Note: this is a pre-print version of the article published in *Brain and Language*, 126 (3), 243–252. [DOI 10.1016/j.bandl.2013.05.020], which can be found at http://www.sciencedirect.com/science/article/pii/S0093934X13001181

Pre-attentive sensitivity to vowel duration reveals native phonology and predicts learning of second-language sounds

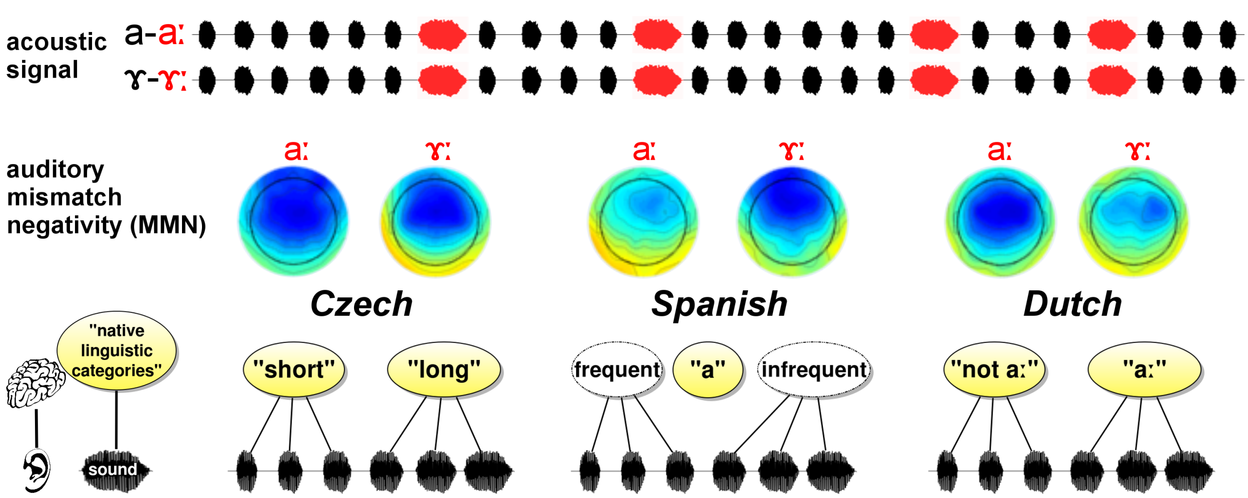
Kateřina Chládková[a], Paola Escudero[b], Silvia C. Lipski[c]

[a] Amsterdam Center for Language and Communication, University of Amsterdam, Spuistraat 210, 1012VT Amsterdam, Netherlands; k.chladkova@uva.nl
[b] MARCS Institute, University of Western Sydney, Bankstown campus, Building 1, Locked Bag 1797, Penrith NSW 2751, Australia; paola.escudero@uws.edu.au
[c] Jean-Uhrmacher-Institute for Clinical ENT Research, Geibelstr. 29-31, 50931 Cologne, Germany; s.lipski@yahoo.de

Corresponding author: Kateřina Chládková
ACLC, University of Amsterdam, Spuistraat 210, 1012VT Amsterdam, Netherlands
email: k.chladkova@uva.nl, phone: +31 20 525 3857

Keywords: native phonology; vowel-duration processing; vowel length; native speech perception; non-native speech perception; mismatch negativity

Graphical abstract



Highlights

- Pre-attentive vowel duration processing depends on the nativeness of the vowel.
- Cross-linguistic measures of brain responses reveal the underlying language system.
- Dutch listeners differ from both Czech and Spanish in vowel duration processing.
- Dutch phonology seems to use duration as a phoneme-specific cue.
- Lack of duration differences in L1 can ease the learning of length contrasts in L2.

Abstract

In some languages (e.g. Czech), changes in vowel duration affect word meaning, while in others (e.g. Spanish) they do not. Yet for other languages (e.g. Dutch), the linguistic role of vowel duration remains unclear. To reveal whether Dutch represents vowel length in its phonology, we compared auditory pre-attentive duration processing in native and non-native vowels across Dutch, Czech, and Spanish. Dutch duration sensitivity patterned with Czech but was larger than Spanish in the native vowel, while it was smaller than Czech and Spanish in the non-native vowel. An interpretation of these findings suggests that in Dutch, duration is used phonemically but it might be relevant for the identity of certain native vowels only. Furthermore, the finding that Spanish listeners are more sensitive to duration in non-native than in native vowels indicates that a lack of duration differences in one's native language could be beneficial for second-language learning.

1 Introduction

Languages differ in their phonemic inventories, that is, in the number of speech sounds that can distinguish word meaning. For instance, the English phonemic inventory includes the two vowels of "sheep" and "ship", namely /i/ and /ɪ/, while Spanish only has /i/. All languages have vowel phonemes that are distinguished in terms of their quality (Maddieson, 1984; Crothers, 1978), as measured by the position of the tongue and jaw or by the acoustic spectral properties of the vowel. However, not all languages distinguish vowel quantity (also known as phonological vowel length), as measured by the duration of the vowel. In quantity languages such as Czech, vowel length is encoded in the phonology so that replacing a long vowel with a short one leads to a change in word meaning such as in the Czech words /saːt/ 'to suck' and /sat/ 'orchard', which are only distinguished by the duration of the vowel. Spanish, on the other hand, is not a quantity language and its phonology does not encode vowel length so that whether the first vowel in the Spanish word /kasa/ 'house' is long or short does not change its meaning.

