The perception of English front vowels by North Holland and Flemish listeners: Acoustic similarity predicts and explains cross-linguistic and L2 perception

Paola Escudero [a,*], Ellen Simon [b], Holger Mitterer [c]

[a] MARCS Auditory Laboratories, University of Western Sydney, Locked Bag 1797, Penrith, NSW 2751, Australia
[b] Linguistics Department, Ghent University, Muinkkaai 42, 9000 Ghent, Belgium
[c] Max-Planck-Institut für Psycholinguistik, Postbus 310, 6500 AH Nijmegen, The Netherlands

ARTICLE INFO

ABSTRACT

We investigated whether regional differences in the native language (L1) influence the perception of second language (L2) sounds. Many cross-language and L2 perception studies have assumed that the degree of acoustic similarity between L1 and L2 sounds predicts cross-linguistic and L2 performance. The present study tests this assumption by examining the perception of the English contrast between /ɛ/ and /æ/ in native speakers of Dutch spoken in North Holland (the Netherlands) and in East- and West-Flanders (Belgium). A Linear Discriminant Analysis on acoustic data from both dialects showed that their differences in vowel production, as reported in and Adank, van Hout, and Van de Velde (2007), should influence the perception of the L2 vowels if listeners focus on the vowels' acoustic/auditory properties. Indeed, the results of categorization tasks with Dutch or English vowels as response options showed that the two listener groups differed as predicted by the discriminant analysis. Moreover, the results of the English categorization task revealed that both groups of Dutch listeners displayed the asymmetric pattern found in previous word recognition studies, i.e. English /æ/ was more frequently confused with English /ɛ/ than the reverse. This suggests a strong link between previous L2 word learning results and the present L2 perceptual assimilation patterns.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Non-native listeners are known to have difficulty perceiving the contrast between vowels or consonants, which do not occur in their native language (Bohn & Munro, 2007; Strange, 1995). Numerous previous studies have examined the effect of different native language backgrounds on non-native speech perception (e.g. Escudero, Benders, & Lipski, 2009; Iverson et al., 2003; Lengeris, 2009; Polka, Colantoni, & Sundara, 2001). Many of these cross-language and L2 perception studies either implicitly or explicitly assume that the degree of acoustic similarity between L1 and L2 sounds predicts listeners' perceptual assimilation of L2 sounds. Flege's Speech Learning Model (SLM, Flege, 1987, 1995), focusing on experienced L2 learners, posits that the smaller the acoustic-phonetic distance between an L1 and an L2 sound, the more difficult it will be for the learner to establish a new category for the L2 sound. However, Flege (1987, 1995) does not propose

a phonetic metric, i.e. it is not specified how the acoustic–phonetic distance should be measured.

According to Best's (1995) Perceptual Assimilation Model for naïve, inexperienced listeners (PAM) as well as Best and Tyler's (2007) Perceptual Assimilation Model for L2 listeners (PAM-L2), listeners assimilate nonnative/L2 sounds to their native categories. As Best and Tyler (2007: 28) point out, this assimilation is based not only on phonetic similarity between the L2 and the L1 sound, as is the case in Flege's SLM (1995), but also on how L2 sounds are perceived as belonging to different phonological categories. As the PAM is based on Articulatory Phonology (Goldstein & Fowler, 2003), it states that the articulatory gestures of the native language, rather than acoustic values, play a crucial role in how L2 sounds are assimilated to native sound categories.

Escudero's Second Language Perception model (L2LP, Escudero & Boersma, 2004; Escudero, 2005, 2009) also proposes a direct link between the perception and production of sounds, but in a different fashion than the SLM. This model states that listeners' sound perception should match the production of sounds in their native language. Escudero and Boersma (2004) demonstrate that /i/ and /ɪ/ have different acoustic properties in Standard Scottish English (SSE) and Standard Southern British English (SSBE) and

that listeners' perception of these vowels closely resembles the specific acoustic properties of their respective dialect. Specifically, the SSE vowels have a larger spectral difference than the SSBE vowels, and vowel duration hardly distinguishes the SSE vowels, while it is an important difference between the SSBE vowels. Escudero (2005) shows that monolingual Spanish listeners perceive the /i/-/ɪ/ contrast differently depending on whether it is produced by an SSE or an SSBE speaker. Specifically, they assimilate the SSE vowel contrast to their Spanish /i/-/e/ contrast, while they classify most tokens of the SSBE vowels as Spanish /i/. Additionally, the authors demonstrate that Spanish learners of English have differential degrees of difficulty for this same English contrast, depending on whether they have SSE or SSBE as their target language. Thus, according to the L2LP model, the acoustic similarity between L1 and L2 sounds influences non-native perception and hence predicts the degree of difficulty that an L2 learner will experience.

By contrast, Strange, Bohn, Trent, and Nishi (2004, 2005) have argued that acoustic similarity is not always a good predictor of cross-language speech perception. They examined the phonetic similarity between the first three formants of North German (NG) and American English (AE) vowels using linear discriminant analysis (Klecka, 1980) and found that the perceptual assimilation could not always be predicted from acoustic similarity. However, a more recent study by Gilichinskaya and Strange (2010), who used the same discriminant analysis technique used in Strange et al. (2004, Strange, Bohn, Nishi, and Trent, 2005), showed that acoustic similarity was a good predictor of cross-language assimilation patterns of American English vowels by Russian listeners, especially for point vowels.

The present study contributes to this debate by examining the perception of the English contrast between /ɛ/ and /æ/ by native speakers of two regional varieties of Dutch, namely the North Holland variety of Dutch spoken in the Netherlands, and the East- and West-Flemish varieties of Dutch spoken in Belgium. These two varieties both lack the English /ɛ/-/æ/ contrast, but differ in the phonetic realization of the Dutch front vowel /ɛ/ (see Section 1.1). Previous studies have shown that dialect differences can have strong effects on the perception of the native language (Floccia, Goslin, Girard, & Konopczynski, 2006; Sumner & Samuel, 2009). This suggests that it is worthwhile to investigate how dialect differences influence the perception of a second language. To this end, we first examine to what extent acoustic similarity between Dutch /ɛ/ in different varieties of Dutch and the two English front vowels can predict patterns of perceptual assimilation. Secondly, we test whether the asymmetric patterns found in the recognition of words containing the /ɛ/-/æ/ contrast (see Section 1.2) by Dutch learners with high English proficiency can also be found in their perception of the same contrast.

