

# Language and Speech

<http://las.sagepub.com/>

---

## **Automaticity and Stability of Adaptation to a Foreign-Accented Speaker**

Marijt J. Witteman, Neil P. Bardhan, Andrea Weber and James M. McQueen

*Language and Speech* published online 6 May 2014

DOI: 10.1177/0023830914528102

The online version of this article can be found at:

<http://las.sagepub.com/content/early/2014/05/05/0023830914528102>

---

Published by:



<http://www.sagepublications.com>

**Additional services and information for *Language and Speech* can be found at:**

**Email Alerts:** <http://las.sagepub.com/cgi/alerts>

**Subscriptions:** <http://las.sagepub.com/subscriptions>

**Reprints:** <http://www.sagepub.com/journalsReprints.nav>

**Permissions:** <http://www.sagepub.com/journalsPermissions.nav>

**Citations:** <http://las.sagepub.com/content/early/2014/05/05/0023830914528102.refs.html>

>> [OnlineFirst Version of Record](#) - May 6, 2014

[What is This?](#)

# Automaticity and Stability of Adaptation to a Foreign- Accented Speaker

Language and Speech

1–22

© The Author(s) 2014

Reprints and permissions:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/0023830914528102

las.sagepub.com



**Marijt J. Witteman**

Max Planck Institute for Psycholinguistics, The Netherlands; Radboud University Nijmegen, The Netherlands

**Neil P. Bardhan**

Max Planck Institute for Psycholinguistics, The Netherlands

**Andrea Weber**

Max Planck Institute for Psycholinguistics, The Netherlands; Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, The Netherlands; Eberhard-Karls Universität Tübingen, Germany

**James M. McQueen**

Max Planck Institute for Psycholinguistics, The Netherlands; Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, The Netherlands; Behavioural Science Institute, Radboud University Nijmegen, The Netherlands

## Abstract

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

---

## Corresponding author:

Neil P. Bardhan, 22 N. 3rd St., Philadelphia, PA 19106, United States of America.

Email: neilpbardhan@gmail.com

## Keywords

Foreign-accented speech, Hebrew-accented Dutch, cross-modal priming, automaticity, long-term adaptation

## 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 speakers' individual variations within our native language without difficulties, probably all of us have sometimes had serious problems coping with foreign-accented speech or regional dialects. 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, that is, accented components of his or her speech become better comprehended. 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, with a focus on one specific but natural feature of foreign-accented speech: a consistent substitution of one phoneme for another. 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., Eisner & McQueen, 2005; Kraljic & Samuel, 2005, 2006; Norris, McQueen, & Cutler, 2003). 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' (Goldstone 1998, p. 586). One example is listeners' adaptation to an artificially created sound, midway between [s] and [f]. Listeners are able to use their lexical knowledge when adapting to such a sound in an exposure phase, that is, interpreting the ambiguous sound 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 has been 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, & 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/, and were then asked to categorize phonemes on a /d/-/t/ continuum and those on a /b/-/p/ continuum. They 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 (Kraljic & Samuel, 2007). Moreover, the process is thought to be automatic (McQueen, Norris, et al., 2006). That is, attention to the acoustic-phonetic detail in the speech, in the form of

metalinguistic judgments about 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 (McQueen, Norris, et al., 2006), perceptual learning effects are observed. Whether or not participants make explicit decisions about the stimuli therefore does not affect the size of the learning effect. Adaptation thus appears to be an automatic (i.e., mandatory and signal-driven) process rather than a controlled (i.e., non-mandatory and attention-driven) process (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977).

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 mainly driven by the native language of the speaker, and in particular by the similarities and differences between the segmental inventories and the suprasegmental properties of the native language and those of the target foreign language. 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, Alexander, & Nygaard, 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, Weber, & McQueen, 2013).

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 induce possibly 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 and McQueen, 2006). Participants listened to an exposure story where either [f] or [s] was replaced by ambiguous sounds, 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 fairly consistently either replace them with phonemes from their L1 or introduce systematic subphonemic changes. Because of these stable properties of the speech signal, it would be beneficial for the perceptual system to adapt to the various categories for a certain speaker over a prolonged period of time. We will investigate here whether this is indeed possible for one specific vowel category, as a first step towards understanding adaptation to the multiple facets of a foreign accent.

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 three 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, in which quality and duration usually covary (e.g., for /i/ and /ɪ/). Note, however, that Dutch does have a vowel contrast (/a–a:/) which is primarily a durational distinction (Escudero, Benders, & Lipski, 2009; Kloots, Verhoeven, Coussé, & Gillis, 2010). 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 and Jang (1997)). In the current experiment, the Hebrew speaker shortened words with [i] to [ɪ] (specific accent marker words; for example Dutch *statief* /stati:f/, ‘tripod’ shortened to \*statif/). These words were contrasted with words spoken by the same Hebrew-accented speaker that canonically contain [ɪ] (globally accented words; e.g., Dutch *verstrikt* /vərstrikt/, ‘entangled’). The words with shortened vowels never created other 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, Manuscript under preparation). Moreover, both the [i] and [ɪ] are possible vowels in Dutch. We will refer to this contrast, for simplicity, as a length contrast, though there are quality differences too (see Table 1).

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 et al. (2013) 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,

**Table 1.** Average formant values (in Hz) at vowel midpoints for the globally accented and specific accent marker prime words, the same words spoken by a female native speaker of Dutch, and average values for female speakers of Southern Standard Dutch.

