Spoken-word recognition in foreign-accented speech by L2 listeners

Andrea Weber [a,b,*], Mirjam Broersma [a,b], Makiko Aoyagi [c]

[a] Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, The Netherlands
[b] Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, The Netherlands
[c] Dokkyo University, Saitama, Japan

ARTICLE INFO

ABSTRACT

Two cross-modal priming studies investigated the recognition of English words spoken with a foreign accent. Auditory English primes were either typical of a Dutch accent or typical of a Japanese accent in English and were presented to both Dutch and Japanese L2 listeners. Lexical-decision times to subsequent visual target words revealed that foreign-accented words can facilitate word recognition for L2 listeners if at least one of two requirements is met: the foreign-accented production is in accordance with the language background of the L2 listener, or the foreign accent is perceptually confusable with the standard pronunciation for the L2 listener. If neither one of the requirements is met, no facilitatory effect of foreign accents on L2 word recognition is found. Taken together, these findings suggest that linguistic experience with a foreign accent affects the ability to recognize words carrying this accent, and there is furthermore a general benefit for L2 listeners for recognizing foreign-accented words that are perceptually confusable with the standard pronunciation.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

"Hop on, dere is noting like seeing Emsterdem from de kennels". With this sentence Dutch businessmen routinely lure English tourists into taking a boat trip to see the marvels of the city. While it may take a moment for the tourist to figure out that he is invited to a boat trip on the canals and not to a stay in a doghouse, a Dutchman passing by may have no difficulties understanding his fellow Dutch speaker immediately. His own accent in English probably sounds similar, and he is familiar with the typical Dutch way of pronouncing English. Over the years, he has had the chance to acquire a detailed representation of Dutch-accented English. But does this indeed entail that he has learned to immediately interpret Dutch-accented words correctly? And is he therefore better at recognizing words that are spoken with a Dutch accent than at recognizing words with a different foreign accent? The present study set out to empirically address these questions in two cross-modal priming studies in which accented words either conformed with the language background of non-native (L2) listeners or not.

Learning (near-) native pronunciation is often seen as the most prominent and persistent difficulty for adult second language learners (e.g., Bongaerts, van Summeren, Planken, & Schils, 1997; Hyltenstam & Abrahamsson, 2003). This implies that even learners with a large vocabulary in the second language and a good command of grammar usually introduce considerable pronunciation variation to the speech signal. The deviations from the pronunciation norms of the target language reflect to a large extent the phonological structure of the speaker's native language, and can range from subtle phonetic deviations, to segmental substitutions, deletions, and insertions, as well as to prosodic differences. A consequence of the native language influence is that the deviations from the target language norm will vary for L2 speakers with different language backgrounds. For example, while Dutch speakers tend to produce the English vowel /æ/ as /ɛ/ (e.g., Tops, Dekeyser, Devriendt, & Geukens, 2001), Japanese speakers typically produce it as /a/ (e.g., Thompson, 2001). In order to understand foreign-accented words, listeners have to learn to handle this type of variation. The fundamental question is then: how do listeners deal with this variability, and how can they arrive at a stable percept from a signal that varies so markedly from the target language norm?

Most things we know about how listeners handle segmental variation in word recognition come from studies on L1 speech with L1 listeners. These studies looked either at arbitrary variation (e.g., Connine, Blasko, & Titone, 1993; Marslen-Wilson, 1993; Marslen-Wilson & Zwitserlood, 1989; Whalen, 1991) or at phonologically ruled variation (e.g., Gaskell & Marslen-Wilson, 1996, 1998; Gow, 2002, Mitterer & Blomert, 2003). Using a battery of tasks, the studies on arbitrary variation have shown, inter alia, that word recognition is sensitive to sub-phonetic mismatches: *soup*, for instance, is recognized more slowly when formant transitions following the /s/ are typical of a different fricative (Whalen, 1991). Mismatches on the phonetic level, such as the non-word *pomato* for English *tomato* (Marslen-Wilson, 1993), can activate the citation form less strongly, or as in the case of Dutch *woning* for *honing*, not at all (Marslen-Wilson & Zwitserlood, 1989). Connine et al. (1993)

* Corresponding author at: Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, The Netherlands. Tel.: +31 24 3521 585; fax: +31 24 3521 213.
E-mail address: andrea.weber@mpi.nl (A. Weber).

have furthermore shown that the more dissimilar the mispronounced sound is to the correct pronunciation, the stronger lexical access is disrupted (see also Marslen-Wilson, Moss, & van Halen, 1996): while *zervice* (voice change) still facilitates the recognition of *service*, *gervice* (voice, place, and manner change) does not. The general outcome of this line of research is that arbitrary mismatches interfere with word recognition, but that word recognition fails only in case of large phonetic dissimilarity between intended and mispronounced sounds.

Research on naturally occurring phonologically ruled variation has shown, on the other hand, that the context in which the variation occurs plays a role. A number of studies, for instance, have found that assimilated word forms, like *greem* for English *green*, are recognized correctly only in assimilation-licensing contexts such as *greem bench* where the /m/ of *greem* and the /b/ of *bench* match in place of articulation (e.g., Coenen, Zwitserlood, & Boelte, 2001; Gaskell & Marslen-Wilson, 1996, 1998; Gow, 2001, 2002; Mitterer & Blomert, 2003). Similarly, reduced word forms like *posman* for *postman* (Ernestus, Baayen, & Schreuder, 2002; Mitterer & Ernestus, 2006; Sumner & Samuel, 2005) or resyllabified forms (e.g., Spinelli, McQueen, & Cutler, 2003; Vroomen & de Gelder, 1999) are recognizable in contexts that license the surface change. All these findings suggest that listeners can recover the underlying form from surface variation, but restoration only happens if the variation occurs in viable contexts.

Recently, a number of studies have started to look at effects of frequency of occurrence on variant recognition in L1. For some types of phonological variation, differences in dominancy exist for variant forms, and listeners have been found to be sensitive to the distribution with which pronunciation variants occur in the environment of their native language (e.g., Connine, Ranbom, & Patterson, 2008; Mitterer & Ernestus, 2006; Pitt, 2009; Ranbom & Connine, 2007; Ranbom, Connine, & Yudman, 2009). For instance, English listeners recognize English words with schwa deletion, like in *corprate* for *corporate*, more easily when the word occurs often with schwa deletion than when it occurs only rarely with schwa deletion in their language (Connine, 2004; Connine et al., 2008). Similarly, Dutch listeners are more likely to infer the presence of a word-final /t/ in words in which t-deletion has a higher likelihood than in words with low likelihood (Mitterer & Ernestus, 2006).

Rather than comparing variant forms with different frequencies of occurrence, some researchers have taken the approach of comparing listener groups with different levels of exposure to variant forms to study experiential effects on immediate variant recognition (e.g., Adank & McQueen, 2007; Scott & Cutler, 1984; Sumner & Samuel, 2009). All these studies looked at variation caused by regional accents, and they showed, among other things, that native listeners who are familiar with a regional accent can recognize accented words quite easily, but native listeners who are unfamiliar with the regional accent have difficulties. For instance, Dutch listeners are slower at making animacy decisions for Dutch words spoken with an unfamiliar regional accent than for words in standard Dutch (Adank & McQueen, 2007). Likewise, Sumner, and Samuel (2009) have shown with a cross-modal priming task that words with a New York City accent (i.e., with r-dropping in words like *beaker*) can facilitate word recognition for native listeners who are familiar with New York city English, but not for native listeners who are unfamiliar with the accent; but even for listeners who are familiar with the New York city accent, words spoken in the General American accent (i.e., without r-dropping) facilitated recognition more strongly than words spoken with the New York city accent.

In the research presented here, we want to expand our knowledge about the influence of linguistic experience on variant word recognition by considering a different type of speech, namely foreign-accented speech, and by testing a different group of

listeners, namely L2 listeners. In contrast to regional accents, foreign accents introduce variability to the speech signal that is driven by the phonology of a second language. In general, two languages overlap less in their phonology than two varieties of one language do. Floccia, Butler, Goslin, and Ellis (2009), for example, have argued that foreign accents tend to affect all phonemes, whereas dialects mainly affect vowels. With respect to perceived accentedness, regional accents are also judged to be more similar to the standard variety of a language than foreign accents are (Floccia, Goslin, Girard, & Konopcynski, 2006). But although foreign-accented speech may deviate more strongly from the standard pronunciation than native accents do, the task of the listener stays the same, namely to recognize intended words despite their variability. It is not unreasonable therefore to assume that the ease of recognition in L2 speech will be modulated by long-term listening experience just as it has been shown for L1 speech.

L2 listeners who have learned a second language later in life happen to naturally form a class of listeners who vary in experience with foreign accents: they usually have ample experience with their own accent and considerably less experience with the accent of L2 speakers with different language backgrounds. Especially, L2 learners who live in their L1 country will have heard their accent in their own L2 speech as well as in the L2 speech of fellow citizens quite regularly. Thus, if listening experience, be it from one's own production or from listening to others, influences the ease of word recognition in foreign-accented speech, then we should find that L2 listeners can recognize words more easily when the words are pronounced in a way typical for their own accent; recognizing words with a foreign accent that does not match with the language background of the listener, on the other hand, should be more difficult.

In contrast to L1 listening, L2 listening is known to be hampered by inaccurate processing of phonemes, a fact that might further influence variant recognition by L2 listeners. The issue of how L2 phonemes are perceived has for many years attracted enormous interest and has produced a huge output of cross-language research (for an overview see for instance Strange (1995) or Bohn and Munro (2007), and last but not least this special issue). The central finding of this line of research is that the phoneme contrasts of the L1 determine how contrasts in the L2 are interpreted. Interpretation is usually inaccurate when the phoneme contrasts required by the L2 differ from the contrasts which distinguish words in the L1: for example, when two non-native sounds are equally close to one native category, or when neither one falls within a native category, discrimination is usually poor, but when two non-native sounds map onto different native categories, discrimination is very good (e.g., Best, 1995; Flege, 1995; Kuhl, 1993). Perception of L2 phonemes can also be less precise than that of L1 phonemes; listeners may be able to recognize them correctly some of the time, but less consistently than native listeners do. The inability to distinguish L2 phoneme categories accurately has been shown to impact L2 word recognition in multiple ways. For example, it can blur the distinction between minimal pairs, such as *flesh* and *fresh* for Japanese listeners (Cutler & Otake, 2004; for Dutch listeners to English see Broersma, 2002, submitted for publication; Cutler & Broersma, 2005; for Spanish listeners to Catalan see Pallier, Colomé, & Sebastián-Gallés, 2001; Sebastián-Gallés, Echeverría, & Bosch, 2005; Sebastián-Gallés, Rodríguez-Fornells, de Diego-Balaguer, & Díaz, 2006), or it can induce temporary ambiguity, with unwanted candidate words not being discarded at the earliest opportunity, when for example Japanese listeners cannot distinguish the onset of *rocket* from the onset of *locker* (Cutler, Weber, & Otake, 2006; for Dutch listeners to English see Broersma, submitted for publication; Weber & Cutler, 2004). The general outcome of these studies is that processing difficulties with L2 sounds can lead to the activation of unintended L2 words (e.g., activation of *flesh* when the speaker says

fresh), and this erroneous activation slows down the recognition of the intended word.[1]

This processing handicap may turn, however, into a benefit when L2 listeners listen to foreign-accented speech: in cases when the L2 speaker mispronounces a segment in a word, and the L2 listener has difficulties distinguishing the deviation from the standard pronunciation, this listener may activate the intended word nevertheless quite strongly, regardless of the deviation. Coming back to the introductory example of a Dutch speaker inviting tourists to see Amsterdam from the kennels, a Japanese tourist passing by may also have no difficulties interpreting the word correctly, since Japanese has neither the first vowel in *kennel* nor the one in *canal* and the Japanese tourist may therefore not notice the mispronunciation. In other words, spoken-word recognition in L2 speech by L2 listeners may not only be facilitated by language-specific experience with accented word forms but also by inaccurate phonemic processing.