1.1. Acoustic differences between L1 dialects

Previous studies have shown that the Dutch vowel /ɛ/ has different acoustic values in different dialects of Dutch. Adank, van Hout, and Smits (2004) and Adank, van Hout, and Van de Velde (2007) measured vowel productions of Dutch speakers from different areas in the Netherlands and Belgium, including North Holland and East- and West-Flanders. The Dutch vowels were produced in the single consonantal context 's–V–s' (V=vowel), so as to avoid any potential dialect-dependent contextual variation. Fig. 1 presents a vowel plot with the F1 and F2 values of /ɛ/ in these three varieties of Dutch and in Standard British English. The figure shows that English /ɛ/ seems closer to North Holland /ɛ/ in terms of F2, but closer to East- and West-Flemish /ɛ/ in terms of F1. Additionally, it can be observed that North Holland /aː/ has a
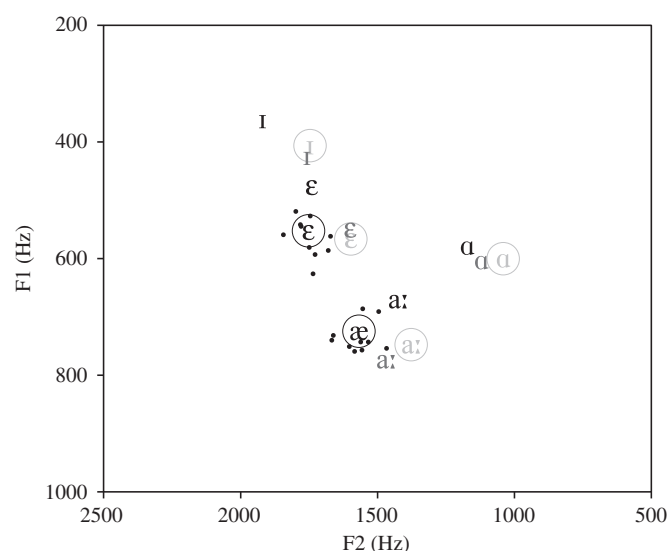

Fig. 1. Vowel plot presenting Average F1 and F2 values (in Hertz) of English /ɛ/ and /æ/ (circled black) produced by a male speaker of Southern British English (taken from Escudero et al., 2008) and of the closest vowels produced by Dutch (black), East Flemish (circled light grey) and West Flemish (dark grey) speakers (taken from Adank et al., 2007). The dots represent the individual tokens of the English vowels used in the present study.

lower F1 than English /æ/, whereas East- and West-Flemish /aː/ has a higher F1 value.

According to models such as Flege's (1995) SLM and Escudero's (2005) L2LP model, the different acoustic realizations of L1 vowels are predicted to lead to different perceptions of L2 sounds by North Holland and East- and West-Flemish listeners. Adank, van Hout, and Van de Velde (2007) show that vowels in Randstad Dutch, which is spoken in the Netherlands and includes North Holland, had significantly different F1 and F2 values from those in East and West Flemish. However, the extent to which these acoustic differences may yield different assimilation patterns for the English vowels /ɛ/ and /æ/ can only be established by means of an appropriate statistical method. While Fig. 1 presents average F1 and F2 values, Linear Discriminant Analysis (Klecka, 1980) can include the vowels of a large number of tokens per dialect and acoustic properties other than F1 and F2, such as vowel duration and F3, which should generate more accurate predictions.

The present study investigates to what extent L1 speakers of different dialects perceive L2 vowels differently, and whether acoustic similarity between native dialect vowels and L2 vowels can explain the differential perceptual assimilation patterns. Following previous studies on the effect of acoustic similarity on non-native vowel perception (Gilichinskaya & Strange, 2010; Strange et al., 2004, 2005), we will use Linear Discriminant Analysis to examine whether dialectal acoustic differences predict different perceptual assimilation patterns.

We then compare the results of the discriminant analysis to those of the Dutch listeners' non-native perception, as tested in a Dutch categorization task of English vowels. According to the L2LP model and to the concept of "language modes" or "language settings" (Grosjean, 2001), a task that promotes the activation of a single language, i.e. an L1 task with L1 options, can tap into the monolingual perception of the stimuli presented, because listeners' other languages are not needed to perform the task. Specifically, a Dutch categorization task, in which listeners are only addressed in Dutch, told that the stimuli are Dutch, and only have Dutch vowels as options, is likely to yield results that are based on monolingual Dutch perception, and may not depend on listeners' experience with other languages. Supporting evidence is found in

Escudero and Boersma (2002), who reported on the results of a categorization task in which Dutch-speaking learners of Spanish were asked to map Spanish vowels to Dutch vowels. These results revealed that Dutch listeners had comparable perception of Spanish vowels, regardless of whether they were beginning, intermediate or advanced learners of Spanish. Similarly, Mayr and Escudero (2010) found comparable results in English learners of German with different levels of L2 experience when presented with an English categorization task, i.e. an L1 task with L1 options (see also Escudero, 2009). The L2LP model's prediction stands in contrast to that posed within Flege's SLM, which states that the L1 phonetic system is not immune to influence from the L2 (e.g., Yeni-Komshian, Flege, & Liu, 2000).

In addition, according to the PAM-L2 and L2LP models, this monolingual perception of the sounds of a new language can be considered the initial state of L2 sound perception. Therefore, the extent to which L1 Dutch regional differences have an effect on the categorization of English vowels provides us with information on the initial state of L2 learning. Since some of these regional differences between learners are predicted to persist in later stages of the learning process, the Dutch task was followed by an English categorization task. We will demonstrate that the differences in non-native perception observed in the Dutch categorization task explain differences in the categorization of the same tokens in an English task, i.e. an L2 task with L2 options.

1.2. Asymmetric mapping of nonnative contrasts

The non-native perception of speech sounds necessarily affects the recognition of words containing those sounds. Therefore, it is not surprising that Dutch listeners have difficulty recognizing minimal pairs containing the English vowels /ɛ/ and /æ/ (such as 'bed' and 'bad') (Broersma, 2005; Weber & Cutler, 2004). Weber and Cutler (2004) examined lexical competition in non-native spoken word recognition with eye-tracking technology, focusing on Dutch natives' recognition of English word pairs differing minimally in the vowels /ɛ/ and /æ/. Their results showed that there was confusion between words containing /ɛ/ (such as 'pencil') and words containing /æ/ (such as 'panda'). Crucially, this confusion was asymmetric: if the auditory target was 'panda', L2 listeners, unlike L1 listeners, also fixated their eyes on a picture of a 'pencil' during the first syllable of the target word. However, if the target word was 'pencil', the word 'panda' yielded very few eye fixations.

Escudero, Hayes-Harb, and Mitterer (2008) confirmed these results with their study of the effect of orthographic information during the learning of novel English words containing /ɛ/ and /æ/. They divided participants into two groups: one group was trained on word-picture pairings with auditory information only; the other group received both auditory information and the written form of the word. During testing, participants heard an auditory form and were asked to click on the corresponding picture. An analysis of their eye gaze revealed that participants who had been trained with orthography showed the same asymmetric pattern observed in Weber and Cutler (2004), while the auditory-only group did not show an asymmetry. The authors suggest that only the group trained with orthography was able to build separate lexical representations for minimal pairs differing only in the vowels /ɛ/ and /æ/.

