tested Dutch listeners with limited experience with German-accented Dutch, but they were now exposed to a four-minute story immediately before the cross-modal priming experiment. The short story was spoken by the same speaker as was used in the priming experiment and would thus function as additional exposure to that speaker. Two versions of the story were recorded, one with 12 strongly-accented items (words containing the [œy]-vowel) not used in the main experiment, and one without strongly-accented items. The goal of this experiment was to investigate whether a short period of familiarization with the speaker would be sufficient to create adaptation to novel words. We expected to replicate the findings of Experiment 1 with respect to the medium- and weakly-accented words (i.e., priming in both halves of the experiment). We also expected to find priming for the strongly-accented words when participants with limited experience had been exposed to a story containing strongly-accented words (in contrast to Experiment 1) at least in the second half and possibly already in the first half of the experiment. We expected participants who listened to the weakly-accented exposure also to show adaptation to the strongly-accented items, but later or to a lesser extent than listeners who were exposed to the story with strongly-accented words.

Method

Participants

Fifty-seven participants took part in Experiment 2. Participants were randomly assigned to one of two story-exposure groups (Group 1: n = 19, 15 females, *M* age 21.18 years; Group 2: n = 19, 16 females, *M* age 22.15 years). Participants from Utrecht were selected using the same criteria as in the limited-exposure group in Experiment 1. As before, therefore, we excluded additional participants who heard German-accented Dutch more than once a week from more than two speakers. In total we excluded 19 participants. As in Experiment 1, participants did not report a hearing disorder and had normal or corrected-to-normal vision. All of them reported basic knowledge of German, with German being their second or third non-native language (after English); none considered themselves fluent in German.

Materials

Two versions of a short story were created. We chose to focus on the effects of exposure to strongly-accented words, because learning for the medium- and weakly-accented words was already at ceiling in Experiment 1. Both stories were recorded by the speaker from Experiment 1. The story was based on the fairytale 'Jorinde and Joringel' by the Brothers Grimm. The two versions were identical except for 12 words. One version contained 12 strongly-accented words with the [œy]-vowel (pronounced by the speaker as [ɔɪ]). None of these words appeared in the main experiment. In the other version, these 12 words where replaced with words without the vowel [œy] (e.g., *duizend,* 'thousand' replaced with *honderd,* 'hundred'). The two stories thus both contained medium- and weakly-accented words. Both versions were recorded from a script using correct Dutch spelling. The speaker was not instructed to change his pronunciation; again, all mispronunciations occurred naturally. The speaker recorded one paragraph at a time. All paragraphs were recorded multiple times, and the best paragraphs were selected to create a story that was spoken without hesitations and misreadings. That is, recordings were selected in which all words were pronounced as intended, including the 12 strongly-accented words consistently pronounced with [ɔɪ]). Recordings were made in a sound-attenuated booth with a Sennheiser microphone and were stored directly onto a computer at a sample rate of 44 kHz.

Design and procedure

Participants were seated in a sound-attenuated booth and informed that they would first listen to a Dutch story spoken by a German speaker and then perform a cross-modal priming experiment. There was no additional task when participants were listening to the story. Group 1 listened to the story with 12 strongly-accented words. Group 2 listened to the story without strongly-accented words. Participants were randomly assigned to a group. Immediately after the story ended, participants saw the instructions for the cross-modal priming experiment on the screen. The cross-modal priming experiment was identical to the one used in Experiment 1. After the experiment, participants filled out the language history questionnaire (identical to Experiment 1).

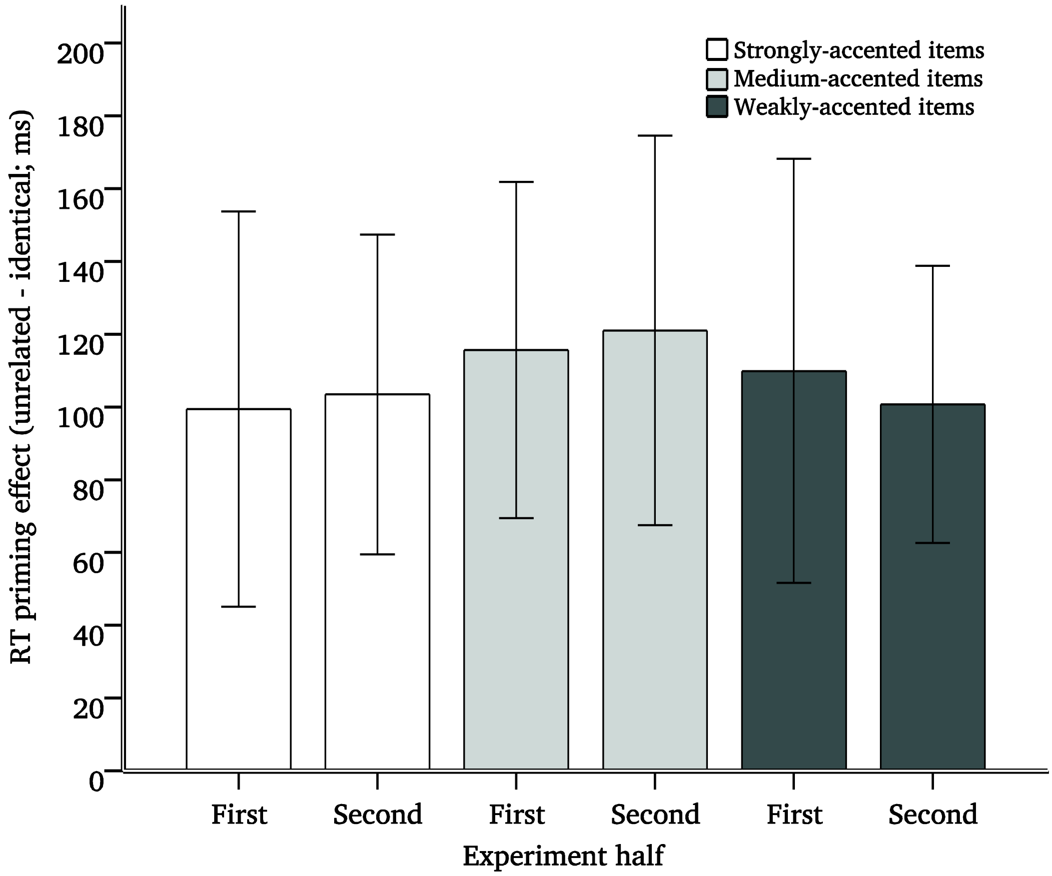
Results

Limited Experience group – Exposure with strongly-accented words

The statistical analyses were identical to those used in Experiment 1. Trials containing errors were excluded (1.3%) from the analyses. We also excluded trials on which the RTs deviated more than 2.5 SD from the condition's overall mean (together $<$ 5% of all trials). Inspection of the errors (see Appendix B) again showed no specific patterns of priming.

The priming effects for participants exposed to the story with strongly accented words are depicted in Figure 4 (mean RTs are given in Appendix B). Participants were faster overall in responding to identical trials compared to unrelated trials ($F1$ (1,18) = 82.771, $p < .001$; $F2$ (1,11) = 132.515, $p < .001$). The participant analysis showed that subjects responded with different speeds to the three word types ($F1$ (2,36) = 3.788, $p = .033$; $F2 < 1$). Participants were faster in the second half of the experiment than in the first half, though this effect was significant only by items ($F1$ (1,18) = 3.638, $p = .074$; $F2$ (1,11) = 6.754, $p = .025$). There were no significant interactions between the factors.

Figure 4. Experiment 2: Priming effects and CIs for limited experience group – Exposure with strongly accented items by experiment halves and accent type.



Planned pair-wise comparisons were used to confirm the priming effects. Table 3 displays the priming effects across participants and items. Participants showed successful adaptation to all accent types in both halves of the experiment.

Table 3. Planned Pairwise Comparisons of Priming Effects for all Accent Types Across Participants and Items for Participants with Exposure to Strongly-accented Items

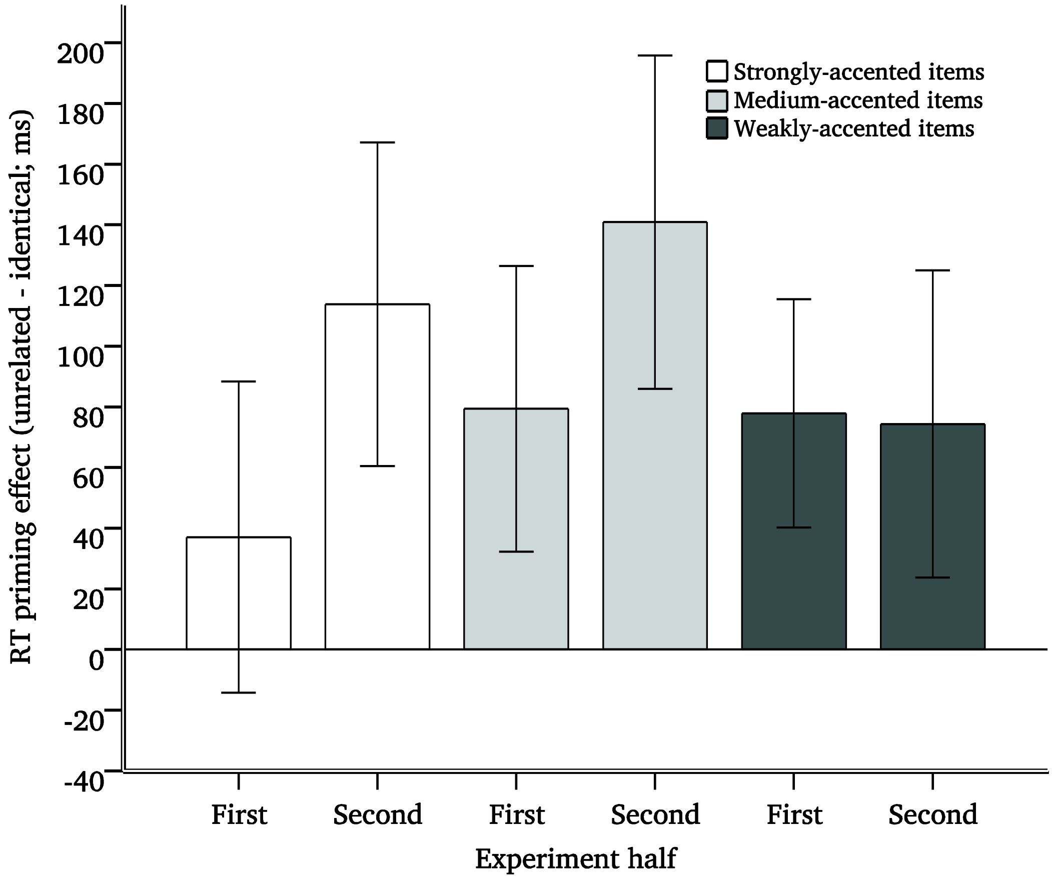
	Participant analysis				Item analysis		
Condition	Half	*df*	*t*1	*p*	*df*	*t*2	*p*
Strongly-accented items	First	1,18	3.858	.001	1,11	3.229	.008
Strongly-accented items	Second	1,18	5.179	.000	1,11	2.576	.026
Medium-accented items	First	1,18	5.373	.000	1,11	5.653	.000
Medium-accented items	Second	1,18	4.906	.000	1,11	6.036	.000
Weakly-accented items	First	1,18	4.204	.001	1,20	4.611	.000
Weakly-accented items	Second	1,18	5.970	.000	1,20	5.378	.000

Limited exposure group – exposure without strongly-accented words

We excluded trials on which participants made errors (1.8%) from the analyses. We also excluded trials on which the RTs deviated more than 2.5 SD from the condition's overall mean (together $< 5\%$ of all trials). The errors (see Appendix B) showed no systematic priming effects.

Priming effects for participants exposed to the story without strongly-accented words are shown in Figure 5 (mean RTs are given in Appendix B). Participants were faster overall in responding to identical trials than unrelated trials ($F1$ (1,18) = 76.162, $p < .001$; $F2$ (1,11) = 35.956, $p < .001$). There was no difference in the way participants responded to the different types of items ($F1$ (2,36) = 1.185, $p = .318$; $F2 < 1$). Participants were not faster overall in the second half compared to the first ($F1 < 1$; $F2 < 1$). There was no interaction between *accent type* and *half* ($F1 < 1$; $F2$ (2,20) = 1.039, $p = .372$), or between *accent type* and *priming* ($F1$ (2,36) = 1.462, $p = .246$; $F2 < 1$). The interaction between *half* and *priming* was significant across participants ($F1$ (1,18) = 5.806, $p = .028$; $F2 < 1$), reflecting that, in general, priming effects were larger in the second half of the experiment than in the first half. The three-way interaction was not significant.

Figure 5. Experiment 2: Priming effects and CIs for limited experience group – Exposure without strongly accented items by experiment halves and accent type.



The results of the planned pairwise comparisons of the priming effects are given in Table 4. They show that participants were able to interpret the weakly-accented items and the medium-accented items throughout the experiment. The strongly-accented items, however, could only be interpreted correctly in the second half of the experiment.

Table 4. Planned Pairwise Comparisons of Priming Effects for all Accent Types Across Participants and Items for Participants Exposed the Story without Strongly-accented Items

		Participant analysis			Item analysis		
Condition	Half	*df*	*t*1	*p*	*df*	*t*2	*p*
Strongly-accented items	First	1,18	1.523	.146	1,11	2.028	.068
Strongly-accented items	Second	1,18	4.731	.000	1,11	3.515	.006
Medium-accented items	First	1,18	3.736	.002	1,11	3.784	.003
Medium-accented items	Second	1,18	4.750	.000	1,11	3.464	.005
Weakly-accented items	First	1,18	4.429	.000	1,20	4.541	.000
Weakly-accented items	Second	1,18	3.386	.003	1,20	4.431	.000

Discussion

When participants were exposed to the same speaker without strongly-accented words, we saw adaptation to the strongly-accented items in the second half of the experiment. Participants thus have some benefit from prior exposure to the speaker even though he did not produce these specific mispronunciations. Moreover, while participants' performance increased for the strongly- and medium-accented items across halves, there was no increase in the priming effect for the weakly-accented items. Since the only difference between the segments in the words in the three sets is in the vowels (we used the same consonants in all three accent types), the differences for the strongly-and medium-accented word must be driven by perceptual adaptation to the vowels.

At this point an interesting comparison can be made between participants with limited experience from Experiment 1 (who did not have a story exposure phase) to the two groups of participants from Experiment 2, who did have an exposure phase. We compared the priming effects for the strongly-accented items, separately for the two halves. Participants who had received strongly-accented exposure showed significantly more priming for the strongly-accented items than participants without prior experience, both in the first ($t(40) = 2.612$, $p = .013$) and second ($t(40) = 2.523$, $p = .025$) half of the experiment. The strongly-accented exposure thus increased priming in both halves of the experiment, and therefore allowed listeners to interpret the accent more easily. Participants who received weakly-accented exposure did not show more priming than participants without exposure in the first half of the experiment ($t(40) = .546$, $p = .588$), but did in the second half of the experiment ($t(40) = 2.721$, $p = .010$). Hearing just the speaker without the strongly-accented items thus gave participants a head start, but was not enough to lead to priming from the start of the experiment.

In Experiment 2 we saw that participants were able to correctly interpret the weakly- and medium-accented items, as in Experiment 1. For the strongly-accented items, however, we saw that if participants listened to the speaker using the strongly-accented items for a short exposure period (four minutes) before starting the cross-modal priming experiment, they could correctly interpret the strongly-accented items from the start of the experiment. This may be evidence for speaker-specific adaptation. Experiment 3 looked further at the effect of speaker. Do the same effects occur when participants with limited exposure listen to different speakers of the same accent in the exposure phase and test phases? Previous research with exposure with multiple talkers does indicate that adaptation to an accent can transfer across speakers (e.g., Bradlow & Bent, 2008; Sidaras, et al., 2009). However, in the Experiment 2 we had only one speaker during exposure, which might not be enough for speaker-independent adaptation to occur. If adaptation to accents is speaker-specific, the exposure phase with a different speaker should not help participants to the same extent as listening to the strongly-accented story spoken by the test speaker. However, when taking into account the results from Experiment 1, where listeners who were very familiar with German-accented Dutch could quickly adapt to a new speaker, it is also possible that adaptation is to some degree speaker-independent in foreign-accented speech. If this is the case, we should see the same results as when participants listened to the strongly-accented exposure from the same speaker.

Experiment 3

Method

Participants

Nineteen participants took part in Experiment 3 (*M* age = 21.48, 17 females). Limited-exposure participants were recruited and selected using the same procedures as in Experiments 1 and 2. Three additional participants were excluded on the basis of the language history questionnaire. All participants reported normal hearing and normal or corrected-to-normal vision. All of them reported basic knowledge of German, with German being their second or third non-native language (after English), and none considered themselves fluent in German.

Materials

In Experiment 3 we used a different speaker to record the story used for the exposure phase. This speaker (Speaker 2) was selected from a pre-test and chosen because he pronounced the [œy] vowel in a similar fashion to the speaker of Experiments 1 and 2 (thus as [ɔɪ]). The speaker was a male native speaker of German, raised in *Nordrhein-Westfalen*. He had lived in The Netherlands for three years while studying in Dutch at the Radboud University Nijmegen. Like the previous speaker (Speaker 1), Speaker 2 was quite fluent in Dutch and knew almost all Dutch words used in the recorded story. Speaker 2 read one Dutch story, identical to the one used in the strongly-accented exposure condition in Experiment 2. The procedure of multiple recordings and of selection of final materials from those recordings was identical to that used in Experiment 2. In particular, recordings were selected in which Speaker 2 consistently pronounced the 12 strongly-accented words with [ɔɪ]). The procedure of the cross-modal priming experiment (with Speaker 1) was the same as in Experiments 1 and 2.

Results

The statistical analyses were identical to those used in Experiments 1 and 2. We excluded trials on which participants made errors (2.5%) and those trials on which the RTs deviated more than 2.5 SD from the condition's overall mean (together < 5% of all trials). The errors (see Appendix B) showed no specific patterns of priming.

As shown in Figure 6 (calculated priming effects) and Appendix B (mean RTs), participants were faster to respond to the identical words than to the unrelated words (*F1* (1,18) = 62.868, $p < .001$; *F2* (1,11) = 64.568, $p < .001$). In addition, participants were faster overall in the second half than the first (*F1* (1,18) = 9.941, $p = .006$; *F2* (1,11) = 22.634, $p = .001$). Participants did not respond differently to the three types of words (*F1* (2,36) = 2.448, $p = .102$; *F2* (2,20) = 1.279, $p = .300$). There were no significant interactions between any of the factors.

Figure 6. Experiment 3: Priming effects and CIs for limited experience group – Exposure with strongly accented items from different speaker by experiment halves and accent type.

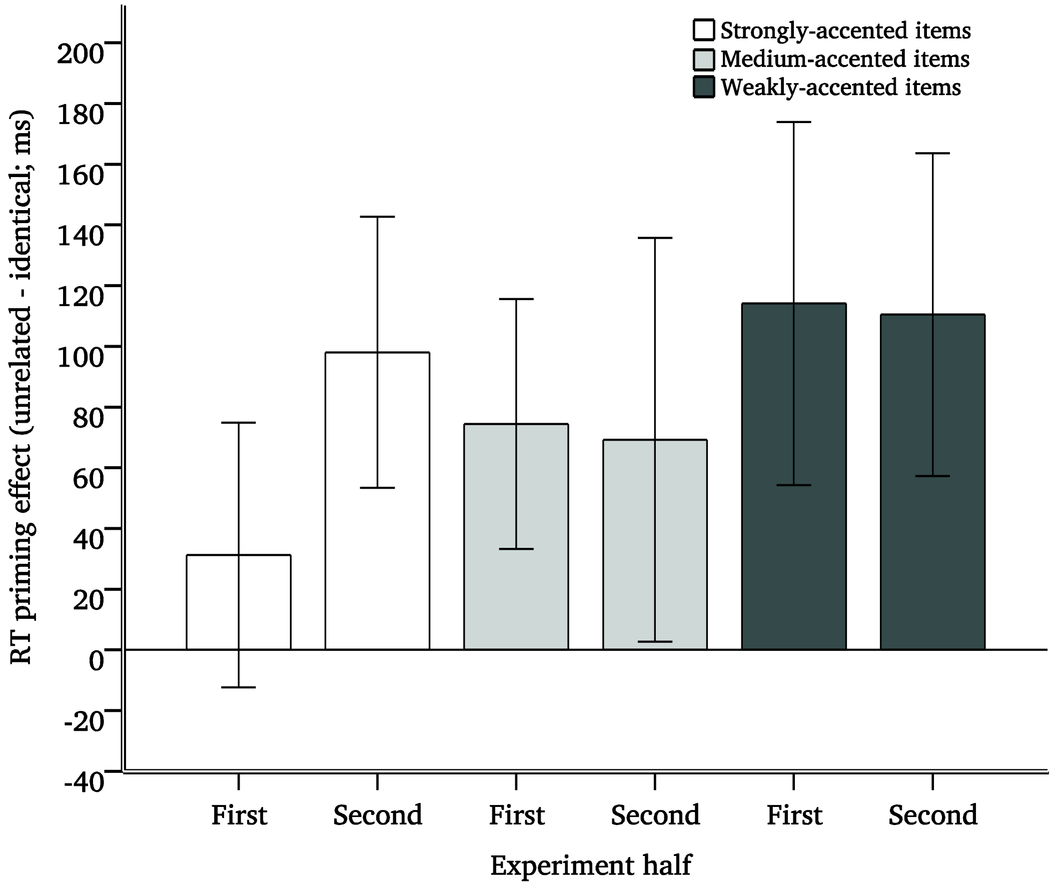


Table 5 displays the planned pair-wise comparisons of the priming effects. They show that participants could understand the weakly-accented and medium-accented words throughout the experiment. The priming effect for the strongly-accented items, however, was present only in the second half of the experiment.

Table 5. Planned Pairwise Comparisons of Priming Effects for all Accent Types Across Participants and Items for Participants with strongly-accented exposure from different speaker

	Participant analysis				Item analysis		
Condition	Half	*df*	*t*1	*p*	*df*	*t*2	*p*
Strongly-accented items	First	1,18	1.748	.098	1,11	1.857	.090
Strongly-accented items	Second	1,18	4.627	.000	1,11	4.446	.001
Medium-accented items	First	1,18	4.147	.001	1,11	4.838	.001
Medium-accented items	Second	1,18	2.142	.046	1,11	3.850	.003
Weakly-accented items	First	1,18	3.836	.001	1,20	4.159	.000
Weakly-accented items	Second	1,18	3.902	.001	1,20	4.333	.000

Discussion

When comparing these results to participants with limited experience who did not receive prior exposure (Experiment 1), we see that participants with exposure to a different speaker of German-accented Dutch did not show more priming for strongly-accented items in the first half of the experiment ($t(40) = .516$, $p = .604$), but did show significantly more priming in the second half of the experiment ($t(40) = 2.116$, $p = .041$). This result suggests that hearing a speaker with the same mispronunciations and the same accent does provide an advantage compared to not hearing the speaker at all (Limited Experience group, Experiment 1), but this advantage is not as large as hearing the same speaker (Limited Experience group – Exposure with strongly-accented items, Experiment 2). We cannot exclude the possibility that the weaker priming effects in Experiment 3 (Speaker 2 exposure) than in Experiment 2 (Speaker 1 exposure) arose because Speaker 2 produced more variable pronunciations of the strongly-accented words and thus was harder to learn from than Speaker 1. This appears unlikely, however, because the two speakers were selected for their similar pronunciation of the critical words, and because recordings of these words in the exposure story were selected in which both speakers consistently mispronounced all the critical vowels as [ɔɪ]).

General Discussion

The present study investigated the effects of different types of experience on word recognition in foreign-accented speech. Dutch listeners with limited prior experience with German-accented Dutch were able to interpret Dutch words with a weak or medium-strength German accent correctly, as measured in a cross-modal priming study, but they had difficulties interpreting strongly-accented words. But when a short exposure phase was added immediately before the cross-modal priming study, these participants' performance on the strongly-accented words improved, even though they had not heard those tokens in the exposure phase. It improved most when listeners had been exposed to the same speaker producing comparable strongly-accented items; in this case the short additional exposure led to equivalent performance to that of listeners with extensive prior experience. Short exposure to the speaker without strongly-accented tokens, as well as short exposure to strongly-accented tokens by a different German-accented speaker, also improved word recognition; improvement in these cases, however, was not observable immediately and emerged only in the second half of the experiment. These findings constitute evidence that short-term adaptation to a naturally-occurring foreign accent generalizes across words and show for the first time that this adaptation depends on the strength of the accented words – specifically on the vowels in those words.

Even though all words used in Experiments 1 and 2 were noticeably accented (as demonstrated in the rating study and the acoustic measurements), native Dutch listeners never had difficulties interpreting the weakly- and medium-accented items. This shows that recognition of variant forms is not necessarily difficult in L2 speech. This is in contrast with research using L1 speech, where it is a robust finding that priming with a word that mismatches with one phoneme from its canonical form does not facilitate target recognition (Marslen-Wilson et al., 1995) and sometimes even causes inhibition (Van Alphen & McQueen, 2006). Possibly, listeners more readily accept variation in pronunciation from L2 speakers; alternatively, deviations from the standard in weakly- and medium-accented words were too small to severely disrupt word recognition (see e.g., Connine, et al., 1993). In the case of the [ɛɪ]-vowel in medium-accented words, it could even be that similarity with a native variation facilitated recognition: e.g., in 'Poldernederlands', the 'polder ij' (Jacobi, 2009) is lowered and closer to [a] in comparison with the standard [ɛɪ], and it is therefore somewhat similar to the German-accented [aɪ]. But even though the German [aɪ] might be close to the 'polder ij', the rating study still indicated that native Dutch listeners considered the medium-accented words to be more accented than the weakly-accented words. In any case, the fact that listeners with limited prior experience could interpret weakly-accented and medium-accented words correctly suggests that extensive experience with an accent is not always required in order to be able to understand the accent. To a certain degree, we can rely on short-term perceptual learning mechanisms for handling variation in foreign-accented speech.