While we do not know much yet about the online processing of foreign-accented words, recently a number of studies have investigated the offline intelligibility of foreign-accented speech. In these studies, L1 and L2 listeners are usually presented with foreign-accented sentences (e.g., Bent & Bradlow, 2003; Hayes-Harb, Smith, Bent, & Bradlow, 2008; Leikin, Ibrahim, Eviatar, & Sapir, 2009; Munro & Derwing, 1999; Munro, Derwing, & Morton, 2006; Stibbard & Lee, 2006; van Wijngaarden, 2001; van Wijngaarden, Steeneken, & Houtgast, 2002) or words (Imai, Walley, & Flege, 2005) and their task is to write down what they heard. The general outcome of these studies is that L1 listeners usually find unaccented L1 speech more intelligible than L2 speech, but L2 listeners can find L2 speech at least as intelligible as L1 speech. Interestingly, some studies have even found that L2 listeners find L2 speech from speakers with varying language background equally intelligible (e.g., Bent & Bradlow, 2003; Munro et al., 2006; van Wijngaarden, 2001; van Wijngaarden et al., 2002), although others studies have failed to find such mutual intelligibility of foreign-accented speech (e.g., Leikin et al., 2009; Stibbard & Lee, 2006). In these studies, segmental characteristics of words to be identified are typically not controlled, and a possible explanation for the discrepancy in findings is therefore that the number of keywords with similar perceptual difficulties for different L2 listeners varied between studies.

In the two cross-modal priming experiments presented here, we control segmental characteristics of foreign-accented words and present these words to L2 listeners whose language background either matches the foreign accent or not. Specifically, we present Dutch and Japanese learners of English with English word forms that are either typical of a Dutch accent in English (e.g., /mu:f/ for English /mu:v/, *move*) or typical of a Japanese accent in English (e.g., /mu:bu/ for English /mu:v/, *move*). In a cross-modal priming task, listeners are presented with an auditory prime word (e.g., /mu:f/), which is followed by the presentation of a visual target (e.g., MOVE). Their task is to decide whether the visual target is a real word or not, and reaction times (RTs) to visual targets are known to be faster when prime and target are phonologically related compared to when they are unrelated. As such, RTs to visual targets indicate how listeners interpret auditory primes. If experience with one's own

accent influences the ease of word recognition in L2, then we should find that Dutch-accented /mu:f/ facilitates recognition of MOVE for Dutch listeners, and Japanese-accented /mu:bu/ facilitates recognition of MOVE for Japanese listeners. If perceptual difficulties furthermore influence word recognition processes in L2 speech, it is possible that Dutch-accented /mu:f/ also has a facilitatory effect for Japanese listeners because for them the deviating /f/ is difficult to distinguish from the target /v/, but Japanese-accented /mu:bu/ should not have the same effect for Dutch listeners.

Because it is very difficult to exactly match two speakers for intelligibility and speaking rate (and both factors influence RTs), we decided to have a Dutch and a Japanese speaker record all primes, and present only primes from one speaker within a given version of an experiment. This means of course that the primes that were untypical of the accent of the speaker (i.e., the Japanese-accented primes for the Dutch speaker and the Dutch-accented primes for the Japanese speaker) did not sound exactly the way they sound when they are produced by a native speaker of that accent. For example, Dutch speakers may exert syllable compression and reduction when a syllable is added in /mu:bu/ in order to maintain a foot-rhythm structure, whereas Japanese speakers tend to maintain vowel durations with or without the added vowel (Tanaka, 2009); therefore those primes were only typical of an accent in the sense that they contained segmental deviations that are accent typical. While there is some evidence that listeners can hear whether an accent is imitated or authentic (Markham, 1999; Tate, 1979; but see Neuhauser & Simpson, 2007), there is to our knowledge no research available on how authenticity affects spoken-word recognition. Predictions on the role of speaker language background were therefore somewhat uncertain. While we assumed that speaker language background would not significantly change the pattern of results, it remained possible that a match between speaker language background and accent form was a prerequisite for any facilitatory effects on word recognition to emerge (and a mismatch was sufficient to disable facilitation).

2. Experiment 1

2.1. Method

2.1.1. Participants

Sixty-six native speakers of Dutch (Experiment 1a) and 30 native speakers of Japanese (Experiment 1b) participated in this experiment. Dutch participants were recruited at Radboud University (the Netherlands), and Japanese participants at Dokkyo University (Japan). All participants had normal hearing, normal or corrected-to-normal vision, and none reported a reading disability; they were volunteers and received a small fee for participation. Dutch participants had received an average of 7 years of training in English as foreign language in secondary education, starting at a mean age of 11. Japanese participants had received an average of 9 years of training in English as foreign language with a mean starting age of 11. At the time of testing, Dutch participants were on average 21 years of age, and Japanese participants 23 years. Following the cross-modal priming experiment, participants completed the Lexical Test for Advanced Learners of English *LexTALE* (Lemhöfer & Broersma, 2009, submitted for publication), a visual lexical-decision test with a high sensitivity for vocabulary size of very advanced learners of English. Dutch participants had an average score of 74% correct and Japanese participants of 63%, confirming their high proficiency in English. Thirty-six of the Dutch participants were presented with the Dutch recordings of the stimuli, and the remaining 30 Dutch participants with the Japanese recordings. All 30 Japanese participants listened to the Dutch recordings of the stimuli.

---

[1] Note that L2 pronunciations often but certainly not always, reflect the speaker's perception difficulties with non-native contrasts: for instance, for both Dutch and German learners of English, /f/ is perceptually most confusable with /θ/, but Germans learners prefer to substitute /θ/ with /s/ and Dutch learners prefer /t/ (see e.g., Hanulikova & Weber, 2010). There is ample research trying to explain why certain substitutions are preferred by L2 learners, and while the phonological structure of the L1 certainly can be an important factor in explaining different substitutions, other factors such as word-dependent characteristics, social factors, varying teaching curricula, and phonetic differences have been proposed to influence the choice of L2 substitutions as well (e.g., Brannen, 2002; Picard, 2002; Teasdale, 1997).

2.1.2. Materials

Twenty-four English mono- and disyllabic words were chosen as visual targets, 12 containing the vowel /æ/ (e.g., *act* /ækt/; henceforth called Vowel items) and the other 12 ending in the consonant /v/ (e.g., *move* /mu:v/, henceforth called Consonant items). Although a number of studies using various methodologies have indicated that vowels and consonants behave differently in language processing (e.g., Costa, Cutler, & Sebastián-Gallés, 1998; van Ooijen, 1996), the distribution in the present study was between items and was not meant to be informative about their role in processing. The main reason for including Vowel and Consonant items in the present study was to test more than one type of variant form. English target words were judged to be familiar for Dutch and Japanese learners of English at university level, and had an average lemma frequency of 125 per million in the *CELEX* database (Baayen, Piepenbrock, & Van Rijn, 1993).

Each visual target word had three auditory prime words: a Dutch-accented prime, a Japanese-accented prime, and an unrelated prime. The first two were phonological variants of the standard pronunciation of the target, with the variant being either typical of a Dutch accent in English or of a Japanese accent in English. For Dutch-accented primes, the vowel /æ/ was replaced with /ɛ/ in Vowel items (*act* pronounced as /ɛkt/), and the word-final /v/ was replaced with /f/ in Consonant items (*move* pronounced as /mu:f/). Dutch lacks the vowel /æ/ in its phoneme inventory (e.g., Gussenhoven, 1999), and it is well known that Dutch learners of English tend to perceive and produce English /æ/ as /ɛ/, a vowel that also exists in Dutch (Broersma, 2005; Cutler & Broersma, 2005; Cutler, Weber, Smits, & Cooper, 2004; Schouten, 1975; Tops et al., 2001; Weber & Cutler, 2004). Similarly, the entire class of word-final obstruents gets devoiced in Dutch, and Dutch learners characteristically devoice final obstruents also in English (Tops et al., 2001). For Japanese-accented primes, the vowel /æ/ was replaced with /a/ in Vowel items, and depending on the final consonant of the target, the vowel /o/, /i/, or /u/ was appended (*act* pronounced as /ˈakto/). For Japanese-accented Consonant items, the word-final /v/ was replaced with /b/, and the vowel /u/ was suffixed (*move* pronounced as /ˈmu:bu/). Just like Dutch, Japanese lacks the English vowel /æ/ in its inventory; unlike Dutch, however, Japanese distinguishes only five vowel qualities (Homma, 1992; Ladefoged, 1993; Shibatani, 1990), and Japanese learners of English tend to perceive and produce English /æ/ as /a/ (Aoyama, Flege, Guion, Akahane-Yamada, & Yamada, 2003; Carruthers, 2006; Nishi & Kewley-Port, 2007). Likewise, Japanese does not have labiodental fricatives, and Japanese learners often produce /b/ instead of English /v/ (Avery & Ehrlich, 1992; Thompson, 2001). Finally, the basic syllable structure in Japanese is CV, and in their attempt to make L2 words conform to the Japanese structure, Japanese learners commonly use vowel paragoge in English (Avery & Ehrlich, 1992). In principle, /u/ is added after most consonants as in *move* /ˈmu:bu/,[2] /o/ after the alveolar stops /t/ and /d/, and /i/ after the palatal affricates /tʃ/ and /dʒ/ (Lovin, 1974 cited in Carruthers, 2006; Tanaka, 2009). Thus, whereas Dutch-accented primes differed from the standard pronunciation of the English target word in either one vowel or one consonant, Japanese-accented primes differed in a vowel or consonant, as well as in the addition of a final vowel. Note that while some of the Japanese-accented primes exist as English loanwords in Japanese, others are not fully naturalized. It is not straightforward to determine loanword status in Japanese. First, there are no suitable dictionaries available providing reliable information about loanword status (Inoue, 2005). More importantly, the

use of loanwords varies heavily between individuals and groups, often related to the use of jargon. Thus, most of the selected items seem to fall within a gray area: they might be common loanwords for some speakers but never be used as such by others. While our predictions are the same for loanwords and non-loanwords, it is possible that Japanese listeners are more familiar with the former and/or accept loanword primes less readily when produced by a non-Japanese speaker; in other words, the strength of observed effects may vary somewhat with loanword status.

As control primes, phonologically and semantically unrelated words, matching in syllable number, were selected for each target word. To avoid obvious accent markers on unrelated primes, care was taken that unrelated primes contained as few phonological substitutions typical for a Dutch or a Japanese accent as possible. For instance, unrelated primes contained almost no consonant clusters, never had an /r/, and ended either with a vowel or with a nasal that did not require vowel paragoge in Japanese-accented English. For a full list of the stimuli used in Experiment 1, see the Appendix.

For filler trials, 48 English words and 72 pseudo-words were selected as visual targets. Visual targets were then combined with different types of auditory primes: 24 of the words and 24 of the pseudo-words had a prime that was identical to the visual target, 8 of the words and 24 of the pseudo-words had a prime that differed from the visual target in either one vowel or one consonant in any position, not corresponding to a common substitution in Dutch-accented or Japanese-accented speech, and 16 of the words and 24 of the pseudo-words had a phonologically and semantically unrelated prime. Target and prime in the filler trials, were mono-, di-, or trisyllabic.