	Hebrew-accented speaker				Native speaker				Dutch [i]	Dutch [i]
	Underlying /i/ (globally accented [i])		Underlying /i/ (specific accent marker; /i/ as [i])		Underlying /i/		Underlying /i/			
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	Mean
F1	407	41	453	34	523	76	405	50	455	317
F2	2288	208	1954	136	2350	172	2741	64	2115	2647
F3	2842	161	2856	119	3170	125	3496	204	2948	3312

Source: Adank et al., 2004.

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 that listeners are very unlikely to be familiar with prior to the experiment.

In experiment 1, we wanted to establish a baseline for the 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 isolated Hebrew-accented primes, and then made lexical decisions to printed Dutch words and non-words. Priming effects were calculated by subtracting the reaction times (RTs) to targets that had been preceded by unrelated primes from those 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., Marslen-Wilson, Nix, & Gaskell, 1995; McQueen, Cutler, et al., 2006). 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, isolated words with statistically significant facilitatory priming, drawing from findings within our group, in particular those of Witteman et al. (2013) (see also e.g., Marslen-Wilson, Moss, & van Halen (1996), Marslen-Wilson & Zwitserlood (1989)). Priming effects will thus be taken as evidence that listeners correctly identified the accented primes as the intended Dutch words, and hence that the listeners had adapted to the Hebrew-accented speaker. The priming task provides a measure of word recognition that does not depend on the participant making a metalinguistic judgment about the accented word (as would be the case with a lexical decision task on the accented words themselves) and thus avoids some potential problems (e.g., with lexical decision, is a mispronunciation a word or not?). On the basis of the findings of Witteman et al. (2013), we expected listeners to be able to recognize the globally accented 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 specific accent marker 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 it, and whether adaptation is present in a delayed test phase. To avoid retesting effects, we designed a short phoneme monitoring task that 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.

## 2 Experiment 1: no prior exposure

### 2.1 Method

**2.1.1 Participants.** We tested 28 native speakers of Dutch (24 females, M age 21.3). These participants were recruited from the Max Planck 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. A language history questionnaire showed that none of the participants had any knowledge of Hebrew. In addition, although all participants identified the speaker as non-native when asked about the auditory stimuli after the experiment, and several noted that, in general, the speaker's vowels were not pronounced in a standard manner, none of the participants reported having noticed anything about the critical vowel substitution in particular. Furthermore, none guessed the native language of the speaker correctly. Their free response guesses, in which multiple answers were permitted, included German (17), French (4), Moroccan (2), Russian (2), Turkish (2), 'Eastern European' (1), Arabic (1) and 'Slavic' (1).

### 2.2 Materials

The cross-modal priming experiment contained 242 trials (50 experimental trials, 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 underlying long /i:/ (e.g., /stati:f/ *statief*, 'tripod'). In the remaining 30 experimental trials the target was a Dutch word with underlying short /ɪ/ (e.g., /vərstrikt/ *verstrikt*, 'entangled'). This difference in number of experimental items across conditions arose because we based our materials on an earlier study (Weber et al., Manuscript under preparation) and needed to keep some items from that study for the phoneme monitoring exposure phase (see experiments 2 and 3). There were sufficient words in both conditions, however, because as few as six items can be enough to show reliable priming effects (e.g., Witteman et al., 2013). Experimental targets were always paired with an identical and an unrelated prime. For targets with long /i:/, the vowel in identical primes was shortened to /ɪ/ (e.g., \*/statɪf/ – specific accent marker pronunciation), and for those with short /ɪ/, identical primes were produced in the global accent with short /ɪ/ (e.g., /vərstrikt/ – globally accented 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 non-word target. Another 64 filler trials also had a word prime and a non-word 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 non-word 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 in parallel. The factor *list* was therefore added to the analyses as a control variable.

### 2.3 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 long vowels rather than lengthen short vowels when we informally analysed 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], 2013). 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 a target /i:/, which was already her natural tendency. The instruction was provided only on the few tokens when she happened to naturally produce a sound closer to /i:/. Other research has indicated that non-native speakers are more variable than native speakers (Wade, Jongman, & Sereno, 2007) and in particular that they can be inconsistent, varying between accent-driven mispronunciations and correct target pronunciations (Hanulíková & Weber, 2012). Some inconsistency in the current speaker's vowels was thus to be expected. The elicitation and selection of pronunciations with consistent shortening of /i:/ to /ɪ/ was for reasons of experimental control and reflected the speaker's strong bias for vowel shortening. All words were checked for other obvious segmental mismatches by a native speaker of Dutch, and re-recorded if necessary in the same session. 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 (Sennheiser K6, Germany). 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 globally accented and the specific accent marker primes, as spoken by the accented speaker and by a native speaker of Southern Standard Dutch (with the correct pronunciations of /i:/ and /ɪ/) recorded in a separate session, as well as reference values for Dutch vowels [ɪ] and [i] for female speakers of Southern Standard Dutch (taken from Adank, Van Hout, & Smits (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 as well.

Note that the current comparison between globally accented and specific accent marker primes is between items, which is why observed differences can always be item-specific differences rather than a true difference between specific accent marker types. The duration of the vowels did not differ significantly for the two item types (M duration for globally accented words was 234 ms (SD



= 109), M duration for specific accent marker words was 271 ms (SD = 123);  $t(48) = 1.122, p > .2$ . When looking at the formants, specific accent marker [ɪ] had a higher first formant and a lower second formant than the globally accented [i], while the values for the third formant are similar. Importantly, however, the  $F2$  and  $F3$  values for our speaker's globally accented and specific accent marker forms are closer to Dutch [ɪ] than to Dutch [i], both for the native speaker's productions of the same sets of words and for a larger set of speakers on different words.