There is, however, also a role for long-term experience in understanding foreign-accented speech. Participants with extensive prior experience with German-accented Dutch had no difficulty recognizing the strongly-accented words, and they could do so even without brief pre-exposure to the speaker (i.e., in Experiment 1). Though it is still possible that listeners with extensive experience may encounter some difficulties interpreting foreign-accented speech, our results suggests that there is speaker-independent adaptation to foreign-accented speech. This is indeed good news for L1 listeners (and L2 speakers) since it implies that we do not have to adapt anew to each L2 speaker of a familiar accent. Since both listener groups showed a basic understanding of the accent (i.e., they could all understand weakly- and medium-accented words), it is likely that additional long-term experience puts listeners a little further ahead of the limited experience listeners and allows them to recognize even the strongly-accented items speaker-independently.

In Experiment 1, strongly-accented items such as /hɔɪs/ for /hœys/ [house] posed difficulties for native Dutch listeners with limited experience throughout the experiment. These listeners thus did not learn to interpret the strongly-accented words correctly during the eight minutes of the experiment. This result differs from earlier studies (e.g. Clarke & Garrett, 2004) that found adaptation to foreign-accented speech within one minute of exposure. One explanation for this difference in findings is that the accent markers in the present study were stronger and hence more difficult to learn. Previous studies on short-term adaptation to foreign accents usually did not control for specific accent markers, and it is feasible that items with varying strength of perceived accentedness were combined in these studies. A second explanation could lie in the contextual presentation of stimuli: Most of the previous studies used sentences, whereas the current experiment used isolated words. Sentence context, of course, provides richer information about the pronunciation habits of a particular speaker than isolated words do, and this additional information could make it easier for participants to tune in to a foreign-accented speaker. The fact that listeners performed better in the cross-modal priming task in Experiment 2 after being exposed to a story featuring the speaker and the accent supports this latter explanation.

In Experiment 3, it was found that short exposure to another speaker with the same accent also aids word recognition for strongly-accented words, but not to the same extent as exposure to the same speaker does. The two German-accented speakers were similar to one another in a number of ways: they were male, approximately the same age, and they mispronounced Dutch words in a similar way. An analysis of the two renditions of each of the twelve critical words from the story revealed that the two speakers were also comparable in terms of pitch (Speaker 1 mean = 133.0 Hz; Speaker 2 mean

= 132.7 Hz; $t(22) = .039$, $p = .969$) and speaking rate (Speaker 1 mean = 425 ms; Speaker 2 mean = 374 ms; $t(22) = 1.229$, $p = .232$). Despite these similarities, listeners recognized the accented words better when they were tested on the forms produced by the same speaker as they had heard during story exposure. This suggests that short-term adaptation to foreign-accented speech is speaker-specific to some extent. But the fact that hearing a different speaker during exposure aided word recognition at all suggests that even the initial stages of learning are also in part speaker-independent. It is possible that participants would be better at learning the accent if they were exposed to more than one or two speakers. This would allow them to learn more about the accent, rather than just about the speaker. Hearing multiple speakers might lead to speaker-independent adaptation. This could be tested by adding multiple accented speakers to the exposure and/or the test phase. Studies that have looked at general accent adaptation (i.e., not to specific accent markers) did show a beneficial effect of exposure to multiple speakers (e.g., Bradlow & Bent, 2008, Sidaras, et al., 2009). In fact, listeners are even able to adapt to an accent when the experiment contains one accented and one native speaker (Trude & Brown-Schmidt, 2011). One possible explanation of this result is that it was easier for participants in the Trude and Brown-Schmidt study to contrast native pronunciations with the accented mispronunciations. More evidence for differential effects of multiple speakers is found in studies on intelligibility of foreign-accented speech (e.g. Munro & Derwing, 1995b), but whether a similar effect is found for online word recognition remains a question for further research.

The results of the present priming studies constrain accounts of how words and their accented variants are represented in the lexicon. Theories on how variants in L1 speech are handled can be divided into two types of accounts: representational and processing-based accounts. The first type of account assumes that variation is encoded in lexical entries, that is, not only are the canonical forms stored, but also other variant forms of these words (Goldinger, 1998). Episodic representational accounts postulate that each word, as well as all its variations, is encoded in the lexicon with fine-grained phonetic detail (e.g., Johnson, 2006, Pierrehumbert, 2001). Another viewpoint is that the lexicon does not have episodic traces, but rather multiple abstract representations for variant forms (e.g., Ranbom & Connine, 2007). These representational accounts of course also make processing assumptions, but their explanation for recognition of variant pronunciations is based on representations. They propose that variant recognition should be easy when listeners have been exposed to the variants before (because the variation would already be encoded in the mental lexicon). In the present experiments, prior storage of variants (in abstract or episodic form) could explain the benefit shown by the participants with extensive exposure to German-accented Dutch and potentially the ability of the participants with limited exposure to recognize the weakly- and medium-accented words. It is less clear, however, how representational accounts could explain the learning shown by the limited-exposure participants on the strongly-accented words, and especially the generalization of learning from the words in the story to the new words in the test phase. A representation-based explanation for this kind of generalization would require that the limited-exposure participants happened to have heard (and stored) German-accented variants of the words used in the test phase prior to the experiment. This is unlikely, but not impossible. However, even if these participants had happened to have stored the pronunciation variants of the test words, a representation-based explanation would still require additional mechanisms that could account for how these stored variants did not immediately influence recognition (i.e., in Experiment 1) but instead only started to do so after in some way being triggered by exposure to other strongly-accented words (i.e. in Experiment 2).

A processing-based account of the adaptation appears to be more plausible. Such accounts assume that only canonical forms are stored in the lexicon, and that listeners learn how to map variant forms onto stored canonical forms through exposure (e.g., Lotto & Holt, 2006; Mitterer, et al., 2006; Gaskell & Marslen-Wilson, 1998). These models thus also make assumptions about representations, but the burden of their explanation for recognition of variant forms is carried by their assumptions about processing. The demonstration of generalization of learning for strongly-accented words from the exposure phase to the test phase speaks for such models. If perceptual adaptation reflects a change in the way a speech sound is mapped onto the lexicon, and that change takes place at a prelexical level of processing, then that learning will be reflected in all words containing that sound (McQueen et al., 2006). In the current situation, if the way the strongly-accented vowel is mapped onto the lexicon is modified during the exposure phase, then in the test phase it should be possible to recognize all words containing that vowel (i.e., whether those specific words have been heard before or not). This processing-based account is thus parsimonious because it is based on the same kind of perceptual-learning mechanism that has been proposed to explain learning about artificial accents and idiosyncratic pronunciations. The mapping that is involved seems to be highly dependent on the type of exposure, as is shown by the different results across the three experiments. The fact that listeners with ample prior experience have less trouble interpreting strongly-accented words than listeners with limited experience may indicate that the accented variants might be more strongly linked to their canonical forms for the listeners with more extensive experience.

Experience with an accent thus plays an important role in adaptation: Different types of experience lead to different kinds of adaptation. Prior experience does not seem to be necessary for adaptation to medium- and weakly-accented items: Re-mapping for these words takes place immediately or during the first few trials. For the strongly-accented items, however, more experience is needed to interpret these items correctly. This experience can either be gained outside the laboratory through extensive exposure to multiple speakers (Experiment 1) or with short-term exposure in the laboratory, but then only under the right circumstances, namely with the same speaker and the same strongly-accented mispronunciations. Only limited exposure of this type appears to be sufficient: In the current experiment, we used a four-minute story with only 12 strongly-accented mispronunciations. This is not the full picture, however, because hearing only the speaker without any strongly-accented words in the story also helped listeners (Experiment 2). A possible explanation could be that adaptation to foreign-accented speech takes place at a general level first before continuing to vowel-specific adaptation. Evidence for this vowel-specific adaptation can be found in Experiment 2: Only participants who have heard both the speaker and the strongly-accented mispronunciations were able to generalize correctly to new items in the first half of the experiment. It is therefore more likely that exposure to the story with weakly-accented words boosted the general adaptation, but not the vowel-specific adaptation.

In summary, though strongly-accented forms disturb online word recognition initially, adaptation to foreign-accented speech occurs within even a couple minutes of exposure. More research is needed to look at how long-lasting this exposure effect is, and whether this is a language-specific phenomenon (e.g., adaptation only occurs when listening to speakers with the same accent) or language-independent (e.g., when listening to accented speech listeners' word recognition becomes more flexible in general). The present study suggests that even within one accent there can be substitutions that differ in their comprehensibility. Recognition of words with vowel substitutions which are judged to be only weakly accented appears to be unproblematic. Vowel substitutions which are judged to be more strongly-accented do create recognition problems, but adaptation to them is rapid, with exposure to as few as 12 such words appearing to be sufficient for successful recognition of other strongly-accented words.

Appendix A - List of Experimental Items

Strongly-accented [œy]-items

Target Dutch spelling (IPA transcription)	*Accented prime (IPA)*	*Unrelated prime*	*Target translation*	*Unrelated prime translation*
buik (bœyk)	bɔɪk	dief	belly	thief
buiten (bœytə)	bɔɪtə	koffie	outside	coffee
duif (dœyf)	dɔɪf	riet	dove	cane
duiker (dœykər)	dɔɪkər	wakker	diver	awake
duim (dœym)	dɔɪm	pink	thumb	pinkie
duivel (dœyvəl)	dɔɪvəl	tuniek	devil	tunic
fruit (frœyt)	frɔɪt	tenue	fruit	uniform
ruiter (rœytər)	rɔɪtər	gewoon	rider	normal
snuit (snœyt)	snɔɪt	kreng	snout	hag
struik (strœyk)	strɔɪk	koorts	shrub	fever
uiterst (œytərst)	ɔɪtərst	binding	final	bond
zuiver (zœyvər)	zɔɪvər	gebaar	pure	gesture

Medium-accented [ɛɪ]-items

Target Dutch spelling (IPA transcription)	*Accented prime (IPA)*	*Unrelated prime*	*Target translation*	*Unrelated prime translation*
cijfer (sɛɪfər)	saɪfər	haven	digit	harbor
lijf (lɛɪf)	laɪf	riem	body	belt
lijst (lɛɪst)	laɪst	naast	list	beside
nijdig (nɛɪdəg)	naɪdəg	diepte	angry	depth
olijf (o:lɛɪf)	o:laɪf	emmer	olive	bucket
pijn (pɛɪn)	paɪn	muur	pain	wall
rijp (rɛɪp)	raɪp	tent	ripe	tent
slijm (slɛɪm)	slaɪm	inkt	slime	ink
spijker (spɛɪkər)	spaɪkər	boeiend	nail	compelling
strijder (strɛɪdər)	straɪdər	inning	warrior	inning
terwijl (tɛrwɛɪl)	tɛrwaɪl	aanzoek	while	proposal
twijfel (twɛɪfəl)	twaɪfəl	prettig	doubt	agreeable

Weakly-accented items

Target Dutch spelling (IPA transcription)	*Accented prime (IPA)*	*Unrelated prime*	*Target translation*	*Unrelated prime translation*
bek (bɛk)	bɛk	pet	beak	cap
bemind (bəmɪnt)	bəmɪnt	immuun	loved	immune
boek (buk)	buk	week	book	week
boete (butə)	butə	effen	fine	plain
code (ko:də)	ko:də	unie	code	union
defect (dəfɛkt)	dəfɛkt	tegoed	defect	credit
dekking (dɛkkɪŋ)	dɛkkɪŋ	katoen	cover	cotton
eb (ɛp)	ɛp	mus	ebb	sparrow
fitting (fɪtɪŋ)	fɪtɪŋ	techniek	fitting	technique
*gala (xa:la:)	xa:la:	jute	gala	jute
hek (hɛk)	hɛk	mes	fence	knife
*hem (hɛm)	hɛm	mik	him	mouth
hik (hɪk)	hɪk	wok	hiccup	wok
hit (hɪt)	hɪt	erg	hit	very
ketting (kɛtɪŋ)	kɛtɪŋ	prikkel	chain	stimulus
*kik (kɪk)	kɪk	pin	peep	pin
koek (kuk)	kuk	deeg	cookie	dough
mep (mɛp)	mɛp	pit	slap	pit
midden (mɪdə)	mɪdə	plicht	middle	duty
min (mɪn)	mɪn	toe	minus	to
minuut (miny:t)	miny:t	bezem	minute	broom
nek (nɛk)	nɛk	baan	neck	lane
niet (nit)	nit	keel	not	throat
uniek (y:nik)	y:nik	limoen	unique	lime

*Items excluded from the analyses of the cross-modal priming experiments

Appendix B – Overview of all Reaction Times and Errors per Experiment by Condition and Half

Overview of RTs (and SDs) for all Conditions in Experiment 1

		Limited experience				Extensive experience			
Accent type	Half	Identical		Unrelated		Identical		Unrelated	
Strongly-accented	First	583	(89)	604	(81)	523	(89)	564	(82)
Strongly-accented	Second	572	(91)	603	(112)	501	(93)	556	(88)
Medium-accented	First	541	(104)	633	(85)	510	(82)	575	(87)
Medium-accented	Second	510	(97)	586	(87)	480	(95)	545	(88)
Weakly-accented	First	563	(71)	637	(87)	507	(87)	595	(103)
Weakly-accented	Second	533	(95)	616	(85)	507	(83)	560	(69)

Overview of RTs (and SDs) for all Conditions in Experiment 2

		Strongly-accented Exposure				Weakly-accented Exposure			
Accent type	Half	Identical		Unrelated		Identical		Unrelated	
Strongly-accented	First	569	(141)	669	(109)	576	(101)	617	(86)
Strongly-accented	Second	554	(100)	666	(67)	517	(116)	642	(101)
Medium-accented	First	572	(91)	700	(107)	550	(85)	642	(99)
Medium-accented	Second	545	(80)	664	(90)	528	(97)	657	(120)
Weakly-accented	First	591	(118)	701	(88)	562	(82)	651	(92)
Weakly-accented	Second	580	(76)	685	(74)	557	(105)	642	(79)

Overview of RTs (and SDs) for all Conditions in Experiment 3

		Strongly-accented Exposure (different speaker)			
Accent type	Half	Identical		Unrelated	
Strongly-accented	First	612	(117)	646	(86)
Strongly-accented	Second	554	(67)	649	(79)
Medium-accented	First	576	(81)	657	(84)
Medium-accented	Second	549	(98)	613	(89)
Weakly-accented	First	584	(89)	691	(104)
Weakly-accented	Second	554	(77)	655	(95)

Overview of Errors across all Conditions and Experiments

			Errors (percentage)			
Experiment	Group	Accent type	Identical		Unrelated	
1	Limited experience	strong	2	(1.6)	5	(4.0)
		medium	0	(0.0)	5	(4.0)
		weak	3	(1.2)	7	(2.8)
1	Extensive experience	strong	2	(0.7)	0	(0.0)
		medium	2	(0.7)	2	(0.7)
		weak	8	(1.7)	4	(0.8)
2	Exposure with strongly-accented words	strong	0	(0.0)	1	(0.4)
		medium	1	(0.4)	1	(0.4)
		weak	4	(1.0)	5	(1.3)
2	Exposure without strongly-accented words	strong	0	(0.0)	2	(0.9)
		medium	1	(0.4)	1	(0.4)
		weak	5	(1.3)	7	(1.8)
3	Exposure from different speaker	strong	4	(1.8)	1	(0.4)
		medium	5	(2.2)	4	(1.8)
		weak	11	(2.8)	15	(3.8)

On the relationship between perceived accentedness, acoustic similarity, and processing difficulty in foreign-accented speech

Chapter 3

A version of this paper appeared as: Witteman, M. J., Weber, A., & McQueen, J. M. (2011). On the relationship between perceived accentedness, acoustic similarity, and processing difficulty in foreign-accented speech. Proceedings of the 12th Annual Conference of the International Speech Communication Association (Interspeech 2011), Florence, Italy (pp. 2229-2232)

Abstract

Foreign-accented speech is often perceived as more difficult to understand than native speech. What causes this potential difficulty, however, remains unknown. In the present study, we compared acoustic similarity and accent ratings of American-accented Dutch with a cross-modal priming task designed to measure online speech processing. We focused on two Dutch diphthongs: ui and ij. Though both diphthongs deviated from standard Dutch to varying degrees and perceptually varied in accent strength, native Dutch listeners recognized words containing the diphthongs easily. Thus, not all foreign-accented speech hinders comprehension, and acoustic similarity and perceived accentedness are not always predictive of processing difficulties.

Introduction

Listening to native speech seems effortless, but understanding foreign-accented speech can be much more difficult (e.g., Adank & McQueen, 2007; Munro & Derwing, 1995b; Witteman, Weber, & McQueen, in press). Foreign accents are common in people who learned a second language (L2) in adulthood, and are mainly caused by language-specific structures from the speaker's native language which influence the production of the L2. Foreign accents can alter the speech signal in a number of ways. For example, phoneme contrasts or allophonic variations can be lacking (e.g., Italy pronounced as Eataly by Italian speakers), subphonemic as well as suprasegmental cues can be realized differently (e.g., deFAULT, pronounced as DEfault by Hungarian speakers), and quite regularly L2 speakers replace speech sounds with those that approximate or match native categories but do not occur in the target language (e.g., Hudson pronounced as Hüdson by Dutch speakers).

While in all the above cases the speech signal deviates from the standard norm of the target language, the amount of deviation likely varies both acoustically and perceptually for different accent markers. That is, foreign-accented speech sounds can vary in how close they are qualitatively and quantitatively to categories of the target language, and at the same time different accent markers can interfere with comprehension to varying degrees. The aim of the current project is to investigate whether acoustic similarity and comprehension go hand in hand, or whether the amount of acoustic deviation is not necessarily predictive of listening performance. Specifically, we are interested in the relationship between acoustic similarity, perceived accentedness, and the degree of processing difficulty. The case will be made with two specific Dutch diphthongs, and processing will be assessed with an online word recognition task.

Research on cross-language acoustic similarity usually compares the acoustic characteristics of speech sounds in different languages (Strange, Levy, & Law, 2009). Based on these comparisons, predictions can then be made about how listeners perceive non-native sounds. Although acoustic similarity indeed often predicts how listeners categorize non-native sounds, acoustic analyses cannot always explain cross-language perceptual difficulties (Flege, 1995). It is therefore unclear whether acoustic analyses in the present study will be able to fully account for processing difficulties in a comprehension task. (Note that we are not testing L2 listeners, rather we are testing L1 listeners on L2 speech, and our task is an online measure of word recognition rather than phoneme categorization.)

Native listeners are generally very sensitive to foreign accents (Munro & Derwing, 1995a). That is, even small traces of foreignness can be detected; consequently, even highly experienced L2 speakers are usually rated as more accented than native speakers. There is, however, also evidence that perceived accentedness is not necessarily predictive of comprehension. As Munro and Derwing (1995a) have shown, for example, it does not take longer for native listeners to assess the truth value of foreign-accented sentences in comparison to L1 sentences. Again, it is therefore unclear whether perceived accentedness in the present study will be in line with processing difficulties.

Our knowledge about how variation in speech is being processed comes mainly from research on variation in native speech. An example of this variation is medial /t/ deletion, where ‘center’ is pronounced as 'senner' or 'sennah' in American English (Pitt, 2009). When participants judged these variants in isolation, they only accepted /t/-deleted variants as words if the phonological environment in which the /t/-deletion occurred was common in production. This effect translated to new variants, but only if the /t/-deletion happened in commonly occurring places. Accepting this type of variation is thus limited to the context in which it occurs.

Sumner and Samuel (2009) looked at another type of variation in speech: variation across dialects. They tested how dialectal variants are recognized and stored in the lexicon, and whether the amount of exposure to phonological variants made a difference. They compared processing of words ending in –er. In the New York City (NYC) dialect, the r is dropped (e.g., (bak[ɚ] is pronounced as bak[ə]). They found that speakers of this dialect and speakers of General American (GA) who were very familiar with the NYC dialect correctly interpreted the r-less forms as the intended form. GA-speakers unfamiliar with the NYC-accent, however, did not show priming for these dialectal forms. The authors conclude that listeners who have (passive or active) familiarity with a dialect are apparently more flexible in form processing than inexperienced listeners are: experienced listeners can deal with more variation to the standard representation when listening to dialectal speech.

Though foreign-accented speech deviates noticeably from the standard, adaptation can take place quickly. Clarke and Garrett (2004), for example, found that native English listeners improved their understanding of a foreign-accented speaker within a couple of sentences. Intelligibility of the speaker also plays an important role. Bradlow and Bent (2008) found that when listeners were exposed to speakers of high, medium, and low intelligibility, listeners needed more exposure to improve their understanding of the less intelligible speakers than of the more intelligible speakers.

This study was designed to shed more light on the driving factors behind processing foreign-accented speech. Specifically, we were interested to see what the roles of acoustic similarity and accentedness rating were in online speech processing.

The target language was Dutch, and the foreign accent was an American English one. Because variation in foreign accents often centers on vowels, these were taken as a starting point for item selection. In particular, we focused on the two Dutch diphthongs [œy] and [ɛɪ]. Both diphthongs are difficult for many learners of Dutch and are likely to deviate substantially from their standard pronunciations. Dutch words containing [œy] and [ɛɪ] were compared with Dutch words without any specific known markers of an American accent. These non-specifically accented words contained varying monophthongs (some of which occur in both English and Dutch), and should be relatively easy for American speakers to produce. The Dutch diphthongs [œy] and [ɛɪ], on the other hand, are typically replaced with diphthongs that resemble American [aʊ] and [aɪ] by American learners of Dutch.

In order to investigate the relationship between acoustic similarity, perceived accentedness, and processing difficulty, we first analyzed the spectral quality of the American-accented Dutch diphthongs, then had Dutch listeners rate the strength of foreign accent, and finally presented the American-accented stimuli to native Dutch listeners in a cross-modal priming study.

Forty-eight Dutch mono- or bisyllabic words were selected. Twelve of them contained the diphthong [œy] as in duif 'dove', 12 the diphthong [ɛɪ] as in rijst 'rice', and the remaining 24 words contained a variety of monophthongs, e.g., [i:] in diep 'deep'. The last group of words contained no sounds that are known to be specifically difficult for English learners of Dutch; it was therefore expected that these words would be perceived as the least accented.

The American speaker of the study consistently substituted Dutch [œy] with a sound that resembles American [aʊ] and Dutch [ɛɪ] with sound that resembles American [aɪ]. He was a native speaker of American English, who moved to the Netherlands less than a year ago. His Dutch proficiency was basic. The Dutch words were recorded together with the filler primes from the cross-modal priming experiment in one session. Recordings were made in pseudo-randomized order from correct Dutch spelling. All mispronunciations occurred naturally.

Method

Acoustic analyses

Vowel durations for the 12 words containing [œy] and for the 12 words containing [ɛɪ] were measured using Praat (Boersma & Weenink, 2009). Duration was labeled from the release of the constriction of the preceding consonant to the formation of the constriction of the following consonant, with labels being placed at zero crossings. Subsequently, F1 and F2 were measured at the 25 and 75 percent time points for each diphthong. These formant values were compared to the average vowel characteristics of native Dutch speakers of Northern Standard Dutch from Adank, Van Hout, and Smits (2004).

Rating experiment

The rating experiment contained the 12 American-accented words with [œy], the 12 American-accented words with [ɛɪ], and 12 of the nonspecifically-accented words. In order to add more variation to the materials, two sets of items spoken by a Dutch native speaker and a native speaker of Italian with a very strong accent in Dutch were added to the task. In the rating study, participants heard one word at a time over closed headphones, immediately followed by a rating scale where they could indicate how accented the word was on a scale ranging from 1 to 10 (1: not accented, 10: very strong accent).

Cross-modal priming experiment

For the priming experiment, all 48 selected Dutch words (12 [œy]-words, 12 [ɛɪ]-words and 24 nonspecifically-accented words) served as visual targets and were each combined with two auditory primes. Primes were either the American-accented variant of the target (identical primes) or phonologically and semantically unrelated Dutch words (unrelated primes). Unrelated primes matched overall in number of phonemes with their target (e.g., prime ketting 'chain' and target prikkel 'incentive'), and the overall lexical lemma frequency of unrelated primes was not different from the frequency of targets (log frequency taken from the CELEX database; Baayen, et al., 1993).

Ninety-six filler items were added to avoid participants using strategic answering methods. Twenty-four of the fillers had a Dutch word as the visual target (18 of them contained [œy] or [ɛɪ] and were preceded by a non-word prime). The remaining 72 fillers had a non-word as their visual target (36 with a word as prime and 36 with a non-word as prime).

Two versions of the experiment were created, so that every participant saw each visual target only once. To control for effects of presentation order, each participant received a different pseudo-randomized list. The first two items of the experiment were always fillers, and there were never two critical items in a row.

Participants were seated in a sound-attenuated cabin and were informed that they would first hear a Dutch word or non-word spoken by an American speaker and then see a Dutch word or non-word on the screen. Their task was to decide as quickly and accurately as possible whether the word presented on the screen was an existing Dutch word or not by pushing one of two buttons on a button box in front of them. Yes-responses were always made with the dominant hand, and reaction times (RTs) were measured from visual target onset.

Auditory primes were presented binaurally over closed headphones at a comfortable listening level. Participants saw the visual targets on a computer screen situated about 50 cm in front of them. Visual targets were presented in white lowercase 24p Tahoma letters on a black background, 500 ms after the acoustic offset of the auditory primes. The visual targets stayed on the screen for 2000 ms, after which the next trial started. The experiment was created in Presentation (version 13. Neurobehavioural Systems Inc.) and used a NESU (Nijmegen Experiment Set-Up) button box.

Participants were 17 native speakers of Dutch (13 females, *M* age 22.11 years). None reported a hearing disorder, and all had normal or corrected-to-normal vision. Participants first completed the cross-modal priming task, before they participated in the rating experiment and completed a language history questionnaire.

Results

Acoustic analyses

Figure 1: Average F1 and F2 values for Dutch /a/, /i/, /o/, /œy/, /ɛɪ/, and /ɔu/ for male Dutch speakers in the Adank et al. corpus (2004) and in bold, F1 and F2 values for /œy/ and /ɛɪ/ for the American-accented speaker in the present study.