It is unclear whether the Dutch language encodes the acoustic dimension of vowel duration as phonological vowel length[1], in other words, whether this language has discrete long and short vowel categories. In his analysis of the Dutch vowel inventory, Moulton (1962) differentiates phonologically short and phonologically long vowels. Similarly, Zonneveld (1993) posits that vowel length is part of the Dutch native phonology. Booij (1995) argues that the vowel /aː/ equals to two units of the vowel /ɑ/ within a syllable. Although Booij claims that vowel length as such is not a phonemic property of Dutch vowels, the proposal of a 'doubling' of an otherwise phonologically identical unit within a syllable implies that language users should have some representation of quantity in their grammar. Van Oostendorp (1995) presents several arguments against phonological vowel length in Dutch and argues that the phonological property of vowel tenseness better accounts for Dutch vowel phonology. Most recently, Botma and van Oostendorp (2012) argue that the Dutch phonology does not at all distinguish between tense and lax (or, long

---

[1] Here we use the term *duration* when we refer to the acoustic dimension, i.e. the phonetic property of the sound, while we use the term *length* when we refer to the abstract linguistic category, i.e. the phonological and contrastive mental representation.

and short) vowel segments, but that the phonetic (i.e. durational) differences between Dutch vowels are due to the structure of the syllable in which a vowel occurs. Botma and van Oostendorp discuss a large number of phonological studies with opposing views, which suggests that the long-lasting phonological debate has not yet lead to a consensus on whether or not Dutch has vowel length.

Phonetic studies on Dutch do not clarify the issue of Dutch vowel length either. In that respect, recent speech production studies (Adank, van Hout & Smits, 2007; Adank, van Hout & van de Velde, 2007; van Leussen, Escudero & Williams, 2011) show that Dutch speakers only use differences in vowel duration to distinguish a limited number of vowels, that is, the use of duration in speech production is inconsistent across vowels. Interestingly, however, speech perception studies suggest that vowel length in Dutch, as compared to English, does have a contrastive role. For instance, two recent studies (Dietrich, Swingley & Werker, 2007; van der Feest & Swingley, 2011) have found that Dutch 18-month-olds and adults are more sensitive to differences in vowel duration than their English counterparts, which the authors attributed to the contrastive role of vowel length in Dutch as opposed to English. Importantly, some studies show that Dutch listeners use vowel duration to distinguish the vowels in the words /mɑn/ 'man' and /maːn/ 'moon' (Nooteboom & Doodeman, 1980), while others demonstrate that Dutch listeners predominantly use vowel spectral properties to distinguish these vowels (Escudero, Benders & Lipski, 2009).

The present study aims to resolve the controversy around the abstract phonological representation of vowel length in the Dutch language. We examined Dutch listeners' pre-attentive processing of vowel duration changes, and compared it to that of Czech listeners, who clearly have short and long vowel phonemes, and to Spanish listeners, whose native phonology treats all vowel durations as equal. Listeners were presented with duration changes in both native and non-native vowels, which enabled the investigation of whether vowel duration processing depends on the listeners' phonemic inventory.

We recorded behavior-independent responses of the auditory system to vowel duration in a categorical oddball-paradigm using electroencephalography (EEG), examining the mismatch negativity (MMN). The MMN is elicited at about 100-250 ms latency when infrequent deviations occur among frequently repeated sound patterns. The MMN is widely accounted as a marker of pre-attentive change detection and is obtained for simple and complex patterns of auditory changes (Näätänen, Tervaniemi, Sussman, Paavilainen & Winkler, 2001; Näätänen, Paavilainen, Rinne & Alho, 2007). What makes the MMN ideally suited for the present investigation of phonological representations is its sensitivity to listeners' linguistic experience: native phonemic contrasts elicit a stronger and often earlier MMN than speech sound contrasts without relation to the listeners' phonology. Crucially, many studies have demonstrated that listeners' native phonology modulates the pre-attentive processing of acoustic information (Näätänen et al., 1997; Sharma & Dorman, 2000; Nenonen, Shestakova, Huotilainen & Näätänen, 2003; Ylinen, Shestakova, Huotilainen, Alku & Näätänen, 2006; Tervaniemi et al., 2006; Kirmse et al., 2008; Menning, Imaizumi, Zwitserlood & Pantev, 2002; Lipski & Mathiak, 2007; Kazanina, Phillips & Idsardi, 2006; Hisagi, Shafer, Strange & Sussman, 2010). As for vowel duration, it has been shown that speakers of quantity languages such as Czech or Finnish, that is, languages which represent vowel duration in terms of abstract phonological categories, have stronger mismatch responses to vowel duration changes than speakers of other languages, including non-quantity languages such as Spanish or Russian (Nenonen et al., 2003; Ylinen et al., 2006, Tervaniemi et al., 2006; Kirmse et al., 2008; Menning et al., 2002; Hisagi et al., 2010; Nenonen, Shestakova, Huotilainen & Näätänen, 2005). Unlike these previous studies, our three-way comparison of pre-attentive processing of vowel duration in Dutch, Czech and Spanish listeners will unravel the phonology underlying Dutch listeners' perception. Specifically, we will be able to show whether Dutch encodes vowel duration in terms of discrete short and long categories, in other words, whether Dutch is like quantity languages such as Czech, or whether it is like non-quantity languages such as Spanish.