The second objective of the present study is to test whether this asymmetric pattern of word recognition is also found in a task that does not involve lexical knowledge, namely sound categorization. One can expect that Dutch listeners classify tokens of the English vowel /æ/ as both /æ/ and /ɛ/, while they classify English /ɛ/ only as /ɛ/. However, no such asymmetry was consistently found by Cutler, Weber, Smits, and Cooper (2004)

in an English sound categorization task. In their study, American English and Dutch native listeners were asked to categorize American English CV and VC syllables by clicking on an English word containing the appropriate sound (the authors also used stimuli with different degrees of background noise but we only refer to the categorization of the clear stimuli, which had no noise added). The results for the English contrast /ɛ/-/æ/ showed that English /ɛ/ and /æ/ showed similar percentage correct classifications and were hence both confused to similar extents, at least in a VC condition.[1] Cutler et al. hence did not confirm that the asymmetry observed in word recognition applied equally to sound categorization. One possible reason for this somehow paradoxical result might be due to the fact that Weber and Cutler (2004), who observed an asymmetry in word recognition, used tokens produced by a SSBE speaker, while Cutler et al. (2004) used tokens produced by an AE speaker. The present study seeks to examine whether the asymmetry can be found in an English sound categorization task that uses stimuli produced by the same SSBE speaker who spoke Weber and Cutler's (2004) stimuli and who was recorded under the same conditions of this previous study. If such an asymmetry were to be found, it would provide evidence for a strong continuity between perception and word recognition.

2. Method

We designed two sound categorization studies, one in Dutch and one in English, to answer the two research questions of the present study. First, we aimed to examine whether acoustic similarity between /ɛ/ in different varieties of Dutch and the two English front vowels can predict patterns of perceptual assimilation. And secondly, we sought to investigate whether an asymmetric pattern of confusion of the two English vowels /æ/ and /ɛ/ is found in speakers of different regional varieties of Dutch in a sound categorization task.

2.1. Participants

In total, 59 informants participated in this study. They were native speakers of Dutch living in the Netherlands or Belgium at the time of testing. They were studying English at university at the time of testing and had received 1.5 years of English instruction and English-medium courses at university. All participants had SSBE as their English target dialect and their model at university. From this larger pool of informants, 42 were selected for analysis on the basis of their regional background and their scores on a general English language comprehension test (DIALANG, Alderson & Huhta, 2005). The selected participants came from two regions, 21 from North Holland and 21 from East- or West-Flanders in Belgium. They were tested in Amsterdam (North Holland) and in Ghent (East-Flanders), respectively. The acoustic values of /ɛ/ in East- and West-Flemish are very similar, as shown in the previous section, and therefore participants from these two areas were treated as one group, which will be referred to as the Flemish listeners.

Prior to the vowel perception tasks, participants performed the general comprehension test, where they heard short speech

---

[1] Dutch listeners correctly classified English /ɛ/ and /æ/ 60% and 52% of the time, respectively, in the CV context and 58% and 56% of the time in the VC context. Additionally, in the VC condition, /ɛ/ was incorrectly classified as /æ/ 25% of the time and /æ/ as /ɛ/ 34% of the time, while in the CV condition /ɛ/ was incorrectly classified as /æ/ 22% of the time and /æ/ was categorized as /ɛ/ 39% of the time. Given that listeners had 15 options to choose from, these classifications are well above chance (6.7%).

fragments (either by one speaker or in the form of a dialogue). Before listening to the fragment, they saw a written question about the content of the fragment at the bottom of the screen, with multiple choice answers. They could listen to each fragment once. Once they had marked their response, they moved on to the next fragment. Only participants who scored C2, C1 or B2 (Flemish: $5 \times$ C2, $12 \times$ C1, $4 \times$ B2; North Holland: $7 \times$ C2, $11 \times$ C1, $3 \times$ B2) were selected for the analysis of their vowel perception. These scores are roughly equivalent to "highly advanced", "advanced" and "upper intermediate" language proficiency levels, respectively.

Participants took part in the study for course credit or received a book voucher for participation. All participants completed a language background questionnaire. Their mean age was 21 for the Flemish participants (range: 19–21, one participant was 38) and 22 for the Dutch participants (range: 19–25). In total, 9 (5 Flemish and 4 Dutch) of the participants were male and 33 were female (16 Flemish and 17 Dutch).

2.2. Stimuli

The target stimuli were five types of English CVC sequences containing the vowels /ɛ/ and /æ/. The sequences were spliced from the disyllabic nonword stimuli used by Escudero et al. (2008), which had been produced by a male native speaker of Standard Southern British English (see Table 1). All nonwords were produced with stress on the first syllable, i.e. on the part that was later spliced for presentation. The advantage of using spliced stimuli is twofold. First, if participants are told that the stimuli are cut out of longer words, the fact that they have phonotactic patterns which do not necessarily exist in Dutch is no longer relevant. For instance, Dutch does not have voiced obstruents in word-final position, but has them in word-medial position. And secondly, if listeners know that the stimuli are not words but parts of words, the distinction between words and nonwords is blurred for purposes of the perception experiment (e.g. 'ten' is not an existing Dutch word, but is part of the Dutch word 'tent'). If any lexical biases had nevertheless cropped up during the experiment, they can be assumed to have been equal for both listener groups.

Each stimulus was produced twice by the same native English speaker, such that 20 acoustically different tokens were used. Table 2 presents the F1 and F2 values for the two productions of each word.

2.3. Design and procedure

Participants performed two tasks, namely a Dutch categorization task, in which they matched English vowels to their closest Dutch equivalents and rated the vowels on a category goodness scale, and an English categorization task, in which they classified English vowels. Both tasks were presented using Praat (Boersma & Weenink, 2010). The stimuli were the same English CVC tokens in both the English and the Dutch categorization tasks, which were spliced from disyllabic English nonwords (see Section 2.2). In total, there were 10 different acoustic realizations of /ɛ/ and /æ/, which were each repeated 6 times, leading to a total of 120 target tokens in both the Dutch and English tasks.[2]

Participants first performed the Dutch categorization task, then, after a pause, the English DIALANG comprehension test, and finally the English categorization task. The order in which the tasks were performed was the same for all participants, because it favored listeners' (L1) Dutch mode during the first task, and activated their English during the English comprehension test, which means that they were likely to be in an English mode for the English categorization task. They were addressed solely in Dutch during the first task and solely in English during the second and third tasks.[3]

For all three tasks, participants sat in front of a computer screen and the audio stimuli were presented to them binaurally over headphones. The session, including the language comprehension test and the two categorization tasks, lasted about an hour. While the categorization tasks were self-timed, i.e. a new audio stimulus was presented only after participants had categorized the previous one, participants were encouraged to complete the tasks as quickly as they could.