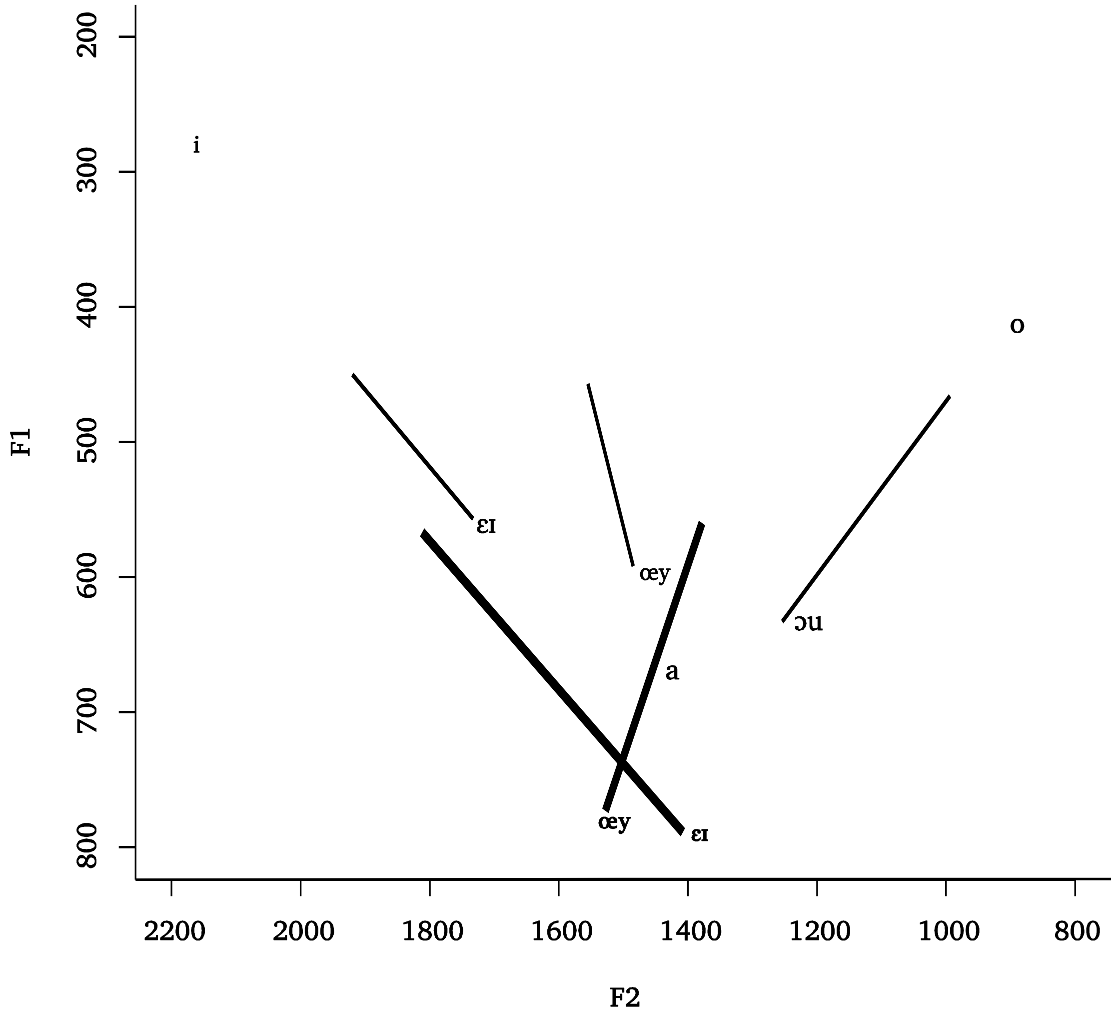


Figure 1 plots the average F1 and F2 values for the Dutch monophthongs /a/, /i/ and /o/, as well as for the three diphthongs /œy/, /ɛɪ/, and /ɔu/ as spoken by male Dutch native speakers (data from Adank et al., 2004); in bold the F1 and F2 values for the two diphthongs /œy/ and /ɛɪ/ are shown as the American speaker of the study produced them.

As one can see, the diphthongs of the American speaker started in comparison with the Dutch values at a higher F1 with an F2 value typical of Dutch /a/. Thus, the beginning of the diphthongs was lower, and in the case of [ɛɪ] more back than standard Dutch pronunciation would predict. In addition, though the [ɛɪ]-vowel produced by the American speaker has a similar trajectory as its Dutch counterpart, the trajectory of the [œy]-vowel is untypical. The F2 value of the [œy]-vowel at the 75% point is lower than at the 25% point, although standard Dutch pronunciation would predict a higher F2 end point. That is, rather than moving to a near-front vowel, the diphthong moved to a near-back vowel. In doing so, the trajectory of the [œy]-vowel resembles more the trajectory of the Dutch [ɔu].

A comparison of the F1 and F2 values in the present study with the average values for American diphthongs (Ladefoged, 1999) suggests that the diphthongs by our American learner of Dutch unsurprisingly approximated the American diphthongs [aʊ] (for [œy]) and [aɪ] (for [ɛɪ]).

In sum, while both Dutch [ɛɪ] and [œy] were produced non-canonically, [ɛɪ] deviated mostly in the first target of the diphthong, but [œy] deviated in the first and second target. In addition to a more distinct deviation, American-accented [œy] productions are spectrally close to another existing Dutch category, namely [ɔu].

Rating experiment

The accentedness ratings of the 17 Dutch participants showed that the Dutch native speaker was rated as least accented and the native Italian speaker as most accented. The ratings for the American speaker were further analyzed with paired-samples t-tests. The average ratings can be found in Table 1.

Table 1: Average ratings for [œy] words, [ɛɪ] words, and words with no specific accent marker produced by American speaker.

Accent type	Mean	*SD*
[œy]	7.24	1.41
[ɛɪ]	6.24	1.18
Non-specific	5.38	1.66

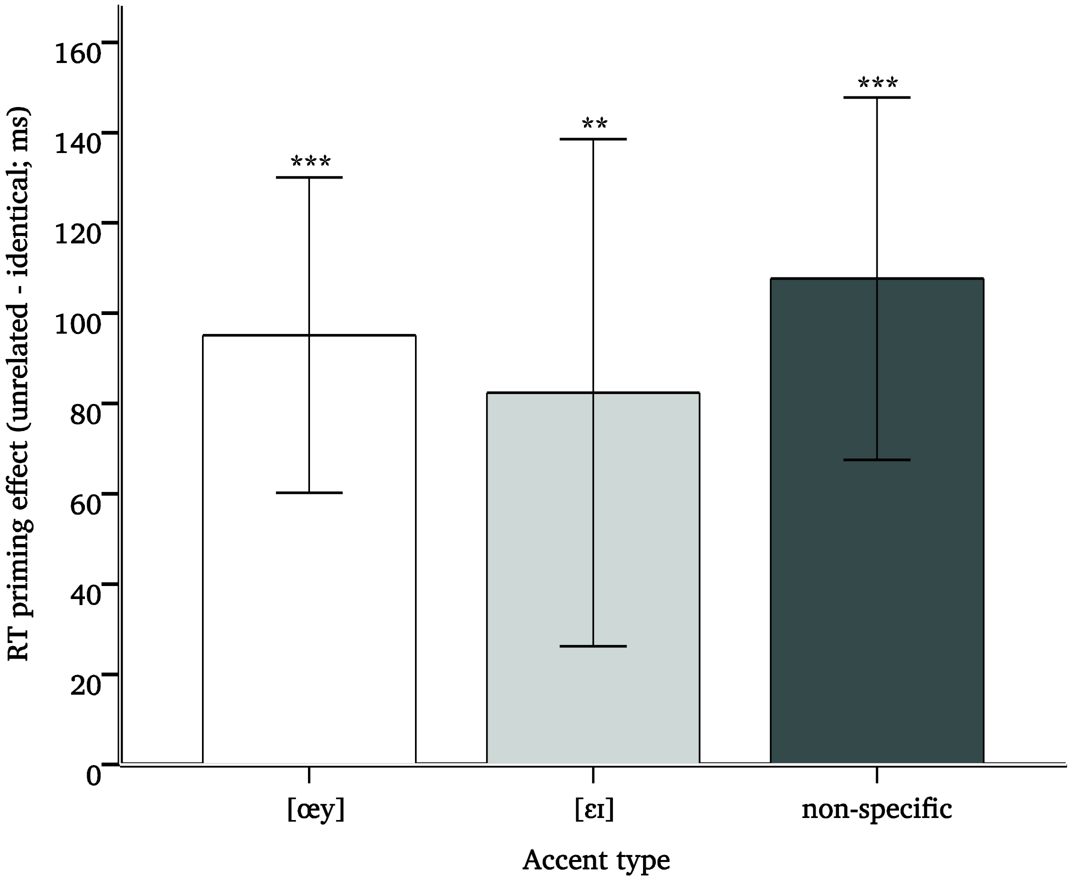
The t-tests indicated that the [œy]-words were rated more accented than the [ɛɪ]-words ($t(16) = 4.712, p < .001$) and the non-specifically accented words ($t(16) = 6.710, p < .001$). In addition, the [ɛɪ]-words were rated more accented than the non-specifically accented words ($t(16) = 4.378, p < .001$). The outcome of the rating study is in line with the acoustic analyses in that the accented productions that deviate the most from the standard pronunciation are also perceived as the most accented.

Cross-modal priming experiment

Three non-specifically accented primes were discarded from the analysis because they had very high error rates (more than 10%). A further 1.2% trials on which participants made errors, as well as trials with RTs that deviated more than 2.5 SD from the condition's mean overall were discarded (together $< 4\%$ of all trials).

The remaining cross-modal priming data was analyzed with a Repeated Measures General Linear Model (GLM) using a 3 (type of accent - strong, medium, weak) x 2 (relatedness - identical vs. unrelated) design. All factors were within-participants. We analyzed the results separately by participants (*F1*) and items (*F2*). The calculated priming effects, i.e. the difference in RTs to targets following unrelated and related primes, are shown in Figure 2.

Figure 2: Average priming effects (and SE) in ms by accent type (** $p < .01$; *** $p < .001$)



Dutch listeners showed a significant main effect of relatedness ($F1$ (1,16) = 83.643, $p < .001$; $F2$ (1,11) = 41.122, $p < .001$) and type ($F1$ (2,32) = 10.800, $p < .001$; $F2$ (2,22) = 4.417, $p = .024$) and no interaction. That is, they responded faster to targets following identical primes compared to targets following unrelated trials. The size of the priming effect differed between the three word types, with the effect being the largest for words with non-specific accent markers. Priming effects for [œy] and [ɛɪ] words were comparable.

Thus, even though American-accented [œy]-words deviated spectrally the most from their standard Dutch pronunciation and were also perceived as the strongest accented, they were recognized quite easily by native Dutch listeners. In fact, there was no difference in ease of interpretation between the [œy]-words and the [ɛɪ]-words, even though the latter words were rated as more weakly accented and deviated less acoustically. Further analyses showed that Dutch listeners did not learn how to interpret American-accented words during the course of the experiment: the priming effects were already present in the first half of the experiment.

Discussion

The purpose of this study was to investigate some of the factors that could cause processing difficulty in the recognition of foreign-accented speech. Native speakers of Dutch performed a cross-modal priming task in which they heard three types of Dutch words spoken by an American speaker: words containing [œy], words containing [ɛɪ], and non-specifically accented words. Moreover, the degree of perceived accentedness of these words and the acoustic similarity of the American-accented diphthongs to the intended Dutch vowels were measured.

The acoustic analyses showed that the American speaker indeed produced diphthongs that clearly deviated from their standard Dutch pronunciations. The speaker's productions resembled existing American diphthongs, and deviated from standard Dutch more substantially in the case of [œy] than in the case of [ɛɪ]. The rating study confirmed this pattern of results with the accent being perceived more strongly in words with the larger acoustic deviation (i.e., words containing [œy]). Although this difference in ratings can also be influenced by other uncontrolled accent markers within the lexical items, vowel quality is very likely an important factor.

Even though the accent markers were thus clearly noticed (also in comparison to L1 Dutch), and furthermore differed in their degree of deviation, the priming study showed that Dutch listeners could interpret words with both accent markers quite well, and with no significant difference in the ease of recognition between them. Thus, the amount of deviation did not measurably affect processing. This is good news for native listeners (and indeed non-native speakers), since acoustic variation and perceived accentedness thus do not necessarily imply difficulties for online comprehension.

There are a number of possible explanations why American-accented words were recognized so easily and why there were no processing differences between variant forms. American-accented words could have been recognized easily because Dutch listeners are familiar with the tested variant forms (either from hearing American-accented Dutch or from other accents that pronounce the Dutch diphthongs similarly). The fact that facilitatory priming was found from the start of the experiment further supports this explanation. Dutch listeners apparently did not need to learn how to interpret American-accented words but could do so right from the start (for converging evidence on familiarity effects, see Witteman, et al., in press). Alternatively, American-accented variant forms could simply not have been accented enough to interfere with processing. The accents tested in Witteman et al. were indeed perceived as even more accented than the accent tested here. Processing difficulties may thus arise only for listeners unfamiliar with the accent and/or for more extreme accents.

The lack of a difference in processing between variant forms either could reflect insufficient sensitivity of the paradigm (though differences have been found before using cross-modal priming; Witteman, et al., in press), or indeed could reflect a dissociation between acoustic similarity, perceived accentedness and processing. In the latter case, it is for example possible that the American-accented variant forms of Dutch [œy] and [ɛɪ] assimilate to the Dutch categories despite their deviations and therefore processing is not hindered. A categorization task with Dutch listeners categorizing the American-accented vowels into Dutch categories could help to clarify this point.

Further research will be needed to further tease apart the characteristics of foreign-accented speech and the subsequent consequences for its processing by native listeners. The current study shows that relying on acoustic measurements or rating tasks might not be sufficient to fully explain all aspects of speech processing.

Automaticity and stability of adaptation to foreign-accented speech

Chapter 4

Witteman, M. J., Bardhan, N. P., Weber, A., and McQueen, J. M. (under revision). Automaticity and stability of adaptation to foreign-accented speech. *Language and Speech*

Abstract

In three cross-modal priming experiments we asked whether adaptation to foreign-accented speech is automatic, and whether adaptation can be seen after a long delay between initial exposure and test. Dutch listeners were exposed to Hebrew-accented Dutch with two types of Dutch words: those that contained [ɪ] (canonical words), and those in which the Dutch [i] was shortened to [ɪ] (variant words). Experiment 1, which served as a baseline, showed that native Dutch participants showed facilitatory priming for canonical, but not variant, words. In Experiment 2, participants performed a 3.5-minute phoneme monitoring task on Day 1, and were tested on their comprehension of the accent 24 hours later using the same cross-modal priming task as in Experiment 1. During the phoneme monitoring task, listeners were asked to detect a consonant that was not strongly accented. In Experiment 3, the delay between exposure and test was extended to one week. Listeners in Experiments 2 and 3 showed facilitatory priming for both canonical and variant words. Together, these results show that adaptation to foreign-accented speech is rapid and automatic, and can be observed after a prolonged delay in testing.

Introduction

Every day, we are confronted with all kinds of variation in speech. For example, we have to adapt to differences because of a speaker's gender, age, and speaking rate, as well as because of background noise and speakers' different emotional states. While we are usually able to handle variation within our native language without difficulties, probably all of us have sometimes had serious problems coping with foreign-accented speech. After listening to a foreign-accented speaker for a little while, however, probably all of us have also experienced that the speaker becomes more intelligible. It is unlikely that the signal has changed under these circumstances, so somehow we must have adapted to the speaker. The present study explores the nature of such adaptation to foreign-accented speech. Specifically, two questions are asked: first, how automatic is this adaptation; and, second, how long-lasting is it?

Listeners are able to adapt to all kinds of variation within native speech, as studies focusing on perceptual learning have shown (e.g., Norris, et al., 2003; Kraljic & Samuel, 2005; 2006; Eisner & McQueen, 2005). Goldstone (1998) defined perceptual learning as "relatively long-lasting changes to an organism's perceptual system that improve its ability to respond to its environment and are caused by its environment" (p. 586). One example is listeners' adaptation to artificially created sounds such as [?], midway between [s] and [f]. Listeners are able to use their lexical knowledge when adapting to these sounds in an exposure phase, i.e., interpreting [?] as either [s] or [f] depending on the word context. This perceptual learning is known to generalize to new words in a test phase (e.g., McQueen, Cutler, et al., 2006) and was found using different tasks for exposure and test (e.g., McQueen, Norris, et al., 2006).

Listeners can also use perceptual learning to adapt to pronunciations in which a sound in a word is consistently replaced with a different native sound (rather than an ambiguous one; Maye, Aslin, and Tanenhaus, 2008). Native English participants listened to a story in which all English front vowels were lowered, after which they made lexical decisions on these words and novel words. Listeners generalized their knowledge of the speaker's 'accent' to new words with the same accent, but did not extend this vowel shift to non-front vowels that did not carry an accent in the exposure phase. Kraljic and Samuel (2006) found that listeners did generalize to new phonemes after exposure, but used stop consonants instead of vowels during the exposure and test phase. Participants were exposed to ambiguous /t/ or /d/ phonemes, and then asked to categorize both these phonemes and phonemes on a /b/-/p/ continuum. Participant showed a small but reliable training effect on both continua. Listeners are even able to accept a familiar non-native sound (English [θ], only known to the Dutch participants as a second language [L2] sound) as a substitute for the native categories [s] or [f] (Sjerps & McQueen, 2010). In fact, the priming effects obtained were comparable in size to those obtained with an ambiguous sound midway between [s] and [f].

Participants are thus able to adapt to artificially induced variation within their native language, even if this variation stems from another language (Sjerps & McQueen, 2010). This adaptation takes place extremely quickly. For example, it is observed after as few as 10 items in the exposure phase (Flege, 1995). Moreover, the process is thought to be automatic (McQueen, Norris, et al., 2006). That is, attention to the accent, in the form of conscious decisions on the mispronunciations, is not a requirement for perceptual learning effects to arise. In fact, even when listeners are instructed to just listen to a story (Eisner & McQueen, 2006) or asked to simply count the number of trials, perceptual learning effects are observed. Whether or not participants made conscious decisions about the stimuli did not affect the size of the priming effect (McQueen, Norris, et al., 2006).

What is less clear, however, is whether foreign-accented speech can count on the same flexibility. Foreign-accented speech brings a different type of variation to the speech signal, namely variation that is driven by the native language of the speaker. Moreover, foreign-accented speech does not create variation on just one phoneme, but rather affects many segments, and to different degrees.

Empirical findings from a cross-modal matching task indicate that general adaptation to foreign-accented speech can indeed be quick (Clarke & Garrett, 2004), in line with results from native perceptual learning studies. Moreover, we get better at understanding foreign-accented speech with additional exposure to it (e.g., Bradlow & Bent, 2008; Sidaras, et al., 2009).

Intelligibility of foreign-accented speech increases as accent strength decreases (Bent & Bradlow, 2003), and listeners benefit from more exposure in understanding low-intelligibility (but not high-intelligibility) speakers (Bradlow & Bent, 2008). Listeners also benefit from short additional exposure when adapting to strongly-accented words, whereas this exposure is not necessary to adapt to less strongly-accented words (Witteman, et al., in press).

These previous studies do indeed show, following Goldstone (1998), that changes to one's perceptual system to improve understanding of language are possible. However, Goldstone also notes that these changes need to be long-lasting. Much less research has focused on this aspect of perceptual learning. It is not obvious that long-lasting changes in the perceptual system are beneficial, because not all changes are stable. Moreover, the perceptual system needs to find a balance between stability and flexibility. If the system is too stable, it might not be able to adapt to temporary changes, or to new speakers. However, if the system is too flexible, it might be constantly re-inventing itself. In order to function optimally, the system needs to find an equilibrium between flexibility and stability.

Foreign-accented speech could provide a situation in which stability is preferred. A speaker who cannot produce a certain non-native phoneme is likely to continue consistently mispronouncing this phoneme. And since most variation in foreign-accented speech is driven by the speaker's first language (L1), this could provide an example of stable pronunciation differences.

Within-language contrasts have been shown to possibly induce long-lasting changes to the perceptual system (Kraljic and Samuel, 2005). Listeners first completed a lexical decision task in which either /s/ or /ʃ/ was replaced by an ambiguous sound, followed by a silent intervening task for 25 minutes. After this delay, listeners categorized /s/-/ʃ/ continua. Not only did the perceptual learning effect remain stable during this delay, it actually increased compared to a no-delay condition.

Evidence of perceptual learning can even be observed after a delay of 12 hours (Eisner & McQueen, 2006). Participants listened to an exposure story where either [f] or [s] was replaced by [?], followed by a categorization task, and, after 12 hours, another categorization task. Listeners showed evidence of perceptual learning after the delay, and effect sizes were equally large for participants first tested in the morning and then again in the evening compared to participants who were tested in the evening and again in the morning and who thus had slept between tests.

Though both these studies indicate that changes to the perceptual system can be observed after a delay, they do not shed much light on how the perceptual system could deal with natural variation in speech. As mentioned above, foreign-accented speech provides an example of speech variation that is likely to remain stable over a longer period of time, because L2 speakers do not have these L2 phonemes in their inventory, and thus replace them with phonemes from their L1. Because of these stable properties of the speech signal, it would be beneficial for the perceptual system to adapt the categories for a certain speaker over a prolonged period of time. Whether this is indeed possible will be investigated here.

The current study thus focuses on two research questions. Firstly, we wanted to see whether adaptation to foreign-accented speech is possible when listeners receive very short and limited exposure, and are not asked to pay attention to the accent. The second question was about the stability of this effect: if participants can adapt to the speaker, is this effect still observable after a day, or even after one week?

Dutch listeners were tested on a natural accent that was unfamiliar to them, namely Hebrew-accented Dutch. One of the differences between Hebrew and Dutch is the richness of the vowel system: whereas Dutch has 13 monophthongs and 3 diphthongs (Gussenhoven, 1999), Hebrew has only five monophthongs (/i,e,a,o,u/; Laufer, 1999). Moreover, Hebrew does not distinguish vowel length phonemically (e.g., Aronson, Rosenhouse, Rosenhouse, & Podoshin, 1996; Laufer, 1999; Most, Amir, & Tobin, 2000), in contrast to Dutch. Hebrew speakers can thus be expected to make durational errors when learning an L2 that does have vowel length differences (for comparable vowel length difficulties for L2 speakers see for example Flege, Bohn, & Jang, 1997). In the current experiment, the Hebrew speaker shortened words with [i] to [ɪ] (variant words; for example Dutch *statief* /sta:ti:f/, 'tripod' shortened to */statɪf/). These words were contrasted with words that naturally contained [ɪ] (canonical words; e.g., Dutch *verstrikt* /vərstrɪkt/, 'entangled'). The words with shortened vowels never created new words in Dutch. This contrast was chosen because vowel shortening in foreign accents is less frequent (and therefore possibly more difficult) than vowel lengthening, which is typical of Italian and Spanish accents, for example (Weber, Di Betta, & McQueen, in prepation). Moreover, both the [i] and [ɪ] are possible vowels in Dutch.

Another reason for choosing Hebrew-accented Dutch was that the Hebrew-speaking population in The Netherlands is small, and therefore it is unlikely that Dutch listeners are already familiar with this accent. This was particularly important because Witteman, Weber, and McQueen (in press) showed that experience with an accent aids Dutch listeners in correctly recognizing foreign-accented Dutch words. Native Dutch listeners with either extensive prior experience with German-accented Dutch (defined as hearing the accent multiple times a week from different speakers) were compared to listeners with limited prior experience (defined as hearing German-accented Dutch less than once a week from no more than one speaker). Participants performed a cross-modal priming task with strongly-, medium-, and weakly-accented words. Listeners with extensive experience with German-accented Dutch recognized all word types correctly, whereas participants with limited prior experience showed adaptation for the weakly- and medium-accented words, but not for the strongly-accented words. However, after a short additional exposure phase (a story from the same speaker with 12 new exemplars of strongly-accented words) immediately before the cross-modal priming test, participants with limited prior experience were able to recognize the strongly-accented words correctly. It is possible that this quick adaptation was present because even the listeners with limited prior experience with German-accented Dutch do have some knowledge of this accent. Therefore, the current study uses an accent listeners are very unlikely to be familiar with prior to the experiment.

In Experiment 1, we wanted to establish a baseline for understanding of Hebrew-accented Dutch without any prior experience or additional exposure. To this end, Dutch listeners completed a cross-modal priming task in which they first listened to Hebrew-accented primes, and then made lexical decisions to printed Dutch words and nonwords. Priming effects were calculated by subtracting the Reaction Times (RTs) to targets which had been preceded by unrelated primes from these to targets preceded by identical primes. RTs to target words are known to be faster in the identical compared to the unrelated condition (see e.g., McQueen, Cutler, et al., 2006; Marslen-Wilson, et al., 1995). Moreover, this facilitatory effect is usually not observed when the prime and target differ by as little as one phoneme (Marslen-Wilson & Zwitserlood, 1989), and sometimes these cases even result in inhibition (e.g., Van Alphen & McQueen, 2006). In the present study, we equate successful recognition of accented words with statistically significant facilitatory priming (see also Marslen-Wilson & Zwitserlood, 1989; Marslen-Wilson, et al., 1996). Priming effects would thus be taken as evidence that listeners correctly identified the accented primes as the intended Dutch words, and hence that they had adapted to the Hebrew-accented speaker. On the basis of the findings of Witteman et al. (in press), we expected listeners to be able to recognize the canonical words successfully. After all, they did not contain segmental substitutions and therefore did not deviate much from how native speakers of Dutch would pronounce them. But we also expected listeners not to be able to adapt to the variant pronunciations, because these words contained segmental mismatches in an accent unfamiliar to the listeners.

The current study was designed to shed more light on whether attention to the accent is necessary to adapt quickly to the accent, and whether adaptation to the accent is present in a delayed test phase. To avoid retesting effects, we designed a short phoneme monitoring task which served as the initial exposure to the accent. In Experiment 2, we tested whether adaptation to the accent could be observed after 24 hours. In Experiment 3, we extended this delay to one week. However, before we could look at these effects, we needed to establish a baseline for adaptation to the accent without exposure or delay. This was done in Experiment 1.