Incidentally, even if vowel duration is not encoded in the phonology of a certain language, native speakers of that language tend to rely on duration to distinguish novel vowels that are present in a foreign language. Using behavioral tasks, a number of studies have shown that Spanish, Catalan, Portuguese, Mandarin, Polish, and Russian learners distinguish English or Dutch vowels through their duration differences, while native listeners predominantly use the vowels' spectral differences (Escudero et al., 2009; Flege, Bohn & Jang, 1997; Escudero & Boersma, 2004; Bogacka, 2004; Cebrian, 2006; Rauber, Escudero, Bion & Baptista, 2005; Kondaurova & Francis, 2008). One explanation for second language learners' reliance on duration states that duration is acoustically highly salient and, therefore, universally accessible to learners regardless of its status in their native phonology (Bohn, 1995). An alternative explanation holds that the processing of duration is always transferred from the learner's native phonology (Escudero & Boersma, 2004). That is, listeners whose language does not employ vowel duration transfer a blank slate for this dimension, which allows them to readily form new length categories in a novel language (Escudero & Boersma, 2004).

Neurophysiological studies have also shown pre-attentive reliance on duration despite its irrelevance in the listeners' native phonology. In Lipski, Escudero and Benders (2012) Spanish learners of Dutch and Dutch natives had similar MMN responses to vowel duration changes for the Dutch vowels /aː/ and /ɑ/. Interestingly, however, speakers of non-quantity languages such as Spanish or Russian seem to process duration changes depending on how close a novel vowel is to their native vowel inventory. Nenonen et al.'s (2005) Russian learners of Finnish had smaller MMNs for vowel duration differences than Finnish natives when they were presented with stimuli that resembled a Russian vowel, while the two groups had similar MMNs for stimuli that did not resemble any Russian vowel. The authors attributed the Russian learners' strong MMN for duration differences in non-native vowels to the fact that they had successfully acquired second-language length categories after considerable exposure to Finnish. Since most previous studies have considered second language learners, it is unclear whether speakers of non-quantity languages such as Russian or Spanish have pre-attentive sensitivity to vowel duration when first exposed to the non-native length contrasts.

Given Nenonen et al.'s (2005) surprising results, the present study aimed at demonstrating whether pre-attentive processing of non-native vowel duration differences is *universal*, that is, independent of how vowel duration is encoded in the listener's native phonology, or *language-specific*, that is, dependent on its encoding within the listener's native phonology. To this end, we presented native and non-native vowels with different durations to Czech, Dutch, and Spanish listeners whose native phonologies are likely to differ on how they encode vowel length. Figure 1 shows the quality properties (first and second formant frequencies) of the native and non-native vowel stimuli used in the present study together with those of the Czech, Spanish and Dutch vowel inventories. It can be observed that [a] (the native vowel quality) falls within /a/ in all three languages, while [ɤ] (the non-native vowel quality) is far from any of the listeners' native vowels.

[Fig. 1 about here]

We tested Czech, Dutch, and Spanish monolingual young adults with very little experience in foreign languages. EEG was recorded in two sessions that took place on different days. In one session, participants listened passively to short and long tokens of the native vowel [a], and in the other, they listened passively to short and long tokens of the non-native vowel [ɤ]. As with Finnish and Russian listeners (Nenonen et al., 2005), if the auditory cortical processing of vowel duration is modulated by whether or not vowel duration can contrast native vowel phonemes, Czech and Spanish listeners should show opposite MMN responses for the native vowel quality. Specifically, Czech listeners should exhibit the largest MMN, while Spanish listeners will have no or the smallest MMN response for [a]. Dutch listeners, for whom the phonological role of vowel duration is unclear, may behave similarly to either the Czech or the Spanish listeners, or, alternatively, show an MMN response that is intermediate between the Czech and Spanish responses.

If non-native perception of vowel duration is phonology-specific and not universal, duration differences in the non-native vowel [ɤ] will elicit the largest MMN in Czech listeners. If the Spanish listeners transfer their native disregarding of duration differences to non-native perception, they will, again, have the smallest MMN in the non-native vowel. Alternatively, based on the findings of numerous L2 perception studies discussed above, duration differences in the non-native vowel [ɤ] may well elicit a large MMN in the Spanish listeners, one comparable to that of quantity language listeners. The latter would demonstrate that not only advanced Russian learners of Finnish (Nenonen et al., 2005), but also non-native listeners with little exposure to a novel language exhibit rapid pre-attentive sensitivity to vowel duration. Crucially, if Dutch phonology encodes vowel duration in terms of abstract length categories as it is in quantity-languages, Dutch listeners will resemble Czechs, and will thus have a large MMN for duration changes in the non-native vowel. If Dutch does not encode vowel duration in its phonology at all, Dutch listeners will resemble the Spanish in the non-native vowel.

Alternatively, if non-native perception of vowel duration is modulated by the universal salience of this acoustic dimension, all three groups should have an equally large MMN for the non-native vowel quality.