## 2.4 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 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. 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 (Sennheiser HD280-13, Germany). Participants saw the visual targets on a computer screen situated about 50 cm in front of them. Visual targets were presented in white lowercase 24-point 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., Berkeley, California) and controlled with NESU hardware (Nijmegen Experiment Set-Up, Nijmegen, The Netherlands). After the cross-modal priming experiment, participants were asked to fill out a language history questionnaire, including a free-response question in which they were asked to try to identify the native language of the speaker.

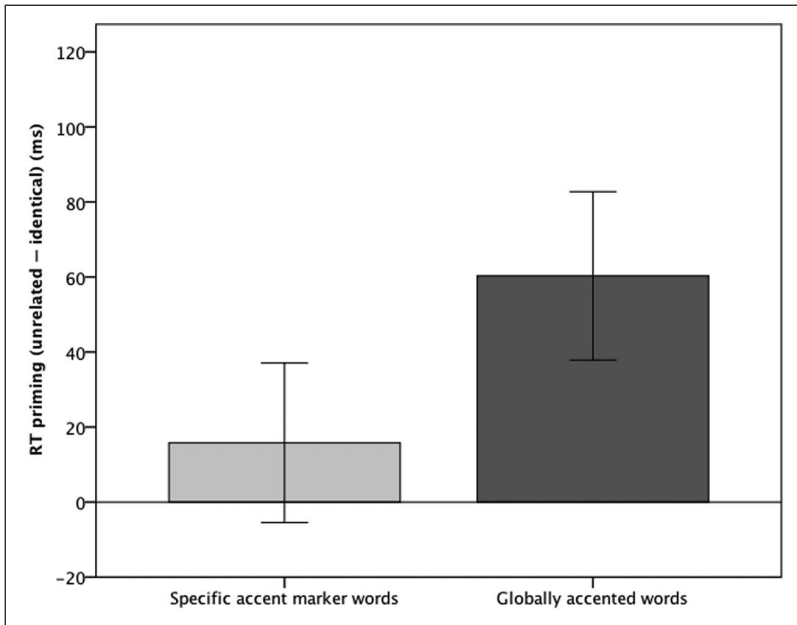
## 2.5 Results

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 analysed with general linear model (GLM) repeated measure analyses of variance (ANOVAs) using a 2 (*accent type* – globally accented, specific accent marker)  $\times$  2 (*priming* – identical, unrelated) design. Both factors were within participants. *List* was added as a between-participant factor. The results were analysed separately by participants ( $F1$ ) and items ( $F2$ ). The analyses with *list* only indicated that listeners were faster overall in one condition compared to the other; there were no significant interactions. Therefore, this factor was not further analysed nor described in the results below.

Of the trials, 3.4% were excluded due to errors or RTs that deviated more than 2.5 SD from the condition's overall mean. Errors were distributed evenly across conditions and items, and were not analysed 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, that is, 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  $F1(1,27) = 22.482, p < .001$ ;  $F2(1,19) = 5.958, p = .026$ ;  $\min F'(1,29) = 4.710, p = .038$ . Furthermore, the participant analysis revealed a main effect of *accent type*  $F1(1,27) = 5.658, p = 0.025$ ;  $F2 < 1$ ;  $\min F' < 1$ , indicating that responses were faster to globally accented than to specific accent marker words. The main effect of priming was further qualified by an interaction between *priming* and *accent type*  $F1(1,27) = 10.110, p = .004$ ;  $F2(1,19) = 3.262, p$



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

**Table 2.** Pairwise comparisons of priming effects for all accent types across participants and items.

Item type	Participant analysis			Item analysis		
	<i>df</i>	<i>t</i> 1	<i>p</i>	<i>df</i>	<i>t</i> 2	<i>p</i>
Specific accent marker words	1,27	1.526	.139	1,19	.793	.438
Globally accented words	1,27	5.515	.000	1,28	3.496	.002

= .087; min  $F'(1,31) = 2.466$ ,  $p > .05$ , indicating that the priming effects differed for the two accent types.

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

## 2.6 Discussion

Experiment 1 showed that without previous exposure to the Hebrew speaker of the experiment, Dutch participants were able to correctly interpret her globally accented words such as *verstrikt*, but not her specific accent marker 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 specific accent marker 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., 2013). 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 explicit, metalinguistic 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 the vowels, 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 specific accent marker 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 specific accent marker words, 10 globally accented 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 contained only 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); though for a variant on this, with only one test phase, see Kraljic & Samuel (2005)). It is thus possible that at least some 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 experiment 2, participants received only exposure on the first day, and were tested only after the delay.

If adaptation to variant words arises during phoneme-monitoring exposure and hence does not depend on sentential context to guide learning, then the specific accent marker words in experiment 2 should show facilitatory priming. Since globally accented words already showed facilitatory priming without any prior exposure in experiment 1, they should also show facilitatory priming in experiment 2 with additional exposure. An interesting question, however, is whether there is an increase in the amount of priming for the globally accented words across experiments. If so (i.e., if there is an increase in priming for both the specific accent marker words and the globally accented words), then this would suggest that adaptation to the Hebrew-accented speaker involves changes in perceptual processing that are not specific to the vowel-shortening feature of the accent. If, in contrast, there is an increase in priming across experiments only for the specific accent marker words, then this would suggest adaptation specific to this feature.