Experiment 1: No prior exposure

Method

Participants

We tested 28 native speakers of Dutch (24 females, *M* age 21.3 years). These participants were recruited from the Max Plank Institute participant pool; the vast majority studied at the Radboud University Nijmegen. All participants volunteered and were paid a small fee for participating. None reported a hearing disorder or language problem, and all had normal or corrected-to-normal vision. The language history questionnaire showed that none of the participants had any knowledge of Hebrew. In addition, though all participants identified the speaker as non-native, none guessed the native language of the speaker correctly when asked after the experiment.

Materials

The cross-modal priming experiment contained 242 trials (50 experimental, 192 fillers). A trial always consisted of an auditory prime followed by a visual target. In 20 of the experimental trials, the target was a Dutch word with half-long /i:/ in its canonical from (e.g., /sta:ti:f/ *statief*, 'tripod' shortened to */statɪf/); in the remaining experimental trials the target was a Dutch word with short /ɪ/ in its canonical form (e.g., /vərstrɪkt/ *verstrikt*, 'entangled'). This difference in number of experimental items across conditions arose because we based our materials on an earlier study (Weber, et al., in prepation) and needed to keep some items from that study for the phoneme monitoring exposure (see Experiments 2 and 3). However, as few as six items can be enough to show reliable priming effects (e.g., Witteman, et al., in press). Experimental targets were always paired with an identical and an unrelated prime. For targets with half-long /i:/, the vowel in identical primes was shortened to /ɪ/ (e.g. */sta:tɪf/ - variant pronunciation), and for those with short /ɪ/, identical primes were produced canonically with short /ɪ/ (e.g., /vərstrɪkt/ - canonical pronunciation). All experimental primes and targets are listed in Appendix A.

The remaining 192 trials were fillers. Of these, 64 were combinations of a word prime and an unrelated nonword target. Another 64 filler trials also had a word prime and a nonword target, but target and prime only differed in one vowel (e.g., prime *ladder*, 'ladder', followed by target LUDDER). Sixteen filler trials were made up of identical word primes and word targets, none of which contained the critical /i:/ or /ɪ/ vowels. Another 32 filler trials consisted of a word prime and an unrelated word target, and the final 16 trials consisted of prime and target word pairs that differed in one vowel (e.g., prime *bol,* 'sphere', followed by target BEL, 'bell'). In total, the experiment had 96 word targets and 96 nonword targets. The ratio of "yes" and "no" responses was therefore 1:1 for errorless participants.

Two counter-balanced lists were created such that every experimental target occurred once in a given list, either in combination with an identical prime or with an unrelated prime. Experimental trials were presented together with filler trials in a list with each experimental trial being preceded and followed by at least one filler trial and the first two trials in a list were always fillers. Due to an error in the list creation for Experiments 1 and 2, the two lists of the experiment were tested consecutively rather than alternated. The factor *list* was therefore added to the analyses as a control variable.

Stimulus recording and acoustic measurements

The speaker was a female native speaker of Hebrew. She was born in Israel and Dutch was her second non-native language, after English. At the time of recording, she had been living in The Netherlands for 13 years and was quite fluent in Dutch but still had a noticeable accent in her pronunciation. She tended to shorten longer vowels rather than lengthen short vowels when we informally analyzed a test recording of her reading a short Dutch text; whether this is indeed representative of Hebrew-accented Dutch in general or just for the speaker cannot be said, as to our knowledge no corpus analysis of Hebrew-accented Dutch exists. It is important to note, however, that Dutch listeners are in general not familiar with Hebrew-accented Dutch. The Hebrew-speaking population in The Netherlands is small. Though there are no official statistics on the number of Hebrew speakers in The Netherlands, in 2011 there were 8367 Israeli citizens in The Netherlands out of more than 1.5 million non-Western immigrants (Centraal Bureau voor de Statistiek [Central Bureau for Statistics], 2011). So, it is unlikely that Dutch listeners would be able to tell whether vowel shortening is typical for Hebrew-accented Dutch in general or not.

The Dutch primes were recorded one by one, separated by a pause, in a clear citation style, recording each word at least two times. In order to have consistent pronunciations for primes with vowel shortening, the speaker was instructed to produce /ɪ/ in words with canonical /i:/. All words were checked for other obvious segmental mismatches by a native speaker of Dutch, and re-recorded if necessary. The recordings were made in a sound-attenuated booth with a Sennheiser microphone and were stored directly onto a computer at a sample rate of 44 kHz. Words were excised from the recording using the speech editor Praat (Boersma & Weenink, 2009), and the best tokens were selected by a native speaker of Dutch.

Table 1 displays the values for the first three formants (measured at the midpoint of the vowel), separately for the canonical and the variant primes, as well as reference values for Dutch vowels [ɪ] and [i] for female speakers of Southern Standard Dutch (taken from Adank, et al., 2004). Southern Standard Dutch was chosen because the speaker lived in an area where this is the dominant dialect, and most participants were recruited in that area too.

Note that the current comparison between canonical and variant primes is between items, which is why observed differences can always be item-specific differences rather than a true difference between variant types. The duration of the vowels did not differ significantly for the two item types (*M* duration for canonical words was 234 ms (*SD* = 109), *M* duration for variant words was 271 ms (*SD* = 123); $t(48) = 1.122$, $p = .267$). When looking at the formants, variant [ɪ] had a higher first formant and a lower second formant than the canonical [ɪ], while the values for the third formant are similar. Importantly however, the values for our speaker's canonical and variant forms are closer to Dutch [ɪ] than Dutch [i].

Table 1. Average formant values (in Hz) at Vowel Midpoints for Canonical and Variant words, Compared to Average Values for Female Speakers of Southern Standard Dutch (taken from Adank, et al., 2004).

	Canonical		Variant		Dutch [ɪ]	Dutch [i]
	Mean	*SD*	Mean	*SD*	Mean	Mean
F1	407	41	453	34	455	317
F2	2288	208	1954	136	2115	2647
F3	2842	161	2856	119	2948	3312

Procedure

Participants were seated in a sound-attenuated booth and informed that they would first hear a Dutch word and then see a Dutch word or nonword on the screen; their task was to decide as quickly and accurately as possible whether the word presented on the screen was an existing Dutch word or not. They responded by pushing one of two buttons on a button box in front of them. Yes-responses were always made with the dominant hand, and RTs were measured from visual target onset. Participants were not told that the speaker was a non-native speaker of Dutch.

Auditory primes were presented binaurally over closed headphones at a comfortable listening level. Participants saw the visual targets on a computer screen situated about 50 cm in front of them. Visual targets were presented in white lowercase 24p Tahoma letters on a black background, 500 ms after the acoustic offset of the auditory primes. The visual targets stayed on the screen for 2000 ms, after which the next trial started. The experiment was created in Presentation (version 13, Neurobehavioural Systems Inc.) and controlled with NESU hardware (Nijmegen Experiment Set-Up). After the cross-modal priming experiment, participants were asked to fill out a language history questionnaire.

Results

One target item with particularly low lexical frequency (see Appendix A) was excluded from the analysis of this and subsequent experiments because of a high error rate (more than 25%).

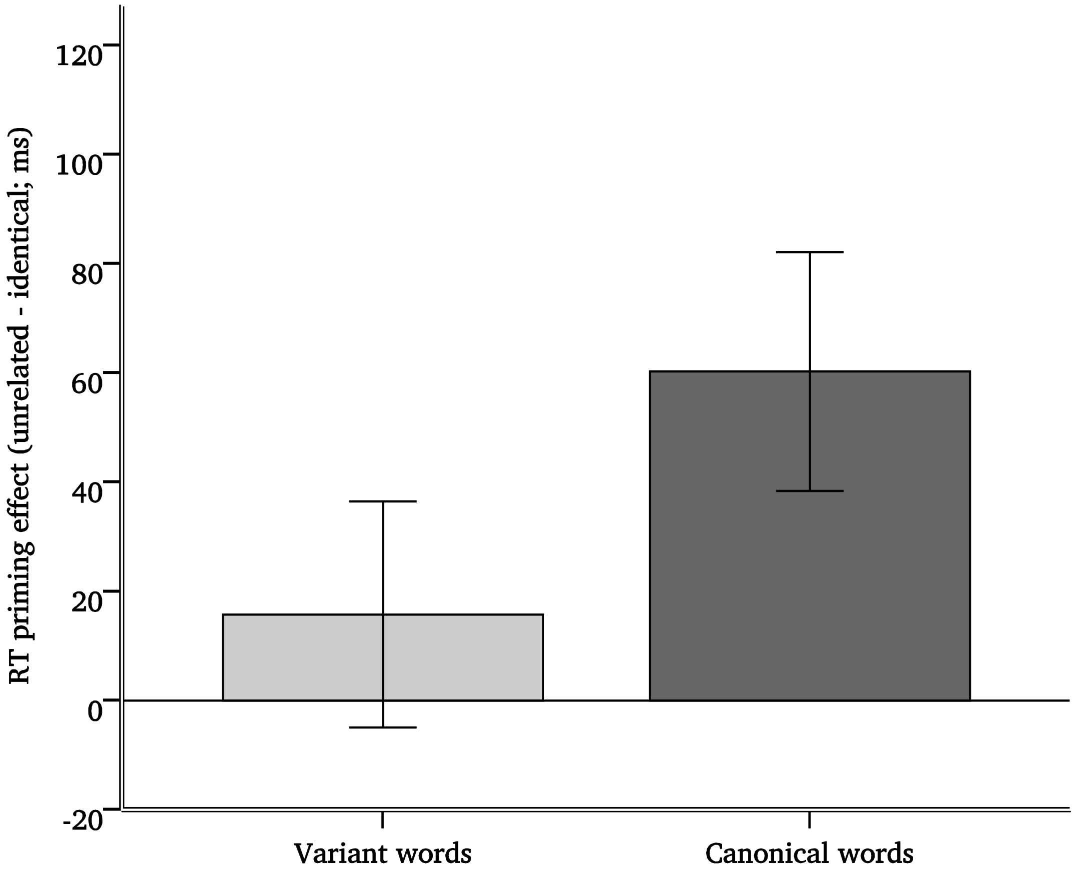
The remaining cross-modal priming data was analyzed with General Linear Model (GLM) Repeated Measures Analyses of Variance (ANOVAs) using a 2 (*accent type* - canonical, variant) x 2 (*priming* – identical, unrelated) design. Both factors were within-participants. *List* was added as a between-participant factor. The results were analyzed separately by participants (*F1*) and items (*F2*). The analysis with *list* only indicated that listeners were faster overall in one condition compared to the other, but there were no significant interactions. Therefore, it was not further analyzed nor described in the results below.

3.4% of trials were excluded due to errors or RTs that deviated more than 2.5 SDs from the condition's overall mean. Errors were distributed evenly across conditions and items, and were not analyzed statistically. Mean RTs and error rates for each condition and for all experiments can be found in Appendix B.

Figure 1 shows the calculated priming effects, i.e., the difference in RTs between responses following related primes and responses following unrelated primes. As can be seen in Figure 1, there was a main effect of *priming*: participants responded more quickly in identical than in unrelated trials ($F1$ (1,27) = 22.482, $p < .001$; $F2$ (1,19) = 5.958, $p = .026$). Furthermore, the participant analysis revealed a main effect of *accent type* (*F1*

(1,27) = 5.658, p = 0.025; *F2* < 1), indicating that responses were faster to canonical than to variant words. The main effect of priming was further qualified by an interaction between *priming* and *accent type* (*F1* (1,27) = 10.110, p = .004; *F2* (1,19) = 3.262, p = .087), indicating that the priming effects differed for the two accent types.

Figure 1: Experiment 1: Priming effects and SEs by accent type for no-exposure participants.



This interaction was investigated further using planned pair-wise comparisons (see Table 2). These showed that participants could not interpret the variant words correctly, but did show adaptation for the canonical words.

Table 2. Pairwise Comparisons of Priming Effects for all Accent Types Across Participants and Items.

	Participant analysis			Item analysis		
Item type	*df*	*t1*	*p*	*df*	*t2*	*p*
Variant words	1,27	1.526	.139	1,19	.793	.438
Canonical words	1,27	5.515	.000	1,28	3.496	.002

Discussion

Experiment 1 showed that without previous exposure to the Hebrew speaker of the experiment, Dutch participants were able to correctly interpret her canonical words such as *verstrikt*, but not her variant words like **statif*. With this baseline finding, we could ask in Experiments 2 and 3 whether a very short exposure phase would be enough to improve recognition of the variant words, and if so, whether this effect would be present without participants paying attention to the mispronunciation and with testing being delayed by one day or one week. Dutch listeners can adapt to German-accented words after having listened to the speaker of the experiment for a short while (Witteman, et al., in press). It was therefore plausible to assume that Dutch listeners could in principle do the same with Hebrew-accented Dutch. However, exposure in Witteman et al. consisted of listening to a read story in which sentential context information could have helped listeners to deduce the intended word form, and listeners were tested for comprehension immediately after exposure.

In Experiments 2 and 3 we wanted to shed more light on the roles of automaticity and stability in the adaptation to foreign-accented speech. To answer the first question, we investigated whether adaptation could be observed even when participants are not required to make conscious decisions about the investigated accent feature. We created a short phoneme monitoring task during exposure, in which participants were asked to detect the consonant /k/ in a list of words; thus, no particular focus was placed on vowel length, and lexical retrieval was not even necessary for the task. Because participants heard only isolated words during the phoneme monitoring task, they were not able to derive further information from the sentential context that a story would provide. That sentential context could provide valuable information on how the variant words need to be interpreted. Moreover, this phoneme monitoring exposure task contained only 20 items that were relevant for the subsequent cross-modal priming task (10 variant words, 10 canonical words). The exposure phase we used to test this automaticity thus was limited in two ways: the overall exposure time to the speaker was very short (3.5 minutes total) and since it only contained isolated words, participants were not able to rely on sentential context to gain more information about the speaker's general pronunciations.

The second question focused on the delay between exposure and test. In most studies using delayed testing, participants were exposed to the accent, tested on it, and after a delay, performed the same test again (e.g., Eisner & McQueen, 2006; Kraljic & Samuel, 2005). It is thus possible that the previous results are influenced by test-retest effects. The additional exposure of the second test may affect results, and listeners may respond differently when they perform a test another time. Therefore, in our experiment, participants received only exposure on the first day, and were tested only after the delay.

If adaptation to variant words is present and does not depend on sentential context to guide learning, then variant words in Experiment 2 should show facilitatory priming. Since canonical words already showed facilitatory priming without any additional exposure in Experiment 1, they should also show facilitatory priming in Experiment 2 with additional exposure.

Experiments 2 and 3 contained a delay phase of one day and one week, respectively. Though several studies have shown it is possible to adapt to foreign-accented speech in the short term (e.g., Bradlow & Bent, 2008; Clarke & Garrett, 2004; Witteman, et al., in press), none have looked whether these effects are also observable after a much longer time period. Research on word learning suggests that memory consolidation, the process of stabilizing a memory trace after initial acquisition (taking place during sleep), plays an important role in the successfulness of word learning (e.g., Davis, Di Betta, Macdonald, & Gaskell, 2009; Dumay & Gaskell, 2012; Tamminen, Payne, Stickgold, Wamsley, & Gaskell, 2010). Dumay and Gaskell (2012) taught participants new words and investigated whether these competed with existing words in the lexicon. Results from a pause detection and a word spotting task revealed no effects immediately after exposure, but a significant inhibition effect after a day and a week, i.e., after a night's sleep in which the knowledge was able to consolidate.

In line with results from memory consolidation, we expected to find that listeners' performance improves after exposure and consolidation. Specifically, we expect that listeners will show priming to the canonical words and the variant words after the delay.

Experiment 2: Exposure 1 day before test

Method

Participants

We tested 20 native speakers of Dutch (16 females, *M* age 22.4). All participants were volunteers recruited from the MPI participant pool; the vast majority studied at the Radboud University Nijmegen. They were paid a small fee for participating. None reported a hearing disorder or language problem, and all had normal or corrected-to-normal vision. The language history questionnaire revealed that none of the participants had any knowledge of Hebrew. Also, no participant guessed the native language of the speaker correctly, though all participants identified the speaker as non-native. All participants indicated that they thought both parts of the experiment were spoken by the same speaker.

Materials

The materials for the cross-modal priming test phase were identical to those used in Experiment 1. The phoneme monitoring exposure phase contained 70 mono- and bisyllabic Dutch words. The majority of these words were nouns

(59), the rest consisted of four adjectives, six verbs, and one number word. None of these words appeared in the main experiment.

All words were recorded in one session, together with the words of the cross-modal priming experiment. Ten of these words had a variant pronunciation, i.e. they contained the half-long vowel /i:/ which was shortened to /ɪ/ (e.g., Dutch /li:f/ ‘sweet’ was shortened to /lɪf/). These mispronunciations did not create other existing Dutch words (e.g., /lɪf/ is not a Dutch word). Another ten words had a canonical pronunciation, i.e. they contained the short vowel /ɪ/ which was produced in its canonical length (e.g., Dutch /fɪlm/ ‘movie’). The remaining 50 words contained no /i:/ or /ɪ/ or any other (half-)long vowel. Therefore, the only exposure participants received to words with (half-)long vowels in their canonical form was to forms with a shortened /i:/. The target phoneme for the phoneme monitoring experiment was /k/; it occurred, in varying word positions, in 28 of the 70 Dutch words (40%).

Design and procedure: phoneme monitoring

Participants were seated in a sound-attenuated booth and informed that they would hear one Dutch word at a time. Their task was to listen for the sound /k/ and press a button whenever they heard it. If a word did not contain the sound /k/, they did not have to press a button. Responses were always made with the dominant hand.

Participants could respond from the onset of the words, with a maximum response time of 2000 ms from word onset. During the time participants could respond, a '+' was shown on a computer screen. After the '+' disappeared, the next trial started. The experiment was created in Presentation (version 13, Neurobehavioural Systems Inc.) and controlled with NESU hardware (Nijmegen Experiment Set-Up). In total, the phoneme monitoring exposure lasted 3.5 minutes.

Design and procedure: cross-modal priming

Participants were asked to come in 24 hours after the phoneme monitoring exposure to take part in the cross-modal priming experiment and fill out a language history questionnaire. This part of the experiment was identical to Experiment 1.

Results

Phoneme monitoring exposure

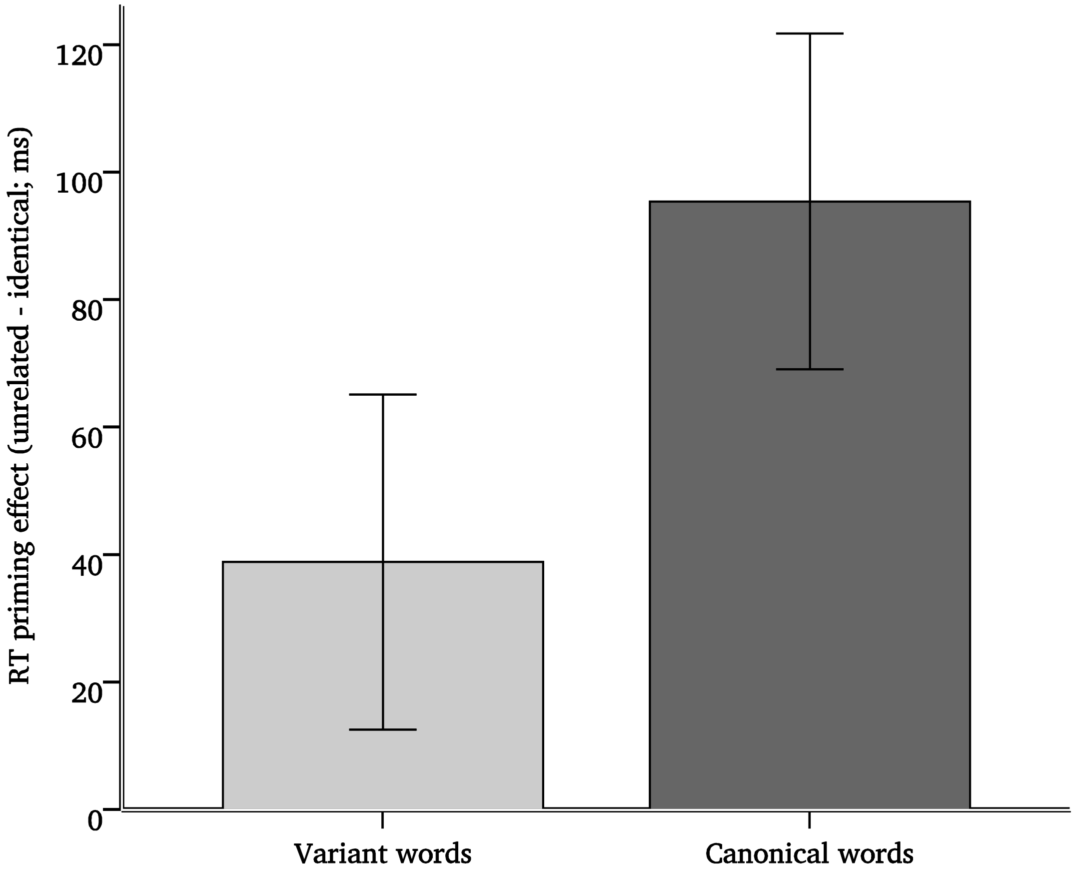
Accuracy for phoneme monitoring was very high: 98.6% correct, with 11 participants making no errors at all and nine participants making one error each. There was no systematic pattern in the errors. The average RT (measured from word onset) for the correct responses was 730 ms (SD = 381). One additional participant was tested but excluded from the analysis due to a low accuracy score (less than 95% correct).

Cross-modal priming test

We excluded 4.3% of trials due to errors or RTs that deviated more than 2.5 SDs from the condition's overall mean. There was no systematic pattern for the errors (distributed evenly across conditions and items, see Appendix B), so these were not analyzed statistically. Results were analyzed in the same way as in Experiment 1.

Calculated priming effects are shown in Figure 2, RTs and error rates per condition and accent type can be found in Appendix B. Participants were faster overall to respond to identical trials compared to unrelated trials ($F1$ $(1,19) = 71.640$, $p < .001$; $F2$ $(1,19) = 31.264$, $p < .001$).

Figure 2: Experiment 2: Priming effects and SEs by accent type for participants with Exposure one day before test.



Participants were equally fast to respond to both accent types (*F1* < 1; *F2* < *1*). There was, however, an interaction between *priming* and *accent type* (*F1* (1,19) = 7.269, *p* = .014; *F2* (1,19) = 7.573, *p* = .013), indicating that priming differed for the two word types. This was investigated further using planned pair-wise comparisons (see Table 3). These showed that, in contrast to Experiment 1, participants were able to interpret both the variant and the canonical words, but as the interaction indicated, priming was larger for canonical forms than for variant forms.

Table 3. Pairwise Comparisons of Priming Effects for all Accent Types Across Participants and Items.

	Participant analysis			Item analysis		
Item type	*df*	*t1*	*p*	*df*	*t2*	*p*
Variant words	1,19	2.961	.008	1,19	2.420	.026
Canonical words	1,19	7.256	.000	1,28	6.894	.000

Discussion

In contrast to Experiment 1, Experiment 2 showed significant priming effects for both the variant words and the canonical words. The priming effects for the variant words indicate that adaptation to the accent is very quick (phoneme monitoring exposure was only 3.5 minutes and contained only 10 tokens of the variant words), can take place when people are not instructed to pay attention to the accent specifically, and that this adaptation is present after at least 24 hours.

In Experiment 3 we wanted to see whether this long-lasting adaptation effect would remain stable over an even longer period of time. Therefore, the delay between the exposure and the test phase was extended to one week. We expected that even after a week's delay, listeners would still be able to interpret both the canonical and the variant words correctly. If the adjustment is to be beneficial for word recognition in foreign-accented speech, it should be stable over time.

Experiment 3: Exposure 1 week before test

Method

Participants

20 native Dutch participants completed Experiment 3 (18 females, M age = 22.1). Participants were recruited from the MPI subject pool and were paid a small fee in return for their participation. None of the participants reported a hearing disorder or language problem. All had normal or corrected-to-normal vision. Participants did not report any knowledge of Hebrew and did not guess the native language of the speaker correctly, but did identify the speaker as non-native.

Procedure

The experimental setup was identical to the one described in Experiment 2, with the only exception being that the delay between the phoneme monitoring exposure and cross-modal priming test was now one week.

Results

Phoneme monitoring exposure

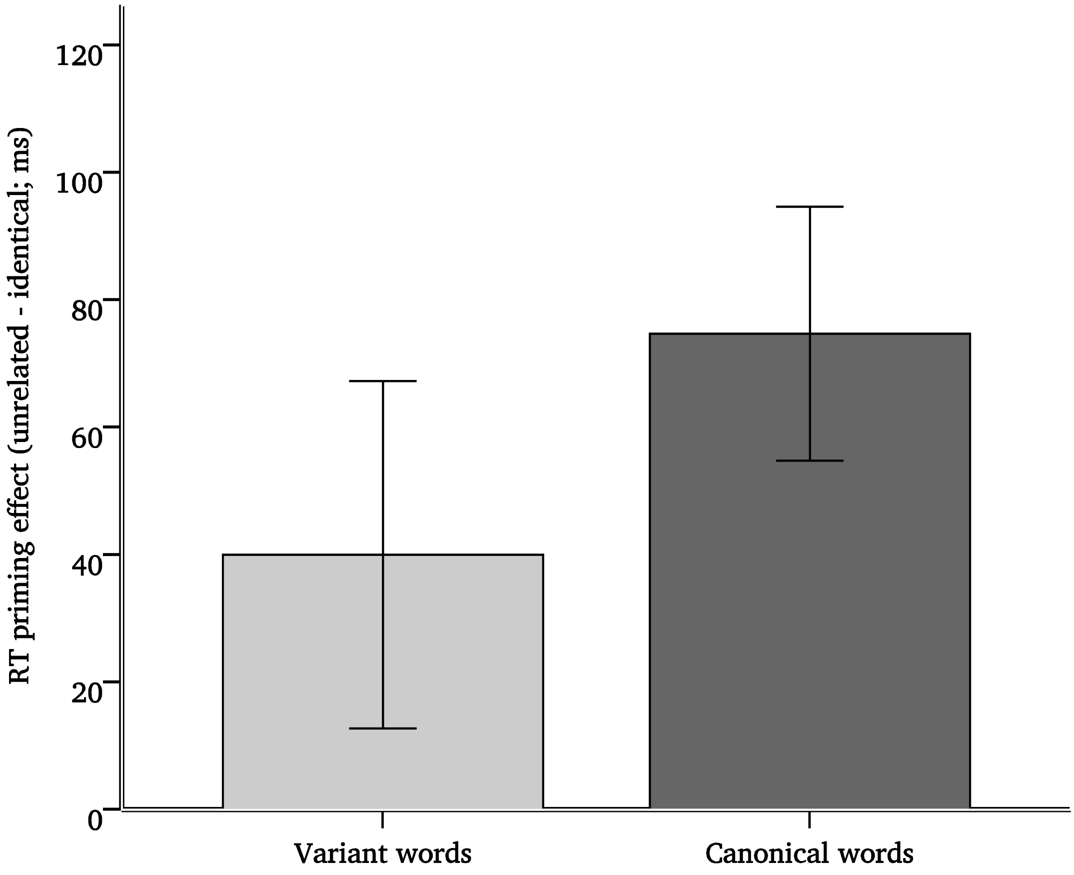
Accuracy for phoneme monitoring was very high: 99.0% correct, comparable to Experiment 2. Ten participants were errorless, five participants made one mistake, and another five participants made two mistakes each. The errors did not reveal a systematic pattern. The average RTs (measured from word onset) for the correct responses were 861 ms (SD = 270 ms).