2 Methods

2.1 Participants

24 Czech, 24 Spanish, and 26 Dutch right-handed listeners, all university students or recent graduates aged 19 to 31 years, took part in the study. They were all monolinguals, who were raised in a monolingual family, had never spent more than 2 months in a foreign country, and had not had exposure to foreign languages above the level of high-school classroom instruction. They rated their knowledge of foreign languages below 4 on a scale from 0 to 7 (where 0 means none, and 7 native-like), and were not linguistics students. None of the participants reported to have had a history of neurological, hearing, or language-related disorders. The Czech participants were from central and southern Moravia in the Czech Republic. The Dutch participants were from the Randstad area in the Netherlands. The Spanish participants were from various regions in Spain. 23 Czech (13 female, mean age = 22.4 years), 22 Spanish (10 female, mean age = 23.0 years) and 24 Dutch (13 female, mean age = 22.6 years) were included in the ERP analysis. Five participants were excluded due to a large number of artifacts (one Spanish and one Dutch), technical errors during data acquisition (one Czech and one Spanish) and ambidexterity revealed after the experiment (one Dutch). Participants gave a written informed consent and were paid for participation. The study was approved by the ethical committee of the Faculty of Humanities, University of Amsterdam and conforms to the guidelines of the Declaration of Helsinki (2008).

2.2 Stimuli

2.2.1 Vowel qualities and duration steps

The stimuli were natural tokens of the Estonian vowels /æ/ and /ɤ/ (henceforth transcribed as [a] and [ɤ], respectively), spoken by a 26-year old native female speaker of standard Estonian, a trained phonetician. The values of the first three formants were 920 Hz, 1634 Hz and 2707 Hz for [a], and 465 Hz, 1462 Hz, and 2920 Hz for [ɤ]. As shown in

Figure 1, the quality of [a] is acoustically close to that of the participants' native vowel category (/a/ in Spanish and Czech, and /aː/ in Dutch and Czech), while the quality of [ɤ] is not close to any vowel in the participants' native languages. Since cross-language acoustic similarity of vowels is a good predictor of their *perceived* similarity (e.g. Escudero & Chládková, 2010; Escudero & Vasiliev, 2011; Escudero & Williams 2011, 2012; Chládková & Podlipský, 2012; Escudero, Simon & Mitterer, 2012), [a] was used as the native vowel quality, and [ɤ] was used as the non-native vowel quality. The two vowels were produced with a flat pitch contour. Formant on-glides and off-glides were discarded so that the duration of the middle vowel portion with stable formants and pitch was 351 ms for [a] and 349 ms for [ɤ]. These tokens were subsequently manipulated using the time-domain pitch-synchronous overlap-and-add algorithm (Moulines & Charpentier, 1990) implemented in the software Praat (Boersma & Weenink, 2011) to yield 6 different durations in psychoacoustically equal steps: 118, 136, 157, 181, 208, and 239 ms; that is, the six stimuli of each of the two vowel qualities differed only in their duration. These 6 duration values were selected on the basis of a pilot behavioral experiment with Czech listeners and the literature on Dutch vowels (described below in section 2.2.2) so that the three short tokens in the present study are in the short category and the three long tokens are in the long category. The stimuli were presented in the categorical oddball paradigm (i.e. many-to-many oddball paradigm; Hisagi et al., 2010; Lipski et al., 2012; Scharinger, Monahan & Idsardi, 2011c) described below in section 2.3. The three shortest items served as the short stimulus category while the three longest items served as the long stimulus category. This paradigm elicits an MMN if listeners perceive the acoustically varied standard stimuli as different from the acoustically varied deviant stimuli.

2.2.2 Behavioral pilot experiment: determining the short-long stimulus boundary

The pilot experiment to determine the short-long boundary was a two-alternative forced-choice identification task. Only Czech listeners participated in this experiment because they are the only group who has explicit labels for the short and long vowel categories. We tested twenty-five young native speakers of Moravian Czech, who did not participate in the EEG experiment. Three stimulus duration continua were created: one with the quality of a high-mid back unrounded vowel [ɤ], one of a low front unrounded [a] and one of a low-mid back rounded [ɑ]. The first and the second vowel quality served as the non-native and the native stimulus quality, respectively, in the EEG experiment. The three different qualities from distinct vowel space regions were used to determine a general short-long boundary that applies across the vowel space. Each continuum ranged from 95 to 245 ms and consisted of 13 duration values equidistantly spaced along the logarithmic scale.

Testing was conducted in a quiet room and the stimuli were presented via circumaural headphones. Participants were instructed to label each stimulus as either a short or a long vowel by clicking on "short" or "long" written in Czech orthography on the computer screen. Each stimulus was repeated 5 times, resulting in a total of 195 stimuli (13 duration values * 3 continua * 5 repetitions) which were randomly shuffled before the experiment. We used logistic regression to obtain an identification function for each participant. Per participant, from the regression function we then computed the location of the short-long boundary. The short-long boundary is located at such a stimulus that would receive each of the two labels "short" and "long" with probability of 0.5. Therefore, the boundary x was computed from the formula: $\ln(0.5/(1-0.5)) = \beta_0 + \beta_1 x$, where β_0 and β_1 are the logistic regression coefficients. Since $\ln(0.5/(1-0.5)) = 0$, $x = -\beta_0/\beta_1$. The pilot experiment detected the average short-long boundary in Czech to lie at 168 ms. Therefore, the durations of the stimuli in the EEG experiment were manipulated so that 168 ms is the boundary between the 3 short tokens and the 3 long tokens.