2.3.1. The Dutch categorization task

In the first task, oral and written instructions were provided solely in Dutch and participants did not know that the stimuli were English words or that they would be performing an English task afterwards. They saw a screen with 12 boxes containing words with each of the 12 Dutch monophthongs in an hVk context. In order to draw the attention to the vowel sound, the vowel letters were capitalized. The orthographic forms of the 12 Dutch nonwords and the phonetic transcriptions of their vowels are shown in Table 3. Only the orthographic forms were presented on the screen.

Participants heard one stimulus at a time and were asked to first click on the box containing the word with the vowel that was most similar to that of the word they had just heard and, secondly, to indicate on a scale from 1 ('slecht', *bad*) to 7 ('goed', *good*) how good an example of such a vowel it was. The instruction 'Choose the vowel that you heard and then say how good it is' (in Dutch: 'Kies de klinker die je gehoord hebt en zeg daarna hoe goed hij is') remained at the top of the screen throughout the experiment. Participants were told that they may need all the different options or just a few of them. As soon as they had clicked both on a word box and a number on the scale, the next audio stimulus was presented.

It is worth mentioning that some stimuli contained phonetic realizations which do not occur in Dutch and which could give the Dutch-speaking listeners a cue about the language in which they were uttered. For instance, voiceless stops are unaspirated in Dutch, and the aspirated [tʰ] in [tʰɛn] and [tʰæn] could thus have been associated with English. Similarly, /ɡ/ is not part of the phonemic inventory of Dutch, but only occurs in coda position as an allophone of /k/ as the result of regressive voice assimilation. The initial [ɡ] in [ɡɛb] and [ɡæb] could thus again have made the listeners presume that the language they heard was English. However, it should be noted that after task completion none of

Table 1
Target stimuli, which were spliced from the nonwords used by Escudero et al. (2008).

/ɛ/			/æ/		
Stimulus	Spliced from		Stimulus	Spliced from	
[tɛn]	Tenzer	[tɛnzə]	[tæn]	Tandek	[tændək]
[mɛs]	meskle	[mɛskəl]	[mæs]	Mastic	[mæstək]
[ɡɛb]	gebbet	[ɡɛbət]	[ɡæb]	Gabble	[ɡæbəl]
[hɛs]	hestel	[hɛstəl]	[hæs]	Haskum	[hæskəm]
[bɛs]	bestet	[bɛstət]	[bæs]	Baskle	[bæskəl]

Table 2
F1 and F2 values in Hz (values in Bark are given between parentheses) for the two tokens of the two vowels produced in each of the contexts. V=vowel. S.D.=standard deviation of the mean.

	V=/ɛ/		V=/æ/	
	F1	F2	F1	F2
tVn, token 1:	515 (5.1)	1746 (12.1)	731 (6.8)	1561 (11.4)
tVn, token 2:	533 (5.2)	1778 (12.2)	739 (6.8)	1602 (11.5)
mVs, token 1:	581 (5.6)	1727 (12)	745 (6.9)	1556 (11)
mVs, token 2:	614 (5.9)	1735 (12.1)	679 (6.4)	1495 (11.1)
ɡVb, token 1:	547 (5.3)	1843 (12.5)	720 (6.7)	1661 (11.8)
ɡVb, token 2:	507 (5)	1798 (12.3)	728 (6.7)	1666 (11.8)
hVs, token 1:	530 (5.2)	1781 (12.2)	675 (6.3)	1553 (11.3)
hVs, token 2:	569 (5.5)	1749 (12.1)	747 (6.9)	1583 (11.5)
bVs, token 1:	550 (5.4)	1671 (11.8)	742 (6.8)	1466 (10.9)
bVs, token 2:	574 (5.5)	1679 (11.8)	731 (6.8)	1533 (11.2)
Mean across tokens and contexts	552 (5.4)	1751 (12.1)	724 (6.7)	1567 (11.4)
S.D.	33 (0.3)	53 (0.2)	26 (0.2)	64.21 (0.3)

Table 3
The 12 choices in the Dutch categorization task.

hIEk	/hik/	hEEk	/heːk/	hIk	/hɪk/	hEk	/hɛk/
hUUk	/hyk/	hEUk	/høk/	hUk	/hʏk/	hAAk	/haːk/
hOEk	/huk/	hOOk	/hoːk/	hOk	/hɔk/	hAk	/hɑk/

Table 4
The 12 choices in the English categorization task.

bee	[biː]	may	[mɛɪ]	tip	[tɪp]	hot	[hɒt]
put	[pʊt]	class	[klɑːs]	cat	[kæt]	mug	[mʌg]
pet	[pɛt]	go	[gəʊ]	call	[kɔːl]	zoo	[zuː]

the participants spontaneously remarked that they had heard English words. We assume that listeners were in a Dutch monolingual perception mode because they were only addressed in Dutch during this part of the experiment, oral and written instructions were in Dutch, and the words on the screen were read aloud in Dutch before the start of the experiment. This assumption is also motivated by previous findings of monolingual perception for Dutch and English listeners, regardless of their L2 proficiencies, when presented with a monolingual task similar to the one used in the present study (Escudero & Boersma, 2002; Mayr & Escudero, 2010).

2.3.2. The English categorization task

For the English categorization task, which was performed after the English general comprehension task, oral and written instructions were given solely in English. The orthographic forms on the screen represented English rather than Dutch vowels and participants did not have to rate the goodness of the tokens on a scale. Since English has an opaque orthographic system, real words rather than nonwords were used and the vowel letters were not capitalized, as the same letters represent different vowels (e.g. the letter ⟨a⟩ represents /æ/ in 'cat', /ɑː/ in 'class' and /ɔː/ in 'call'). The orthographic forms and their phonetic transcriptions are presented in Table 4.

During the task, the instruction 'Choose the vowel that you heard' remained on top of the screen throughout the experiment. The participants were asked to read the words to the experimenter before the experiment started, so that the experimenter could confirm that the participants had the intended Southern British English vowels in mind when performing the task (e.g. that they did not produce the word 'class' with an /æ/ instead of an /ɑː/).