Cross-modal priming test

We excluded 3.8% of trials due to errors or RTs that deviated more than 2.5 SDs from the condition's overall mean. Error rates were distributed evenly across conditions and items (see Appendix B). The data was analyzed in the same way as Experiments 1 and 2.

The calculated priming effects are pictured in Figure 3; mean RTs and error rates per condition are described in Appendix B. Participants were faster to respond to the identical trials compared to the unrelated trials ($F1$ $(1,19) = 37.582$, $p < .001$; $F2$ $(1,19) = 19.667$, $p < .001$). There was no main effect of *accent type* ($F1 < 1$; $F2 < 1$), indicating that participants overall RTs to these two item types did not differ. There was an interaction between *priming* and *accent type* ($F1$ $(1,19) = 5.482$, $p = .030$; $F2$ $(1,19) = 3.042$, $p = .097$), indicating that priming effects were larger for the canonical words compared to the variant words.

Figure 2: Experiment 3: Priming effects and SEs by accent type for participants with Exposure 1 week before test.



The planned comparisons (Table 4) revealed that participants showed significant priming for both the variant and the canonical words.

Table 4. Pairwise Comparisons of Priming Effects for all Accent Types Across Participants and Items.

	Participant analysis			Item analysis		
Item type	*df*	*t1*	*p*	*df*	*t2*	*p*
Variant words	1,19	2.938	.008	1,19	2.256	.036
Canonical words	1,19	7.495	.000	1,28	7.738	.000

Discussion

In Experiment 3, we found that participants could interpret both the variant words and the canonical words, thereby replicating the result we found in Experiment 2. We thus showed that the adaptation the accented speaker remains stable for at least one week, even when the initial exposure to the speaker has been very limited.

General Discussion

The present study investigated whether adaptation to foreign-accented speech is automatic, and whether it is stable over time. Native Dutch listeners performed a cross-modal priming task in which they showed adaptation to canonical Hebrew-accented words (without specific mispronunciations), but not to variant words in which Dutch [i] was shortened to [ɪ] (Experiment 1). However, after a 3.5 minute exposure task performed 24 hours before the same cross-modal priming test, native Dutch listeners were able to interpret both word types correctly (Experiment 2). This effect remained stable even when the delay between exposure and test was extended to one week (Experiment 3). These findings indicate that adaptation to foreign-accented speech is not only quick, but also automatic and long-lasting.

Foreign-accented speech without substantial segmental mismatches does not seem to interfere with understanding, which is good news for L2 listeners and speakers. Apparently the perceptual system is flexible enough to deal with smaller deviations almost instantly. When there are segmental substitutions (like in the variant items), a short exposure phase can be enough to adapt to these words, which is in line with previous research (Witteman, et al., in press). The fact that listeners can interpret variant words after a short phoneme monitoring exposure also indicates that their knowledge of the accent can be transferred across different tasks and does not require conscious attention to the mispronunciations typical for the accent. Previous experiments (e.g., Eisner & McQueen, 2006; Kraljic & Samuel, 2005) also made use of different tasks between exposure and test, but repeated the same test. Not only did participants receive more exposure overall, they were also already trained on and familiarized with the test paradigm already. However, these factors did not seem to affect the results, as the present study indicates that even without a paradigm with repeated testing, listeners are able to adapt to foreign-accented speech after a delay.

Several perceptual learning studies have made use of a story during exposure (e.g., Eisner & McQueen, 2006; Maye, et al., 2008; Witteman, et al., in press), which differs in a number of ways from phoneme monitoring. First, a story provides a rich sentential context from which a listener can gain much more information about the speaker's accent than from isolated words. Moreover, if a foreign-accented word is difficult to understand, a sentence context makes it much easier to decide what word was uttered. Second, phoneme monitoring and listening to a story differ in the attention spent on the task. Because the only task during the story was to listen to it, participants were able to attend to all aspects of the accent. In the phoneme monitoring exposure, however, participants were asked to pay attention to a different phoneme, one not specific to the accent. Even after this type of exposure, however, listeners' performance on the variant words improved. This indicates that attention to specific mispronunciations is not required in order to adapt to foreign-accented speech. This is more good news for listeners of foreign-accented speech: just listening to a foreign-accented speaker is enough for listeners to adapt, so it is not necessary to actively think about the mispronunciations of the speaker. This leaves a listener free to focus on the message a speaker wants to convey.

Adaptation to foreign-accented speech seems to be automatic. This is in line with perceptual learning results in the L1 domain (McQueen, Norris, et al., 2006). However, it might be the case that automatic perceptual learning effects hold only when the contrast listeners have to learn is present in their phoneme inventory, as perception of contrasts that do not exist in one's native language (like the /r/-/l/ contrast for Japanese speakers) often requires prolonged and explicit training to be mastered (e.g., Bradlow, Akahane-Yamada, Pisoni, & Tohkura, 1999; Logan, Lively, & Pisoni, 1991). However, when listeners need to map a non-native sound onto an existing category (e.g., Sjerps & McQueen, 2010; Witteman, et al., in press), or retune their existing category boundaries, like in the present study, adaptation is possible even after a short time of exposure. In a way, it is even more surprising that listeners can adapt to foreign-accented speech when the accented sound maps onto an existing other sound in the lexicon, rather than mapping a non-native sound onto a native sound. In the latter case, all listeners need to do is extend one category, whereas in the current experiment, listeners need to extend one category (the category of the variants), while keeping the category of the canonical vowels the same. Listeners are able to do this, as evidenced by the by the priming effects for the canonical words. If listeners would apply a shift in category boundaries not just for the variant, but also for the canonical words, interpretation difficulties for the canonical words would arise. However, the interpretation of these words does not seem to be affected, even after exposure. This can be taken as evidence for finely nuanced adaptation mechanisms.

We also found evidence that adaptation to foreign-accented speech is long-lasting, as it can be observed after one day (Experiment 2) or one week (Experiment 3). Because listeners were exposed to all kinds of speech outside the lab between exposure and test, but still managed to show adaptation during the exposure phase, it is likely that they adapted to this specific speaker, rather than broadened their categories in a speaker-general fashion. If this type of adaptation is indeed speaker-specific, this could explain how the perceptual system is able to find a balance between flexibility on the one hand, and stability on the other. How long-lasting these effects are exactly is not within the scope of the current experiment, but it is likely that they will not last infinitely. Moreover, it is possible that listeners in Experiments 2 and 3 were quicker to adapt because they knew they were scheduled to come back to the same laboratory, even though they were never informed of the details of the test task until right before they performed it. Whether these adaptation effects could be generalized to other speakers of the same accent remains a question for further research.

The present study also provides valuable information for models of spoken word recognition. There are two main theories about how listeners are able to deal with deviations in the speech signal: representational and processing theories. Representational models assume that the lexicon has entries for every single word, as well as for every variation on these words (e.g., Goldinger, 1998). Episodic representational models also state that all variation is encoded in the lexicon, but then in terms of fine-grained phonetic detail (e.g., Johnson, 2006; Pierrehumbert, 2001). Finally, it has been proposed that the lexicon contains multiple abstract representations for variant forms (Ranbom & Connine, 2007).

One way in which representational accounts can explain how listeners can deal with the added variation of non-standard speech (e.g., foreign-accented speech) is that listeners store all variation in their lexicon. Upon hearing these variants a second time, adaptation could be achieved by re-accessing these forms. For the present experiment, this would mean that listeners could only adapt to accented speech if they have heard these specific variants before, that is, had some experience with the accent. Experience with the accent is unlikely in the current experiment: Hebrew-accented Dutch is not common in The Netherlands, and none of the participants indicated that they heard the accent before. It is of course possible that listeners had heard this accent (or a similar one) previously, but it is improbable that they would already have stored representations for (most of the) experimental words. Moreover, even if some listeners heard the accent before, we expect the level of experience to be the same across experiments. This premise is somewhat difficult to combine with the findings presented here. In Experiment 1, listeners did not show priming to variant words, whereas in Experiments 2 and 3, listeners did. The exposure phase contained variant words, but these tokens were different from the ones presented during test. A representational account based on earlier exposure (and subsequent storage) of the variant words would have great difficulty explaining how listeners in Experiments 2 and 3 were able to adapt to the variant words, while listeners in Experiment 1 could not do so.

A processing account might provide a better explanation for the data at hand. One possible account might assume that the lexicon contains only the canonical representations of the critical words examined here, and that variation due to foreign-accented speech is resolved at a pre-lexical level. Listeners learn from exposure with an accent how variations should be mapped on the (stored) canonical form (Lotto & Holt, 2006; Mitterer, et al., 2006; Gaskell & Marslen-Wilson, 1998). Thus, information on how the speaker of this experiment pronounces words (or specifically for this experiment: shortens a vowel), can be carried over from the exposure to the test phase. Perceptual learning studies in L1 have already proposed that if a sound is mapped onto the lexicon, at a pre-lexical level of processing, this learning can be generalized to all words in the lexicon that contain that sound (McQueen, Cutler, et al., 2006; Sjerps & McQueen, 2010). This could explain why listeners are able to adapt to the variant words after they have been exposed to the accent, but not before. Moreover, this would also mean that adaptation to foreign-accented speech would use the same mechanisms as adaptation to artificial and within-language variation.

In sum, the perceptual system is highly flexible: listeners do not show any sign of difficulty when listening to foreign-accented speech without clear segmental substitutions. Foreign-accented speech with segmental substitutions can provide some initial problems, but a short exposure phase is enough for listeners to adapt even to these words. Moreover, though being exposed to foreign-accented speech is enough to improve performance, it is not necessary for listeners to pay specific attention to the mispronunciations at hand. Finally, once listeners have adapted to a foreign-accented speaker, this effect remains stable for at least a week.

Appendix A - List of Experimental Items

Variant words

Target (Dutch spelling)	*Accented prime (IPA)*	*Unrelated prime*	*Target translation*	*Unrelated prime translation*
advies	ɑtfɪs	fornuis	advice	stove
bier	bɪr	huid	beer	skin
brief	brɪf	kers	letter	cherry
dier	dɪr	kroeg	animal	bar
kies	kɪs	homp	molar	chunk
kiezen	kɪzə	emmer	choose	bucket
lief	lɪf	spuit	sweet	syringe
mier	mɪr	wraak	ant	revenge
mythe	mɪtə	afwas	myth	dishes
niezen	nɪzə	erker	sneeze	oriel
papier	papɪr	talent	paper	talent
plezier	pləzɪr	abdij	fun	abbey
riem	rɪm	kaft	belt	cover
servies	sɛrvɪs	kameel	service	camel
sierlijk	sɪrlək	plaksel	graceful	adhesive
statief	statɪf	boeket	tripod	bouquet
textiel	tɛkstɪl	roman	textile	novel
ventiel	vɛntɪl	tapijt	valve	carpet
vierkant	vɪrkɑnt	bloesem	square	blossom
vriezen	vrɪzə	foto	freeze	photo

Canonical words

Target (Dutch spelling)	*Accented prime (IPA)*	*Unrelated prime*	*Target translation*	*Unrelated prime translation*
bacil	basıl	soldaat	bacillus	soldier
bassist	bɑssıst	factuur	bass player	invoice
blik	blık	staan	stand	can
bliksem	blıksəm	eikel	acorn	lightning
blind	blınt	rouw	mourning	blind
cirkel	sırkəl	hertog	duke	circle
conflict	kɔnflıkt	sopraan	conflict	soprano
cursist	kʏrsıst	tabak	student	tobacco
delict	dəlıkt	buffet	offence	buffet
dicht	dıxt	jas	coat	closed
ding	dıŋ	kruis	thing	cross
gravin	xravın	dolfijn	countess	dolphin
hitte	hıtə	mosterd	mustard	heat
inkt	ıŋkt	luik	hatch	ink
klimmen	klımə	omroep	broadcasting	climb
*klip	klıp	hand	cliff	hand
knikker	knıkər	wortel	carrot	marble
kokkin	kɔkın	natuur	(female) chef	nature
lift	lıft	markt	elevator	market
mist	mıst	koorts	fog	fever
pinda	pında	rooster	schedule	peanut
prins	prıns	storm	prince	storm
pupil	py:pıl	framboos	pupil	raspberry
rimpel	rımpəl	stempel	stamp	wrinkle
rits	rıts	herfst	fall	zipper
stilte	stıltə	kaneel	cinnamon	silence
toerist	turıst	lakei	tourist	lackey
verstrikt	vərstrıkt	bestuur	entangled	government
vinger	vıŋər	koepel	dome	finger
winter	wıntə	keuren	examine	winter

*Discarded from all analyses

Appendix B – Overview of RTs and Error Rates

Overview of RTs (SDs) and Error Rates for all Conditions and Experiments

		RTs (*SD*) CMP		Errors (percentage)	
Delay	Item type	Identical	Unrelated	Identical	Unrelated
No delay	Variant	614 (93)	629 (79)	2 (0.7)	4 (1.4)
No delay	Canonical	578 (92)	638 (77)	4 (1.0)	10 (2.5)
24 hours	Variant	582 (90)	621 (110)	2 (1.0)	5 (2.5)
24 hours	Canonical	554 (93)	649 (93)	11 (3.8)	7 (2.4)
1 week	Variant	584 (72)	624 (63)	4 (2.0)	3 (1.5)
1 week	Canonical	560 (69)	634 (63)	8 (2.8)	7 (2.4)

Tolerance for inconsistency in foreign-accented speech

Chapter 5

Witteman, M. J., Weber, A., and McQueen, J. M. (under revision). Tolerance for inconsistency in foreign-accented speech. *Psychonomic Bulletin and Review.*

Abstract

Adaptation to foreign-accented speech can be quick, but is this also the case when a speaker's accent is inconsistent? In this study we investigated whether listeners are able to adapt to a foreign-accented speaker who has an inconsistent accent. Two groups of native Dutch listeners participated in a cross-modal priming experiment, either in a consistent-accent condition (German-accented items only) or an inconsistent-accent condition (German-accented and native-like pronunciations intermixed). The experimental items were identical for both groups (words with vocalic substitutions characteristic of German-accented speech and words without categorical vocalic substitutions), while the fillers differed in accentedness (German-accented or native-like words). All items were spoken by the same speaker: a German native who could pass for a Dutch native speaker. Listeners in the consistent-accent group were able to adapt quickly to the speaker (i.e., showed facilitatory priming), both for items with and without vocalic substitutions. Listeners in the inconsistent-accent condition also recognized words without vocalic substitutions quickly throughout the experiment. Although these listeners did not show adaptation to words with vocalic substitutions in the first half of the experiment, they did in the second half. Together, these results show that adaptation to foreign-accented speech is rapid. Accent inconsistency does slow listeners down initially, but just a short period of additional exposure is enough for them to adapt to the speaker. Listeners can therefore tolerate inconsistency in foreign-accented speech.

Introduction

Foreign-accented speech deviates noticeably from native speech. Nevertheless, recent research has shown that listeners are able to deal with this variation remarkably well: just a few minutes of exposure, or a couple of sentences, can be enough to 'tune in' to a foreign-accented speaker (e.g., Bradlow & Bent, 2008; Clarke & Garrett, 2004; Witteman, et al., in press). The perceptual system thus seems flexible enough to handle the considerable variation that foreign accents introduce to the speech signal. The present paper investigates the boundaries of this adaptation. Specifically, it asks whether listeners are able to adapt to mispronounced words of a speaker with an inconsistent accent.

Listeners can adapt quickly to different kinds of variation in native speech (L1): for example, when a sound in a word is replaced by an ambiguous native sound (e.g., Norris, et al., 2003; McQueen, Norris, et al., 2006), by another native sound (e.g., Maye, et al., 2008), or even by a non-native sound (Sjerps & McQueen, 2010). This adaptation can be speaker-specific (e.g., Nygaard & Pisoni, 1998; Eisner & McQueen, 2005). Adaptation is seen by listeners' adjustments of their category boundaries in categorization tasks, or their correct interpretation of words with mispronunciations in lexical decision tasks. The perceptual system is thus highly flexible, but this flexibility must also be limited. While adapting to a speaker's lisp is useful for future encounters with that speaker, adapting to their drunken slurred speech is not. Somehow the perceptual system must find a balance between flexibility and stability.

One type of variation arises from differences in pronunciation due to a foreign accent. Pronunciation is one of the most difficult domains of a second language (L2) to master, and very few late L2 learners achieve native-like pronunciation (Flege, et al., 1995). Foreign-accented speech is driven by the speaker's native language, and segmental variation, for example, typically arises when a target language phoneme does not exist in the speaker's native language. In that case, foreign-accented speakers regularly substitute the L2 sound with a native language sound, as in Dutch learners of English saying *penda* for *panda*. Although only few L2 learners will ever pass as native speakers, there is variation in how consistent learners are in their segmental substitutions. In particular, pronunciations of highly proficient learners can often vary as a result of their attempts to approach canonical sounds (Hanulíková & Weber, 2012). Even though foreign-accentedness itself thus remains a speaker characteristic that can be detected easily, the speaker's exact mispronunciation may vary, making the signal inconsistent.

It is possible that when listeners identify a speaker as being non-native, they more readily adapt to that speaker than they would for a native speaker, simply because regular deviations from the target language can be anticipated for this group. Listeners can take their knowledge of a speaker's idiosyncrasies into account when adapting to them: they are more forgiving in accepting grammatical errors when the errors are made by an L2 speaker rather than an L1 speaker (Hanulíková, van Alphen, van Goch, & Weber, 2011), and they relax their vowel categories more readily for L2 speakers than for native speakers (Hay, Nolan, & Drager, 2006). But listeners also put boundaries on their adaptation depending on the nature of the mispronunciations: they generalize what they have learned about certain deviations if they are spoken by foreign-accented speakers but not by native speakers (Eisner, Melinger, & Weber, under revision). Moreover, when listeners learn that an L1 speaker's mispronunciations are incidental (e.g., uncharacteristic for the speaker), they do not show perceptual learning effects, whereas they do if this information is not provided (Kraljic, et al., 2008; Kraljic & Samuel, 2011). But it is not yet clear whether adaptation to foreign-accented speech depends on how consistent that accent is. Are listeners more or less inclined to adapt to a specific pronunciation variant depending on whether the L2 speaker's accent is a consistent trait or not?

In the present experiment, Dutch listeners were either exposed to Dutch experimental primes with a typical German-accented segmental substitution and German-accented fillers (consistent-accent condition), or by the same German-accented experimental primes and native-like fillers (inconsistent-accent condition). All listeners thus heard the same German-accented experimental items, but filler accentedness varied between groups. To control for speaker effects, all materials were spoken by a native speaker of German who could easily pass as a Dutch native speaker. All listeners had limited prior experience with German-accented Dutch, such that they could still improve in their recognition of the German accent (see Witteman, et al., in press).

In a cross-modal priming study, Dutch participants listened to primes and subsequently made lexical decisions on visually presented targets. Reaction Times (RTs) are known to be faster when the prime and target word are identical, compared to unrelated pairs (e.g., Marslen-Wilson, et al., 1995), and even small differences between the auditory prime and visual target will prevent significant facilitatory priming (e.g., Van Alphen & McQueen, 2006). Significant facilitatory priming effects will therefore be taken as our measure for successful online word recognition and hence adaptation to the variant forms (e.g., Marslen-Wilson, et al., 1996).

We expected that listeners in the consistent-accent condition would learn to adapt to the speaker's accent, either immediately or during the course of the experiment. Whether the listeners in the inconsistent-accent condition would adapt, however, was less clear from the previous literature. While listeners are quick to adapt to foreign-accented speech in general (e.g., Bradlow & Bent, 2008; Clarke & Garrett, 2004; Witteman, et al., in press), they are not known to do so in native speech when confronted with incidental (i.e., inconsistent) mispronunciations (e.g., Kraljic, et al., 2008; Kraljic & Samuel, 2011). Moreover, if listeners take information about the speaker's pronunciations into account, they might be less inclined to adapt to a native-like speaker, for whom these mispronunciations are more likely to be incidental than to a speaker who appears to mispronounce words consistently.

Thus, if listeners in the inconsistent-accent condition are not affected by the inconsistent information, we would expect to see no differences between the consistent- and inconsistent-accent conditions. If the conflicting information does interfere with adaptation, we would expect that listeners in the inconsistent-accent condition show slower adaptation compared to the consistent-accent condition, or maybe even no adaptation at all.

Method

Participants

We tested 48 native speakers of Dutch: half participated in the consistent-accent condition (22 females, *M* age 21.2 years), and half in the inconsistent-accent condition (21 females, *M* age 22.3 years). All participants were recruited through the subject database of Utrecht University and received a small fee for their participation. Participants reported normal hearing, normal or corrected-to-normal vision, and no language problems. None were fluent in German, and all said English was their most fluent non-native language. On the basis of a language history questionnaire, we selected the listeners who only had limited experience with German-accented Dutch, i.e. heard German-accented Dutch less than once a week (for a similar selection procedure, see Witteman et al., in press). Because of this requirement, 16 additional participants were excluded.

Materials

The cross-modal priming experiment contained 144 trials (48 experimental, 92 fillers; all Dutch words and nonwords). A trial consisted of an auditory prime followed by a visual target. Experimental targets were always paired with an identical and an unrelated prime (for an overview, see Appendix A). Half of the experimental targets contained the Dutch vowel /œy/ (e.g., /dœym/ [thumb]) which is commonly mispronounced as /ɔɪ/ by German speakers and serves as a strong marker of a German accent in Dutch (see Witteman, et al., in press). The remaining experimental targets served as control items and contained a range of vowels and consonants that are shared between German and Dutch (e.g., /dɛkkɪŋ/ [cover]), and were therefore not expected to contain obvious segmental substitutions in German-accented Dutch. Because of their lack of substitutions and thus their proximity to Dutch pronunciation, we expected these items to induce comparable priming effects in both listener groups.

The filler items contained a variety of segments that carry a noticeable accent when produced by German speakers. Of all filler primes, 25% contained /œy/, all with nonword targets, so that not every prime containing /œy/ would require a yes response. Another 37.5% of the primes was made up of words with different accent markers: for example, German-accented Dutch strongly aspirates word initial /p,k,t/, deletes word final schwas (e.g., Dutch /lopə/ pronounced as /lopn/), and replaces Dutch /ɛɪ/ with German /aɪ/ (Doeleman, 1998). The remaining filler primes also contained these segments, but in nonwords. Overall, half the trials had a word target, thus resulting in 50% yes-responses for errorless participants.

We created two versions of the experiment. The experimental items were identical in the two versions, but while one contained German-accented filler primes (consistent-accent condition), the other version had the same fillers with native-like pronunciations (inconsistent-accent condition). For both versions of the experiment we created two counterbalanced lists, so that every experimental item appeared once in a given list, either with an identical or unrelated prime. Lists were pseudo-randomized: experimental trials were always preceded and followed by at least one filler. Moreover, the first two trials of the experiment were always fillers.

Stimulus recording

The speaker was a male native speaker of German, who was extremely fluent in Dutch, and judged to sound like a native speaker of Dutch by Dutch native speakers. He grew up in *Nordrhein-Westfalen* and started learning Dutch at age twenty, when he moved to The Netherlands. At the time of recording, he had lived in The Netherlands for six years and spoke both Dutch and German on a daily basis.

Multiple tokens of the primes were recorded one by one, in clear citation style. All filler primes and the control items were first recorded in the speaker's natural (native-like) accent. After that, we recorded the experimental primes with /œy/ with the German-accented vowel substitution /ɔɪ/ (e.g., /dœym as /dɔɪm/), as well as the filler primes in an instructed German accent. For those items, a strongly-accented German speaker served as a role model. All recordings were done with unmarked Dutch orthography. The native-like recordings were checked by two native speakers of Dutch, who corrected the speaker only when a German accent could be heard in the recordings, which occurred very rarely. The German speaker confirmed that all German-accented items were pronounced as intended.

All recordings were done in a sound-attenuated booth using a Sennheiser microphone and stored directly onto a computer at a sample rate of 44 kHz. Subsequently, we excised the primes using Praat (Boersma & Weenink, 2009). The best tokens were selected by the first author (a Dutch native speaker).