The complete set of auditory primes was recorded twice, once by a female native speaker of Dutch and once by a female native speaker of Japanese. The speakers were students at Radboud University and Dokkyo University respectively, and they were selected by the authors on the basis of their moderate accent in English and their basic knowledge of phonetics. For the Dutch speaker, primes in the Dutch condition were presented in regular English spelling, and the speaker consistently produced /ɛ/ rather than /æ/, and /f/ rather than /v/ without further instructions. Primes in the Japanese Condition were presented in both regular English spelling and in modified spelling (e.g., *act* and *Acto*, *move* and *mooboo*), and the speaker was asked to produce the modified spelling; the regular spelling was consulted when the grapheme-to-phoneme conversion was unclear, and the speaker received further explanations about the desired pronunciation where necessary. For the Japanese speaker, primes in the Japanese Condition were presented in regular English spelling. In most cases, the speaker produced the desired phonological variants spontaneously; where necessary (for a number of items that did not seem to be pronounced with a strong Japanese accent) she was further instructed about the desired pronunciation.[3] Primes for the Dutch Condition were now presented in modified spelling, (e.g., *ect* and *moof*). The two speakers produced the primes one by one, separated by a pause, in a clear citation style, recording each prime

---

[2] The correct Japanese vowel would be unrounded /ɯ/. Since this is neither a vowel of English nor of Dutch, we decided to transcribe the vowel as /u/ throughout the article.

[3] We can think of at least two reasons for why the Japanese speaker was not fully consistent in her production of Japanese-accented variant forms: one, the Japanese speaker was more proficient in English than the Dutch speaker. As a result, the Japanese speaker may have been better at times to avoid a strong accent in her pronunciation. Two, Japanese learners of English are usually aware of their pronunciation deviating from the canonical form (in line with the form of loanwords), but Dutch learners are often not. Thus, possibly the Japanese speaker purposefully aimed for a less strongly accented pronunciation. In principle, correcting the Japanese speaker's pronunciation of some items could have had two effects: either the instructed items could be more "clear" instances of Japanese-accented English and therefore prime more strongly, or they could be less natural and therefore prime less strongly.

at least three times. The recordings were made in a soundproof booth with a Sennheiser microphone and were stored directly onto a computer at a sample rate of 41.5 kHz. Primes were excised from the recording using the speech editor Praat (Boersma & Weenink, 2009), and the best tokens were selected by the three authors. Acoustic analysis of the selected tokens confirmed that the Dutch-accented Consonant items from both speakers always ended in a voiceless obstruent (e.g., no voice bar was visible in /mu:f/), however, the vowel preceding word-final /f/ was, in proportion to the complete word, somewhat longer for the Dutch speaker (49%) than for the Japanese speaker (35%). F1, F2, and F3 formants were measured at midpoint for the vowel /ɛ/ in Dutch-accented Vowel items and for the vowel /a/ in Japanese-accented Vowel items: they were on average 718, 2026, 2879 Hz for /ɛ/ by the Dutch speaker, 719, 2245, 3186 Hz for /ɛ/ by the Japanese speaker, 719, 1509, 2728 Hz for /a/ by the Dutch speaker, and 749, 1657, 3036 Hz for /a/ by the Japanese speaker (Broersma, Aoyagi, & Weber, 2010).

Three versions of the experiment were constructed, one for each prime condition of any experimental item. Each version began with seven practice trials which were representative of the experiment, and further contained all 24 experimental trials and all 120 filler trials in pseudo-random order such that before each experimental trial there was at least one filler trial. Experimental trials appeared once in each version, and prime condition (Japanese, Dutch, and unrelated) was counterbalanced across versions, resulting in eight experimental trials per condition and version. Vowel items and Consonant items were equally distributed across versions. Each version of the experiment was presented to an equal number of participants, with the order of trial presentation being different for each participant. The distribution of experimental trials and filler trials within each version was such that the number of visual targets with identical primes, of visual targets with unrelated primes, and of visual targets with related primes was equal for both real-word targets and for pseudo-word targets: that is, each participant saw 24 pseudo-word targets and 24 word targets with identical auditory primes, 24 pseudo-word targets and 24 word targets with unrelated primes, and 24 pseudo-word targets and 24 word targets with primes similar but not identical to the target.

2.1.3. Procedure

Participants were tested one at a time in a quiet room. At the beginning of a session, they received instructions in English informing them that they would hear an English word, directly after which a real English word or a non-word would appear on the computer screen in front of them. They were asked to press with their dominant hand a response button labeled "yes" if they thought the word on the screen was a real English word, and to press with their non-dominant hand a response button labeled "no" if they thought it was not an existing English word. Participants were asked to respond both as fast and as accurately as possible. The instructions were presented auditorily, recorded by the experimental speaker, and were presented simultaneously in written form on the computer screen. After the seven practice trials, participants could ask questions before the experiment started.

The experiment was controlled with the NESU (Nijmegen Experiment Set-Up) software. Auditory stimuli were presented binaurally over closed headphones, and at their offset a visual stimulus was presented on the computer screen. Participants responded by pressing one of two response buttons on a box in front of them. Reaction times (RTs) were measured from the onset of the presentation of the visual stimuli. There was no time limit for responses. 1000 ms after a button response, the following auditory stimulus was played.

2.2. Results and discussion

2.2.1. Experiment 1a

Fig. 1 shows the mean RTs, measured from visual target onset, for correct responses in the three prime conditions, collapsed over Vowel and Consonant items, for Dutch participants listening to the recordings of the Dutch speaker and of the Japanese speaker. 11 responses with RTs shorter than 200 ms or longer than 2000 ms (.7% of all experimental trials) were considered as outliers and have been removed from these results. On average, Dutch participants had 96.3% correct responses when listening to the Dutch speaker, and 93.5% correct when listening to the Japanese speaker ($F_1[1, 64]=3.19$, $p > .07$; $F_2[1, 22]=5.97$, $p < .05$; with no significant difference between conditions: F_1 and $F_2 < 1$), demonstrating that they had little trouble performing the lexical-decision task in their L2. Average RTs per condition and average percentages correct responses are listed in Table 1.

Analyses of Variance (ANOVAs) were conducted for RTs across participants (F_1) and across items (F_2) to examine the effects of *substitution type*, *prime condition*, and *speaker*. The first two comparisons were within-participant and the last two within-item. There was a main effect of *prime condition* ($F_1[2, 128]=48.67$, $p < .001$; $F_2[2, 44]=27.02$, $p < .001$), no main effect of *speaker* ($F_1 < 1$; F_2 $[1, 22]=7.33$, $p < .05$), no main effect of *substitution type* (F_1 $[1, 64]=2.29$, $p > .1$; $F_2 < 1$), a significant interaction between *prime condition* and *speaker* (F_1 $[2, 128]=8.25$, $p < .001$; $F_2[2, 44]=3.79$, $p < .05$), but no significant three-way interaction (F_1 $[2,128]=1.32, p > .1; F_2 < 1$). Thus, responses to the two speakers were not significantly different, Vowel items and Consonant items were responded to equally fast, and crucially response times differed between the three prime conditions.

Further analyses showed a facilitatory effect in the Dutch condition (i.e., faster RTs following Dutch-accented primes than
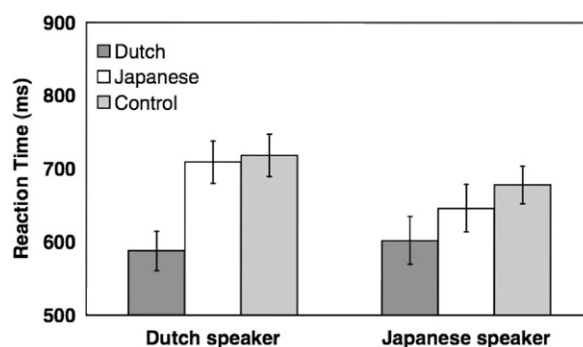

Fig. 1. Average RTs for correct responses after Dutch-accented primes, Japanese-accented primes, and unrelated primes from Dutch participants in Experiment 1a. Error bars represent the 95% confidence interval.

Table 1
Average RTs of correct responses and percentages correct responses for Dutch and Japanese participants in Experiment 1a and b.

Condition	Dutch speaker		Japanese speaker	
	Correct (%)	RT (ms)	Correct (%)	RT (ms)
Experiment 1a				
Dutch	98.2	588	93.6	602
Japanese	96.5	709	92.9	646
Control	94.1	718	94.1	678
Experiment 1b				
Dutch	94.6	710		
Japanese	92.9	751		
Control	88.1	819		

following unrelated primes), both when the stimuli were spoken by a Dutch speaker ($F_1[1, 35]=119.12$, $p < .001$; $F_2[1, 22]=83.17$, $p < .001$] and when they were spoken by a Japanese speaker ($F_1[1, 29]=19.73$, $p < .001$; $F_2[1, 22]=10.10$, $p < .01$). Thus, regardless of speaker, Dutch-accented /ɛkt/ and /mu:f/ facilitated recognition of *act* and *move* for Dutch listeners. This finding is in line with Broersma and Cutler (2008, in press), who also found significant priming of *act* and *groove* for Dutch listeners when an English speaker deliberately mispronounced it as /ɛkt/ or /gru:f/, respectively. The priming effect in the present study was furthermore larger for the Dutch speaker than for the Japanese speaker, as the interaction between *prime condition* (with the two levels 'Dutch-accented' and 'unrelated') and *speaker* showed ($F_1[1, 64]=12.66$, $p < .001$; $F_2[1, 22]=4.18$, $p > .05$).

For the Japanese condition, no facilitatory effect was found; that is, RTs for *act* following Japanese-accented /ˈakto/ were not significantly faster than RTs following unrelated primes ($F_1[1, 64]=3.67$, $p > .05$; $F_2[1, 22]=1.31$, $p > .1$). This time, the pattern was the same for both speakers, with no interaction between *speaker* and *prime condition* (F_1 and $F_2 < 1$). In other words, the main influence of *speaker* was on primes that were typical for the accent of the L2 listeners: when the language background of the speaker matched with that of the listeners, facilitatory priming effects were stronger than when there was a mismatch between language background of speaker and listeners. No such speaker influences were found for primes that were not typical for the accent of the L2 listeners: in this case, primes did not facilitate word recognition, regardless of the language background of the speaker. The recordings from the Dutch and the Japanese speakers differed in numerous phonetic details, mostly revealing a transfer from the speakers' L1 phonology to their L2 (VOT in stops, use of glottal stops, rounding of /u/, consonant gemination, and rhythm, to name a few), but importantly this variation was not enough to create or undo priming effects from accent-typical phonological variants, at best it was able to modulate the strength of these effects: Dutch-accented primes facilitated target word recognition for Dutch listeners but Japanese-accented primes did not, even when primes were spoken by a Japanese speaker. Clearly, accented word forms that matched with the language background of the listener were processed differently from accented forms that mismatched; only the former had a facilitatory effect on word recognition.

2.2.2. Experiment 1b

Fig. 2 shows the mean RTs for correct responses in the three prime conditions, collapsed over Vowel and Consonant items, for Japanese participants listening to the recordings of the Dutch speaker. Six outliers (.8% of all experimental trials) were removed from the results. On average, Japanese participants responded correctly 91.9% of the time (with no significant differences between conditions). Overall, the Japanese listeners in Experiment 1b, were 88 ms slower in their responses than the Dutch listeners in Experiment 1a; the high number of correct responses shows, however, that Japanese listeners too had no difficulties performing the task in their L2. Average RTs per condition and average percentages correct responses are listed in Table 1.

ANOVAs with the two factors *prime condition* (within-participant and within-item) and *substitution type* (within-participant and between-item) were conducted for RTs. There was a main effect of *prime condition* ($F_1[2, 54]=11.80$, $p < .001$; $F_2[2, 44]=8.98$, $p < .001$), a trend towards a main effect of *substitution type* ($F_1[1, 27]=5.44$, $p < .05$; $F_2[1, 22]=2.31$, $p > .1$) with responses to Vowel items being slower than responses to Consonant items, and no interaction between the two.