3. Results

3.1. Discriminant analysis and predictions for perceptual assimilation

We used Linear Discriminant Analysis (LDA, Klecka, 1980) to predict assimilation results from the acoustic properties of the English vowels and those of the closest North Holland and Flemish vowels. We followed the procedure reported in Strange et al. (2004, 2005) and Gilichinskaya and Strange (2010). Two different native discriminant analysis models, one for North Holland and one for Flemish,[4] were trained on the F1, F2, F3 (in Bark) and duration values of the Dutch vowels /ɪ, ɛ, aː and ɑ/. These are the vowels that are acoustically closest to Southern British English /ɛ/ and /æ/ and thus the main candidates for assimilation of the English tokens into Dutch categories. The values were taken from Adank et al.'s (2007) corpus for Northern Standard Dutch, East and West Flemish. The formant measures entered in each of the models correspond to those of the two repetitions of each vowel by 10 male speakers of each dialect, which were measured at the midpoint of the vowel.[5] Only Dutch male speakers were used because the English tokens were produced by a male speaker, and therefore including both male and female speakers for the native Dutch model would add an extra parameter that could make the cross-language model more difficult to interpret. We used the cross-validation method to test the accuracy of the models' classifications: four tokens of each vowel were randomly assigned to the cross-validation set (20%) and the remaining 80% were assigned to the training set for North Holland. For Flemish, which had twice the number of tokens, eight tokens were used for cross-validation. On the basis of the vowels' F1, F2, F3 and duration values, the models yielded 98.4% correct classifications for the training sets in both models, and 100% and 96.9% for the cross-validation set in the North Holland and Flemish models, respectively. This good performance of the LDA model is partly due to our choice of only male Dutch speakers, which reduced between-speaker variance. Note that this is justified by the fact that listeners seem to normalize for between-speaker differences at a pre-phonetic stage of processing (Sjerps, Mitterer, & McQueen, 2011; Watkins & Makin, 1996).

---

[4] A single Flemish model was used, with both East- and West-Flemish tokens, since the Flemish participant group consisted of both East- and West-Flemish listeners (see Section 2.1).
[5] Many thanks to Patti Adank for sharing her SPSS file with the values for all their individual tokens.

Table 5
Categorization (in %) of English /ɛ/ and /æ/ as Dutch vowels by North Holland (NH) and East- and West-Flemish (FL) listeners as predicted by an auditory LDA model and observed responses from listeners.

Stimuli	Dutch vowel choice									
	/aː/		/ɑ/		/ʏ/		/ɛ/		/ɪ/	
	NH	FL	NH	FL	NH	FL	NH	FL	NH	FL
LDA models										
/ɛ/	0	0	0	0			100	70	0	30
/æ/	0	0	50	0			50	100	0	0
Listeners										
/ɛ/	0	0	1	0	0	1	87	58	12	41
/æ/	2	2	35	16	0	5	62	75	0	2

The trained models hence seem to capture the vowel classifications well. This allows us to use them to predict how Dutch and Flemish listeners may perceive Southern British English vowels, by using measurements of the first three formants and vowel duration and feed these values into the existing native models. Each of the native models, North Holland versus Flemish, classified 10 tokens of the English /ɛ/ and /æ/ vowel, which were the tokens presented to the listeners in our study. The formant values for these tokens are reported in the Stimuli section. The two top rows in Table 5 show the extent (in percentages) to which /ɛ/ and /æ/ tokens were classified as each of the Dutch vowels, which were used for North Holland and Flemish discriminant analysis models. Note that the Linear Discriminant Analysis model works on the assumption that the similarity of *auditory* parameters can predict how L2 vowels are perceived in the L1. This assumption motivates the bark transformation for formant frequencies, since the bark scale more closely represents how the human auditory system processes acoustic stimuli. As such, linear discriminant analysis is a way of implementing the predictions of the L2LP model.

As can be seen in the table, the classification scores are quite different for the two varieties. For English /ɛ/, the North Holland model yields 100% Dutch /ɛ/ classifications, while the Flemish model classifies the same tokens as either /ɛ/ (70%) or /ɪ/ (30%). For English /æ/, the situation is the opposite: while the Flemish model classified all tokens as Dutch /ɛ/, the North Holland model yielded /ɛ/ (50%) and /ɑ/ (50%) classifications. Importantly, neither model yielded /aː/ classifications for the English vowels. Recall that the native models also included duration as a classification factor, and therefore the duration of the English vowels was compared to that of the Dutch vowels. It seems that both English /ɛ/ and /æ/ have durations that compare well to Dutch short vowels, and therefore a long vowel, such as /aː/, does not seem to be a good match for these English vowels.

In sum, the results of the discriminant analysis clearly predict a difference for the perceptual assimilation of English vowels by listeners from two different Dutch varieties. Specifically, following the results of this analysis, we first predict that Flemish listeners will classify English /ɛ/ as Dutch /ɪ/ more frequently than North Holland listeners. Secondly, North Holland listeners will more frequently assimilate English /æ/ to /ɑ/ than the Flemish listeners. And finally, listeners in both groups will not select Dutch /aː/ for neither of the two English vowels because both models classified them as short Dutch vowels, such as /ɪ, ɛ and ɑ/. If these predictions are confirmed, they would provide strong evidence that perceptual assimilation patterns can be predicted on the basis of acoustic measurements, since the discriminant analysis was based purely on F1, F2, F3 and vowel duration values of North Holland versus Flemish vowels.

3.2. North Holland and Flemish categorization of English vowels

Table 5 shows the percentages in which English tokens of /ɛ/ and /æ/ were classified as any of the twelve Dutch vowel options by the North Holland (NH) and East- and West-Flemish (FL) listeners. Only Dutch vowels which occurred as responses in any of the tasks for any of the stimuli are presented in the table.

Table 5 reveals that, in the overall majority of cases, English /ɛ/ was categorized as Dutch /ɛ/. The second preferred choice for both listener groups was the vowel /ɪ/. The mapping /ɛ/-to-/ɛ/ was more frequent for the North Holland than for the Flemish listeners, who chose Dutch /ɛ/ 87 and 58% of the time, respectively. In line with the predictions based on the discriminant analysis, Flemish listeners chose Dutch /ɪ/ to a larger extent than the North Holland listeners (41% versus 12%, respectively).

To test the differences in the percentage of time the listener groups chose /ɛ/ and /ɪ/, we first filtered the data to exclude the few responses that were neither /ɪ/ nor /ɛ/. With this new data set, we calculated a mixed-effect model with subjects and item as random effects and Dialect as a fixed effect. The dependent variable was whether the response was /ɛ/ or not, using a binomial linking function, which takes into account that proportions are bound to be between zero and one, thus avoiding the possibility of false inferences that can occur if a linear linking function is used (cf. Jaeger, 2008). The linear mixed effect allows us to use the single trial data by also estimating the crossed random effect from subjects and items (see, Jaeger, 2008, for additional information). For categorical fixed-effect predictors, the models map one condition on the intercept and generate regression weights for all other levels. In the current case, the Flemish group was assigned to the intercept, and the regression weight for the North Holland group ($B=4.6$, $p < 0.001$) revealed that the likelihood of an /ɛ/ response was significantly higher in the latter group. Importantly, since we limited the data analysis to the two most common options /ɛ/ and /ɪ/, an elevated likelihood of /ɛ/ responses also means a diminished likelihood of /ɪ/ responses.