Rating experiment

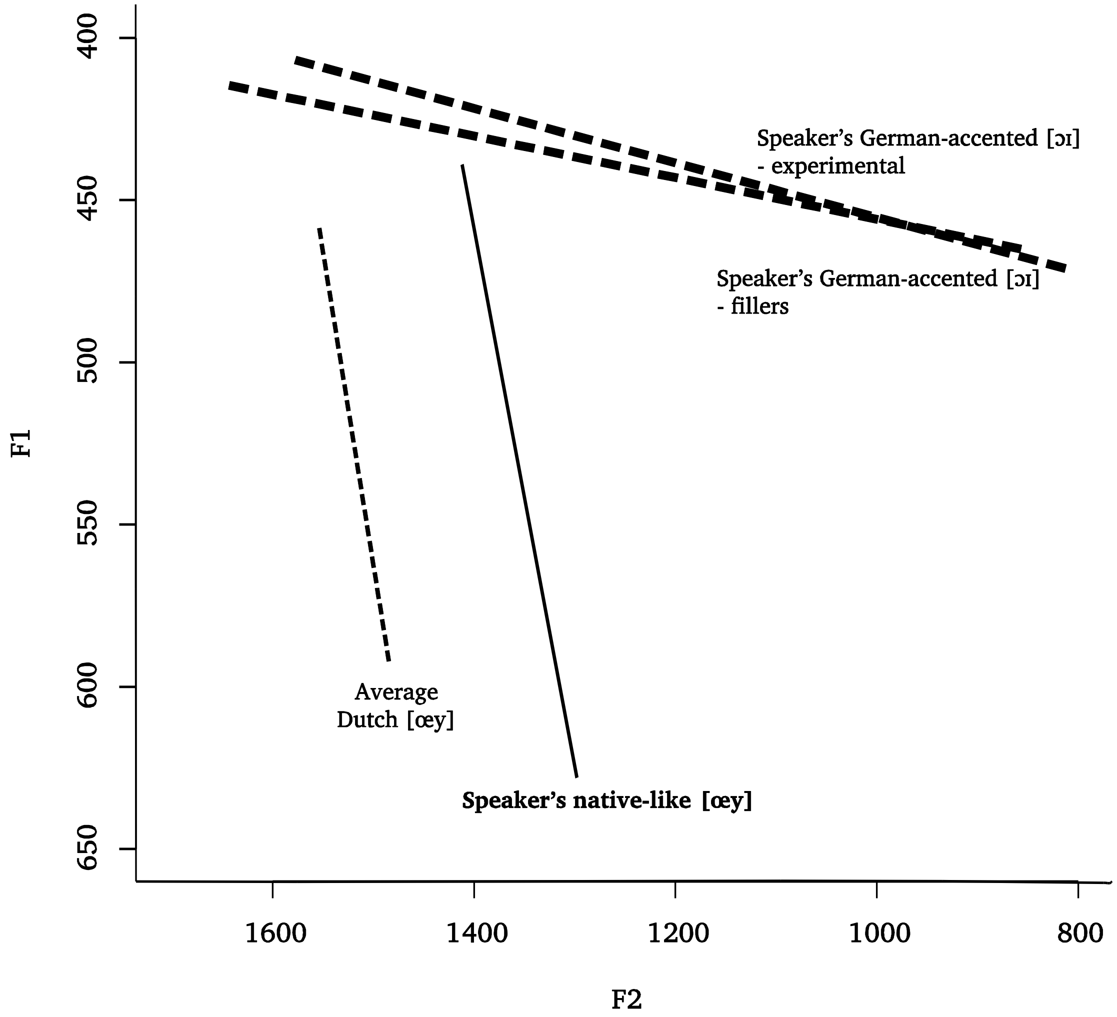
To make sure that the fillers used for the two conditions of this experiment indeed differed in terms of accentedness, ten native Dutch speakers who did not participate in the priming experiment rated the items. The rating experiment consisted of 108 words from the main experiment: 54 German-accented fillers and 54 native-like fillers. Furthermore, we added 36 Dutch words produced by an American speaker and 36 Dutch words produced by an Italian speaker, to add more variation to the materials. Listeners heard one word at a time, over closed headphones, immediately followed by a visual rating scale on which they indicated with the keyboard how accented the word was on a scale of 0 (no accent at all) to 9 (very strongly accented). The data were analyzed with paired-samples t-tests that indicated that the German-accented fillers ($M = 6.456$, $SD = 1.006$) were clearly rated as more accented than the native-like fillers ($M = 1.793$, $SD = .683$; $t(9) = 13.831$, $p < .001$). This pattern also held true when looking separately at fillers with

/œy/ (M = 7.350 vs. M = 1.708; $t(9) = 15.283$, $p < .001$) and at fillers without /œy/ (M = 6.200 vs. M = 1.817; $t(9) = 12.854$, $p < .001$).

Acoustic measurements

We also analyzed the speaker's [œy]-vowels by measuring 12 native-like filler primes, 12 German-accented filler primes and 24 German-accented experimental primes. We measured the first two formants at the 25 and 75 percent points of the vowels. In Figure 1, we plotted our speaker’s pronunciations contrasted with averaged values of Dutch [œy] taken from Adank, Van Hout, & Smits (2004).

Figure 1: Average F1 and F2 formant values of the speaker’s native-like [œy] (solid line) and German-accented [ɔɪ] in each type of item (think dashed lines), and average Dutch male [œy] pronunciations (thin dashed line).



Though our speaker’s F1 and F2 are lower overall than those of the average Dutch male speaker, the trajectory of the vowel is very similar. Importantly, the trajectory of the speaker’s native-like [œy] is very similar to average Dutch trajectories and differs substantially from his pronunciation of German [ɔɪ]. Finally, the speaker's German-accented vowels for the experimental and filler primes are almost identical.

Procedure

The experiment was administered in a sound-attenuated booth. Participants were informed they would hear a single speaker pronouncing Dutch words and nonwords and then see a Dutch word or nonword on a computer screen. Participants made lexical decisions to the visually-presented words as quickly and accurately as possible. Participants got a button box with two response buttons and always made yes responses with their dominant hand. RTs were measured from target onset. Participants received no information about the speaker's native language.

Auditory primes were presented binaurally over closed headphones at a comfortable listening level. Visual targets were presented in white lowercase 24p Tahoma letters on a black background, 500 ms after the acoustic offset of the auditory primes. The maximum response time for each target was 2000 ms, after which the next trial started. The experiment was created in Presentation (version 13, Neurobehavioural Systems Inc.) and controlled with NESU hardware (Nijmegen Experiment Set-Up). After the cross-modal priming experiment, participants filled out a language history questionnaire.

Results

Results were analyzed separately for the two versions of the experiment. For the participant analyses, we used a GLM Repeated Measures Analysis of Variance (ANOVA) design with a two (*accent type* – experimental, control) by two (*priming* – identical, unrelated) by two (*half* – first, second) design. All were within-participant factors. The item analyses were done with a GLM Univariate analysis, in which *accent type*, *priming* and *half* were between-item factors. We also conducted planned comparisons to look at the priming effects. These effects were examined with paired sample t-tests across participants, and independent sample t-tests across items, and both were split up by *half* and *accent type*. Mean RTs and error rates for both conditions are displayed in Table 1.

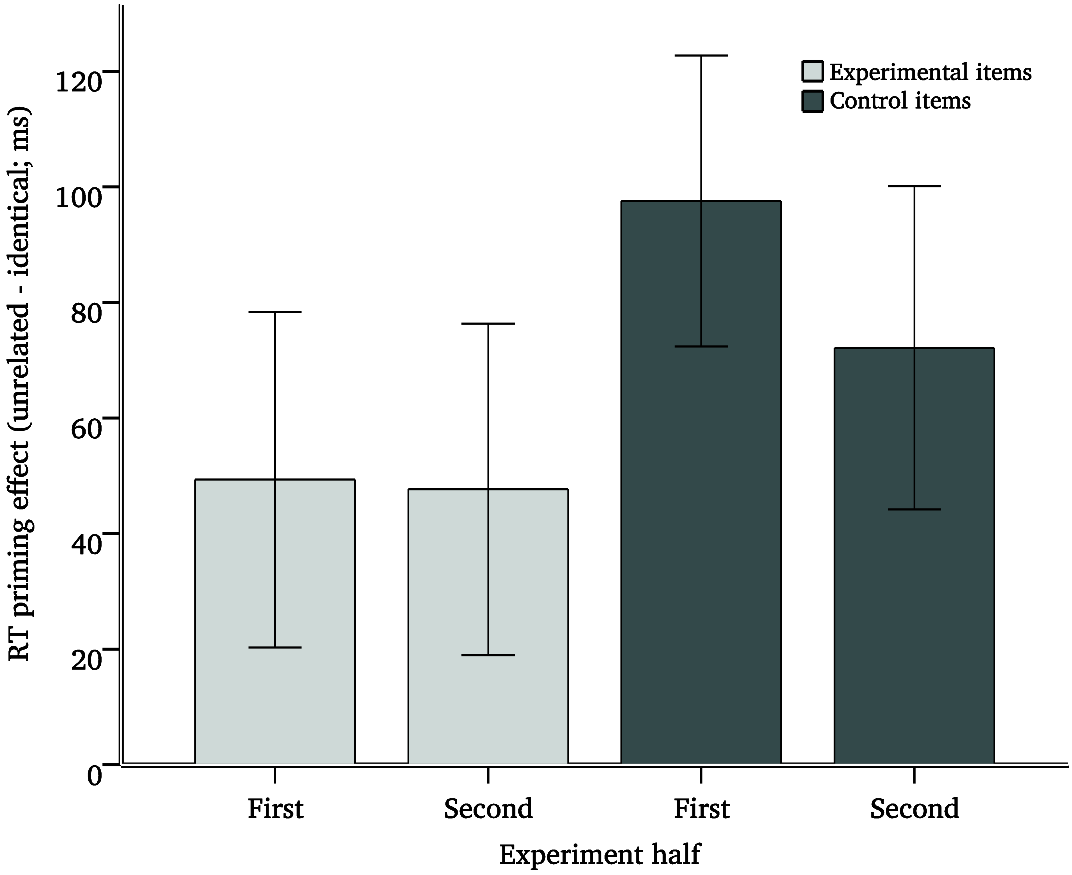
Table 1. Overview of RTs (SDs) and Error Rates for all Conditions and Experiments

		RTs (*SD*) CMP		Errors (percentage)	
Condition	Accent type	Identical	Unrelated	Identical	Unrelated
Consistent	Experimental	661 (194)	704 (154)	2 (0.7)	3 (1.0)
Consistent	Control	641 (168)	714 (164)	11 (3.8)	14 (4.9)
Inconsistent	Experimental	608 (168)	656 (122)	1 (0.3)	4 (1.4)
Inconsistent	Control	572 (133)	655 (117)	9 (3.1)	8 (2.8)

Consistent-accent condition

Figure 2 shows priming effects separately for conditions and halves. Listeners responded faster to identical than to unrelated items ($F1$ (1,23) = 123.171, p < .001; $F2$ (1,96) = 34.246, p < .001). Across participants, there was also an effect of *accent type*, indicating that listeners were faster overall to respond to the control items ($F1$ (1,23) = 8.159, p = .009; $F2$ (1,96) = 1.698, p = .196).

Figure 2: Priming Effects and SEs for Participants in the Consistent-Accent Condition



Moreover, priming effects were smaller for the experimental items than for the control items (*F1* (1,23) = 5.005, *p* = .035; *F2* (1,96) = 2.825, *p* = .096), but did not differ across the two halves (*F1* > 1; *F2* > 1), and there was no three-way interaction between *priming*, *accent type*, and *half* (*F1* (1,23) = 1.054, *p* = .315; *F2* > 1). Planned comparisons of the priming effects are displayed in Table 2.

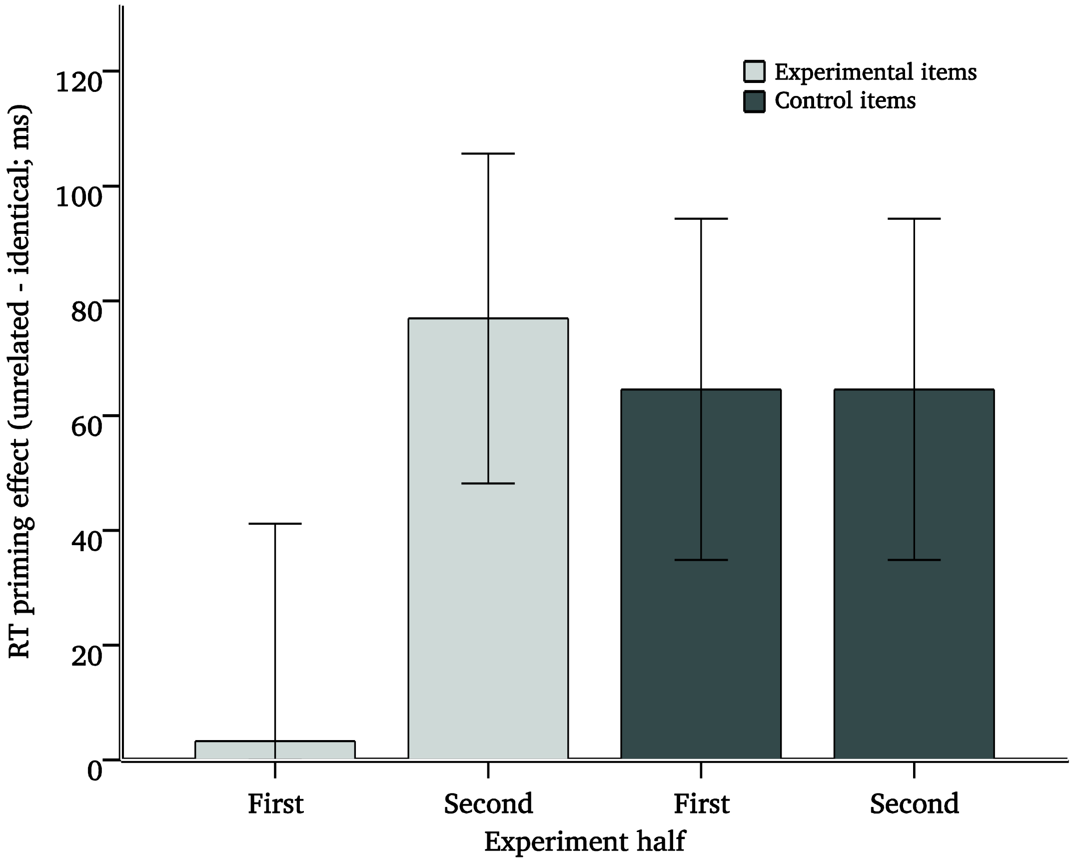
Table 2. Planned Pairwise Comparisons of Priming Effects for all Accent Types Across Participants and Items for Participants in the Consistent-Accent Condition

		Participant analysis			Item analysis		
Accent type	Half	*df*	*t1*	*p*	*df*	*t2*	*p*
Experimental	First	1,23	3.398	.002	1,22	2.336	.029
Experimental	Second	1,23	3.322	.003	1,22	2.222	.037
Control	First	1,23	7.746	.000	1,22	4.303	.000
Control	Second	1,23	5.162	.000	1,22	2.886	.009

Inconsistent-accent condition

Figure 3 displays priming effects separately for conditions and halves. Listeners responded faster overall to identical compared to unrelated trials ($F1$ (1,23) = 64.579, $p < .001$; $F2$ (1,96) = 20.372, $p < .001$). Listeners were equally fast for both accent types ($F1$ (1,23) = 1.217, $p = .281$; $F2 >$ 1), and faster overall in the second half of the experiment ($F1$ (1,23) = 26.604, $p < .001$; $F2$ (1,96) = 11.750, $p = .001$).

Figure 3: Priming Effects and SEs for Participants in the Inconsistent-Accent Condition



Participants were faster overall for the experimental items in the second half compared to the first, but not for the control items (*F1* (1,23) = 10.975, *p* = .003; *F2* (1,96) = 4.579, *p* = .032). Moreover, across participants there was a marginally significant interaction between *half* and *priming,* indicating that priming effects differed for the two halves (*F1* (1,23) = 3.620, *p* = .069; *F2* (1,96) = 1.784, *p* = .185). Finally, the significant three-way interaction across participants between *priming, accent type* and *half* suggests the priming effects for the experimental items in the first half were smaller compared to the other priming effects (*F1* (1,23) = 7.580, *p* = .011; *F2* (1,96) = 1.922, *p* = .169).

To further investigate the interaction between *half* and *priming*, as well as the three-way interaction, we ran analyses separately for the experimental items (words with segmental mismatches), and found a main effect of *priming* (*F1* (1,23) = 9.330, p = .006; *F2* (1,96) = 5.758, p = .021) and *half* (*F1* (1,23) = 54.527, p < .001, *F2* (1,96) = 18.286, p < .001), as well as an interaction between *priming* and *half* (*F1* (1,23) = 54.527, p < .001; *F2* (1,96) = 4.305, p = .044), indicating that priming effects for the experimental items were larger in the second half compared to the first.

A separate analysis for the control words revealed a main effect of *priming* (*F1* (1,23) = 51.192, p < .001; *F2* (1,96) = 15.168, p < .001), but no effect of *half* (*F1* (1,23) = 1.491, p = .234; *F2* > 1), nor an interaction between *priming* and *half* (*F1* > 1; *F2* > 1). Table 3 shows the planned comparisons for the priming effects.

Table 3. Planned Pairwise Comparisons of Priming Effects for all Accent Types Across Participants and Items for Participants in the Consistent-Accent Condition

		Participant analysis			Item analysis		
Accent type	Half	*df*	*t1*	*p*	*df*	*t2*	*p*
Experimental	First	1,23	.172	.865	1,22	.224	.825
Experimental	Second	1,23	5.352	.000	1,22	3.243	.004
Control	First	1,23	4.960	.000	1,22	2.404	.025
Control	Second	1,23	4.347	.000	1,22	3.349	.003

Taken together, these results show that while both listener groups were able to adapt to the control items immediately, only listeners in the consistent-accent condition showed facilitatory priming for the experimental items from the start of the experiment. The listeners in the inconsistent condition were able to adapt to the experimental items, but only in the second half of the experiment. This pattern was confirmed in a comparison across listener groups, done separately for the experimental words. The participant analysis for the experimental words showed that priming effects in the consistent-accent condition did not differ between the two halves, while they did for the inconsistent-accent condition ($F1$ (1,46) = 7.433, p = .009; $F2$ (1,96) = 2.425, p = .123).

Discussion

The present study looked at the boundaries of adaptation to foreign-accented speech by investigating whether listeners can adapt to a native-like speaker with a variable accent. One group of listeners heard a speaker with a consistent German accent, whereas a second group heard the same speaker producing a specific accent marker inconsistently.

Both listener groups interpreted control words without vocalic mismatches correctly. This was expected, because these words did not differ substantially from canonical Dutch pronunciations. Moreover, listeners in the consistent-accent condition were able to adapt to words with German-accented vocalic mismatches from the start of the experiment, which is evidence for extremely rapid adaptation to the speaker. Note that adaptation to the same vocalic mispronunciation from a different speaker in Witteman et al. (in press) was not immediate, but this could be due to the current speaker's mispronunciations being rated as less strongly accented.

Nonetheless, this quick adaptation is good news for L2 speakers and listeners, as listeners did not show any interpretation difficulties, even when words contained noticeable mispronunciations.

Listeners in the inconsistent-accent condition did not recognize the words with vocalic mismatches initially, but could in the second half of the experiment. Remember that these listeners also heard the speaker pronounce filler primes with the same vowel correctly, and all other filler primes sounded native-like too. They thus got very little indication that the speaker had a foreign accent, but still adapted to the mispronunciations. Previous research indicated that listeners will not adapt to a native speaker when they first heard them pronounce items correctly (Kraljic, et al., 2008; Kraljic & Samuel, 2011). One difference is that here native-like and foreign-accented words were intermixed rather than presented consecutively, but this does not explain why listeners adapted in the second half of the experiment. A possible explanation is that listeners are aware that foreign-accented speech is naturally variable, even within speakers (Hanulíková & Weber, 2012), while native speech is much more constant, and they therefore adapt even in the face of inconsistent input. This is in line with Eisner et al. (under revision) who found that English listeners generalize devoicing mispronunciations across word positions for non-native Dutch speakers, but not for native English speakers.

Our results also shed more light on the finding that adaptation to foreign-accented speech takes longer as a function of the strength of a speaker's accent (Bradlow & Bent, 2008). Though participants were exposed to identical experimental items, listeners in the inconsistent condition took longer to adapt. Accent inconsistency can thus also play a role in the speed of adaptation to foreign-accented speech, together with accent strength.

In summary, the perceptual system for speech has been shown to be flexible and rapidly able to adjust, as listeners are able to adapt to a foreign accent quickly, even when the speaker has an inconsistent accent. Inconsistency in foreign-accented speech creates no permanent problems for listeners: though it slows down adaptation, listeners need only a little longer to catch up.

Appendix A - List of Experimental Items

Experimental items

Target Dutch spelling (IPA transcription)	Accented prime (IPA)	Unrelated prime	Target translation	Unrelated prime translation
buik (bœyk)	bɔɪk	dief	belly	thief
buiten (bœytə)	bɔɪtə	koffie	outside	coffee
duif (dœyf	dɔɪf	riet	dove	reed
duiker (dœykər)	dɔɪkər	wakker	diver	awake
duim (dœym)	dɔɪm	pink	thumb	pinkie
duivel (dœyvəl)	dɔɪvəl	tuniek	devil	tunic
fornuis (fɔrnœys)	fɔrnɔɪs	type	stove	type
fruit (frœyt)	frɔɪt	tenue	fruit	uniform
hoofdhuid (ho:fthœyt)	ho:fthɔɪt	engerd	scalp	creep
kruipen (krœypə)	krɔɪpə	aanzoek	crawl	proposal
pluizig (plœyzəx)	plɔɪzəx	emmer	fluffy	bucket
pruik (prœyk)	prɔɪk	riem	wig	belt
ruiter (rœytər)	rɔɪtər	gewoon	rider	normal
snuit (snœyt)	snɔɪt	kreng	snout	hag
spuit (spœyt)	spɔɪt	tent	syringe	tent
struik (strœyk)	strɔɪk	koorts	shrub	fever
trui (trœy)	trɔɪ	naast	sweater	beside
tuinman (tœynmɑn)	tɔɪnmɑn	inning	gardener	inning
uiterst (œytərst)	ɔɪtərst	binding	final	bond
uitspraak (œytspra:k)	ɔɪtspra:k	boeiend	pronunciation	compelling
vuil (vœyl)	vɔɪl	inkt	dirty	ink
zuinig (zœynəx)	zɔɪnəx	hitte	thrifty	heat
zuivel (zœyvəl)	zɔɪvəl	haven	dairy	harbor
zuiver (zœyvər)	zɔɪvər	diepte	pure	depth

Control items

Target Dutch spelling (IPA transcription)	Accented prime (IPA)	Unrelated prime	Target translation	Unrelated prime translation
bek (bɛk)	bɛk	pet	beak	cap
bemind (bəmɪnt)	bəmɪnt	immuun	loved	immune
boek (buk)	buk	week	book	week
boete (butə)	butə	effen	fine	plain
code (ko:də)	ko:də	unie	code	union
defect (dəfɛkt)	dəfɛkt	tegoed	defect	credit
dekking (dɛkkɪŋ)	dɛkkɪŋ	katoen	cover	cotton
eb (ɛp)	ɛp	mus	ebb	sparrow
fitting (fɪtɪŋ)	fɪtɪŋ	techniek	fitting	technique
hek (hɛk)	hɛk	mes	fence	knife
hik (hɪk)	hɪk	wok	hiccup	wok
hit (hɪt)	hɪt	erg	hit	very
ketting (kɛtɪŋ)	kɛtɪŋ	prikkel	chain	stimulus
koek (kuk)	kuk	deeg	cookie	dough
mep (mɛp)	mɛp	pit	slap	pit
midden (mɪdə)	mɪdə	plicht	middle	duty
min (mɪn)	mɪn	toe	minus	to
minuut (miny:t)	miny:t	bezem	minute	broom
nek (nɛk)	nɛk	baan	neck	lane
niet (nit)	nit	keel	not	throat
oma (o:ma:)	a:dər	ader	grandmother	vein
uniek (y:nik)	y:nik	limoen	unique	lime
werk (wɛrk)	ja:r	jaar	work	year
wet(wɛt)	y:r	uur	law	hour

Summary and conclusions

Chapter 6

Summary of the results

The aim of the series of experiments described in this thesis was to investigate how listeners are able to adapt to foreign-accented speech. Just listening to foreign-accented speakers immediately reveals that this type of speech is much more variable than native speech. Although initially it may seem impossible to comprehend such speakers, often simply listening to them for a short time already helps. I tried to shed more light on this process of adaptation by looking at when it takes place, and what circumstances aid and prohibit adaptation to foreign-accented speech (or, more specifically for this thesis: foreign-accented Dutch).

Throughout this thesis, I used a cross-modal priming task and facilitatory priming as a measure of successful online word recognition (see also Marslen-Wilson & Zwitserlood, 1989; Marslen-Wilson, et al., 1996). In cross-modal priming, listeners first hear a word or nonword prime – in the present experiments spoken with a foreign accent – and then make a lexical decision about a visual target word or nonword, usually presented at the acoustic offset of the spoken prime. When listeners are exposed to the same word twice, they will be faster to respond than when the word they hear is different from the one they see. This is called a facilitatory priming effect. Previous work showed that this facilitatory effect is very specific: it is usually not observed when the prime and target differ by as little as one phoneme (Marslen-Wilson & Zwitserlood, 1989), and sometimes these cases even result in inhibition (e.g., Van Alphen & McQueen, 2006). Priming effects can thus be taken as evidence that listeners correctly identified the accented primes as the intended Dutch words, and hence that they had adapted to the speaker. In addition to cross-modal priming, I used acoustic measures and accent rating studies to establish the accentedness of words.

Chapter 2 focused on the role of experience with a foreign accent on adaptation to a speaker of that accent. Previous research showed that listeners can improve their understanding of a foreign-accented speaker within a few sentences or minutes (e.g., Bradlow & Bent, 2008, Clarke & Garrett, 2004), but these studies typically looked at the sentential level of understanding and did not control for specific accent markers. In my experiments, I wanted to investigate whether some specific segmental mispronunciations would be more detrimental to online word recognition than others, and whether difficulties with understanding these mispronounced words varied with prior experience with the accent.

Experiment 2.1 looked at whether listeners with extensive experience with an accent would behave differently compared to listeners with very limited experience with that same accent. I tested two groups of native Dutch listeners in a cross-modal priming experiment with German-accented Dutch: one group of listeners who in their daily lives heard German-accented Dutch multiple times a week from more than one speaker (extensive experience group), and another group who heard German-accented Dutch less than once a week from no more than one speaker (limited experience group). Three different types of mispronunciations were selected to test these groups on: strongly-, medium-, and weakly-accented words. The strongly-accented words all contained the Dutch vowel [œy], mispronounced by the German-accented speaker as [ɔɪ]. Medium-accented words contained Dutch [ɛɪ] mispronounced as [aɪ]. Weakly-accented words contained only segments shared between Dutch and German, so that they would not include segmental mismatches. The level of accentedness of these words was confirmed with acoustic analyses and a rating experiment.