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., 2013), at the time of writing, none have asked 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 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 pause detection

and a word-spotting task revealed no effects immediately after exposure, but a significant inhibition effect after a day and a week, that is, after at least one night's sleep in which the knowledge was able to consolidate. In line with these results from memory consolidation, we expected to find that listeners' performance would improve after exposure and consolidation. Specifically, we expected that listeners would show priming to the globally accented words and the specific accent marker words after the delay.

## 3 Experiment 2: exposure 1 day before test

### 3.1 Method

**3.1.1 Participants.** We tested 20 native speakers of Dutch (16 females, M age 22.4). All participants were volunteers recruited from the Max Planck Institute for Psycholinguistics (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. As in experiment 1, no participant reported noticing the critical vowel substitution. Also, no participant guessed the native language of the speaker correctly, though all participants identified the speaker as non-native. Participants' free response guesses included French (7), German (4), Turkish (including 'Turkish/Moroccan') (3), Polish (2), Russian (2), 'Afghanistan' (1) and 'not Dutch, don't know which country' (1). All participants indicated that they thought both parts of the experiment were spoken by the same speaker.

### 3.2 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 specific accent marker pronunciation, that is, they contained the long vowel /i:/, which was shortened to /i/ (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 10 words had a globally accented pronunciation, that is, they contained the short vowel /ɪ/, which was produced in its canonical length (e.g., Dutch /film/ 'movie'). The remaining 50 words contained no /i:/ or /ɪ/ or any other long vowel. Therefore, the only exposure participants received to words with long vowels in their globally accented 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%).

### 3.3 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.

### 3.4 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 the language history questionnaire. This part of the experiment was identical to experiment 1.

## 3.5 Results

**3.5.1 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, with four errors (less than 95% correct), which was substantial compared to the other 20 participants.

**3.5.2 Cross-modal priming test.** We excluded 4.3% of trials due to errors or RTs that deviated more than 2.5 SD 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 analysed statistically. Results were analysed in the same way as in experiment 1.

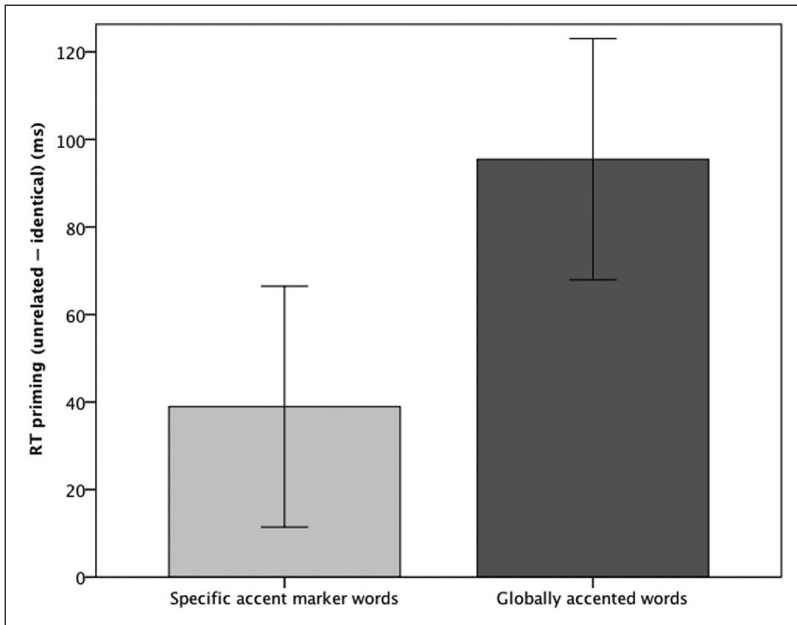
Calculated priming effects are shown in Figure 2; RTs and error rates per condition and accent type can be found in Appendix B. Participants were faster overall to respond to identical trials compared to unrelated trials  $F_1(1,19) = 71.640, p < .001$ ;  $F_2(1,19) = 31.264, p < .001$ ;  $\min F'(1,33) = 21.765, p < .001$ .

Participants were equally fast to respond to both accent types ( $F_1 < 1$ ;  $F_2 < 1$ ;  $\min F' < 1$ ). There was, however, an interaction between *priming* and *accent type*  $F_1(1,19) = 7.269, p = .014$ ;  $F_2(1,19) = 7.573, p = .013$ ;  $\min F'(1,38) = 3.709, p > .05$ , 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 specific accent marker and the globally accented words, but as the interaction indicated, priming was larger for globally accented forms than for specific accent marker forms.

## 3.6 Discussion

In contrast to experiment 1, experiment 2 showed significant priming effects for both the specific accent marker words and the globally accented words. The priming effects for the specific accent marker 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 specific accent marker 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 globally accented and the specific accent marker words correctly. If