With respect to English /æ/, the table shows that it was mostly categorized as either /ɛ/ or /ɑ/. The Dutch /ɛ/ category attracted the overall majority of responses but, as predicted by the discriminant analysis, this time more so for Flemish (75%) than for North Holland listeners (62%). Additionally, and again as predicted by the discriminant analysis, North Holland listeners chose another vowel more frequently than Flemish listeners, i.e. they chose Dutch /ɑ/ 35% of the time, while Flemish listeners chose the same vowel only 16% of the time. This differential perception for the two listener groups matches the classification percentages of the models that compare the acoustic properties in the two Dutch varieties with those of the English vowels. Importantly, as predicted by the discriminant analysis, the vowel /aː/ was not chosen by either of the listener groups, despite it having similar formant values to those of English /æ/ in both Dutch varieties (as shown in Fig. 1). The longer duration of /aː/ indeed makes it an unlikely response for Dutch listeners because both English vowels are short. Boersma and Escudero (2008) found a similar result for Dutch listeners of the five Spanish vowel monophthongs, for which they avoided the Dutch long vowels, including /aː/, in a Dutch categorization task similar to the one used in the present study.

The differences between listener groups in the classification of English /æ/ also proved significant in a multi-level model with a binomial linking function. The model predicted whether the response was /ɛ/ or /ɑ/ (mapped onto 0 and 1, respectively), and revealed a significant effect of listener group ($B=2.46$, $p < 0.01$). The positive regression weight means that North Holland listeners categorized /æ/ more often as /ɑ/ and less often as /ɛ/ than Flemish listeners.

Table 6
Categorization (in %) of English /ɛ/ and /æ/ as English vowels by North Holland and Flemish listeners.

Stimuli	Categorized									
	/æ/		/ɑ/		/ʌ/		/ɛ/		/ɪ/	
	NH	FL	NH	FL	NH	FL	NH	FL	NH	FL
/ɛ/	9	8	3	0	0	1	73	61	15	30
/æ/	45	61	21	7	2	5	32	26	1	0

Table 7
Effect of native dialect on the English categorization of /æ/ and /ɛ/ (based on the multi-level models with a binomial linking function). B=Beta coefficient from the multi-level models for the difference between the Flemish and North Holland groups. Positive values indicate better performance by the Flemish. p=p-value for statistical significance.

Stimuli	/æ/		/ɑ/		/ɛ/		/ɪ/	
	B	p	B	p	B	p	B	p
/ɛ/	0.35	n.s.	na	na	−2.62	< 0.05	2.48	< 0.05
/æ/	1.76	< 0.05	−1.45	=0.14	−1.67	=0.12	na	na

The results of the goodness-of-fit scores also support the observation that the English vowels are perceived differently by North Holland and Flemish listeners. The results of two-tailed paired *t*-tests revealed that North Holland participants rated Dutch /ɛ/ as a better match to English /ɛ/ than the Flemish listeners (5.2 versus 4.7 points on a scale from 1 to 7, $p < 0.001$). In contrast, the difference in how the Flemish and North Holland listeners rated /ɪ/ as a match to English /ɛ/ (4.2 versus 3.9 points) was not significant, probably due to the fact that only half of the listeners ever chose this option.[6]

In sum, North Holland and Flemish listeners have different perceptual assimilation patterns for English /ɛ/ and /æ/, as predicted by the LDA. Specifically, although the majority of English /ɛ/ and /æ/ tokens were assigned to the same Dutch category /ɛ/ by both groups of listeners, the largest difference between listeners was on the second most frequently chosen category, i.e. /ɪ/ for English /ɛ/ and /ɑ/ for English /æ/. The comparison of the listener results and those of the discriminant analysis show that the acoustic similarity between the vowels of the two dialects and those of the English vowels is a reliable predictor of the differential perceptual assimilation patterns.

3.3. English categorization

The results of the English categorization task shown in Table 6 reveal that both the North Holland and the Flemish listeners correctly classified most English /ɛ/ and /æ/ stimuli, and thus confirm the participants' high proficiency in the perception of English sounds (see Section 2). The table shows differences in classification percentages between the two listener groups. That is, although the initial stage of L2 learning, as reflected in the Dutch categorization task, may eventually be overcome by learners, clear traces of regional influence can be observed when comparing Tables 5 and 6. That is, for English /ɛ/, both groups mainly use the options /ɛ/ and /ɪ/, but, again, the Flemish listeners chose /ɪ/ more often than the Northern-Dutch listeners (in fact, twice as often, i.e. 30% versus 15%). Importantly, the presence of the new response category in English, namely /æ/, only influenced their classification of English /æ/, where there appears to be a group difference, while this option was hardly chosen when classifying English /ɛ/.

To test the significance of the difference between the two groups in the English task, we examined each English vowel separately and focused on the major response categories, i.e. the responses /æ/, /ɛ/ and /ɪ/ for English /ɛ/, and the responses /æ/, /ɛ/ and /ɑ/ for English /æ/. Each response category was used as the "correct" response in a multi-level model with a binomial linking function. In this analysis, the North Holland listeners were mapped on the intercept, so that a positive regression weight indicates that Flemish listeners use a given category more often

than their North Holland peers, while a negative regression weight indicates that they use a category less frequently.

Table 7 shows that in line with their Dutch categorization results, both groups chose the correct category for both vowels more often than they chose any other category, and the Flemish listeners correctly classified the majority of the /æ/ tokens. This confirms the participants' high proficiency in the perception of English sounds (see Section 2.1). The table also shows that Flemish listeners are more successful at correctly classifying English /æ/ than North Holland listeners.

In sum, the results of the English categorization task indicate that the Dutch categorization results, which were directly predictable from the discriminant analysis, leave traces in the English categorization results. Moreover, English /æ/ was more frequently categorized as Dutch /ɑ/ by the North Holland than by the Flemish listeners in the Dutch categorization task. Again, the same pattern can be found in the English task, in which the North Holland listeners, as opposed to the Flemish listeners, frequently mapped English /æ/ onto /ɑ/. As for the second objective of our study, the results in Table 6 clearly show that there is an asymmetry in the classification of English /ɛ/ and /æ/ in both listener groups, i.e. while English /æ/ was incorrectly mapped onto /ɛ/ in 32% and 26% of the tokens, English /ɛ/ was incorrectly mapped onto English /æ/ in only 9% and 8% of the tokens. These results demonstrate that the asymmetry observed in word recognition by Weber and Cutler (2004) and Escudero et al. (2008) is also found in the categorization of Southern British English vowels. For both types of listeners, a phonetic explanation for the asymmetry is plausible. In Section 4, we will further address the phonetic versus orthographic explanations for the current findings in comparison to those of Escudero et al. (2008).