Subsequent pair wise comparisons of prime conditions (collapsed over Vowel and Consonant Items), revealed a facilitatory
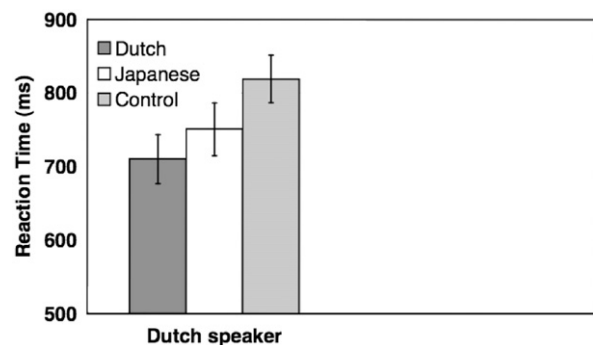

Fig. 2. Average RTs for correct responses after Dutch-accented primes, Japanese-accented primes, and unrelated primes from Japanese participants in Experiment 1b. Error bars represent the 95% confidence interval.

effect of Japanese-accented primes ($F_1[1, 27]=15.73$, $p < .001$; $F_2[1, 22]=5.75$, $p < .05$); that is, for Japanese listeners, RTs for *act* were faster following Japanese-accented /ˈakto/ than following unrelated primes. However, Dutch-accented /ɛkt/ also caused facilitation for this listener group ($F_1[1, 28]=28.52$, $p < .001$; $F_2[1, 22]=35.68$, $p < .001$).[4]

The fact that facilitation was found for Japanese-accented primes is in line with the above findings from Dutch listeners, showing that accented word forms that match with the language background of the listeners facilitate word recognition. However, in contrast to the above findings, accented forms that mismatched with the language background of the listeners also caused facilitation in Experiment 1b. In an additional analysis we compared responses of Japanese listeners in Experiment 1b and Dutch listeners listening to the Dutch speaker in Experiment 1a and found indeed a significant interaction between *native language* and *prime condition* ($F_1[2, 124]=5.66$, $p < .01$; $F_2[2, 44]=4.37$, $p < .05$).

Since Japanese participants were tested in Japan, had never lived in the Netherlands, and had no knowledge of Dutch, ample experience with Dutch cannot be the explanation for this facilitatory effect for Dutch-accented primes in Experiment 1b. Alternatively, Japanese participants may have learned to interpret Dutch-accented words correctly during the course of the experiment. Note that Dutch participants in Experiment 1a, had not shown comparable learning for Japanese-accented words, but in light of recent evidence for short-term adaptation to ambiguously produced sounds during the course of an experiment (e.g., Kraljic, Samuel, & Brennan, 2008; Norris, McQueen, & Cutler, 2003) and for native listeners becoming faster at identifying words in foreign-accented sentences after only a few trials (Clarke & Garrett, 2004; Clarke, 2002), the possibility of learning remained for Japanese listeners. To test this possibility, we analyzed RTs during the first and second half of the experiment separately. The pattern of results was, however, the same in both halves of the experiment: priming from Japanese- and Dutch-accented primes for Japanese listeners, with no main effect of experimental half (F_1 & $F_2 < 1$) and no interactions.[5] Thus, Japanese listeners showed priming from Dutch-accented words already early on in the experiment.

---

[4] In an attempt to assess the effect of loanword status in Japanese, we asked four native speakers of Japanese to judge for each item whether it was commonly used, never used, or sometimes used as a loanword. There were only five items that all four judges agreed were commonly used as loanwords (only two of which were also considered common loanwords by Inoue, 2005) and four items that were never used as loanwords, indicating again that there is no clear cut between loanwords and non-loanwords. Statistical analyses of those items was impossible due to a lack of statistical power, but the pattern of results did not show any obvious differences between the perception of those five loanwords and four non-loanwords.

[5] Neither was there an effect of experimental half for Dutch listeners in Experiment 1a.

The other possibility is that Dutch-accented primes were perceptually similar to the English standard pronunciation for Japanese listeners, and this is why they caused significant priming. The Japanese phoneme inventory contains neither the vowels /ɛ/ and /æ/, nor the consonants /v/ and /f/. Research on cross-language speech perception has repeatedly shown an increase in phoneme confusions for sounds that are not part of the listeners' L1 inventory (for an overview, see for example Strange, 1995; Bohn & Munro, 2007). For English vowels, Nishi, and Kewley-Port (2007) have recently shown that Japanese listeners correctly identify English /ɛ/ only 50% of the time, and they confuse it most often with English /æ/. Although, to our knowledge, for English consonants no equivalent confusion matrix for Japanese listeners exists, Takata and Nabelek (1990) report frequent errors in consonant identification for /v/ and /f/.[6] Due to these perceptual confusions, Dutch-accented /ɛkt/ and /mu:f/ may be indistinguishable from the correct English pronunciations /ækt/ and /mu:v/ for Japanese listeners, causing the primes, that were perceptually close to the target, to facilitate target recognition for the Japanese listeners. Second, variability within the English pronunciation of the phonemes may have added to the Japanese listeners' perceptual difficulties. The two vowels /ɛ/ and /æ/ are rather close to each other in the acoustic vowel space of some dialects of English (e.g., Clopper, Pisoni, & de Jong, 2005), and voiced obstruents are often pronounced as phonetically unvoiced in native English speech (Haggard, 1978; Stevens, Blumstein, Glicksman, Burton, & Kurowski, 1992). Third, Japanese listeners have been shown to be sensitive to the duration of the preceding vowel as a perceptual cue to final voicing in English, even though they have no native language experience with it (Crowther & Mann, 1992); they may thus have picked up that the Dutch speaker was lengthening the vowel in the Consonant items. All these factors may have contributed to the result that the Dutch-accented primes were perceptually so close to the target that they facilitated target recognition for the Japanese listeners.[7]

While for Consonant items, the absence of a voice bar in the spectrogram indicated that the Dutch speaker had indeed devoiced the final consonant in Dutch-accented /mu:f/, formant values for the vowel in /ɛkt/ could not ascertain category membership in the same way. It remained possible therefore, that the Dutch speaker had produced a vowel that was closer to the standard /æ/, and this "correct" pronunciation caused priming for Japanese listeners. However, we had also recorded the speaker's production of English vowels in the disyllable /hVbə/ (real words were used to aid in pronouncing vowels with ambiguous spelling), and formant values averaged over three tokens indicated that the speaker did not differentiate between English /ɛ/ in *hebba* (F1 572 Hz, F2 1934 Hz) and /æ/ in *habba* (F1 586 Hz, F2 1914 Hz). Since /ɛ/, but not /æ/, is an existing phoneme in the Dutch speaker's native repertoire, and the observed formant values better match the citation style values for

English /ɛ/ than for /æ/ (see e.g., Ladefoged, 1993; Strange et al., 2007), it is more likely, however, that the Dutch speaker indeed produced a vowel that was closer to /ɛ/ than to /æ/ (and the authors who selected the best tokens of the recordings of course also perceived it like that).

Note that priming effects for Japanese-accented word forms, on the other hand, cannot be easily explained by perceptual confusions. First of all, the vowel /a/ in Japanese-accented /ˈakto/ is not a phoneme of English (e.g., Ladefoged, 1999), and Japanese listeners perceive it as most similar to the English category /ɑ/ (Strange et al., 1998); and while English /b/ and /v/ are confusable at least for mid and low proficiency Japanese learners (Guion et al., 2000), Japanese-accented forms always contained an inserted vowel in coda position, thereby making the syllable structure of the English words permissible in Japanese. Listeners are known to be sensitive to phonological structure in their native language (e.g., Cutler & Otake, 1994, 2002; Otake, Hatano, & Yoneyama, 1996), and it is unlikely therefore that Japanese listeners could not perceptually distinguish between the non-permissible CVC in /mu:v/ and the permissible CVs in /mu:bu/. As Fais et al. (2005) have shown, Japanese listeners judge indeed the non-permissible form *keet* on a scale of goodness as significantly less good than the permissible form *keeto*. Rather, we suggest that Japanese-accented word forms primed for Japanese listeners because Japanese listeners have learned through prior exposure how to correctly interpret their own accent in English.

Perceptual confusions could of course be part of the underlying mechanism for facilitatory effects of Dutch-accented primes for Dutch listeners too. For Dutch listeners in Experiment 1a, Dutch-accented /ɛkt/ and /mu:f/ were perceptually closer to the correct English pronunciation than Japanese-accented primes were. Note however, that Dutch listeners are less likely to confuse English /ɛ/ with /æ/ than they are to confuse /æ/ with /ɛ/ (Cutler et al., 2004), and they can categorize English final voicing contrasts as accurately as native English listeners can (Broersma, 2005, 2008, 2010); they just seem to be less good at using this information for online word recognition (e.g., Broersma & Cutler, 2008; Cutler et al., 2006; Weber & Cutler, 2004). While we do not want to claim that perceptual difficulties did not modulate the results in Experiment 1a, it is simply not possible to tease apart the influence of perceptual confusions from experiential effects for Dutch-accented primes and Dutch listeners.

The lack of a facilitatory effect of Japanese-accented primes for Dutch listeners in Experiment 1a, on the other hand, could have been the result of a complete comprehension break-down: /ˈmu:bu/ is for Dutch listeners so far from the correct English pronunciation /mu:v/ that listeners may have failed to recover the intended word. In fact, when we played the Japanese-accented words to Dutch colleagues, and asked them to tell us which word this could be, they had difficulties recognizing the underlying English words. Maybe it was not so much the lack of experience with Japanese-accented primes but more the failure to understand them that resulted in a lack of facilitation for Dutch listeners. In Experiment 2, a new set of material was selected to address this issue. The manipulation was the same as in Experiment 1, only this time the Japanese-accented primes differed less strongly from the correct English pronunciation.

In Japanese-accented English, the final phoneme of a word determines whether the phonological rule of vowel paragoge will be applied or not. No vowel will usually be added when the word ends already in a vowel or when it ends in a nasal (Lovin, 1974 cited in Carruthers, 2006; Tanaka, 2009). Thus, in *happy* and *action* the vowel /æ/ will be replaced with /a/ (/ˈhapi/ and /ˈakʃən/ respectively) by Japanese speakers, but no other phonological changes will be made. When the final consonant is /d/, on the other hand, a vowel will be added to the word but the consonant will not be changed (e.g., *indeed* is pronounced as /ɪnˈdi:do/ by Japanese speakers).

---

[6] Most studies on English phoneme identification by Japanese listeners report identification in terms of Japanese categories and not in terms of English categories (e.g., for vowels, Strange et al., 1998; for consonants, Guion, Flege, Akahane-Yamada, & Pruitt, 2000).

[7] In 1999, Dupoux, Kakehi, Hirose, Pallier, and Mehler found that Japanese listeners report an epenthetic [u] vowel between consonants in VCCV syllables and also have difficulties distinguishing VCCV from VCVCV syllables. The authors concluded that phonotactic constraints of Japanese induced Japanese listeners to perceive illusionary vowels. From this it seems possible that Japanese listeners in Experiment 1 also added an illusionary vowel at the end of Dutch-accented words in order to fulfill Japanese phonotactic requirements. *Moof* would then have been perceived as *moofoo* for instance. Priming effects for the Dutch-accented forms may then in part have been due to *moofoo* being quite similar to the Japanese-accented form *mooboo*. Note, however, that there is evidence (Fais, Kajikawa, Werker, & Amano, 2005) for vowels in word-final position that Japanese listeners can reliably hear the difference between CVC and CVCo. This makes perceived similarity with Japanese-accented forms a less likely explanation for facilitatory priming effects of Dutch-accented forms.