The short-long boundary used to separate short and long stimuli in the present EEG study reliably separates phonetically short and long Dutch vowels. The phonetically long Dutch vowels (e.g. /aː/, diphthongized vowels, and true diphthongs) are longer than the present short-long boundary of 168 ms, while phonetically short vowels (e.g. /ɑ/, /ɪ/, /i/) are shorter (Adank et al., 2007a, 2007b). The average Spanish vowel is (slightly) shorter than the present short-long boundary (Zimmerman & Sapon, 1958; Chládková, Escudero & Boersma, 2011).

2.3 Procedure

EEG was recorded in two sessions, one for the native and one for the non-native vowel stimuli, in two different days within a week. Native and non-native vowel stimuli were presented in separate sessions to avoid the influence of their differential status within the listeners' phonemic inventory. The order of the two sessions was counterbalanced across subjects so that in the first session (first day) half of the subjects of each language group listened to native vowels, while the other half listened to non-native vowels. Each session consisted of two 30-minute blocks of EEG-recording (block 1, block 2), with a 15-minute break between blocks.

In one block, short vowels were the standard stimuli and long vowels were the deviants, while in the other long vowels were standards and short vowels deviants. The order of blocks was counterbalanced across subjects but was kept identical across a participant's two sessions. Within a block, the deviant category occurred with a probability of 15.2%. All three deviants and standards were evenly represented in the deviant and the standard category, respectively. Each block started with 20 standards, followed by the oddball sequence which contained 300 deviants (100 deviants of each type), for a total of 2022 stimuli per block. A deviant was always followed by 3 to 8 standards. The inter-stimulus interval was varied randomly in 5 steps between 800 and 932 ms. Stimuli were presented at 60 dB SPL via a single loudspeaker placed in front of the participant at a distance of 1 m at chin level.

Testing took place in sound-attenuated speech laboratories at the University of Amsterdam and at the Palacký University in Olomouc. Eighteen Czech participants were tested in Olomouc, while the Dutch, Spanish and six Czech participants were tested in Amsterdam (the Spanish and the 6 Czechs were exchange university students who had arrived in Amsterdam less than 2 weeks prior to the time of their second session to ensure they had as little foreign-language exposure as possible). All participants were tested with the same equipment. During stimulus presentation, participants watched a muted movie of their choice (originally spoken in their native language) with subtitles in their native language. At the beginning of each session, participants were given information about the sounds to be played (either sounds from their native language or sounds from a foreign unknown language, depending on the session) and were instructed to disregard the sounds and just watch the movie.

2.4 EEG recording and pre-processing

EEG was recorded from 64 active Ag-AgCl electrodes placed according to the International 10/20 placement in a cap (BioSemi) fitted to participant's head size. Seven external electrodes were used: placed on the nose (offline reference), below and above the right eye, on the left and right temple (ocular activity), and on the right and left mastoid. The input/output gain was 31.25 nV/bit, the EEG signal was recorded at 8kHz and later downsampled to 512 Hz.

The EEG was offline referenced to the nose channel. Slow drifts were removed by subtracting from each channel a line so that the first and the last sample become zero. The data were band-pass filtered in the frequency domain with a low cut-off of 1 Hz (0.5 Hz bandwidth) and a high cut-off of 30 Hz (15 Hz bandwidth). The data were epoched from -100 ms to 700 ms relative to stimulus onset. For subsequent baseline correction the mean voltage in the 100-ms pre-stimulus interval was subtracted from each sample in the epoch. Artifact correction was done automatically (rejection of epochs with +/- 75 μV at any channel) and by subsequent visual inspection. Participants (one Spanish and one Dutch) with more than 50% of artifact-contaminated epochs were excluded from further analysis.

Per participant per block, the epochs of the three short stimuli and the epochs of the three long stimuli were averaged. Per participant, two difference waves were derived by subtracting (1) the average waveform of short standards (from one block) from the average waveform of short deviants (from the other block), and (2) the average waveform of long standards from the average waveform of long deviants. There was thus a within-subject factor "duration-type" with two levels, namely short and long, referring to the comparison of short standards with short deviants from reversed blocks, and long standards with long deviants from reversed blocks, respectively. Previous studies have reported an asymmetry in MMN to duration decrements versus duration increments: duration decrements (equal to the duration-type short in the present study) often yield smaller MMN than duration increments (i.e. duration-type long) (Kirmse et al., 2008; Hisagi et al., 2010; Lipski et al., 2012). Therefore, the factor duration-type was included to test whether asymmetries between MMN to short and MMN to long stimuli were also present in this study.

In the first block of EEG recording, half of the participants per language were presented with short deviants among long standards, while the other half of participants were presented with long deviants among short standards. It has been suggested that MMN to speech stimuli may be reduced over time due to habituation (McGee et al., 2001). To control for any habituation effects, our analyses also included the between-subjects factor "first-deviant-duration" with two levels: short and long, which refers to the duration-type of deviants from the first block.

We searched for a negative peak ("group-peak") between 200 and 360 ms post stimulus-onset for each channel in the grand-average difference waveforms per language, first-deviant-duration, vowel-quality, and duration-type. Subsequently, per participant, we computed the mean amplitude over a 40-ms time window centered at the group-peak, which was our measure of MMN amplitude. Statistical tests were done with the alpha level of 0.05.