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

**Table 3.** Pairwise comparisons of priming effects for all accent types across participants and items.

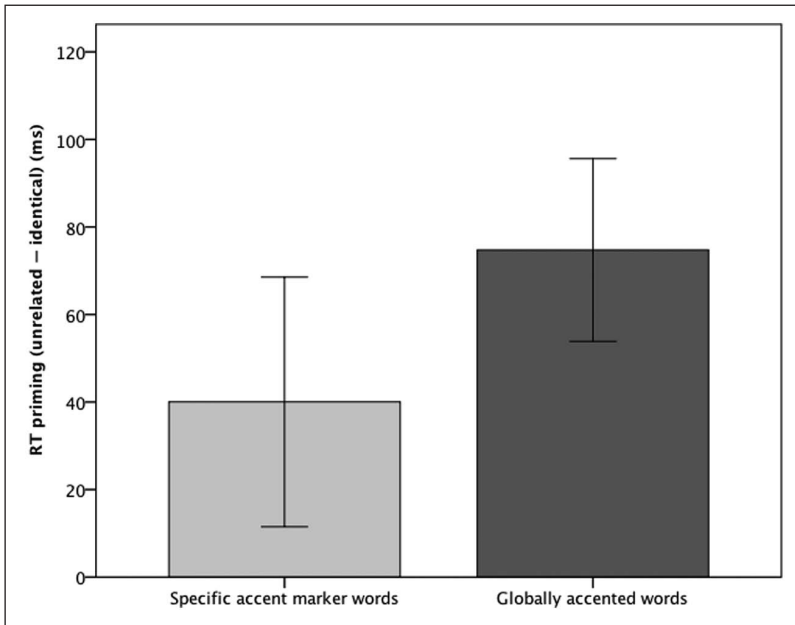
Item type	Participant analysis			Item analysis		
	<i>df</i>	<i>t</i> 1	<i>p</i>	<i>df</i>	<i>t</i> 2	<i>p</i>
Specific accent marker words	1,19	2.961	.008	1,19	2.420	.026
Globally accented words	1,19	7.256	.000	1,28	6.894	.000

the adjustment is to be beneficial for word recognition in foreign-accented speech, it should be stable over time.

## 4 Experiment 3: exposure 1 week before test

### 4.1 Method

**4.1.1 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. Participants' free response guesses included French (4), German (4), Turkish (including 'Turkish/Moroccan/something Arabic') (4), Polish (2), Russian (1), 'Asian' (1) and 'Eastern European' (1). As in both previous experiments, although most participants noted that the speaker's vowels were not standard pronunciations, none reported that they had noticed the critical vowel substitution.



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

## 4.2 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.

## 4.3 Results

**4.3.1 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).

**4.3.2 Cross-modal priming test.** We excluded 3.8% of trials due to errors or RTs that deviated more than 2.5 SD from the condition's overall mean. Error rates were distributed evenly across conditions and items (see Appendix B). Errors were again not analysed. The RT data was analysed in the same way as experiments 1 and 2.

The calculated priming effects are shown 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  $F_1(1,19) = 37.582, p < .001$ ;  $F_2(1,19) = 19.667, p < .001$ ;  $\text{min } F'(1,35) = 12.911, p < .001$ . There was no main effect of *accent type* ( $F_1 < 1$ ;  $F_2 < 1$ ;  $\text{min } F' < 1$ ), indicating that participants overall RTs to these two item types did not differ. There was an interaction between *priming* and *accent type* across participants  $F_1(1,19) = 5.482, p = .030$ ;  $F_2(1,19) = 3.042, p = .097$ ;  $\text{min } F'(1,35) = 1.956, p > .05$ , indicating that priming effects were larger for the globally accented words compared to the specific accent marker words.

**Table 4.** Pairwise comparisons of priming effects for all accent types across participants and items.

Item type	Participant analysis			Item analysis		
	<i>df</i>	<i>t</i> / <i>F</i>	<i>p</i>	<i>df</i>	<i>t</i> / <i>F</i>	<i>p</i>
Specific accent marker words	1,19	2.938	.008	1,19	2.256	.036
Globally accented words	1,19	7.495	.000	1,28	7.738	.000

The planned comparisons (Table 4) revealed that participants showed significant priming for both the specific accent marker and the globally accented words.

**4.3.3 Cross-experiment analysis.** A comparison of the results across all three experiments was then carried out to confirm that there was an effect of exposure, to ask whether the length of the exposure-test delay had an effect, and to test whether adaptation was limited to the specific accent marker words. *Experiment* was a between-subjects factor, while *accent type* and *priming* were within-subjects factors. *Experiment* and *priming* were within-item factors and *accent type* was a between-item factor. When all three experiments were included, the interaction between *priming* and *experiment* was significant  $F(1, 2, 65) = 3.267, p = .044, F(2, 46) = 5.896, p = .005, \text{min } F'(2, 109) = 2.102, p > .05$ . In contrast, when only experiments 2 and 3 were included in the analysis, there were no significant interactions involving *experiment*. This suggests first that the interaction in the overall analyses with three experiments reflects a performance difference between experiment 1 and the other two experiments, as the separate analyses for each experiment have also indicated. In other words, exposure enhanced the priming effect. Second, the lack of a difference between the latter two experiments shows that the length of the exposure-test delay had no effect on the extent of adaptation.

The analyses of all three experiments revealed no three-way interaction of *experiment*, *accent type* and *priming*, ( $F1 < 1, F2 < 1, \text{min } F' < 1$ ). This suggests that there was no difference in the increase in priming due to exposure between the specific accent marker words and the globally accented words. Both types of word show evidence of perceptual adaptation due to phoneme-monitoring exposure. While these effects are of similar magnitude, as the lack of an interaction suggests, there is a critical difference between the two conditions. The priming effect for the globally accented words was present in experiment 1 and became stronger after prior exposure, but for the specific accent marker words priming was absent in experiment 1 and emerged only after exposure.

#### 4.4 Discussion

In experiment 3, we found that participants could interpret both the specific accent marker words and the globally accented words, thereby replicating the result we found in experiment 2. We thus showed that the adaptation to the accented speaker remained stable for at least one week, even when initial exposure to the speaker was very limited. The cross-experiment analyses confirmed this view, and showed in addition that there was adaptation not only to specific features of Hebrew-accented Dutch but also to more global properties of the accent.