The Dutch-accented forms of *happy*, *action*, and *indeed* will diverge from the English standard in the same way that Dutch-accented forms in Experiment 1 did. If a match between accent variant in the speech signal and accent of the listener is crucial, then Dutch listeners should again show only priming for Dutch-accented primes in Experiment 2, even though the Japanese-accented primes now diverge less from the English standard and Dutch listeners should be able to understand the intended word more easily. Since the Dutch-accented primes will again be perceptually confusing for Japanese listeners, we predict that Japanese listeners will show priming not only for primes that converge with their own accent (i.e., Japanese-accented primes), but also for Dutch-accented primes. Recordings were again made with a Dutch speaker and with a Japanese speaker to assess possible differences in results for the two speakers.

3. Experiment 2

3.1. Method

3.1.1. Participants

Sixty native speakers of Dutch (Experiment 2a) and 42 native speakers of Japanese (Experiment 2b) participated in Experiment 2. None had taken part in the previous experiment. Participants came from the same population as participants in Experiment 1, with the Dutch students having received an average of 7 years of training in English as foreign language, starting at a mean age of 11, and the Japanese students with an average of 8 years of training in English, with a mean starting age of 12. At the time of testing, Dutch participants were on average 21 years of age, and Japanese participants 25 years. In the *LexTALE* test (Lemhöfer & Broersma, 2009, sub-mitted for publication), Dutch participants had an average score of 79% correct and Japanese participants of 67%.[8] Half of the Dutch participants were presented with the Dutch recordings of the stimuli, and the other half with the Japanese recordings. In contrast to Experiment 1, all 42 Japanese participants now listened to the Japanese recordings of the stimuli.

3.1.2. Materials

Twenty-four new English words were selected as visual targets. Again, half of the target words contained the vowel /æ/ (Vowel items), but this time they all ended in either a vowel, /r/, or /n/ (e.g., *happy* /ˈhæpɪ/), in which case Japanese speakers do not typically add a vowel at the end (e.g., Carruthers, 2006). As before, the other half of the target words ended in a voiced obstruent (Consonant items), but this time the obstruent was /d/ rather than /v/ (e.g., *indeed* /inˈdiːd/). In Japanese-accented English, a vowel will be suffixed to these words, but the obstruent will not be changed (e.g., Carruthers, 2006). Target words were disyllabic, and had an average lemma frequency of 123 per million in the *CELEX* database (Baayen et al., 1993).

As in Experiment 1, each target word had a Dutch-accented prime, a Japanese-accented prime, and an unrelated prime. Dutch-accented primes had the same form as in Experiment 1. Thus, English /æ/ was replaced with /ɛ/ (e.g., *happy* pronounced as /ˈhɛpɪ/), and the word-final /d/ was devoiced as /t/ (e.g., *indeed* pronounced as /inˈdiːt/). For the Japanese-accented primes, in Vowel items, /æ/ was again replaced with /a/ (e.g., *happy* pronounced as /ˈhapɪ/), and in Consonant items, /o/ was added after the word-final /d/ (e.g., *indeed* pronounced as /inˈdiːdo/). In contrast to Experiment 1, thus, both Dutch- and Japanese-accented primes now differed from the standard

---

pronunciation in only one phoneme. As unrelated primes, phonologically and semantically unrelated words, matching in syllable number with a target, were selected (see Appendix). Filler trials were the same as in Experiment 1.

The materials of Experiment 2 were recorded with the speakers of Experiment 1 in the same session. Recordings were processed accordingly. Three versions of Experiment 2, comparable to the versions of Experiment 1, were constructed. Note that in Experiment 2, some unrelated prime words were used for more than one target word. Experimental lists were created such that each participant still heard each unrelated prime only once.

3.1.3. Procedure

The procedure was the same as in Experiment 1.

3.2. Results and discussion

3.2.1. Experiment 2a

Mean RTs for Dutch participants, listening to the recordings of the Dutch speaker and of the Japanese speaker, are shown in Fig. 3a and b, respectively. 10 responses with RTs shorter than 200 ms or longer than 2000 ms (.7% of all experimental trials) were considered as outliers and were removed from further analyses. High percentages of correct responses demonstrate again that Dutch listeners had no difficulties performing the task. On average, Dutch participants had 96.6% correct responses when listening to the Dutch speaker, and 94.3% correct when listening to the Japanese speaker (with marginally significant differences between speakers, $F_1[1, 58] = 1.75$, $p > .1$; $F_2[1, 22] = 6.83$, $p < .05$, and between conditions, $F_1[2, 116] = 2.84$, $p > .06$, $F_2[2, 44] = 4.71$, $p < .05$). Average RTs per condition and average percentages correct responses are listed in Table 2.

ANOVAs, comparable to the ones in Experiment 1, showed a main effect of *prime condition* ($F_1[2, 116] = 30.32$, $p < .001$; $F_2[2, 44] = 38.63$, $p < .001$), a main effect of *substitution type* significant by participants ($F_1[1, 58] = 21.41$, $p < .001$; $F_2[1, 22] = 2.10$, $p > .1$) with Vowel items being responded to more quickly, no main effect of *speaker* (F_1 and $F_2 < 1$), and no significant interaction between any of the factors. Thus, as in Experiment 1, response times differed between the three prime conditions; in contrast to Experiment 1, however, speaker identity had no measurable influence on the response pattern.

Subsequent pair wise comparisons of prime conditions (collapsed over *speaker* and *substitution type*) revealed a facilitatory effect of Dutch-accented primes ($F_1[1, 58] = 72.63$, $p < .001$; $F_2[1, 22] = 70.78$, $p < .001$); that is, for Dutch listeners, RTs for *happy* were faster following Dutch-accented /ˈhɛpɪ/ than following unrelated primes. No significant priming was found for Japanese-accented /ˈhapɪ/ ($F_1[1, 58] = 1.86$, $p > .1$; $F_2[1, 22] = 1.55$, $p > .1$). Thus, even though Japanese primes varied less drastically from the
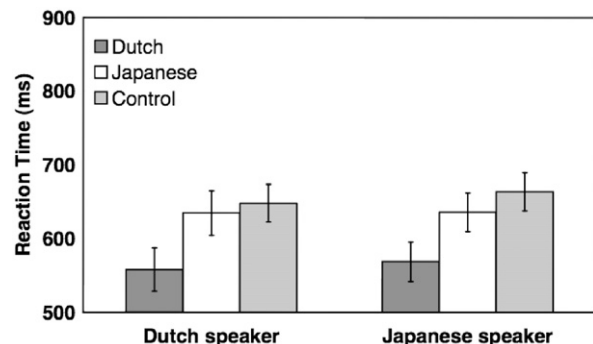

Fig. 3. Average RTs for correct responses after Dutch-accented primes, Japanese-accented primes, and unrelated primes from Dutch participants in Experiment 2a. Error bars represent the 95% confidence interval.

Table 2
Average RTs of correct responses and percentages correct responses for Dutch and Japanese participants in Experiment 2a and b.

| Condition	Dutch speaker		Japanese speaker	
	Correct (%)	RT (ms)	Correct (%)	RT (ms)
Experiment 2a				
Dutch	98.3	558	96.6	569
Japanese	96.6	635	93.7	636
Control	95.0	648	92.5	664
Experiment 2b				
Dutch			98.2	726
Japanese			97.9	679
Control			95.8	796

Fig. 4. Average RTs for correct responses after Dutch-accented primes, Japanese-accented primes, and unrelated primes from Japanese participants in Experiment 2b. Error bars represent the 95% confidence interval.

correct English pronunciation in Experiment 2, they still did not facilitate recognition of target words. As in Experiment 1, only accented word forms that matched with the language background of the listener had a facilitatory effect on word recognition for Dutch listeners.

3.2.2. Experiment 2b

Fig. 4 shows the mean RTs for correct responses in the three prime conditions for Japanese participants, listening to the recordings of the Japanese speaker. Sixteen outliers were removed from the results (1.6% of all experimental trials). On average, Japanese participants responded correctly 97.3% of the time (with no significant differences between conditions ($F_1[2, 82]=1.85$, $p > .1$; $F_2[2, 44]= 1.12$, $p > .1$). Average RTs per condition and average percentages correct responses are listed in Table 2.

As in Experiment 1, ANOVAs with the two factors *prime condition* and *substitution type* were conducted for RTs. There was a main effect of *prime condition* ($F_1[2, 82]=22.89$, $p < .001$; $F_2[2, 44]=9.92$, $p < .001$), no main effect of *substitution type* (F_1 and $F_2 < 1$) and no interaction between the two. Subsequent pair wise comparisons of prime conditions revealed a facilitatory effect of Japanese-accented primes ($F_1[1, 41]=42.80$, $p < .001$; $F_2[1, 22]=19.40$, $p < .001$); that is, for Japanese listeners, RTs for *happy* were faster following Japanese-accented /ˈhapɪ/ than following unrelated primes. However, Dutch-accented /ˈhɛpɪ/ also caused facilitation for this listener group ($F_1[1, 41]=17.00$, $p < .001$; $F_2[1, 22]=8.39$, $p < .01$). Again, there was no difference in response patterns between the first and the second half of the experiment, and a direct comparison with the Dutch listeners listening to the Japanese speaker in Experiment 2a revealed a significant interaction between *native language* and *prime condition* ($F_1[2, 140]=12.75$, $p < .001$; $F_2[2, 44]=8.21$, $p < .001$). The results are therefore a full replication of the findings in Experiment 1.[9]

Again, we would like to argue that Japanese-accented primes facilitated word recognition for Japanese listeners because the accent in the primes conformed with the language background of the listeners, and Dutch-accented primes facilitated word recognition because the primes were perceptually confusable with the English standard pronunciation for Japanese listeners. As in Experiment 1, Dutch-accented Vowel items contained the vowel /ɛ/ (/ˈhɛpɪ/) which is difficult to distinguish from /æ/ for Japanese listeners (e.g., Nishi & Kewley-Port, 2007); in contrast to
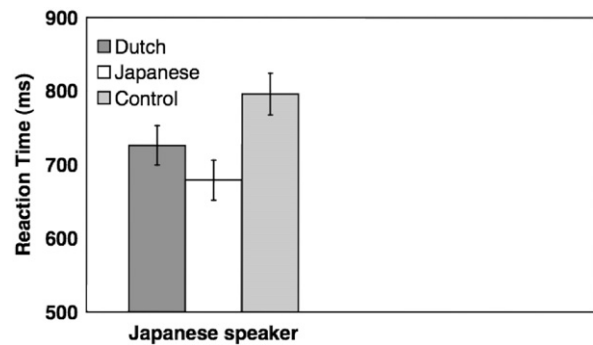
Experiment 1, however, Consonant items in Experiment 2 involved /t/ (/inˈdiːt/), a phoneme that is contrasted in Japanese with /d/, though not in final position. While Crowther and Mann (1992, 1994) have shown that Japanese listeners can nevertheless distinguish final /d/ quite well from /t/, other studies have failed to show this ability (Morrison, 2002). In light of these contrasting results it remains possible therefore that for Consonant items experience with variability in voicing from L1 speakers or other cues to voicing were mainly responsible for the observed facilitatory effects.

4. General discussion

In order to recognize words spoken with a foreign accent, listeners have to deal with the variability introduced to the speech signal by L2 speakers. While we know already that naturally occurring variation in L1 speech usually poses little difficulties for lexical recognition, especially when listeners have experience with the variant forms (e.g., Connine et al., 2008; Mitterer & Ernestus, 2006; Pitt, 2009; Ranbom & Connine, 2007; Ranbom et al., 2009), the main goal of the present study was to test whether immediate lexical recognition can also occur in L2 speech. To this end we conducted two cross-modal priming experiments with Dutch-accented and Japanese-accented English words. If experience plays a similar role for word recognition in L2 speech as it does in L1 speech, then listeners with the same language background as the L2 speaker should recognize accented words quite easily, while listeners with a different language background should find it more difficult. This is indeed what was found: Dutch-accented /ɛkt/ facilitated recognition of English *act* for Dutch listeners, and Japanese-accented /ˈakto/ facilitated recognition of English *act* for Japanese listeners. Naturally, Dutch learners of English have ample experience with the Dutch accent and Japanese learners with the Japanese accent, and this experience allowed listeners to recognize those variant forms that are conform with their language background quite easily.