4. Discussion

The analysis of the Dutch and English categorization tasks by listeners from North Holland and East- and West-Flanders yielded the following findings which directly answered our research questions: (1) the two listener groups differed in their non-native perception of the English vowels /ɛ/ and /æ/, (2) both listener groups showed an asymmetric pattern of L2 perception in that English /æ/ was categorized as /ɛ/ more often than /ɛ/ was categorized as /æ/ and (3) only North Holland listeners classified English /ɛ/ more accurately than /æ/ (73% versus 45% correct, respectively), while Flemish listeners had a 61% correct classification for both vowels.

With respect to the first finding, it was shown that dialectal variation in the acoustic properties of Dutch /ɛ/ seems to account for differences in non-native vowel perception. These differences were shown in the percentages with which the two listener groups assimilated the English vowels to a Dutch vowel other than /ɛ/. That is, for English /ɛ/, Flemish listeners chose Dutch /ɪ/ significantly more often than North Holland listeners, while for

---

[6] There was no significant difference between listener groups in their goodness of fit for English /æ/.

English /æ/, North Holland listeners chose Dutch /ɑ/ significantly more frequently than Flemish listeners. A linear discriminant analysis showed that these differences can be explained on the basis of the acoustic/auditory properties of the vowels. Two models, one for North Holland and one for Flemish Dutch, were trained on the basis of auditory parameters of the L1 tokens alone and they were, unsurprisingly, successful in classifying new Dutch tokens from the respective Dutch variety withheld from the training set. The models were then used to predict how listeners from the two variants of Dutch should perceive English vowels. These predictions fitted well with the observed data (Table 5). As such, our results are not in line with those of Strange et al. (2004, 2005), who have argued that acoustic similarity is not always a good predictor of cross-language speech perception. By contrast, they confirm the predictions made by the L2LP model (Escudero & Boersma, 2004; Escudero, 2005, 2009), according to which the acoustic similarity between L1 and L2 sounds explains the patterns found in non-native vowel perception.

Our second objective was to investigate whether an asymmetric confusion pattern of the two English vowels /æ/ and /ɛ/ is observed in speakers of different varieties of Dutch in a task that does not require lexical knowledge, i.e. in a sound categorization task. Weber and Cutler (2004) had previously reported such an asymmetry in a word-recognition task using eye-tracking. Words containing the /æ/ vowel led to activation of words with /ɛ/, but words containing /ɛ/ did not activate words with the /æ/-vowel. The results of the English categorization task revealed a similar pattern. Both groups of Dutch listeners displayed the asymmetric pattern, i.e. English /æ/ was more frequently confused with English /ɛ/ than the reverse. This finding runs counter to the results of Cutler et al. (2004), as described in Section 1.1, who found no asymmetry in their English categorization task. One difference between the studies is that the stimuli used in Weber and Cutler (2004) and in the present study are Standard Southern British English (SSBE), while Cutler et al. (2004) used American English (AE) stimuli (see Section 1.2). The acoustic properties of AE /ɛ/ and /æ/ from Hillenbrand, Getty, Clark, and Wheeler (1995), who report on vowels produced by male speakers of the same dialect as that considered in Cutler et al. (2004), show little difference in F1 (5.6 versus 5.7 Bark) or F2 (12.3 versus 12.9 Bark) for the two vowels. It is well known that AE makes a clear duration distinction between /ɛ/ and /æ/ (Hillenbrand et al., 1995, Table V). In addition, unlike the values of the SSBE vowels, both AE vowels have values that match equally well with the F1 and F2 values for Dutch /ɛ/. Therefore, it seems that phonetic closeness explains why both North Holland and Flemish listeners have an asymmetry in their perception of the SSBE vowels but not in their perception of the same AE vowels. As stated in Section 1, this provides evidence for continuity between perception and word recognition.

An important alternative for explaining the patterns found in both listener groups is the orthographic explanation suggested in Cutler, Weber, and Otake (2006) and demonstrated in Escudero et al. (2008). According to this explanation, Dutch learners of English associate syllables with the front vowels /ɛ/ and /æ/ with words containing the letter ⟨e⟩, but are hesitant to map the same front vowels to the letter ⟨a⟩, which in Dutch corresponds to /ɑ/. Given that in the task of the present study orthographic information was always available, i.e. the response options were the spelled forms of common English words, there is no way of disentangling the two explanations. Further research should show whether the asymmetry still occurs in both listener groups in the absence of orthographic information.

A way to further test the origin of the asymmetric pattern in Dutch learners of English would be to use Escudero et al.'s (2008) word learning paradigm (see also Simon, Chambless, & Alves,

2010) or Escudero and Wanrooij's (2010) auditory-only versus auditory and orthographic XAB categorization tasks with North Holland and Flemish listeners. This would allow us to investigate whether both groups of listeners exhibit the asymmetry in the absence of orthography or whether they show symmetric responses for both vowels, as was the case for the Dutch listeners in Escudero et al. (2008). It is worth mentioning that Escudero et al.'s results suggest that the Dutch variety of their listeners may be closer to that of the Flemish than to that of the North Holland listeners of the present study.

In sum, the present study finds important dialectal differences between North Holland and Flemish listeners' categorization of English /ɛ/ and /æ/ as Dutch vowels, which lead to differences in their non-native perception of the same vowels in an English sound categorization task. This shows that dialect differences not only influence the perception of the first language (Floccia et al., 2006; Sumner & Samuel, 2009), but also affect the learning of a second language.

Our data show that the exact acoustic vowel properties in the variety of the native language can predict how vowels in a second language will be perceived. This indicates that, in contrast to previous findings (Strange et al., 2004, 2005), the acoustic (and hence auditory) properties are important in determining similarity across languages. It is worth mentioning that our LDA models only included a subset of the vowel inventories of English and Dutch, while Strange et al. (2004, 2005) included whole inventories, which could account for the higher predictive success in our model. However, Gilichinskaya and Strange (2010), who used the eight monophthongs of New Yorkian English, report that the results of their LDA predicted Russian listener's classifications for seven out of the eight American English vowels.

As such, our results run counter to assumptions that focus on the similarity in the abstract phonological features of speech sounds (see, e.g., Frisch, 1996), and suggest that similarity is tightly related to the detailed acoustic properties of these sounds. This raises the question as to whether the critical features for similarity comparisons are the acoustic properties themselves (as assumed by L2LP), or the articulatory gestures which give rise to them (as assumed by PAM-L2). Our data do not allow us to distinguish between these claims, because auditory similarity is strongly related to articulatory similarity for vowels (see Mitterer & Ernestus, 2008, for an example of how to disentangle acoustic and articulatory properties). Nevertheless, our results emphasize that detailed phonetic properties are critical to the investigation of similarity across languages and dialects, and that a focus on abstract phonological features is not the most appropriate.