3 Results

We first ran an exploratory repeated-measures ANOVA on the MMN amplitude measured at Fz with language and first-deviant-duration as the between-subjects factors, and vowel-quality and duration-type as the within-subject factors. This analysis yielded a significant two-way interaction of first-deviant-duration and duration-type ($F[1,63]=67.027$; $p<0.001$). Inspection of this two-way interaction revealed that the average MMN in participants who were first presented with *long* deviants was -1.653 μV for *long* stimuli and -0.357 μV for *short* stimuli. The average MMN in participants who were first presented with *short* deviants was -1.117 μV for *short* stimuli and -0.169 μV for *long* stimuli, with the latter not being significantly different from 0. Thus, the MMN was approximately five times larger for deviants from the first block than for deviants from the second block. Importantly, the considerable attenuation of MMN to deviants presented in the second block is independent of duration-type and might be a result of habituation to the frequently repeated standards in the first block (McGee et al., 2001). This means that the small MMN from the second block likely represents a less reliable measure of the subjects' sensitivity to duration changes. Therefore, all further cross-language and cross-vowel comparisons were conducted on the MMN elicited by the deviants from the first block. This resulted in a single difference wave, or duration-type, per participant, which is the duration-type of the deviant presented in the first block. That is, duration-type was treated as a between-subjects factor in the subsequent analysis. Accordingly, Table 1 and Figures 2-4 show the data that were compared across groups, namely the MMN for deviants from the first block.

The subsequent repeated-measures ANOVA thus had language (Czech vs. Spanish vs. Dutch) and duration-type (short vs. long) as the between-subjects factors, and vowel-quality (native vs. non-native) as the within-subject factor. This analysis was run on the MMN amplitude measured at 9 channels (Fz, FCz, Cz, F3, F4, FC3, FC4, C3, C4), and also included the within-subject factors anteriority (frontal: Fz, F3, F4; fronto-central: FCz, FC3, FC4, central: Cz, C3, C4) and laterality (midline: Fz, FCz, Cz; left: F3, FC3, C3; right: F4, FC4, C4). The mean MMN amplitudes averaged across the 9 sites (Fz, FCz, Cz, F3, F4, FC3, FC4, C3, C4) are listed in Table 1. Figure 2 shows the grand-average difference waveform at FCz and the topographical MMN distributions for each language, vowel-quality and duration-type; Figure 3 then shows the standard and deviant waveforms at FCz.

[Table 1 about here]

[Fig. 2 about here]

[Fig. 3 about here]

The ANOVA revealed significant main effects of duration-type ($F[1,63]=8.618$, $p=0.005$), anteriority ($F[2,126]=12.916$, $p<0.001$) and laterality ($F[2,126]=14.018$, $p<0.001$), as well as a two-way interaction of language and vowel-quality ($F[2,63]=5.247$, $p=0.008$).[2]

As for the main effect of duration-type, long deviants elicited a larger MMN than short deviants by on average 0.524 μV. Regarding the main effects of laterality and

---

[2] A similar ANOVA that was run on MMN amplitude elicited by deviants in block 2 did not yield any main effects of vowel-quality, language, or duration-type, neither any interactions involving at least two of these factors (all *p*'s > 0.1). This provides further evidence for the fact that all listeners, irrespective of their language, vowel-quality or duration-type, habituated to the standards form block 1, which did not yield an MMN when they were presented as deviants in block 2.

anteriority, MMN amplitude was largest at frontal and fronto-central sites (by, on average, 0.173 and 0.164 μV, respectively), and it was more prominent at the midline and the right hemisphere than at the left hemisphere for all groups (by, on average, 0.207 and 0.188 μV, respectively); see Figure 2. Below we inspect the significant two-way interaction of language and vowel-quality.

Independent-samples *t*-tests were carried out to compare the MMN amplitude (averaged across the 9 sites) across languages separately for the native and for the non-native vowel. The comparisons showed that for duration changes in the native vowel, Spanish listeners had significantly smaller MMN amplitude than Czech listeners by, on average, 0.563 μV ($t[43]=2.019$, $p=0.025$, 95% confidence interval [c.i.] of the difference = 0.001..1.126), and smaller MMN than Dutch listeners by, on average, 0.492 μV (although only nearly-significant with $\alpha = 0.05$; $t[44]=1.540$, $p=0.065$, c.i. = -0.152..1.136); no significant differences were found between Dutch and Czech. In contrast, in the non-native vowel quality, Dutch listeners had a significantly smaller MMN amplitude than Czech listeners by, on average, 0.599 μV ($t[45]=2.674$, $p=0.005$, c.i. = 0.148..1.050) and smaller MMN than Spanish listeners by, on average, 0.582 μV ($t[44]=2.674$, $p=0.020$, c.i. = 0.030..1.133), while no significant difference was found between Spanish and Czech.

Subsequently, paired-samples *t*-tests were run to compare the MMN amplitude (averaged across the 9 sites) between the native and the non-native vowel within each language. These comparisons showed that, in Spanish listeners, the MMN amplitude in the non-native vowel was larger than in the native vowel by, on average, 0.780 μV ($t[21]=2.756$, $p=0.006$, c.i. = 0.191..1.369). Although in Czech and Dutch, the difference does not reach significance, it can be observed that Czech listeners follow a similar trend to that of Spanish listeners in that their MMN in the non-native vowel also appears to be larger than in the native vowel (by, on average, -0.235 μV; $p=0.095$). Conversely, the Dutch MMN follows the opposite trend in that their MMN tends to be larger in the native than in the non-native vowel (with the average difference being +0.293 μV; $p=0.134$). Figure 4 shows the mean MMN amplitude per language and vowel-quality averaged across the two duration-types and across the 9 sites.