The experience the two learner groups had with the variant forms most likely stemmed from a combination of producing the forms themselves and from hearing them being produced by other learners. This brings up the question of whether both types of experience influenced word recognition in a comparable way. Previous studies have indeed argued for an influence on speech perception from both types. Bell-Berti, Raphael, Pisoni, and Sawusch (1979) for example, claimed that speech perception is influenced by how listeners, as speakers, produce speech themselves. Massaro (1987) and Diehl and Kluender (1989), on the other hand, argued that prior experience with hearing other people speak creates acoustic memories of phonemes and syllables that are

---

[9] As for Experiment 1, four native speakers of Japanese again decided for each item whether it was used as a loanword in Japanese. There were seven items that all four judges agreed were commonly used as loanwords (five of which were also considered common loanwords by Inoue, 2005) and seven items all four agreed were never used as loanwords. Statistical analyses of those 14 items showed that there was no significant main effect of loanword status and no interaction between loanword status and condition.

associated with their underlying mental representations. Within the scope of the present study, this question cannot be answered since it is impossible to determine the amount of prior experience participants had with producing and hearing the variant forms. For example, even if we would have evidence that participants were not producing the variant forms at the point of testing, they still may have produced them at an earlier stage of second language learning. Similarly, it is not feasible to exactly establish how often and from how many different speakers participants have heard the variant forms before. Recent evidence from Hanulikova and Weber (submitted for publication) as well as from Di Betta, McQueen, and Weber (2010), however, suggests that producing variant forms is not a prerequisite for their immediate recognition. Hanulikova and Weber (submitted for publication), for example, found no correlation between individual production frequencies and looking behavior in an eye-tracking study on the recognition of English words with /θ/-substitutions by Dutch and German L2 listeners; and Di Betta et al. (2010) found in a cross-modal priming study that immediate recognition of Italian-accented words was possible for native listeners of English, who were familiar with the Italian accent but of course did not produce the variant forms themselves.

Linguistic experience was, however, not the only factor influencing the ease of word recognition. In both Experiments 1 and 2, Japanese listeners also rapidly recognized words that were typical of a Dutch accent in English. In this case, we argued that perceptual difficulties of Japanese listeners with English phoneme contrasts made the Dutch-accented pronunciation and the correct English pronunciation perceptually indistinguishable and hence immediate word recognition was not disturbed. Dutch listeners, on the other hand, were not expected to have perceptual difficulties with the Japanese accent, which was clearly distinct from the correct English pronunciation. Accordingly, Japanese-accented primes did not facilitate target recognition for Dutch listeners.

L2 learners' perception of L2 sounds is often inaccurate, and L2 listeners may have experience with L2 speakers with varying proficiency and language backgrounds. Thus, one could expect L2 listeners to immediately recognize accented words even when the accent does not match their own language background. This was, however, not what we found. For Dutch listeners, Japanese-accented primes did not facilitate target recognition, neither when the primes were perceptually very different from the correct English pronunciation (Experiment 1a), nor when they were closer to it (Experiment 2a). L2 listeners are therefore not generally insensitive to segmental deviations in non-standard pronunciation. On the other hand, when the L2 speech contained specific contrasts that are known to be difficult to perceive for L2 listeners, L2 listeners failed to notice a non-standard pronunciation even when the pronunciation was untypical of their own accent (Experiments 1b and 2b). This is not surprising. We know from research on L2 listening with native speech, that perceptual difficulties can blur the distinction between words, which leads to the activation of unintended words (e.g., Broersma & Cutler, 2008, in press; Pallier et al., 2001; Weber & Cutler, 2004): for example, Dutch listeners who have difficulties distinguishing English /æ/ from /ɛ/, activate erroneously the word *pencil* when hearing the onset of the word *panda* (Weber & Cutler, 2004). The same obliteration is happening in non-native speech, except that in this case, rather than hindering recognition, perceptual difficulties of L2 listeners allow the recognition of intended words despite their non-standard pronunciation: in the present study, both Japanese and Dutch listeners recognized the intended word *act* correctly when hearing the Dutch-accented from *ect*. Note, however, that this benefit of L1-imposed perceptual assimilation may turn into a detriment when the non-standard pronunciation matches another existing word: for example, the form *bet* could be the non-standard pronunciation of *bat* or be intended as *bet*. In the latter case, interpreting *bet* as *bat* (as Dutch listeners have

been shown to do; Broersma, submitted for publication; Cutler & Broersma, 2005) would be detrimental.

When translated to intelligibility, our results may offer a possible explanation for why some studies find mutual intelligibility of foreign-accented speech for L2 listeners (e.g., Bent & Bradlow, 2003; Munro et al., 2006, van Wijngaarden, 2001), while other studies do not (e.g., Leikin et al., 2009; Stibbard & Lee, 2006). Based on the present findings it could be that foreign-accented sentences are mutually intelligible when the sentences contain many contrasts that are difficult to perceive for all the different L2 listeners; when there are few such contrasts that all the L2 listeners find difficult to perceive, intelligibility would not be mutual.

The present results furthermore show that effects of linguistic experience can be separated from effects of perceptual confusion. The Japanese-accented primes were perceptually different from the standard pronunciation, not only for Dutch listeners but also for Japanese listeners. But while the Japanese accent did not facilitate word recognition for Dutch listeners, it did facilitate recognition for Japanese listeners. Regular experience with their own L2 accent made Japanese-accented words easily recognizable for Japanese listeners. This finding is in line with recent evidence from Catalan–Spanish bilinguals indicating that long-term exposure to Spanish-accented Catalan influences word recognition in the bilinguals dominant language Catalan (Sebastián-Gallés et al., 2005, 2006). In a lexical-decision task, Catalan-dominant bilinguals had more difficulties rejecting non-words made from real words by changing a vowel involving a Catalan-specific contrast (/e/–/ɛ/), than rejecting non-words made by changing a vowel involving a contrast common in Catalan and Spanish (/i/–/u/). Perceptually, Catalan-dominant listeners can distinguish between Catalan /e/ and /ɛ/ as well as between /i/ and /u/, but only non-words involving the former contrast were frequently accepted as words. Sebastián-Gallés et al. (2006) argue that regular exposure to Spanish speakers of Catalan who fail to produce the Catalan vowel contrast made Catalan–Spanish bilinguals accept the variant forms as words.

A further interesting outcome of the presented research is that priming effects were not drastically affected by whether the speaker was an authentic speaker of the accent or not: for Dutch listeners, primes typical of a Dutch accent facilitated English word recognition and primes typical of a Japanese accent did not, regardless of whether the primes were actually spoken by a Dutch or Japanese speaker. This suggests that imitating an accent may be sufficient to ease word recognition for L2 listeners. Note though that speakers in our experiments did not consciously imitate an accent; that is, the Dutch speaker did not try to sound Japanese when reading primes that contained segmental substitutions typical for a Japanese accent, just as the Japanese speaker did not try to sound Dutch when reading Dutch-accented primes. As a consequence, numerous finer phonetic cues gave away the speaker as being non-authentic. But apparently a coarse segmental match was enough to make the accented words recognizable for L2 listeners. We do not want to argue that fine phonetic detail is irrelevant, given the rich evidence for the importance of phonetic detail in lexical activation (e.g., Dahan, Magnuson, Tanenhaus, & Hogan, 2001; Marslen-Wilson & Warren, 1994; McQueen, Norris, & Cutler, 1999; Whalen, 1991), rather we simply want to make the point that lexical recognition can succeed in L2 speech in spite of mismatching fine phonetic details.

Although our research clearly shows that experience with a foreign accent, as well as perceptual difficulties, can influence processing of that accent, we can only speculate at this point about the mechanisms underlying this effect. With respect to how phonetic perception relates to lexical storage for L2 listeners, recent evidence from eye-tracking studies suggests that perception alone does not determine lexical storage (e.g., Cutler et al., 2006; Escudero, Hayes-Harb, & Mitterer, 2008; Weber & Cutler, 2004);

that is, a phonetic distinction that cannot be reliably perceived in the speech input can nevertheless be established in the lexicon. In an English eye-tracking study, Dutch listeners were likely to look first at the picture of a *pencil* when instructed by a native speaker of English to look at a *panda,* but instructions to look at a *pencil* did induce few early looks to *panda* (Weber & Cutler, 2004); likewise Japanese learners' of English looked at a *locker* when instructed to look at a *rocket,* but the reverse did not apply (Cutler et al., 2006). In both cases, the L2 learners were known to have perceptual difficulties with the tested vowel and consonant contrasts, but the asymmetry in the results indicates that the two initial syllables of *pencil* and *panda,* and of *locker* and *rocket,* were not represented as interchangeable homophones in the learners' lexicon.[10]

Thus, when we find with a priming task that /ɛkt/ primes recognition of *act* for Dutch listeners (see also Broersma & Cutler, 2008, in press), this does not necessarily mean that the stored representation for *act* contains an /ɛ/ and not an /æ/. This is even easier to imagine for the Japanese listeners; because /ɛ/ is not even a sound of Japanese, it is likely that for Japanese listeners the representation of *act* contains a vowel that approximates the English /æ/, based on listening experience with native English speech. Japanese listeners' phonetic processing, then, is not precise enough to notice the phonetic-to-lexical mismatch from /ɛkt/ to *act*, explaining why /ɛkt/ primes *act* for those listeners. In contrast to the Japanese listeners, Dutch listeners had ample experience with Dutch-accented /ɛkt/, and this accumulated experience may have altered the stored representation of *act*, to include more than one realization of the vowel. Thus it is possible that for Dutch listeners multiple representations of *act* exist, and /ɛkt/ may prime *act* not because of inaccurate phonetic-to-lexical mapping, but due to mapping /ɛkt/ to the stored variant *ekt*.

If this were true, it would mean that more than one mechanism can underlie successful variant recognition for L2 listeners: Variant forms can either be recognized by undoing the effects of speech production before contacting the mental lexicon, or by disambiguating variant forms after lexical access. Both mechanisms have also been proposed for variant recognition in L1 speech with L1 listeners. While lexical storage accounts for L1 speech vary in their assumptions about the nature of stored representations, episodic (e.g., Johnson, 2006; Pierrehumbert, 2001) versus abstract (Ranbom & Connine, 2007), they all assume that our mental lexicon reflects the variability of the speech signal by storing multiple variant forms. Processing-based accounts, on the other hand, have in common that phonologically altered forms are assumed to be recognized without explicit storage of the alternatives in the lexicon, either via general auditory mechanisms (e.g., Lotto & Holt, 2006; Mitterer, Csepe, & Blomert, 2006) or via more abstract pattern recognition (e.g., Gaskell & Marslen-Wilson, 1998; Gow, 2002; Smits, 2001).

What the underlying mechanisms for successful variant recognition for L2 listeners are, we can only speculate at this point. Many other questions also remain to be addressed for a fuller understanding of spoken-word recognition in foreign-accented speech. For example, are foreign-accented words recognized as easily as unaccented words, and is there a difference in whether the listeners are more familiar with one form or the other? How much exposure is necessary before lexical entries are being added? And do the effects of perceptual confusions imply that there is less mutual recognizability of foreign-accented words for highly proficient L2 learners than for learners with lower proficiency? And last but not least, does increasing contact with L2 speech entail that native listeners will erroneously activate *pan* when hearing *pen* in L1 speech? All those questions require further research, and we hope

to address some of them in future work. The present research has taken a first step towards unraveling them by showing how linguistic experience and perceptual difficulties can influence word recognition for L2 listeners in foreign-accented speech.