Listeners with extensive experience with German-accented Dutch showed facilitatory priming for all word types from the start of the experiment, that is, they were able to interpret all word types correctly and had successfully adapted to this novel speaker. Listeners with limited experience with German-accented Dutch also immediately adapted to the speaker for the weakly- and medium-accented items, but could not interpret the strongly-accented words correctly. This experiment thus showed that if listeners have had extensive prior experience with an accent through frequent interactions with German-accented speakers, they can very rapidly adapt to a novel speaker of that accent and do so without observable problems, even when these pronunciations are strongly accented. But the picture is positive even for listeners with limited prior experience with a foreign accent: such listeners can also immediately adapt to medium- and weakly-accented words, and only have difficulty with the strongly-accented words.

In Experiment 2.2, I wanted to see whether listeners with limited prior experience with German-accented Dutch could learn to interpret the strongly-accented words with a brief additional exposure to the speaker. To do that, I recorded two versions of a short story, using the same speaker as in the cross-modal priming experiment. The first story contained twelve strongly-accented words (all with the vowel [œy], mispronounced as [ɔɪ]; strongly-accented exposure). These words were not in the cross-modal priming experiment. In the second story these words were replaced by weakly-accented words (weakly-accented exposure). Both stories lasted about four minutes and were played to the participants directly before the cross-modal priming experiment. I was interested to see whether such a short exposure phase would suffice for participants to adapt to the strongly-accented words, and whether just listening to the speaker (without any strongly-accented words) would yield the same effects.

Both exposure groups correctly interpreted the medium- and weakly-accented words, as was expected, because listeners had been able to do that already without additional exposure in Experiment 2.1. Moreover, listeners in the strongly-accented exposure condition immediately adapted to the speaker on the strongly-accented words. Listeners in the weakly-accented exposure condition did not interpret the strongly-accented words correctly in the first half of the experiment, but were able to do so in the second half of the experiment. Experiment 2.2 thus showed that very little additional exposure to the speaker is sufficient for listeners to adapt to a speaker, even if the words are strongly accented and the listeners have limited experience with the accent of the non-native speaker. Moreover, listeners benefit most from hearing the speaker and the mispronunciations, but even without the specific mispronunciations, listeners profit from additional exposure to the speaker. This suggests that there might be two levels of adaptation: a general level, at which listeners adjust to the speaker, and a specific level, where adaptation to specific mispronunciations takes place.

In Experiment 2.3, I was interested whether listeners would also be able to generalize across speakers, that is, to adapt to one accented speaker after having heard another speaker of the same accent. Most research done on native language (L1) suggests that this is not likely: adaptation is usually specific to one speaker (e.g., Eisner & McQueen, 2005; Nygaard & Pisoni, 1998), which makes sense because when one speaker has a lisp for example, it is not likely that the next speaker will pronounce words in a similar way. For foreign-accented speech, however, this situation is somewhat different: since the deviations are driven by the native language of a speaker, all foreign-accented speakers with the same native language will sound somewhat alike and tend to make similar mistakes. To see whether listeners were able to apply their knowledge of one speaker's accent to another speaker, the strongly-accented exposure story was recorded with a different

German-accented speaker who mispronounced words in a similar way as the speaker of the cross-modal priming experiment. Like in Experiment 2.2, listeners first heard the strongly-accented exposure story (but this time by the new speaker), immediately followed by the cross-modal priming experiment with the original speaker.

As in Experiments 2.1 and 2.2, listeners were able to adapt to the speaker for the weakly- and medium-accented words. However, listeners were not able to correctly interpret the strongly-accented words in the first half of the experiment, but were able to do so in the second half. These results thus show that it does help somewhat to be exposed to a speaker of the same accent, but not to the same extent as exposure to the same foreign-accented speaker does.

Taken together, the experiments in Chapter 2 showed that experience with an accent plays an important role in adaptation: different types of experience lead to different kinds of adaptation. Moreover, not all accented words interfere with successful word recognition; only the strongly-accented words require more experience to be interpreted correctly. This experience can either be gained outside the laboratory through extensive exposure to multiple speakers (Experiment 2.1) or with short-term exposure in the laboratory (Experiment 2.2). Even hearing a different speaker of the accent helps listeners adapt more quickly to the strongly-accented words (Experiment 2.3).

In Chapter 3 I further investigated what is more important for adaptation to foreign-accented speech: the size of acoustic deviation from the target language or the degree of perceived accentedness. I thus tried to relate processing difficulties (i.e., difficulty to adapt to a foreign-accented speaker) to acoustic analysis and accentedness ratings. The same materials and procedure as in the cross-modal priming experiment of Chapter 2 were used, but the primes were now recorded by an American-accented speaker. On the basis of the acoustic analyses and the accentedness ratings, items were divided up in three categories: [œy]-items, [ɛɪ]-items, and nonspecifically accented items. In contrast to German speakers, American speakers typically replace Dutch [œy] with [aʊ], but in line with German speakers they replace Dutch [ɛɪ] with [aɪ].

The acoustic analyses revealed that though both diphthongs differed from their canonical targets, the speaker's pronunciation of [œy] differed more than his pronunciation of [ɛɪ]. The rating study confirmed this pattern: the words with [œy] were rated most accented, followed by the words with [ɛɪ] and the nonspecifically accented items. Despite the differences between the rating scores, all words were still noticeably accented.

The cross-modal priming experiment served as the measure of processing fluency and showed that native Dutch listeners were able to interpret the speaker for all three word types; that is, they all showed significant facilitatory priming. But the priming effects were largest for the non-specifically accented words, while they were comparable for the other two word types. The fact that the [œy]-items in Chapter 3 did not interfere with word recognition, while they had interfered in Chapter 2, most likely indicates that the different segmental deviations in the [œy]-items of the American speaker in Chapter 3 were not large enough to prevent adaptation to the speaker, in contrast to the deviations from the German speaker in Chapter 2.

These results thus show that neither the size of acoustic deviations nor the degree of perceived accentedness can fully predict the extent of processing difficulties, and that different degrees of accentedness do not necessarily interfere with word recognition processes.

In Chapter 4 I focused on the process of adaptation to foreign-accented speech, and investigated whether this adaptation can also be observed on the long term. Experiment 2.1 had indicated that long-term adaptation is possible after prolonged experience with an accent outside the lab. In Chapter 4, I investigated whether such long-term effects could also be found after only brief exposure to an accented speaker. To ensure that listeners did not have prior experience with the accent, I chose an accent not common in The Netherlands: Hebrew-accented Dutch. There were two types of items: variant items, in which the speaker shortened half-long [i] to [ɪ], and canonical items which contained [ɪ] in their standard form.

Experiment 4.1 served as a baseline to see whether listeners could adapt to the speaker without any additional exposure. Native Dutch listeners completed a cross-modal priming experiment and showed successful word recognition for the canonical, but not the variant items. In Experiment 4.2, participants did the same priming experiment but were first familiarized with the speaker and her accent using a phoneme monitoring task. In this task, participants heard isolated words and were asked to respond whenever they heard /k/ (a phoneme not specific for the accent). The task was very short: it lasted only 3.5 minutes and participants heard 10 variant words, 10 canonical words, and 50 other words without [i], [ɪ], or (half-)long vowels. None of the words appeared in the cross-modal priming experiment.

Participants returned 24 hours later for the cross-modal priming experiment, which was identical to the one used in Experiment 4.1. Results showed that listeners were now able to correctly interpret both the canonical and the variant items. The extremely short phoneme monitoring exposure thus proved to be sufficient for participants to adapt to the speaker, even with 24 hours between exposure and test.

In Experiment 4.3, the delay between exposure and test was extended to one week, while keeping everything else identical to Experiment 4.2. I replicated the results of Experiment 4.2: listeners were able to adapt to the speaker for both the canonical and the variant items.

The listeners in these experiments thus needed very little exposure to adapt to the speaker, possibly even less than demonstrated in Chapter 2. Though the exposure phases of Experiments 2.2 and 2.3 and those of Experiments 4.2 and 4.3 were roughly the same length, listeners in Chapter 2 got much more information about the speaker, because they had listened to a story and not just isolated words. Apart from the fact that they listened to many more words overall, a story also provides a much richer sentential context from which one can deduce what the speaker meant. In Chapter 4, participants heard only isolated words, yet this turned out to be enough for them to be able to adapt to the speaker.

Secondly, the phoneme monitoring exposure indicated that it is not even necessary to pay attention to specific mispronunciations or consciously try to understand what a speaker is saying for listeners to be able to adapt. After all, the participants' only task was to listen for /k/ in isolated words, so everything else could have been ignored. Adaptation to foreign-accented speech thus appears to take place automatically.

Finally, Chapter 4 showed that long-term adaptation to a foreign-accented speaker is possible even when initial exposure is short. This demonstrates that the perceptual system is not only flexible enough to adapt to a speaker, it is also stable enough to maintain these adaptations for at least a week.

Chapters 2, 3, and 4 all indicated that the perceptual system is very flexible when it comes to foreign-accented speech. But where does this flexibility end? This question was posed in Chapter 5, where I investigated the boundaries of adaptation to foreign-accented speech. Specifically, I was interested to see whether listeners would also adapt to foreign-accented speech when it is inconsistent. A German native speaker was recorded who was so fluent in Dutch that he could pass as a native speaker. Two types of experimental items were used: control items, in which all phonemes were shared between German and Dutch so that there were no obvious segmental substitutions, and experimental items, which did contain segmental substitutions, for example German-accented [ɔɪ] instead of Dutch [œy]. The fillers were made up of a variety of segments that typically carry a German accent, and were recorded by the same speaker in both a native-like Dutch version and a German-accented version.

I created two conditions of a cross-modal priming experiment: a consistent-accent version (Experiment 5.1), in which both the experimental items and the fillers were German-accented, and an inconsistent-accent version (Experiment 5.2), in which the experimental items were German-accented, but the fillers were native-like. Crucially, the experimental items for both conditions were identical, only the accentedness of the fillers differed.

Listeners in the consistent-accent condition were able to adapt to both the control items (without segmental mismatches) and the experimental items (with segmental mismatches). Listeners in the inconsistent-accent condition, however, showed significant priming to the control items from the start of the experiment, but could only adapt to the experimental items in the second half of the experiment. This result thus shows that (in)consistency is an important factor in interpreting foreign-accented speech. Even though listeners are slowed down by inconsistently foreign-accented speech initially, they catch up quickly, indicating that they can tolerate such inconsistencies.

Conclusions

From the results presented in this thesis important conclusions can be drawn about adaptation to foreign-accented speech. The main conclusions are outlined here.

Flexibility of the perceptual system

One important conclusion that can be drawn is that although it may not always feel this way when listening to a foreign-accented speaker, the perceptual system is able to deal with the variation added to the speech signal by foreign-accented speakers remarkably well. Many of the accented words tested in this thesis did not hinder successful online word recognition, even though they were all perceived as accented.

Another finding is that accentedness differs between words; that is, not every word spoken by the same foreign-accented speaker is equally difficult to understand. While previous research mostly focused on a global measure of accentedness (by looking at intelligibility of different speakers, mainly at a sentence level), I have shown that there is also a lot of variation at the word level, even when looking at a single speaker. By distinguishing between different types of words, it was possible to learn more about the way the perceptual system deals with variation in speech.

Through looking at specific accent markers, I found that words without segmental mismatches never led to observable processing difficulties in any of the nine experiments in this thesis, even though listeners did realize that the words were spoken by a non-native speaker. Moreover, even words with segmental mismatches do not necessarily interfere with successful word recognition. Even though native segments were substituted by non-native ones, listeners still did not have observable problems understanding the speaker. This is very good news for L2 speakers and listeners, because it shows that the perceptual system is flexible enough to deal with these smaller deviations almost instantly.

The perceptual system is even flexible enough to deal with a speaker with an inconsistent accent, as demonstrated in Chapter 5. Though this accent inconsistency slows listeners down initially, listeners are able to adapt within minutes. This result does not follow findings in L1 research, where listeners did not adapt to a speaker's mispronunciations when they also heard the speaker say similar items correctly (Kraljic, Samuel, & Brennan, 2008; Kraljic & Samuel, 2011). A possible explanation for this difference is that foreign-accented speech is naturally variable, even within speakers (Hanulíková & Weber, 2012; Wade, et al., 2007), while native speech is much less variable. While native speech varies within categories (i.e., an /i/, though differently pronounced, will usually remain within the /i/ category), a foreign-accented speaker might sometimes pronounce a sound that falls within the intended category, and at other times a sound that falls in a different category (e.g., /i/ pronounced as /ɪ/). Listeners might have picked up on the greater inconsistency and variability in foreign-accented speech, and may thus have learned to adapt more flexibly to foreign-accented than to native-accented speech.

Adaptation is quick and automatic

Even when listeners are not able to adapt to foreign-accented words initially, it usually does not take long for listeners to catch up. The exposure phases in Chapters 2 and 4 were less than 5 minutes, but this was enough for listeners to achieve successful word recognition. In the inconsistent-accent condition of Chapter 5 adaptation was even achieved without an additional exposure phase: listeners adapted already during the experiment.

Chapter 4 indicated that adaptation is not only quick, but it is also automatic. Listeners did not have to fully understand which words the speaker intended to produce during the exposure phase; that is, they did not have to access their mental lexicon in order to complete the exposure task (phoneme monitoring). Moreover, they were not asked to pay attention to the mispronunciations, rather, they were asked to focus on a segment that is not specific to the accent. Despite all of this, listeners did show adaptation after brief exposure. It can therefore be concluded that adaptation to foreign-accented speech, like adaptation in L1 speech (e.g., McQueen, Norris, et al., 2006), is automatic.

This automaticity is also seen in listener's ability to transfer their knowledge of the speaker's accent across tasks: from listening to a story (Experiments 2.2 and 2.3) or phoneme monitoring (Experiments 4.2 and 4.3) to a cross-modal priming task.

The role of experience

I also demonstrated that experience with an accent is an important factor for adaptation, especially for words with large deviations from their canonical forms. When listeners are regularly exposed to speakers of an accent, they have no problems adapting to a new speaker of the same accent (Experiment 2.1). Even when listeners are not familiar with an accent, they can learn to adapt to a speaker very quickly, and this adaptation remains stable for at least a week (Experiments 4.2 and 4.3). Whether this form of adaptation also transfers to other speakers of the same accent was not tested in Chapter 4, but the results of Experiment 2.3 indicate that exposure to one speaker does help listeners adapt to another speaker of the same accent.

Moreover, listening to the same speaker, but not the specific strongly-accented mispronunciations, also helped listeners (Experiment 2.2). A possible explanation for this phenomenon is that adaptation to foreign-accented speech takes place at a general level first and then continues with further adjustments in response to specific mispronunciations. This could also explain why listeners are able to improve their word recognition even when they do not attend to the mispronunciations specifically (Experiments 2.2, 2.3, 4.2 and 4.3).

Models of speech recognition

The results presented in this thesis also have implications for models of spoken word recognition. As described in the introduction of this thesis, these models can be roughly divided into two camps: representational and processing-based accounts, but there are many variations of these two types.

Generally speaking, representational models assume that the mental lexicon is large in that all variation is stored there in separate entries (e.g., Goldinger, 1998). Multiple pronunciations are either stored as separate abstract level representations (e.g., Ranbom & Connine, 2007) or as separate episodic memories, rich in fine-grained phonetic detail (e.g., Johnson, 2006; Pierrehumbert, 2001). These accounts of course also make processing assumptions, but their explanation for recognition of variant pronunciations is based on representations. One way in which representational accounts could explain how listeners deal with the added variation of foreign-accented speech is that listeners store all of this variation in their lexicon. Upon hearing these variants a second time, adaptation could be achieved by re-accessing these forms. Prior storage of variants (in abstract or episodic form) could explain the benefit that participants with extensive experience had compared to listeners with limited experience with German-accented Dutch (Experiment 2.1).

It is less clear, however, how representational accounts would explain the short-term learning described in this thesis (Experiments 2.2, 2.3, 4.2, 4.3, and 5.2), and especially the generalization of learning from exposure to test phases (since different words were used in these parts of the experiments). A representation-based explanation for this kind of generalization would require that the participants happened to have heard (and stored) the accented forms prior to the experiment. While this is possible for the German-accented Dutch words used in Chapter 2 and 5, it is very improbable for the Hebrew-accented Dutch words in Chapter 4, since Dutch listeners are typically not familiar with the Hebrew accent. Also, it is not clear how representational models could explain how certain listener groups were not able to adapt in the first half of the experiment, but were able to do so in the second Experiments 2.2, 2.3, and 5.2). If adaptation is explained by previous exposure to items, this means that listeners had previous exposure with some items, but not others. Moreover, similar problems arise with the result that listeners from the same population were only able to adapt to the speaker after an exposure phase (Experiments 2.2, 2.3, 4.2, and 4.3). Representational models would require additional mechanisms to explain how the representations heard during exposure or in the beginning of the experiment did not immediately influence recognition but instead only started to do so after in some way being triggered by exposure to other accented words. Finally, it would be hard for representational models to explain why listeners responded differently to the same experimental items depending on the accentedness of the fillers (Chapter 5).

A processing-based account provides a better explanation for the data presented in this thesis. Generally, these models propose that variation is resolved at a pre-lexical level, and that the mental lexicon contains only canonical representations of words (Lotto & Holt, 2006; Mitterer, et al., 2006;

Gaskell & Marslen-Wilson, 1998). Listeners learn from exposure to an accent how variations should be mapped on the (stored) canonical forms. These models thus also make assumptions about representations, but the crux of their explanation for recognition of variant forms is carried by their assumptions about processing.

Information on how the speaker of an experiment pronounces words can be carried over from the exposure to the test phase, or can be established during the course of the experiment. The demonstration of generalization of learning from the exposure phase to the test phase, or during the experiment, speaks for such processing-based models. Perceptual learning studies in L1 have already proposed that if adaptation entails a change in the way a sound is mapped onto the lexicon, and this occurs at a pre-lexical level of processing, then this learning will generalize to all words in the lexicon that contain that sound (McQueen, Cutler, et al., 2006; Sjerps & McQueen, 2010). One possible way in which this process is achieved could be through a re-mapping of the accented vowels onto the lexicon. If this mapping is modified because of experience with the accent (which could be achieved through long-term exposure outside the laboratory, or short-term exposure during the experiment), then it is possible that after hearing several accented vowels, listeners can recognize all words containing that vowel. If this pre-lexical adjustment were to be the source of adaptation to foreign-accented speech, this could explain why listeners are able to adapt to the words after they have had some exposure to similar mispronunciations, but not before.

Moreover, this would also mean that adaptation to foreign-accented speech would use the same mechanisms as adaptation to artificial and within-language variation, making a processing-based account more parsimonious than a representational account.

This thesis took an important step towards understanding how listeners are able to process foreign-accented speech. The results presented here provide evidence that even though it may seem impossible at times, we are able to tune in to another person's foreign accent remarkably well. This finding also has implications for the way second language courses are taught around the world. Previous research has indicated that L2 listeners often have noticeable problems adapting to foreign-accented speech (e.g., Weber, et al., 2011; Broersma & Cutler, 2011), while the research presented in this thesis suggests that native speakers do not have insurmountable problems understanding foreign-accented speech. Since pronunciation is one of the hardest aspects to master when learning a second language (e.g., Flege, et al., 1995), and because, as shown here, speech perception has great flexibility, second language learners' energies could be more effectively allocated towards improving listening skills rather than achieving perfect pronunciation.

For native listeners, however, the main message of this thesis is a very positive one: if you are listening to a foreign-accented speaker who is hard to understand, just sit back, listen, and relax: within minutes, your perceptual system will do the work for you.

References

Adank, P., & McQueen, J. M. (2007). *The effect of an unfamiliar regional accent on spoken word comprehension.* Paper presented at the International Congress of Phonetic Sciences, Saarbrücken, Germany.

Adank, P., Van Hout, R., & Smits, R. (2004). An acoustic description of the vowels of Northern and Southern Standard Dutch. *Journal of the Acoustical Society of America, 116*(3), 1729-1738.

Aronson, L., Rosenhouse, J., Rosenhouse, G., & Podoshin, L. (1996). An acoustic analysis of modern Hebrew vowels and voiced consonants. *Journal of Phonetics, 24*(2), 283-293.

Baayen, R. H., Piepenbrock, R., & Van Rijn, H. (1993). The CELEX lexical database—Dutch, English, German. *Philadelphia, PA: Linguistics Data Consortium.*

Bent, T., & Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. *Journal of the Acoustical Society of America, 114*(3), 1600-1610.

Boersma, P., & Weenink, D. (2009). Praat: doing phonetics by computer (Version 5.1. 05)[Computer program](2009). *Last retrieved Oct, 31*.

Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., & Tohkura, Y. (1999). Training Japanese listeners to identify English/r/and/l: Long-term retention of learning in perception and production. *Attention, Perception, & Psychophysics, 61*(5), 977-985.

Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition, 106*(2), 707-729.

Broersma, M., & Cutler, A. (2011). Competition dynamics of second-language listening. *The Quarterly Journal of Experimental Psychology, 64*(1), 74-95.

Clarke, C. M., & Garrett, M. (2004). Rapid adaptation to foreign accented speech. *Journal of the Acoustical Society of America, 116*(6), 3647-3658.

Connine, C. M., Blasko, D. G., & Titone, D. (1993). Do the beginnings of spoken words have a special status in auditory word recognition? *Journal of Memory and Language, 32*(2), 193-210.

Connine, C. M., Ranbom, L. J., & Patterson, D. J. (2008). Processing variant forms in spoken word recognition: The role of variant frequency. *Perception & psychophysics, 70*(3), 403-411.

Dahan, D., Drucker, S. J., & Scarborough, R. A. (2008). Talker adaptation in speech perception: Adjusting the signal or the representations? *Cognition, 108*(3), 710-718.

Davis, M. H., Di Betta, A. M., Macdonald, M. J. E., & Gaskell, M. G. (2009). Learning and consolidation of novel spoken words. *Journal of Cognitive Neuroscience, 21*(4), 803-820.

Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility. *Studies in second language acquisition, 19*(1), 1-16.

Doeleman, R. (1998). *Native reactions to nonnative speech.* Unpublished PhD thesis, University of Tilburg, The Netherlands.

Dumay, N., & Gaskell, M. (2012). Overnight lexical consolidation revealed by speech segmentation. *Cognition, 123*(1), 119-132.

Eisner, F., & McQueen, J. M. (2005). The specificity of perceptual learning in speech processing. *Attention, Perception, & Psychophysics, 67*(2), 224-238.

Eisner, F., & McQueen, J. M. (2006). Perceptual learning in speech: Stability over time. *Journal of the Acoustical Society of America, 119*(4), 1950-1953.

Eisner, F., Melinger, A., & Weber, A. (under revision). Constraints on the transfer of perceptual learning in accented speech.

Ernestus, M., Baayen, R. H., & Schreuder, R. (2002). The recognition of reduced word forms. *Brain and Language, 81*(1-3), 162-173.

Flege, J. E. (1984). The detection of French accent by American listeners. *Journal of the Acoustical Society of America, 76*(3), 692-707.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: theoretical and methodological issues* (pp. 229-273): York Press.

Flege, J. E., Bohn, O. S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics, 25*(4), 437-470.

Flege, J. E., Munro, M. J., & MacKay, I. R. A. (1995). Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America, 97*(5), 3125-3134.

Flege, J. E., Schirru, C., & MacKay, I. R. A. (2003). Interaction between the native and second language phonetic subsystems. *Speech Communication, 40*(4), 467-492.

Gaskell, M. G., & Marslen-Wilson, W. D. (1996). Phonological variation and inference in lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 22*(1), 144-158.

Gaskell, M. G., & Marslen-Wilson, W. D. (1998). Mechanisms of phonological inference in speech perception. *Journal of Experimental Psychology: Human Perception and Performance, 24*(2), 380-396.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological review, 105*(2), 251-279.

Goldstone, R. L. (1998). Perceptual learning. *Annual review of psychology, 49*(1), 585-612.

Gow, D. W. (2002). Does English coronal place assimilation create lexical ambiguity? *Journal of Experimental Psychology: Human Perception and Performance, 28*(1), 163-179.

Grosjean, F. (2010). *Bilingual: Life and reality*. Cambridge, MA: Harvard Univ Press.

Gussenhoven, C. (1999). Illustrations of the IPA: Dutch. *Handbook of the International Phonetic Association*, 74-77.

Hanulíková, A., van Alphen, P. M., van Goch, M. M., & Weber, A. (2011). When one person's mistake is another's standard usage: The effect of foreign accent on syntactic processing. *Journal of Cognitive Neuroscience, 24*(4), 878-887.

Hanulíková, A., & Weber, A. (2012). Sink positive: Linguistic experience with th substitutions influences nonnative word recognition. *Attention, Perception & Psychophysics, 74*(3), 613-629.

Hay, J., Nolan, A., & Drager, K. (2006). From fush to feesh: Exemplar priming in speech perception. *Linguistic review, 23*(3), 351-379.

Hayes-Harb, R., Smith, B. L., Bent, T., & Bradlow, A. R. (2008). The interlanguage speech intelligibility benefit for native speakers of

Mandarin: Production and perception of English word-final voicing contrasts. *Journal of Phonetics, 36*(4), 664-679.

Hongyan, W., & van Heuven, V. J. (2007). *Quantifying the interlanguage speech intelligibility benefit.* Paper presented at the 16th International Congress of Phonetic Sciences, Saarbrücken, Germany.

Jacobi, I. (2009). *On variation and change in diphthongs and long vowels of spoken Dutch.* Unpublished PhD thesis, University of Amsterdam, The Netherlands.

Jenkins, J. J., Strange, W., Nishi, K., Fitzgerald, B. H., Trent, S. A., & Thornton, D. H. (1997). Acoustic comparison of the effects of coarticulation on the production of Japanese and American English vowels. *The Journal of the Acoustical Society of America, 102*, 3134.