[Fig. 4 about here]

In sum, our results show that in the native vowel [a], Czech and Dutch listeners have a larger MMN to duration changes than Spanish listeners. In contrast, in the non-native vowel [ɤ], Czech and Spanish listeners have a larger MMN than Dutch listeners. Spanish listeners have a reliably larger MMN in the non-native than in the native vowel. The same, although non-significant, trend is observed in Czech listeners, while the opposite trend is seen in Dutch listeners.

4 Discussion

Neurophysiological research demonstrated that experience with native phonology shapes the early pre-attentive speech sound processing. In that respect, both the phonemic status of the sound within one's native language or dialect (Näätänen et al., 1997; Nenonen et al., 2005; Kazanina et al., 2006; Scharinger et al., 2011c) and the specific phonological structure of the sound (the contrastive role of its various acoustic dimensions) (Obleser, Lahiri & Eulitz, 2004; Ylinen et al., 2006; Lipski, Lahiri & Eulitz, 2007; Hisagi et al., 2010; Scharinger, Idsardi & Poe, 2011a; Scharinger, Merickel, Riley & Idsardi, 2011b) can modulate listeners' pre-attentive response to speech sounds. To this date, however, the phonological structure of some languages remains unclear (e.g. status of vowel length in languages such as Dutch). Therefore, we used cross-linguistic and cross-stimulus comparisons of the mismatch response – a measure of pre-attentive processing of speech sounds – to reveal phonological structure. With this approach, our study represents a first step towards resolving the unclear phonological status of vowel length in Dutch, for which speech production and perception studies gave conflicting evidence.

We compared Dutch listeners' pre-attentive receptiveness for vowel duration to that of Czech listeners whose native phonology unequivocally encodes vowel length, and to that of Spanish listeners whose native phonology unequivocally does not encode vowel length. We assessed duration processing in two types of stimuli: a native and a non-native vowel quality, because if a language encodes a particular phonetic dimension (i.e. vowel duration) in terms of discrete phonological categories (i.e. vowel length categories) in its phonology, then this encoding should be generalized across vowel phonemes.

4.1 Phonological role of vowel duration in the native phonology

Our results demonstrate that Dutch listeners' MMN response to vowel duration in the native vowel quality [a] does not differ from that of quantity-language listeners (Czechs) and is larger than that of non-quantity language listeners (Spanish). This indicates that duration differences between phonetically short and long tokens of [a] may signal a category boundary in Dutch: that is, Dutch listeners do not perceive short [a] as the same category as long [aː]. However, our results also show that Dutch listeners' receptiveness to duration in the non-native vowel quality [ɤ] was smaller than that of Czech listeners. This indicates that the change between a short [ɤ] and a long [ɤː] is more likely to represent a phonemic, i.e. linguistically relevant, change for Czech than for Dutch listeners, despite the fact that the spectral quality of [ɤ] is not phonemic in either language.

We thus find that Dutch does not pattern with either Czech or Spanish, which suggests that next to quantity and non-quantity phonologies, there is (at least) one other type of phonological system with respect to phonetic vowel duration and phonological length. Specifically, the Dutch quantity-like large MMN to duration in [a] indicates that the phonetic dimension of duration is used phonemically in some way in Dutch. However, the Dutch MMN to duration in [ɤ], which was smaller than that of quantity-language listeners, indicates that the phonological encoding of vowel duration in Dutch differs from that of quantity languages.

Here we put forward that in Dutch, phonetic vowel duration is not encoded in terms of abstract phonological categories for vowel length, but it is used to define the identity of certain vowel qualities, such as the vowel /aː/ which has to have a long duration.

This is illustrated in Figure 5 where it can be seen that in Dutch, an [a]-like stimulus with short duration is not perceived as the /aː/ category. Note that Dutch also has diphthongs such as /eː/ (realized as [eⁱ]) or /ɛɪ/, which are phonetically longer than most monophthongs (Adank et al., 2007a). If duration is a cue to diphthongs, perception of an [e]-like vowel may then work similarly to an [a]-like vowel: a short tokens of [e] will not be perceived as /eː/, but possibly as a different vowel category.

[Fig. 5 about here]

Recall however that we did not detect a significant difference within Dutch listeners between their MMN to the native versus the non-native vowel quality. However, we believe that this within-language finding for Dutch should only be interpreted with caution. This is because it is well known that in speech production, different vowel qualities have different intrinsic durations: high or mid vowels such as /i/ or /e/ are intrinsically shorter than low vowels such as /a/ (Lehiste, 1970). In line with that, a recent experiment by Meister, Werner and Meister (2011) demonstrated that vowel height affects the perceived short-long boundary for listeners of quantity languages: high vowels require less duration difference to be perceived as long than low vowels. This implies that a physically identical duration difference is likely to be perceived as relatively large in a mid (or high) vowel, and as relatively small in a low vowel. Since our stimuli included a mid (non-native) and a low (native) vowel, a quantity-language listener should exhibit a (slightly) different MMN for each of these vowels, i.e. the MMN for the non-native mid vowel [ɤ] should be larger than that of the native low vowel [a]. And, in fact, the Czechs appear to follow this trend, while Dutch listeners seem to follow the opposite trend.