Table A1
Experiment 1.

Target words	Prime words Dutch condition	Prime words Japanese condition	Prime words unrelated condition
Vowel items			
act /ækt/	/ɛkt/	/akto/	sky
damage /ˈdæmɪdʒ/	/ˈdɛmɪdʒ/	/ˈdamɪdʒi/	tension
exam /ɪgˈzæm/	/ɪgˈzɛm/	/ɪgˈzamu/	monsoon
habit /ˈhæbɪt/	/ˈhɛbɪt/	/ˈhabɪto/	mission
hazard /ˈhæzəd/	/ˈhɛzəd/	/ˈhazədo/	veto
magic /ˈmædʒɪk/	/ˈmɛdʒɪk/	/ˈmadʒɪku/	lemon
manage /ˈmænɪdʒ/	/ˈmɛnɪdʒ/	/ˈmanɪdʒi/	Fourteen
panic /ˈpænɪk/	/ˈpɛnɪk/	/ˈpanɪku/	Whisky
patch /pætʃ/	/pɛtʃ/	/patʃi/	Key
Spanish /ˈspænɪʃ/	/ˈspɛnɪʃ/	/ˈspanɪʃu/	eighty
static /ˈstætɪk/	/ˈstɛtɪk/	/ˈstatɪku/	pony
tactic /ˈtæktɪk/	/ˈtɛktɪk/	/ˈtaktɪku/	daily
Consonant items			
above /əˈbʌv/	/əˈbʌf/	/əˈbʌbu/	possess
achieve /əˈtʃiːv/	/əˈtʃiːf/	/əˈtʃiːbu/	typhoon
behave /bɪˈheɪv/	/bɪˈheɪf/	/bɪˈheɪbu/	Canoe
cave /keɪv/	/keɪf/	/keɪbu/	Fame
dive /daɪv/	/daɪf/	/daɪbu/	Noon
give /gɪv/	/gɪf/	/gɪbu/	Tie
motive /ˈmoːtɪv/	/ˈmoːtɪf/	/ˈmoːtɪbu/	Item
move /muːv/	/muːf/	/muːbu/	Snow
naïve /naˈiːv/	/naˈiːf/	/naˈiːbu/	Cartoon
native /ˈneɪtɪv/	/ˈneɪtɪf/	/ˈneɪtɪbu/	Money
perceive /pəˈsiːv/	/pəˈsiːf/	/pəˈsiːbu/	Esteem
shave /ʃeɪv/	/ʃeɪf/	/ʃeɪbu/	Dawn

Table A2
Experiment 2.

Target words	Prime words Dutch condition	Prime words Japanese condition	Prime words control condition
Vowel items			
action /ˈækʃən/	/ˈɛkʃən/	/ˈakʃən/	money
champion /ˈtʃæmpjən/	/ˈtʃɛmpjən/	/ˈtʃampjən/	veto
chapter /ˈtʃæptər/	/ˈtʃɛptər/	/ˈtʃaptər/	monkey
fancy /ˈfænsɪ/	/ˈfɛnsɪ/	/ˈfansɪ/	item
fashion /ˈfæʃən/	/ˈfɛʃən/	/ˈfaʃən/	whisky
hammer /ˈhæmər/	/ˈhɛmər/	/ˈhamər/	fourteen
happy /ˈhæpɪ/	/ˈhɛpɪ/	/ˈhapɪ/	mission
matter /ˈmætər/	/ˈmɛtər/	/ˈmatər/	tension
passion /ˈpæʃən/	/ˈpɛʃən/	/ˈpaʃən/	lemon
pattern /ˈpætən/	/ˈpɛtən/	/ˈpatən/	eighty
shadow /ˈʃædoː/	/ˈʃɛdoː/	/ˈʃadoː/	pony
statue /ˈstætʃuː/	/ˈstɛtʃuː/	/ˈstatʃuː/	daily
Consonant items			
afford /əˈfɔːd/	/əˈfɔːt/	/əˈfɔːdo/	typhoon
award /əˈwɔːd/	/əˈwɔːt/	/əˈwɔːdo/	canoe
awkward /ˈɔːkwəd/	/ˈɔːkwət/	/ˈɔːkwədo/	mountain
behind /bɪˈhaɪnd/	/bɪˈhaɪnt/	/bɪˈhaɪndo/	possess
beyond /bɪˈjɒnd/	/bɪˈjɒnt/	/bɪˈjɒndo/	cartoon
decide /dɪˈsaɪd/	/dɪˈsaɪt/	/dɪˈsaɪdo/	typhoon
demand /dɪˈmaːnd/	/dɪˈmaːnt/	/dɪˈmaːndo/	cartoon
indeed /inˈdiːd/	/inˈdiːt/	/inˈdiːdo/	possess
inside /inˈsaɪd/	/inˈsaɪt/	/inˈsaɪdo/	possess
persuade /pəˈsweɪd/	/pəˈsweɪt/	/pəˈsweɪdo/	canoe
succeed /səkˈsiːd/	/səkˈsiːt/	/səkˈsiːdo/	canoe
toward /təˈwɔːd/	/təˈwɔːt/	/təˈwɔːdo/	cartoon

---

[10] Escudero et al. (2008) have shown that spelling training is one possibility how listeners can construct a lexical distinction they cannot perceive reliably.

Acknowledgments

Appendix A. Experimental stimuli used in Experiments 1 and 2

See Tables A1 and A2.

References

Adank, P., & McQueen, J. M. (2007). The effect of an unfamiliar regional accent on spoken-word comprehension. In *Proceedings of the 16th international conference of phonetic sciences* (pp. 1926–1928), Saarland University, Germany.

Aoyama, K., Flege, J. E., Guion, S. G., Akahane-Yamada, R., & Yamada, T. (2003). Foreign accent in English words produced by Japanese children and adults. In *Proceedings of the 15th international congress of phonetic sciences* (pp. 3201–3204), Barcelona, Spain.

Avery, P., & Ehrlich, S. (1992). *Teaching American English pronunciation*. Oxford, UK: Oxford University Press.

Baayen, H., Piepenbrock, R., & Van Rijn, H. (1993). The CELEX lexical database (CDROM). Philadelphia: Linguistic Data Consortium, University of Pennsylvania.

Bell-Berti, F., Raphael, L. J., Pisoni, D. B., & Sawusch, J. R. (1979). Some relationships between speech production and perception. *Phonetica*, 36, 373–383.

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

Best, C. T. (1995). A direct realist perspective on cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171–204). Timonium, MD: York Press.

Boersma, P., & Weenink, D. (2009). *Praat: Doing phonetics by computer* (Version 5.1.16) [Computer program].

Bohn, O. -S., & Munro, M. J. (Eds.) (2007). Language experience in second language speech learning: In honor of James Emil Flege. Amsterdam: John Benjamins.

Bongaerts, T., van Summeren, C., Planken, B., & Schils, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition*, 19, 447–465.

Brannen, K. (2002). The role of perception in differential substitution. *Canadian Journal of Linguistics—Revue Canadienne de Linguistique*, 47, 1–20.

Broersma, M. (2002). Comprehension of non-native speech: Inaccurate phoneme processing and activation of lexical competitors. In *Proceedings of the seventh international conference on spoken language processing* (pp. 261–264), University of Colorado, Boulder.

Broersma, M. (2005). Perception of familiar contrasts in unfamiliar positions. *Journal of the Acoustical Society of America*, 117, 3890–3901.

Broersma, M. (2008). Flexible cue use in nonnative phonetic categorization. *Journal of the Acoustical Society of America*, 124, 712–715.

Broersma, M. (2010). Perception of final fricative voicing: Native and nonnative listeners' use of vowel duration. *Journal of the Acoustical Society of America*, 127, 1636–1644.

Broersma, M. (submitted for publication). Increased lexical activation and reduced competition in second-language listening.

Broersma, M., Aoyagi, M., & Weber, A. (2010). Cross-linguistic production and perception of Japanese- and Dutch-accented English. *Journal of the Phonetic Society of Japan*, 14, 60–75.

Broersma, M., & Cutler, A. (2008). Phantom word activation in L2. *System: An International Journal of Educational Technology and Applied Linguistics*, 36, 22–34.

Broersma, M., & Cutler, A. (in press). Competition dynamics of second-language listening. *Quarterly Journal of Experimental Psychology*. Advance online publication. doi:10.1080/17470218.2010.499174.

Carruthers, S. W. (2006). Pronunciation difficulties of Japanese speakers of English: Predictions based on a contrastive analysis. *HPU TESL working paper series*, Vol. 4, pp. 17–23.

Clarke, C. M. (2002). Perceptual adjustment to foreign-accented English with short-term exposure. In *Proceedings of the seventh international conference on spoken language processing* (pp. 253–256), Denver.

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

Clopper, C., Pisoni, D., & de Jong, K. (2005). Acoustic characteristics of the vowel systems of six regional varieties of American English. *Journal of the Acoustical Society of America*, 118, 1661–1676.

Coenen, E., Zwitserlood, P., & Boelte, J. (2001). Variation and assimilation in German: Consequences for lexical access and representation. *Language and Cognitive Processes*, 16, 535–564.

Connine, C. M. (2004). It's not what you hear but how often you hear it: On the neglected role of phonological variant frequency in auditory word recognition. *Psychonomic Bulletin & Review*, 11, 1084–1089.

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, 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*, 7BI0, 403–411.

Costa, A., Cutler, A., & Sebastián-Gallés, N. (1998). Effects of phoneme repertoire on phoneme decision. *Perception & Psychophysics*, 60, 1022–1031.

Crowther, C. S., & Mann, V. (1992). Native language factors affecting use of vocalic cues to final consonant voicing in English. *Journal of the Acoustical Society of America*, 92, 711–722.

Crowther, C. S., & Mann, V. (1994). Use of vocalic cues to consonant voicing and native language background: The influence of experimental design. *Perception & Psychophysics*, 55, 13–525.

Cutler, A., & Broersma, M. (2005). Phonetic precision in listening. In W. Hardcastle, & J. Beck (Eds.), *A figure of speech* (pp. 63–91). Mahwah, NJ: Erlbaum.

Cutler, A., & Otake, T. (1994). Mora or phoneme? Further evidence for language-specific listening. *Journal of Memory and Language*, 33, 824–844.

Cutler, A., & Otake, T. (2002). Rhythmic categories in spoken-word recognition. *Journal of Memory and Language*, 46, 296–322.

Cutler, A., & Otake, T. (2004). Pseudo-homophony in non-native listening. *Journal of the Acoustical Society of America*, 115, 2392.

Cutler, A., Weber, A., & Otake, T. (2006). Asymmetric mapping from phonetic to lexical representations in second-language listening. *Journal of Phonetics*, 34, 269–284.

Cutler, A., Weber, A., Smits, R., & Cooper, N. (2004). Patterns of English phoneme confusions by native and non-native listeners. *Journal of the Acoustical Society of America*, 116, 3668–3678.

Dahan, D., Magnuson, J. S., Tanenhaus, M. K., & Hogan, E. M. (2001C). Subcategorical mismatches and the time course of lexical access: Evidence for lexical competition. *Language and Cognitive Processes*, 16, 507–534.

Di Betta, A. -M., McQueen, J. M., & Weber, A. (2010). Adaptation to Italian-accented English: A comparison of native and nonnative listeners. *Poster presented at the workshop for psycholinguistic approaches to speech recognition in adverse conditions*, Bristol.