## 5 Discussion

The present study investigated whether adaptation to foreign-accented speech can be automatic, without a listener explicitly attending to it, and whether the adaptation is stable over time. Native



Dutch listeners performed a cross-modal priming task in which they showed that they could recognize globally accented Hebrew-accented words (without specific mispronunciations), but not specific accent marker words in which the 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 at least one speaker's foreign-accented speech can not only be quick, but also automatic and long-lasting.

Foreign-accented speech without substantial segmental mismatches does not seem to interfere substantially with understanding, at least as measured in the present study. This 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 specific accent marker items), a short exposure phase can be enough to adapt to these words, which is in line with previous research (Witteman et al., 2013). In that study, phonemes were replaced with other, non-native phonemes (the Dutch diphthong /œy/ substituted by the German diphthong [ɔɪ]). The present study shows for the first time that there can also be adaptation when the substitution involves a different phoneme from within the target language.

The fact that listeners can interpret specific accent marker words after a short phoneme monitoring exposure phase also indicates that their knowledge of the accent can be transferred across different tasks and does not require 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, and some experiments even 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. The present study indicates, however, 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., 2013), which differs in a number of ways from phoneme monitoring. First, a story provides a rich sentential context with more information than is available in isolated words. In particular, 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 required for 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 specific accent marker 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 those who listen to 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 the listener free to focus on the message the speaker wants to convey.

Adaptation to this type of foreign-accented speech seems to be automatic (i.e., mandatory and signal-driven, not dependent on attention; Schneider & Shiffrin (1977), Shiffrin & Schneider (1977)). 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., 2013), or retune their

existing category boundaries, like in the present study, adaptation is possible even after a short period 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 specific accent marker), while keeping the category of the globally accented vowels the same. Listeners are able to do this, as shown by the priming effects for the globally accented words. If listeners would apply a shift in category boundaries not just for the specific accent marker, but also for the globally accented words, interpretation difficulties for the globally accented 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 indefinitely. 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, or a similar, accent also 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: lexical-representational accounts and processing accounts. Lexical-representational models assume that variation is dealt with through lexical storage. They assume that the lexicon has entries not only for each word, but also for every variation on these words (e.g., Goldinger, 1998). Episodic representational models state that all variation is encoded in the lexicon including fine-grained phonetic detail (e.g., Johnson, 2006; Pierrehumbert, 2001). Alternatively, it has been proposed that the lexicon contains multiple abstract representations for variant forms (Ranbom & Connine, 2007).

The way in which lexical-representational accounts explain how listeners deal with the added variation of non-standard speech (e.g., foreign-accented speech) is thus that listeners store that variation in their lexicons. Upon hearing the variants a second time, adaptation could be achieved by re-accessing these forms. For the present experiment, this would mean that listeners could adapt to accented speech only if they had heard these specific variants before, that is, if they had had some prior experience with the accent. Experience with this accent is unlikely, however: Hebrew-accented Dutch is not common in The Netherlands, and none of the participants indicated that they had 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 had heard the accent before, we can assume that the level of experience was the same across experiments. This premise is difficult to combine with the findings presented here. In experiment 1, listeners did not show priming to specific accent marker words, whereas in experiments 2 and 3, they did. If the adaptation observed in the latter two experiments were due to exposure to Hebrew-accented Dutch prior to the experiment, then there ought to have been evidence of adaptation in experiment 1.

It is more plausible that the adaptation to the specific accent marker words arose due to the exposure phase in experiments 2 and 3 rather than prior experience. Lexical-representational models also cannot easily provide an explanation for this shorter-term learning effect, because although the exposure phase contained specific accent marker words, they were different words from those presented during the test phase. Without additional mechanisms, storage of one set of words with a given mispronunciation will not lead to generalization to other words with the same mispronunciation (Cutler, Eisner, McQueen, & Norris, 2010; McQueen, Cutler, et al., 2006). Minimally, a representational account would require phonological abstraction over stored lexical representations specifying that the mispronounced vowel in the exposure words is the same vowel as that in the test words.

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 to an accent how variations should be mapped on the (stored) canonical form (for similar accounts in other domains, see e.g., Lotto & Holt (2006), Mitterer, Csépe, Honbolygo, & Blomert (2006) and Gaskell & Marslen-Wilson (1998)). In the present case, lexical knowledge would be used to adjust the way vowels are mapped onto the lexicon (specifically, that the short vowel [ɪ] in the input needs to be mapped onto lexical representations of words with the long vowel [i]). Thus, information on how the speaker for this experiment pronounces words can be carried over from the exposure to the test phase, and does not depend on experience with the accent prior to the experiment. Perceptual learning studies in L1 have already demonstrated lexically-guided retuning of pre-lexical processing (Norris et al., 2003). If these lexically-driven adjustments are made to the way a sound is mapped onto the lexicon at a pre-lexical level of processing, this learning will generalize to all words in the lexicon that contain that sound (Cutler et al., 2010; McQueen, Cutler, et al., 2006; Sjerps & McQueen, 2010). This pre-lexical processing account thus provides the phonological abstraction that would need to be added to a lexical-representational account, at a stage of processing where lexical generalization will follow automatically, and without the need to store multiple pronunciation variants. This account also has the advantage that adaptation to foreign-accented speech would use the same mechanisms as have been proposed to account for adaptation to artificial and within-language variation.