References

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

Adank, van Hout, P., & Van de Velde, H. (2007). An acoustic description of the vowels of northern and southern standard Dutch II: regional varieties. *Journal of the Acoustical Society of America*, 121, 1130–1141.

Alderson, J. C., & Huhta, A. (2005). The development of a suite of computer-based diagnostic tests based on the Common European Framework. *Language Testing, 22*, 301–320.

Best, C. (1995). A direct realist view of cross-language speech perception. In: Strange W. (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues* (pp. 171–204). Baltimore: York Press.

Best, C., & Tyler, M. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In: O.-S. Bohn, & M. Munro (Eds.), *Language experience in second language speech learning. In honor of James Emil Flege* (pp. 13–34). Amsterdam: John Benjamins.

Boersma, P., & Escudero, P. (2008). Learning to perceive a smaller L2 vowel inventory: An optimality theory account. In: P. Avery, E. Dresher, & K. Rice (Eds.), *Contrast in phonology: Theory, perception, acquisition* (pp. 271–301). Berlin: Mouton de Gruyter.

Boersma, P., & Weenink, D. (2010). Praat: doing phonetics by computer, version 5.0.27 [computer programme], retrieved from ⟨http://www.praat.org⟩.

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

Broersma, M. (2005). *Phonetic and lexical processing in a second language*. Doctoral dissertation, MPI Nijmegen.

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.

Escudero, P. (2005). *Linguistic perception and second-language acquisition: explaining the attainment of optimal phonological categorization*. PhD dissertation, Utrecht University, LOT Dissertation Series 113.

Escudero, P. (2009). Linguistic Perception of "similar" L2 sounds. In: P. Boersma, & S. Hamann (Eds.), *Phonology in perception* (pp. 151–190). Berlin: Mouton de Gruyter.

Escudero, P., & Boersma, P. (2002). The subset problem in L2 perceptual development: Multiple-category assimilation by Dutch learners of Spanish. In: Barbora Skarabela, Sarah Fish, & Anna H.-J. Do (Eds.), *Proceedings of the 26th annual Boston University conference on language development* (pp. 208–221). Somerville, MA: Cascadilla.

Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition, 26*, 551–585.

Escudero, P., Hayes-Harb, R., & Mitterer, H. (2008). Novel L2 words and assymetric lexical access. *Journal of Phonetics, 36*, 345–360.

Escudero, P., & Wanrooij., K. (2010). The effect of L1 orthography on L2 vowel perception. In: E. Simon, & M. Van Herreweghe (Eds.), *The relation between orthography and phonology, Language and Speech*, 53.3.

Escudero, Paola, Benders, T. & Lipski, S.(2009). Native, non-native and L2 perceptual cue weighting for Dutch vowels: The case of Dutch, German, and Spanish listeners. Journal of Phonetics, 37, 452-466.

Flege, J. E. (1987). The production of new and similar phones in a foreign-language—Evidence for the effect of equivalence classification. *Journal of Phonetics, 15*(1), 47–65.

Flege, J. E. (1995). Second language speech learning: theory, findings, and problems. In: W. Strange (Ed.), *Speech perception and linguistic experience: issues in cross-language research* (pp. 233–277). Timonium, MD: York Press.

Floccia, C., Goslin, J., Girard, F., & Konopczynski, G. (2006). Does a regional accent perturb speech processing? *Journal of Experimental Psychology: Human Perception and Performance, 32*(5), 1276–1293.

Frisch, S. A. (1996). *Similarity and frequency in phonology [dissertation]*. Evanston, IL: Northwestern University.

Gilichinskaya, Y. D., & Strange, W. (2010). Perceptual assimilation of American English vowels by naïve Russian listeners. *Journal of the Acoustical Society of America, 128*, EL80–EL85.

Goldstein, Louis M., & Fowler, Carol (2003). Articulatory phonology: a phonology for public language use. In: Antje S. Meyer, & O.Schiller Niels (Eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities*. Mouton de Gruyter.

Grosjean, F. (2001). The bilingual's language modes. In: J. Nicol (Ed.), One mind, two languages: Bilingual language processing (pp. 1–22). Oxford: Blackwell.

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

Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A., & Siebert, C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition, 87*(1), B47–B57.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*, 434–446.

Klecka, W. R. (1980). *Discriminant analysis, quantitative applications in the social sciences*, Vol. 19. Newbury Park, CA: Sage.

Lengeris, A. (2009). Perceptual assimilation and L2 learning: Evidence from the perception of Southern British English vowels by native speakers of Greek and Japanese. *Phonetica, 66*(3), 169–187.

Mayr, R., & Escudero, P. (2010). Explaining individual variation in L2 perception: rounded vowels in English learners of German. *Bilingualism: Language and Cognition, 13*, 279–297.

Mitterer, H., & Ernestus, M. (2008). The link between speech perception and production is phonological and abstract: Evidence from the shadowing task. *Cognition, 109*, 168–173.

Polka, L., Colantonio, C., & Sundara, M. (2001). A cross-language comparison of /d/-/th/ perception: evidence for a new developmental pattern. *Journal of the Acoustical Society of America, 109*, 2190–2201.

Simon, E., Chambless, D., & Alves, U. K. (2010). Understanding the role of orthography in the acquisition of a non-native vowel contrast. *Language Sciences, 32.3*, 380–394.

Sjerps, M., Mitterer, H., & McQueen, J. M. (2011). Constraints on the processes responsible for the extrinsic normalization of vowels. *Attention, Perception, & Psychophysics, 73*, 1195–1215.

Strange, W. (Ed.). (1995). *Speech perception and linguistic experience*. Baltimore: York Press.

Strange, W., Bohn, O. S., Nishi, K., & Trent, S. A. (2005). Contextual variation in the acoustic and perceptual similarity of North German and American English vowels. *Journal of the Acoustical Society of America, 118*, 1751–1762.

Strange, W., Bohn, O. S., Trent, S. A., & Nishi, K. (2004). Acoustic and perceptual similarity of North German and American English vowels. *Journal of the Acoustical Society of America, 115*, 1791–1807.

Sumner, M., & Samuel, A. G. (2009). The role of experience in the processing of cross-dialectal variation. *Journal of Memory and Language, 60*, 487–501.

Watkins, A. J., & Makin, S. J. (1996). Effects of spectral contrast on perceptual compensation for spectral-envelope distortion. *Journal of the Acoustical Society of America, 99*, 3749–3757.

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

Yeni-Komshian, G. H., Flege, J., & Liu, S. (2000). Pronunciation proficiency in the first and second languages of Korean-English bilinguals. *Bilingualism: Language and Cognition, 3*, 131–150.