Diehl, R. L., & Kluender, K. R. (1989). On the objects of speech perception. *Ecological Psychology*, 1, 121–144.

Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., & Mehler, J. (1999). Epenthetic vowels in Japanese: A perceptual illusion?. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 1568–1578.

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

Escudero, P., Hayes-Harb, R., & Mitterer, H. (2008). Novel second-language words and asymmetric lexical access. *Journal of Phonetics*, 36, 345–360.

Fais, L., Kajikawa, S., Werker, J., & Amano, S. (2005). Japanese listeners' perceptions of phonotactic violations. *Language and Speech*, 48, 185–201.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–276). Timonium, MD: York Press.

Floccia, C., Butler, J, Goslin, J., & Ellis, L. (2009). Regional and foreign accent processing in English: Can listeners adapt?. *Journal of Psycholinguistic Research*, 38, 379–412.

Floccia, C., Goslin, J., Girard, F., & Konopcynski, G. (2006). Does a regional accent perturb speech processing?. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 1276–1293.

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

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

Gow, D. W. (2001). Assimilation and anticipation in continuous spoken word recognition. *Journal of Memory and Language*, 45, 133–159.

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

Guion, S. G., Flege, J. E., Akahane-Yamada, R., & Pruitt, J. C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America*, 107, 2711–2724.

Gussenhoven, C. (1999). Dutch. In *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet* (pp. 74–77). Cambridge: University Press.

Haggard, M. (1978). The devoicing of voiced fricatives. *Journal of Phonetics*, 6, 95–102.

Hanulikova, A., & Weber, A. (2010). Production of English interdental fricatives by Dutch, German, and English speakers. In *Proceedings of the sixth international symposium on the acquisition of second language speech, new sounds* (pp. 173–178), Poznan, Poland.

Hanulikova, A., & Weber, A. (submitted for publication). Sink positive: Linguistic experience with th-substitutions influences non-native word recognition.

Hayes-Harb, R., Smith, B. L., Bent, T., & Bradlow, A. (2008). The interlanguage speech intelligibility benefit for native speakers of Mandarin: Production and perception of English word-final voicing contrasts. *Journal of Phonetics, 36*, 664–679.

Homma, Y. (1992). *Acoustic phonetics in English and Japanese*. Tokyo: Yamaguchi-shoten.

Hyltenstam, K., & Abrahamsson, N. (2003). Maturational constraints in SLA. In Doughty, & Long (Eds.), *The handbook of second language acquisition*. Rowley, MA: Blackwell.

Imai, S., Walley, A. C., & Flege, J. E. (2005). Lexical frequency and neighborhood density effects on the recognition of native and Spanish-accented words by native English and Spanish listeners. *Journal of the Acoustical Society of America, 117*, 896–907.

Inoue, M. (2005). Validity of selecting the most important words from a vocabulary list of 550 Katakana loan words. *Journal of Kobe Yamate University, 7*, 1–14.

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

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.

Kuhl, P. K. (1993). Innate predispositions and the effects of experience in speech perception: The native language magnet theory. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. McNeilage, & J. Morton (Eds.), *Developmental neurocognition: Speech and face processing in the first year of life* (pp. 259–274). Dordrecht: Kluwer Academic Publishers.

Ladefoged, P. (1993). *A course in phonetics* (3rd ed.). Orlando: Harcourt Brace.

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

Leikin, M., Ibrahim, R., Eviatar, Z., & Sapir, S. (2009). Listening with an accent: Speech perception in a second language by late bilinguals. *Journal of Psycholinguistic Research, 38*, 447–457.

Lemhöfer, K., & Broersma, M. (2009). LexTALE: A quick, but valid measure for English proficiency. *Poster presented at the 15th annual conference on architectures and mechanisms for language processing* (AMLaP 2009), Barcelona, Spain.

Lemhöfer, K., & Broersma, M. (submitted for publication). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English.

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

Lovin, J. B. (1974). Loanwords and the phonological structure in Japanese. *Unpublished Ph.D. dissertation*, University of Chicago.

Markham, D. (1999). Listeners and disguised voices: The imitation and perception of dialectal accent. *Forensic Linguistics: The International Journal of Speech, Language and the Law, 6*, 289–299.

Marslen-Wilson, W. D. (1993). Issues of process and representation in lexical access. In G. Altmann, & R. Shillcock (Eds.), *Cognitive models of language processes: The second Sperlonga meeting* (pp. 187–210). Hillsdale: Lawrence Erlbaum.

Marslen-Wilson, W. D., 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., & Warren, P. (1994). Levels of perceptual representation and process in lexical access. *Psychological Review, 101*, 653–675.

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.

Massaro, D. (1987). *Speech perception by ear and eye: A paradigm for psychological inquiry*. Hillsdale, NJ: Erlbaum.

McQueen, J. M., Norris, D., & Cutler, A. (1999). Lexical influence in phonetic decision making: Evidence from subcategorical mismatches. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 1363–1389.

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

Mitterer, H., Csepe, V., & Blomert, L. (2006). The role of perceptual integration in the recognition of assimilated word forms. *The Quarterly Journal of Experimental Psychology, 59*, 1395–1424.

Mitterer, H., & Ernestus, M. (2006). Listeners recover /t/'s that speakers reduce: Evidence from /t/-lenition in Dutch. *Journal of Phonetics, 34*, 73–103.

Morrison, G. S. (2002). Interaction of cues to vowel identity and consonant voicing: Cross-language perception. *Journal of the Acoustical Society of America, 111*, 2362.

Munro, M. J., & Derwing, T. M. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning, 49*, 285–310.

Munro, M. J., Derwing, T. M., & Morton, S. L. (2006). The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition, 28*, 111–131.

Neuhauser, S., & Simpson, A. (2007). Imitated or authentic? Listeners' judgment of foreign accents. In *Proceedings of the 16th international congress of phonetic sciences* (pp. 1805-1808), Saarland University, Germany.

Nishi, K., & Kewley-Port, D. (2007). Training Japanese listeners to perceive American English vowels: Influence of training sets. *Journal of Speech, Language and Hearing Research, 50*, 1496–1509.

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

Otake, T., Hatano, G., & Yoneyama, K. (1996). Speech segmentation by Japanese listeners. In T. Otake, & A. Cutler (Eds.), *Phonological structure and language processing: Crosslinguistic studies* (pp. 183–201). Berlin: Mouton de Gruyter.

Pallier, C., Colomé, A., & Sebastián-Gallés, N. (2001). The influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries. *Psychological Science, 12*, 445–449.

Picard, M. (2002). The differential substitution of English /θ ð/ in French: The case against underspecification in L2 phonology. *Lingvisticæ Investigationes, 25*, 87–96.

Pierrehumbert, J. B. (2001). Exemplar dynamics: Word frequency, lenition, and contrast. In J. Bybee, & P. Hopper (Eds.), *Frequency effects and emergent grammar* (pp. 137–157). Amsterdam: John Benjamins.

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*, 19–36.

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

Ranbom, L. J., Connine, C. M., & Yudman, E. M. (2009). Is phonological context always used to recognize variant forms in spoken word recognition? The role of variant frequency and context distribution. *Journal of Experimental Psychology: Human Perception and Performance, 35*, 1205–1220.

Schouten, M. E. H. (1975). *Native-language interference in the perception of second-language vowels: An investigation of certain aspects of the acquisition of a second language*. Unpublished doctoral dissertation, Utrecht University, The Netherlands.

Scott, D. R., & Cutler, A. (1984). Segmental phonology and the perception of syntactic structure. *Journal of Verbal Learning and Verbal Behavior, 23*, 450–466.

Sebastián-Gallés, N., Echeverría, S., & Bosch, L. (2005). The influence of initial exposure on lexical representation: Comparing early and simultaneous bilinguals. *Journal of Memory and Language, 52*, 240–255.

Sebastián-Gallés, N., Rodríguez-Fornells, A., de Diego-Balaguer, R., & Díaz, B. (2006). First- and second-language phonological representations in the mental lexicon. *Journal of Cognitive Neuroscience, 18*, 1277–1291.

Shibatani, M. (1990). *The languages of Japan*. New York: Cambridge University Press.

Smits, R. (2001). Hierarchical categorization of coarticulated phonemes: A theoretical analysis. *Perception & Psychophysics, 63*, 1109–1139.

Spinelli, E., McQueen, J. M., & Cutler, A. (2003). Processing resyllabified words in French. *Journal of Memory and Language, 48*, 233–254.

Stevens, K. N., Blumstein, S. E., Glicksman, L., Burton, M., & Kurowski, K. (1992). Acoustic and perceptual characteristics of voicing in fricatives and fricative clusters. *Journal of the Acoustical Society of America, 91*, 2979–3000.

Stibbard, R. M., & Lee, J.-I. (2006). Evidence against the mismatched interlanguage intelligibility benefit hypothesis. *Journal of the Acoustical Society of America, 120*, 433–442.

Strange, W. (Ed.) (1995). Speech perception and linguistic experience: Issues in cross-language research. Baltimore: York Press.

Strange, W., Akahane-Yamada, R., Kubo, R., Trent, S., Nishi, K., & Jenkins, J. (1998). Perceptual assimilation of American English vowels by Japanese listeners. *Journal of Phonetics, 26*, 311–344.

Strange, W., Weber, A., Levy, E. S., Shafiro, V., Hisagi, M., & Nishi, K. (2007). Acoustic variability within and across German, French, and American English vowels: Phonetic context effects. *Journal of the Acoustical Society of America, 122*, 1111–1129.

Sumner, M., & Samuel, A. G. (2005). Perception and representation of regular variation: The case of final /t/. *Journal of Memory and Language, 52*, 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*, 487–501.

Takata, Y., & Nabelek, A. K. (1990). English consonant recognition in noise and in reverberation by Japanese and American listeners. *Journal of the Acoustical Society of America, 88*, 663–666.

Tanaka, S. (2009). *Nitizyou gengo ni hisomu on-housoku no sekai [Phonological laws and patterns hidden in our language]*. Tokyo: Kaitakusha Publisher.

Tate, D. A. (1979). Preliminary data on dialect in speech disguise. In *Current issues in the phonetic sciences: Proceedings of the IPS-77 congress* (pp. 847–850). Amsterdam: Benjamins.

Teasdale, A. M. (1997). On the differential substitution of English [θ]: A phonetic approach. *Calgary Working Papers in Linguistics, 19*, 71–85.

Thompson, I. (2001). Japanese speakers. In M. Swan, & B. Smith (Eds.), *Learner English: A teacher's guide to interference and other problems* (pp. 296–309). Cambridge, UK: Cambridge University Press.

Tops, G. A. J., Dekeyser, X., Devriendt, B., & Geukens, S. (2001). Dutch speakers. In M. Swan, & B. Smith (Eds.), *Learner English: A teacher's guide to interference and other problems* (pp. 1–20). Cambridge, UK: Cambridge University Press.

van Ooijen, B. (1996). Vowel mutability and lexical selection in English: Evidence from a word reconstruction task. *Memory & Cognition, 24*, 573–583.

van Wijngaarden, S. J. (2001). Intelligibility of native and non-native Dutch speech. *Speech Communication, 35*, 103–113.

van Wijngaarden, S. J., Steeneken, H. J. M., & Houtgast, T. (2002). Quantifying the intelligibility of speech in noise for non-native listeners. *Journal of the Acoustical Society of America, 111*, 1906–1916.

Vroomen, J., & de Gelder, B. (1999). Lexical access of resyllabified words: Evidence from phoneme monitoring. *Memory & Cognition, 27*, 413–421.

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Learning, 50*, 1–25.

Whalen, D. H. (1991). Perception of the English [s] [S] distinction relies on fricative noises and transitions, not on brief spectral slices. *Journal of the Acoustical Society of America, 90*, 1776–1785.