Although this study was designed to examine adaptation to one specific vowel substitution in foreign-accented speech, it is important to note that the exposure manipulation had effects not only on the way the specific accent marker words were processed but also on the processing of the globally-accented words. Indeed, the increase in priming after exposure (i.e., from experiment 1 to experiments 2 and 3) was equally strong across the two conditions. This suggests that the participants were not tuning in only to the specific accent marker, but also to other features of Hebrew-accented Dutch. We suggest that the perceptual learning mechanism proposed for the shortened vowels (pre-lexical adjustments in vowel mappings) is also likely to apply to the adjustments to other accent characteristics. That is, the increase in priming with the globally accented words reflects adjustments in the mappings for a variety of segments. There is thus a qualitative change to the words with the specific accent marker (these words could not be recognized reliably without pre-exposure, but could after pre-exposure) and a quantitative change to the globally accented words (more robust recognition after pre-exposure), but both reflect the same kind of pre-lexical learning processes. We propose that these processes (and their characteristics of automaticity and stability) apply not only to the specific and global characteristics of Hebrew-accented Dutch studies here, but also to foreign-accented speech more generally. These processes do not seem, however, to apply when listeners hear native speakers produce vowel substitutions

characteristic of foreign-accented speech (Eisner, Weber, & Melinger, 2013) or when native speakers appear to have temporary disruptions of their speech (Kraljic, Samuel, & Brennan, 2008).

In sum, the perceptual system is highly flexible. The listeners in the current experiments did not show any measurable sign of difficulty when listening to foreign-accented speech without clear segmental substitutions. Foreign-accented speech with a segmental substitution caused some initial problems, but a short exposure phase was enough for the listeners to adapt even to these words. Moreover, though being exposed to foreign-accented speech was enough to improve performance, it was not necessary for the listeners to pay specific attention to a specific mispronunciation in order to adjust to it. Finally, once the listeners had adapted to a foreign-accented speaker, this effect remained stable for at least a week.

## Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

## References

- Adank, P., Van Hout, R., & Smits, R. (2004). An acoustic description of the vowels of Northern and Southern Standard Dutch. *The Journal of the Acoustical Society of America*, *116*, 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*, 283–293.
- Bent, T., & Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. *Journal of the Acoustical Society of America*, *114*, 1600–1610.
- Boersma, P., & Weenink, D. (2009). Praat: doing phonetics by computer (Version 5.1. 05) [Computer program]. Last retrieved on 31 October 2009 from <http://praat.org>.
- 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*, 977–985.
- Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition*, *106*, 707–729.
- Central Bureau voor de Statistiek (Central Bureau for Statistics, 21 May, 2013). *Bevolking: generatie, geslacht, leeftijd en herkomstgroepering, 1 januari*. Retrieved from <http://statline.cbs.nl/StatWeb/publication/?DM=SLNL&PA=37325&D1=a&D2=a&D3=0&D4=0&D5=97&D6=9-17&VW=T>
- Clarke, C. M., & Garrett, M. (2004). Rapid adaptation to foreign accented speech. *Journal of the Acoustical Society of America*, *116*, 3647–3658.
- Cutler, A., Eisner, F., McQueen, J.M., & Norris, D. (2010). *How abstract phonemic categories are necessary for coping with speaker-related variation*. In C. Fougerson, B. Kühnert, M. D'Imperio, & N. Vallée (Eds.), *Laboratory phonology 10* (pp. 91–111). Berlin: de Gruyter.
- 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*, 803–820.
- Dumay, N., & Gaskell, M. (2012). Overnight lexical consolidation revealed by speech segmentation. *Cognition*, *123*, 119–132.
- Eisner, F., & McQueen, J. M. (2005). The specificity of perceptual learning in speech processing. *Attention, Perception, & Psychophysics*, *67*, 224–238.
- Eisner, F., & McQueen, J. M. (2006). Perceptual learning in speech: Stability over time. *The Journal of the Acoustical Society of America*, *119*, 1950–1953.
- Eisner, F., Weber, A., & Melinger, A. (2013). Constraints on the transfer of perceptual learning in accented speech. *Frontiers in Psychology*, *4*, 148. doi: 10.3389/fpsyg.2013.00148
- Escudero, P., Benders, T., & Lipski, S. (2009). Native, non-native and L2 perceptual cue weighting for Dutch vowels: The case of Dutch, German, and Spanish listeners. *Journal of Phonetics*, *37*, 452–466.