Johnson, K. (2006). Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of Phonetics, 34*(4), 485-499.

Jongman, A., Wade, T., & Sereno, J. (2003). *On improving the perception of foreign-accented speech.* Paper presented at the 15th International Congress of Phonetic Sciences, Barcelona, Spain.

Kohler, K. (1999). Illustrations of the IPA: German. *Handbook of the International Phonetic Association*, 86-89.

Kraljic, T., & Samuel, A. G. (2005). Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology, 51*(2), 141-178.

Kraljic, T., & Samuel, A. G. (2006). Generalization in perceptual learning for speech. *Psychonomic Bulletin & Review, 13*(2), 262-268.

Kraljic, T., & Samuel, A. G. (2011). Perceptual learning evidence for contextually-specific representations. *Cognition, 121*, 459-465.

Kraljic, T., Samuel, A. G., & Brennan, S. E. (2008). First impressions and last resorts. *Psychological Science, 19*(4), 332.

Ladefoged, P. (1999). American English *Handbook of the International Phonetic Association* (pp. 41–44). Cambridge: Cambridge University Press.

Lane, H. (1963). Foreign accent and speech distortion. *Journal of the Acoustical Society of America, 35*, 451-453.

Laufer, A. (1999). Illustrations of the IPA: Hebrew. *Handbook of the International Phonetic Association*, 96-99.

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English/r/and/l: A first report. *Journal of the Acoustical Society of America, 89*(2), 874-886.

Lotto, A. J., & Holt, L. L. (2006). Putting phonetic context effects into context: A commentary on Fowler (2006). *Attention, Perception, & Psychophysics, 68*(2), 178-183.

Marslen-Wilson, W. (1993). Issues of process and representation in lexical access. In G. T. M. Altmann & R. Shillcock (Eds.), *Cognitive models of speech processing: the Second Sperlonga Meeting* (pp. 187-210).

Marslen-Wilson, W., Moss, H. E., & van Halen, S. (1996). Perceptual distance and competition in lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 22*(6), 1376-1392.

Marslen-Wilson, W., & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception and Performance, 15*(3), 576-585.

Marslen-Wilson, W. D., Nix, A., & Gaskell, G. (1995). Phonological variation in lexical access: Abstractness, inference and English place assimilation. *Language and Cognitive Processes, 10*(3-4), 285-308.

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.

McQueen, J. M., Cutler, A., & Norris, D. (2006). Phonological abstraction in the mental lexicon. *Cognitive Science, 30*(6), 1113-1126.

McQueen, J. M., Norris, D., & Cutler, A. (2006). The dynamic nature of speech perception. *Language and Speech, 49*(1), 101-112.

Mitterer, H., & Blomert, L. (2003). Coping with phonological assimilation in speech perception: Evidence for early compensation. *Perception & psychophysics, 65*(6), 956-969.

Mitterer, H., Csépe, V., Honbolygo, F., & Blomert, L. (2006). The recognition of phonologically assimilated words does not depend on specific language experience. *Cognitive Science, 30*(3), 451-479.

Mitterer, H., & Ernestus, M. (2006). Listeners recover/t/s that speakers reduce: Evidence from/t/-lenition in Dutch. *Journal of Phonetics, 34*(1), 73-103.

Most, T., Amir, O., & Tobin, Y. (2000). The Hebrew vowel system: Raw and normalized acoustic data. *Language and Speech, 43*(3), 295-308.

Munro, M. J., & Derwing, T. M. (1995a). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning, 45*(1), 73-97.

Munro, M. J., & Derwing, T. M. (1995b). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech, 38*(3), 289-306.

Munro, M. J., Flege, J. E., & MacKay, I. R. A. (1996). The effects of age of second language learning on the production of English vowels. *Applied Psycholinguistics, 17*(03), 313-334.

Niedzielski, N. (1999). The effect of social information on the perception of sociolinguistic variables. *Journal of language and social psychology, 18*(1), 62-85.

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology, 47*(2), 204-238.

Nygaard, L. C., & Pisoni, D. B. (1998). Talker-specific learning in speech perception. *Attention, Perception, & Psychophysics, 60*(3), 355-376.

Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *The Journal of the Acoustical Society of America, 24*, 175.

Pierrehumbert, J. (2001). Why phonological constraints are so coarse-grained. *Language and Cognitive Processes, 16*(5-6), 691-698.

Pitt, M. A. (2009). How are pronunciation variants of spoken words recognized? A test of generalization to newly learned words. *Journal of Memory and Language, 61*(1), 19-36.

Ranbom, L. J., & Connine, C. M. (2007). Lexical representation of phonological variation in spoken word recognition. *Journal of Memory and Language, 57*(2), 273-298.

Scott, D. R., & Cutler, A. (1984). Segmental phonology and the perception of syntactic structure. *Journal of verbal learning and verbal behavior, 23*(4), 450-466.

Sidaras, S. K., Alexander, J. E. D., & Nygaard, L. C. (2009). Perceptual learning of systematic variation in Spanish-accented speech. *Journal of the Acoustical Society of America, 125*(5), 3306-3316.

Sjerps, M. J., & McQueen, J. M. (2010). The bounds on flexibility in speech perception. *Journal of Experimental Psychology: Human Perception and Performance, 36*(1), 195-211.

Strange, W., Levy, E. S., & Law, F. F. (2009). Cross-language categorization of French and German vowels by naïve American listeners. *The Journal of the Acoustical Society of America, 126*(3), 1461-1476.

Sumner, M., & Samuel, A. G. (2005). Perception and representation of regular variation: The case of final/t. *Journal of Memory and Language, 52*(3), 322-338.

Sumner, M., & Samuel, A. G. (2009). The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language, 60*(4), 487-501.

Tamminen, J., Payne, J. D., Stickgold, R., Wamsley, E. J., & Gaskell, M. G. (2010). Sleep spindle activity is associated with the integration of new memories and existing knowledge. *The Journal of Neuroscience, 30*(43), 14356-14360.

Trude, A. M., & Brown-Schmidt, S. (2011). Talker-specific perceptual adaptation during online speech perception. *Language and Cognitive Processes*.

Van Alphen, P. M., & McQueen, J. M. (2006). The effect of voice onset time differences on lexical access in Dutch. *Journal of Experimental Psychology: Human Perception and Performance, 32*(1), 178-196.

Wade, T., Jongman, A., & Sereno, J. (2007). Effects of acoustic variability in the perceptual learning of non-native-accented speech sounds. *Phonetica, 64*(2-3), 122-144.

Weber, A., Broersma, M., & Aoyagi, M. (2011). Spoken-word recognition in foreign-accented speech by L2 listeners. *Journal of Phonetics, 39*(4), 479-491.

Weber, A., Di Betta, A. M., & McQueen, J. M. (in prepation). Treak or trit: Adaptation to genuine and arbitrary foreign accents.

Witteman, M. J., Weber, A., & McQueen, J. M. (in press). Strength of a foreign accent and listener familiarity with it co-determine speed of perceptual adaptation.

Samenvatting

Begrijpen wat iemand zegt kan zo simpel zijn, maar als die persoon een andere moedertaal heeft en met een accent spreekt, wordt deze simpele taak vaak een heel stuk lastiger. Als je echter een tijdje luistert naar iemand met een accent, wordt deze persoon vaak een stuk begrijpelijker. Hoe kan dat? Het is onwaarschijnlijk dat de uitspraak van de persoon ineens verandert. Kennelijk kan ons brein zich aanpassen aan een afwijkende uitspraak. Dit proefschrift gaat over dat proces: hoe kunnen mensen taal met een accent begrijpen?

Om dit te onderzoeken heb ik verschillende experimenten gedaan, die allemaal gaan over taal met een buitenlands accent (bijvoorbeeld Nederlands met een Duits accent). Hoewel Nederlandse accenten (zoals Twents) soms ook lastig te begrijpen zijn voor niet-sprekers van dat accent en de grens tussen een dialect en een taal arbitrair is, heb ik er voor gekozen om in dit proefschrift alleen buitenlandse accenten te onderzoeken, omdat die accenten vaak consistenter zijn en vaak meer afwijken van de doeltaal. Als in deze samenvatting dus 'accent' gebruikt wordt, zal dit altijd refereren aan een buitenlands accent.

Hoofdstuk 2, het eerste experimentele hoofdstuk van dit proefschrift, onderzoekt hoe belangrijk ervaring met een accent is om een spreker met datzelfde accent te begrijpen. Eerder onderzoek heeft aangetoond dat het makkelijker wordt om een spreker met een accent te begrijpen naarmate je langer naar deze persoon luistert, maar er is al verbetering in het begrijpen te zien na enkele zinnen of minuten. Hoewel deze onderzoeken zeer waardevol zijn, kijken ze niet naar of een spreker altijd even makkelijk te begrijpen is. Misschien zijn zelfs van één en dezelfde spreker bepaalde woorden heel makkelijk te begrijpen, terwijl andere woorden een stuk lastiger zijn. Dit komt doordat sommige klanken zowel voorkomen in de moedertaal van de spreker als in de taal die hij probeert te spreken (alle klanken in het woord 'dekking' komen bijvoorbeeld zowel in het Nederlands als het Duits voor) en andere klanken daarentegen, maar in één taal voorkomen. Een bekend voorbeeld is de klank 'ui', waarmee bijna alle sprekers met een andere moedertaal dan Nederlands moeite hebben.

Of woorden met verschillende accentsterkte anders worden verwerkt, ook al komen ze van dezelfde spreker, wordt onderzocht in Hoofdstuk 2. In Experiment 2.1 heb ik twee groepen luisteraars vergeleken, namelijk Nederlanders die meerdere malen per week Nederlands met een Duits accent horen, en Nederlanders die dit minder dan eens per week horen. Deze twee groepen proefpersonen hoorden drie verschillende typen Nederlandse woorden, gesproken door één Duitse moedertaalspreker: licht-accent woorden (zoals 'dekking', waar de klanken allemaal overlappen), gemiddeld-accent woorden (deze bevatten de klank 'ij', door de Duitse spreker uitgesproken als 'ai') en sterk-accent woorden (met de klank 'ui', door de Duitse spreker uitgesproken als 'oi').

Het experiment liet zien dat de mate van ervaring met het accent niet van belang was voor het begrijpen van de licht-accent en de gemiddeld-accent woorden: beide groepen konden deze woorden meteen begrijpen. Het verschil zat in de sterk-accent woorden: de luisteraars met veel ervaring hadden geen problemen met deze woorden, maar de luisteraars zonder uitgebreide ervaring konden deze woorden niet begrijpen.

Dit experiment liet dus zien, dat zelfs woorden die door een en dezelfde spreker uitgesproken worden, niet altijd allemaal even goed begrepen kunnen worden.

Experiment 2.2 onderzocht of luisteraars met weinig ervaring met Nederlands met een Duits accent even goed konden worden als de luisteraars met veel ervaring met dat accent, alleen door iets langer naar de spreker te luisteren. Proefpersonen hoorden een verhaaltje van vier minuten gesproken door dezelfde spreker als de rest van het experiment, voor zij hetzelfde experiment deden als de luisteraars van Experiment 2.1. Dit korte verhaaltje hielp enorm, vooral als er in dat verhaaltje ook sterk-accent woorden zaten. Als deze er niet in zaten en proefpersonen dus wél dezelfde spreker alvast hadden gehoord, maar zonder dat deze ook sterk-accent woorden uitsprak, hielp dit pas in de tweede helft van het experiment.

Experiment 2.3 stelde een andere vraag: kun je ook de kennis van het accent van de ene spreker toepassen op een andere spreker, als deze sprekers hetzelfde accent hebben? Hoewel een dergelijk mechanisme meestal niet handig is om nieuwe sprekers te begrijpen (bijvoorbeeld als iemand slist, dan is het niet handig om er vanuit te gaan dat iedereen dat doet), is dat wèl heel handig bij accenten, want omdat de variatie gedreven wordt door de klanken van de moedertaal, kun je er vanuit gaan dat alle mensen met dezelfde moedertaal ongeveer dezelfde versprekingen maken. Het bleek dat luisteraars inderdaad de kennis van één accentspreker kunnen toepassen op het begrijpen van een andere spreker met hetzelfde accent, maar dat het toch wel moeilijker is dan wanneer ze het accent van één en dezelfde spreker afkomstig is.

Hoofdstuk 2 liet dus zien dat verschillende vormen van ervaring en accent zorgen voor verschillende manieren van adaptatie. Maar dit geldt alleen voor de woorden met een sterk accent: de andere woorden konden door iedereen zonder problemen begrepen worden.

Maar wat maakt dat een woord moeilijk te begrijpen is? Is dat de sterkte van het accent, of hoeveel een woord akoestisch verschilt van het doelwoord? Dit is onderzocht in Hoofdstuk 3. Sterk-accent, gemiddeld-accent en licht-accent Nederlandse woorden werden opgenomen met een Amerikaanse spreker. Vervolgens heb ik gemeten hoeveel de woorden akoestisch verschilden van de doelwoorden, en aan Nederlandse proefpersonen gevraagd hoe sterk het accent van de woorden was. Bij een andere groep proefpersonen heb ik bekeken hoe moeilijk de woorden te begrijpen waren. De geteste woorden van Hoofdstuk 3 bleken even makkelijk (of moeilijk) te begrijpen, ongeacht hoe sterk het accent of hoeveel afwijking er was. Dit betekent waarschijnlijk dat de verschillen tussen de woordtypes in dit experiment niet groot genoeg waren, omdat alle woorden zonder problemen begrepen konden worden.

Tot nu toe hebben we dus gezien dat mensen zich vrij makkelijk kunnen aanpassen aan taal met een accent. Maar hoe lang blijft dit effect zichtbaar? Verdwijnt het alweer na een paar minuten, of zijn er langere effecten te zien? Dit was de vraag in Hoofdstuk 4. Nu onderzocht ik woorden met een Hebreeuws accent zonder grote afwijking van het doelwoord (canonieke woorden) en woorden met een grote afwijking (variantwoorden). De keuze voor Hebreeuws is gemaakt omdat het onwaarschijnlijk is dat onze Nederlandse proefpersonen dit accent eerder hebben gehoord.

In Experiment 4.1 keek ik naar wat luisteraars konden begrijpen van het accent, zonder ervaring met het accent. Het bleek dat luisteraars geen moeite hadden met de canonieke woorden (die geen sterk accent hadden), maar wel met de variantwoorden (met een duidelijke afwijking). In Experiment 4.2 gaf ik luisteraars wat extra hulp: ze mochten vantevoren al even heel kort naar de spreker luisteren, zowel naar canonieke als variantwoorden. Dit was weliswaar 24 uur voor het experiment, maar het bleek toch al genoeg: nu hadden ze geen problemen, noch met de canonieke, noch met de variantwoorden. Dit effect bleef hetzelfde als de tijd tussen de eerste blootstelling aan de spreker en de test verlengd werd naar een week (Experiment 4.3).

Hoofdstuk 4 liet dus zien dat je echt heel weinig naar een spreker hoeft te luisteren om deze te kunnen leren begrijpen. Dit effect blijft zelfs lang bestaan: er zijn in ieder geval nog effecten te zien na een week.

Hoofdstukken 2, 3 en 4 hebben een gemeenschappelijke boodschap: luisteren naar taal met een accent is niet zo moeilijk. Eventjes opletten en dan ben je er eigenlijk al. Maar zijn er grenzen aan dit effect? Dit was de vraag in Hoofdstuk 5. Nu keek ik naar wat luisteraars doen als een accent inconsistent is, als een spreker bijvoorbeeld soms 'ui' zegt en soms 'oi'. Daarom nam ik een Duitse spreker op die niet te onderscheiden was van een Nederlander als hij Nederlands sprak. Ik vroeg deze spreker zowel de woorden op te nemen in zijn ‘normale’ accent (dus eigenlijk zonder accent) en met een geïmiteerd Duits accent. Een groep proefpersonen luisterde naar zijn 'Duitse' uitspraak (consistente groep), de andere groep hoorde Nederlandse en Duitse uitspraken door elkaar (inconsistente groep) en luisteraars werden getest op licht-accent en sterk-accent woorden. Proefpersonen in de consistente groep konden alle woorden begrijpen, ook de sterk-accent woorden, maar luisteraars in de inconsistente groep begrepen de sterk-accent woorden alleen in de tweede helft van het experiment. Dus hoewel de inconsistente groep wel problemen had met de sterk-accent woorden, konden ze dit al heel snel overwinnen.

Ons perceptuele systeem, het systeem waarmee we spraak begrijpen, blijkt dus enorm flexibel. Een paar minuten spraak, een paar zinnen of zelfs maar enkele woorden helpen ons al enorm om ons aan te passen aan afwijkende spraak. Sterker nog, voor de meeste woorden gaat aanpassing zelfs nog sneller: alleen de woorden met een sterk accent zorgen voor begripsproblemen, maar dit kan snel verholpen worden door extra ervaring met een accent. Dit is bijzonder goed nieuws voor sprekers met een accent en de mensen die naar hen luisteren. Als je moedertaalspreker bent, kun je je snel en automatisch aanpassen aan een spreker met een accent: het is niet nodig om speciale aandacht te besteden aan de afwijkingen. Deze informatie is belangrijk voor iedereen die een tweede taal wil leren: moedertaalsprekers kunnen je makkelijk begrijpen, dus het is belangrijker te werken aan je eigen luistercapaciteit dan te proberen je accent te verbeteren of zelfs weg te krijgen. En voor moedertaalsprekers: mocht je iemand tegenkomen die onmogelijk te begrijpen lijkt: wacht gewoon af en ontspan, je perceptuele systeem zal zich vanzelf aanpassen aan de spreker en het geheel begrijpelijk maken.

Curriculum vitae

Marijt J. Witteman was born in 1984 in Rotterdam, The Netherlands. She obtained a Bachelor's degree in Social Sciences (cum laude), with specializations in psychology, linguistics and statistics, from University College Utrecht. In 2008, she obtained a Master's degree in Behavioral Science (cum laude), with a specialization in psychology, from the Radboud University Nijmegen. In that same year, she received a Prins Bernhard Cultuurfonds scholarship and a VSB fund scholarship to continue the research of her Master's at the Pennsylvania State University (USA) for six months. In 2009, she was awarded a three-year scholarship from the Max Planck Society to do her PhD research at the Max Planck Institute for Psycholinguistics, where she joined the Adaptive Listening Group.

List of publications

Witteman, M. J., Weber, A., & McQueen, J. M. (in press). Foreign accent strength and listener familiarity with an accent co-determine speed of perceptual adaptation. *Attention, Perception, and Psychophysics*.

Scharenborg, O., Witteman, M. J., Weber, A. (2012). Computational modelling of the recognition of foreign-accented speech. *Proceedings of Interspeech: 13th Annual Conference of the International Speech Communication Association.*

Witteman, M. J., Bardhan, N. P., Weber, A., & McQueen, J. M. (under revision). Automaticity and stability of adaptation to foreign-accented speech.

Witteman, M. J., Weber, A., & McQueen, J. M. (under revision). Tolerance for inconsistency in foreign-accented speech.

Witteman, M. J., Weber, A., & McQueen, J. M. (2011). On the relationship between perceived accentedness, acoustic similarity, and processing difficulty in foreign-accented speech. *Proceedings of Interspeech 2011: 12th Annual Conference of the International Speech Communication Association.*

Witteman, M. J., & Segers, E. (2010). The modality effect tested in children in a user-paced multimedia environment. Journal of Computer Assisted Learning, 26, 132-142. doi:10.1111/j.1365-2729.2009.00335.x.

MPI Series in Psycholinguistics

1. The electrophysiology of speaking: Investigations on the time course of semantic, syntactic, and phonological processing. *Miranda van Turennout*
2. The role of the syllable in speech production: Evidence from lexical statistics, metalinguistics, masked priming, and electromagnetic midsagittal articulography. *Niels O. Schiller*
3. Lexical access in the production of ellipsis and pronouns. *Bernadette M. Schmitt*
4. The open-/closed-class distinction in spoken-word recognition. *Alette Haveman*
5. The acquisition of phonetic categories in young infants: A self-organising artificial neural network approach. *Kay Behnke*
6. Gesture and speech production. *Jan-Peter de Ruiter*
7. Comparative intonational phonology: English and German. *Esther Grabe*
8. Finiteness in adult and child German. *Ingeborg Lasser*
9. Language input for word discovery. *Joost van de Weijer*
10. Inherent complement verbs revisited: Towards an understanding of argument structure in Ewe. *James Essegbey*
11. Producing past and plural inflections. *Dirk Janssen*
12. Valence and transitivity in Saliba: An Oceanic language of Papua New Guinea. *Anna Margetts*
13. From speech to words. *Arie van der Lugt*
14. Simple and complex verbs in Jaminjung: A study of event categorisation in an Australian language. *Eva Schultze-Berndt*
15. Interpreting indefinites: An experimental study of children's language comprehension. *Irene Krämer*
16. Language-specific listening: The case of phonetic sequences. *Andrea Weber*
17. Moving eyes and naming objects. *Femke van der Meulen*

18. Analogy in morphology: The selection of linking elements in Dutch compounds. *Andrea Krott*
19. Morphology in speech comprehension. *Kerstin Mauth*
20. Morphological families in the mental lexicon. *Nivja H. de Jong*
21. Fixed expressions and the production of idioms. *Simone A. Sprenger*
22. The grammatical coding of postural semantics in Goemai (a West Chadic language of Nigeria). *Birgit Hellwig*
23. Paradigmatic structures in morphological processing: Computational and cross-linguistic experimental studies. *Fermín Moscoso del Prado Martín*
24. Contextual influences on spoken-word processing: An electrophysiological approach. *Daniëlle van den Brink*
25. Perceptual relevance of prevoicing in Dutch. *Petra M. van Alphen*
26. Syllables in speech production: Effects of syllable preparation and syllable frequency. *Joana Cholin*
27. Producing complex spoken numerals for time and space. *Marjolein Meeuwissen*
28. Morphology in auditory lexical processing: Sensitivity to fine phonetic detail and insensitivity to suffix reduction. *Rachèl J. J. K. Kemps*
29. At the same time...: The expression of simultaneity in learner varieties. *Barbara Schmiedtová*
30. A grammar of Jalonke argument structure. *Friederike Lüpke*
31. Agrammatic comprehension: An electrophysiological approach. *Marlies Wassenaar*
32. The structure and use of shape-based noun classes in Miraña (North West Amazon). *Frank Seifart*
33. Prosodically-conditioned detail in the recognition of spoken words. *Anne Pier Salverda*
34. Phonetic and lexical processing in a second language. *Mirjam Broersma*
35. Retrieving semantic and syntactic word properties. *Oliver Müller*
36. Lexically-guided perceptual learning in speech processing. *Frank Eisner*
37. Sensitivity to detailed acoustic information in word recognition. *Keren B. Shatzman*
38. The relationship between spoken word production and comprehension. *Rebecca Özdemir*
39. Disfluency: Interrupting speech and gesture. *Mandana Seyfeddinipur*

40. The acquisition of phonological structure: Distinguishing contrastive from non-contrastive variation. *Christiane Dietrich*
41. Cognitive cladistics and the relativity of spatial cognition. *Daniel B.M. Haun*
42. The acquisition of auditory categories. *Martijn Goudbeek*
43. Affix reduction in spoken Dutch. *Mark Pluymaekers*
44. Continuous-speech segmentation at the beginning of language acquisition: Electrophysiological evidence. *Valesca Kooijman*
45. Space and iconicity in German Sign Language (DGS). *Pamela Perniss*
46. On the production of morphologically complex words with special attention to effects of frequency. *Heidrun Bien*
47. Crosslinguistic influence in first and second languages: Convergence in speech and gesture. *Amanda Brown*
48. The acquisition of verb compounding in Mandarin Chinese. *Jidong Chen*
49. Phoneme inventories and patterns of speech sound perception. *Anita Wagner*
50. Lexical processing of morphologically complex words: An information-theoretical perspective. *Victor Kuperman*
51. A grammar of Savosavo, a Papuan language of the Solomon Islands. *Claudia Wegener*
52. Prosodic structure in speech production and perception. *Claudia Kuzla*
53. The acquisition of finiteness by Turkish learners of German and Turkish learners of French: Investigating knowledge of forms and functions in production and comprehension. *Sarah Schimke*
54. Studies on intonation and information structure in child and adult German. *Laura de Ruiter*
55. Processing the fine temporal structure of spoken words. *Eva Reinisch*
56. Semantics and (ir)regular inflection in morphological processing. *Wieke Tabak*
57. Processing strongly reduced forms in casual speech. *Susanne Brouwer*
58. Ambiguous pronoun resolution in L1 and L2 German and Dutch. *Miriam Ellert*
59. Lexical interactions in non-native speech comprehension: Evidence from electro-encephalography, eye-tracking, and functional magnetic resonance imaging. *Ian FitzPatrick*

60. Processing casual speech in native and non-native language. *Annelie Tuinman*
61. Split intransitivity in Rotokas, a Papuan language of Bougainville. *Stuart Robinson*
62. Evidentiality and intersubjectivity in Yurakaré: An interactional account. *Sonja Gipper*
63. The influence of information structure on language comprehension: A neurocognitive perspective. *Lin Wang*
64. The meaning and use of ideophones in Siwu. *Mark Dingemanse*
65. The role of acoustic detail and context in the comprehension of reduced pronunciation variants. *Marco van de Ven*
66. Speech reduction in spontaneous French and Spanish. *Francisco Torreira*
67. The relevance of early word recognition: Insights from the infant brain. *Caroline Junge*
68. Adjusting to different speakers: Extrinsic normalization in vowel perception. *Matthias J. Sjerps*
69. Structuring language: contributions to the neurocognition of syntax. *Katrien R. Segaert*
70. Infants' appreciation of others' mental states in prelinguistic communication: a second person approach to mindreading. *Birgit Knudsen*
71. Gaze behavior in face-to-face interaction. *Federico Rossano*
72. Sign-spatiality in Kata Kolok: how a village sign language of Bali inscribes its signing space. *Connie de Vos*
73. Who is talking? Behavioural and neural evidence for norm-based coding in voice identity learning. *Attila Andics*
74. Lexical processing of foreign-accented speech: Rapid and flexible adaptation. *Marijt J. Witteman*
75. The use of deictic versus representational gestures in infancy. *Daniel Puccini*