- 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, 437–470.
- Gaskell, M. G., & Marslen-Wilson, W. D. (1998). Mechanisms of phonological inference in speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 380–396.
- Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological review*, 105, 251–279.
- Goldstone, R. L. (1998). Perceptual learning. *Annual review of psychology*, 49, 585–612.
- Gussenhoven, C. (1999). Illustrations of the IPA: Dutch. In *Handbook of the International Phonetic Association* (pp. 74–77). Cambridge, UK: Cambridge University Press.
- Hanulíková, A., & Weber, A. (2012). Sink positive: Linguistic experience with th substitutions influences nonnative word recognition. *Attention, Perception & Psychophysics*, 74, 613–629.
- Johnson, K. (2006). Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of Phonetics*, 34, 485–499.
- Kloots, H., Verhoeven, J., Coussé, E., & Gillis, S. (2010). The role of duration in vowel categorization: Flemish vs. Dutch listeners: a case study. *Studies van de BKL*, 5, 1–14.
- Kraljic, T., & Samuel, A. G. (2005). Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology*, 51, 141–178.
- Kraljic, T., & Samuel, A. G. (2006). Generalization in perceptual learning for speech. *Psychonomic Bulletin & Review*, 13, 262–268.
- Kraljic, T., & Samuel, A. G. (2007). Perceptual adjustments to multiple speakers. *Journal of Memory and Language*, 56, 1–15.
- Kraljic, T., Samuel, A. G., & Brennan, S. E. (2008). “First impressions and last resorts: How listeners adjust to speaker variability.” *Psychological Science* 19, 332–338.
- Laufer, A. (1999). Illustrations of the IPA: Hebrew. *Handbook of the International Phonetic Association* (pp. 96–99). Cambridge, UK: Cambridge University Press.
- 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, 874–886.
- Lotto, A. J., & Holt, L. L. (2006). Putting phonetic context effects into context: A commentary on Fowler (2006). *Attention, Perception, & Psychophysics*, 68, 178–183.
- 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, 1376–1392.
- 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, 285–308.
- Marslen-Wilson, W. D., & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 576–585.
- Maye, J., Aslin, R. N., & Tanenhaus, M. K. (2008). The weckud wetch of the wast: Lexical adaptation to a novel accent. *Cognitive Science*, 32, 543–562.
- McQueen, J. M., Cutler, A., & Norris, D. (2006). Phonological abstraction in the mental lexicon. *Cognitive Science*, 30, 1113–1126.
- McQueen, J. M., Norris, D., & Cutler, A. (2006). The dynamic nature of speech perception. *Language and Speech*, 49, 101–112.
- 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, 451–479.
- Most, T., Amir, O., & Tobin, Y. (2000). The Hebrew vowel system: Raw and normalized acoustic data. *Language and Speech*, 43, 295–308.
- Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology*, 47, 204–238.
- Pierrehumbert, J. (2001). Why phonological constraints are so coarse-grained. *Language and Cognitive Processes*, 16, 691–698.
- Ranbom, L. J., & Connine, C. M. (2007). Lexical representation of phonological variation in spoken word recognition. *Journal of Memory and Language*, 57, 273–298.

- Schneider, W., & Shiffrin, R. M. (1977) Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, 84, 1–66.
- Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, 84, 127–190.
- Sidasas, S. K., Alexander, J. E. D., & Nygaard, L. C. (2009). Perceptual learning of systematic variation in Spanish-accented speech. *The Journal of the Acoustical Society of America*, 125, 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, 17.
- 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, 14356–14360.
- 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, 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, 122–144.
- Weber, A., Di Betta, A. M., & McQueen, J. M. (Manuscript under preparation). Treak or trit: Adaptation to genuine and arbitrary foreign accents.
- Witteman, M. J., Weber, A., & McQueen, J. M. (2013). Foreign accent strength and listener familiarity with an accent co-determine speed of perceptual adaptation. *Attention, Perception & Psychophysics*. Advance online publication. doi: 10.3758/s13414-012-0404-y

### Appendix A. List of experimental items

Specific accent marker words.

Target (Dutch spelling)	Identical prime IPA	Unrelated prime	Target translation	Prime translation
advies	ɑtʃis	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	spuut	sweet	syringe
mier	mɪr	wraak	ant	revenge
mythe	mɪtə	afwas	myth	dishes
niezen	nɪzə	erker	sneeze	oriel
papier	pɑpɪr	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	boekket	tripod	bouquet
textiel	tɛkstɪl	roman	textile	novel
ventiel	vɛntɪl	tapijt	valve	carpet
vierkant	vɪrkɑnt	bloesem	square	blossom
vriezen	vri:zə	foto	freeze	photo

(Continued)

**Appendix A.** (Continued)

Globally accented words.

Target (Dutch spelling)	Identical prime IPA	Unrelated prime	Target translation	Prime translation
bacil	basil	soldaat	bacillus	soldier
bassist	bassɪst	factuur	bass player	invoice
blik	blik	staan	stand	can
bliksem	bliksem	eikel	acorn	lightning
blind	blɪnt	rouw	mourning	blind
cirkel	sɪrkəl	hertog	duke	circle
conflict	kɔnflɪkt	sopraan	conflict	soprano
cursist	kʏrsɪst	tabak	student	tobacco
delict	dəlɪkt	buffet	offence	buffet
dicht	dɪxt	jas	coat	closed
ding	dɪŋ	kruis	thing	cross
gravin	xravn	dolfijn	countess	dolphin
hitte	hɪtə	mosterd	mustard	heat
inkt	ɪŋkt	luik	hatch	ink
klimmen	klimə	omroep	broadcasting	climb
*klip	klɪp	hand	cliff	hand
knikker	knɪkər	wortel	carrot	marble
kokkin	kɔkɪn	natuur	(female) chef	nature
lift	lɪft	markt	elevator	market
mist	mɪst	koorts	fog	fever
pinda	pɪnda	rooster	schedule	peanut
prins	pɪns	storm	prince	storm
pupil	pʏ:pɪl	framboos	pupil	raspberry
rimpel	rɪmpəl	stempel	stamp	wrinkle
rits	rɪts	herfst	fall	zipper
stilte	stɪltə	kaneel	cinnamon	silence
toerist	tʊrɪst	lakei	tourist	lackey
verstrikt	vərstrɪkt	bestuur	entangled	government
vinger	vɪŋər	koepel	dome	finger
winter	wɪntə	keuren	examine	winter

\*Discarded from all analyses.

**Appendix B.** Overview of RTs and error rates

Overview of RTs (SD) and error rates for all conditions and experiments.

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