It thus seems that Dutch listeners' MMN for the native versus the non-native conditions may have been influenced by their differential sensitivity for duration across different vowel qualities (i.e. different vowel heights). However, this explanation is only preliminary since further research should compare Dutch /aː/ to another low vowel, e.g. /ɑ/, in order to conclusively determine Dutch duration sensitivity across vowels.

4.2 Non-native use of vowel duration

As can be seen in Figure 1, Dutch differentiates twelve spectral qualities, while Czech and Spanish only five. In line with that, Escudero et al. (2009) have shown that Dutch listeners weigh spectral properties heavier than durational properties (while Spanish learners of Dutch weighted duration heavier than spectrum). One might therefore argue that Dutch listeners will only show high sensitivity to spectral and less so to durational changes for all vowels. This may explain why they had the smallest MMN to duration differences in the non-native condition, but it cannot explain why they had a similarly large MMN to that of Czechs in the native condition, since neither condition contained spectral differences that could be used to distinguish the vowels.

Regarding Spanish listeners, their duration sensitivity in the non-native condition was surprisingly similar to that of the Czech, who are quantity-language speakers, and larger than that of the Dutch. In that respect, recall that Nenonen et al. (2005) also showed that pre-attentive sensitivity to duration is larger in non-quantity listeners for non-native than for native vowel qualities. Unlike Nenonen et al. whose participants were advanced learners of a quantity language, our Spanish monolinguals' strong duration sensitivity in non-native vowels cannot be due to experience with second-language length categories. Since the system underlying MMN generation is affected by linguistic experience (e.g. Näätänen et al., 1997), it is plausible that the complete lack of meaningful duration differences in Spanish signifies that duration is phonologically a blank slate dimension and that therefore, processing of non-native vowel duration is not affected by native phonemic representations.

Importantly, these findings for Dutch and Spanish listeners' duration processing in non-native vowels speak to theories of second-language speech development. Bohn (1995) proposed that in areas of vowel space where spectral properties do not contrast any native phonemes, listeners would attend to duration to differentiate novel vowel contrasts, which explains the larger MMN to duration in the non-native versus the native condition for Spanish listeners, but does not explain why Dutch listeners do not use duration to the same extent as Czech and Spanish listeners in the non-native condition. Additionally, one could interpret Bohn's hypothesis as suggesting that listeners with few spectral contrasts (Spanish) in their L1 would be more sensitive to duration than listeners with many spectral contrasts (Dutch), which also explains why Spanish listeners resort to duration in the non-native condition. However, such an interpretation would predict that Dutch listeners are not sensitive to duration in any vowel quality, which runs contrary to the results for the native condition.

Alternatively, our data seems to be in line with Escudero and Boersma (2004) and Escudero et al. (2009), who proposed that sensitivity to vowel duration in foreign vowels depends on the learners' native phonology. Specifically, our results show that Czech listeners, who encode phonetic vowel duration in terms of abstract phonological length categories (i.e. 'short' and 'long') in their phonology, apply the duration cue equally in native and non-native perception. That is, they generalize their use of the *phonological* *length contrast* to non-native vowel inventories. Spanish listeners either have a blank slate for duration (i.e. they do not encode phonetic duration at all in their phonology) and are able to learn different categories along this dimension in the same way infants learn the sound categories of their first language (Escudero & Boersma, 2004)[3], or automatically use duration when confronted with any non-native vowel quality (Bohn 1995).

In Dutch, vowel duration is neither a blank slate dimension (as it is in Spanish) nor is it divided in phonological length contrasts such as 'short' vs. 'long' (as it is in Czech), but it is an acoustic dimension that contributes to the identity of only certain vowel phonemes, e.g. /aː/ (Adank et al., 2007a; Nooteboom & Doodeman, 1980). This explains why Dutch listeners have a lower sensitivity to duration for the non-native condition than Spanish and Czech listeners. This low sensitivity to non-native duration-based contrasts might also help to explain Dutch listeners' substantial difficulty in distinguishing English words such as *buzz* and *bus*, whose difference is primarily cued by vowel duration in native English listeners (Elsendoorn, 1985; Broersma, 2005; Broersma, 2010).

Beside the between-language differences, we found that long deviants elicited larger mismatch responses than short deviants, independently of vowel type or listeners' native language. This asymmetry has been reported previously for various languages (Kirmse et al., 2008; Hisagi et al., 2010; Lipski et al., 2012). This seems to result from general processing mechanisms in that an unexpected extra portion of acoustic signal (long deviant among short standards) evokes stronger responses than an unexpected absence of acoustic signal (short deviant among long standards).

5. Conclusions

[3] Infants learn to discriminate novel speech sounds through a learning mechanism known as statistical or distributional learning (Maye, Werker & Gerken, 2002; Maye, Weiss & Aslin, 2008) and adults seem to acquire L2 phonemes via the same mechanism (Maye & Gerken, 2001; Gulian, Escudero & Boersma, 2007; Hayes-Harb, 2007; Escudero, Benders & Wanrooij, 2011). If this learning mechanism remains active throughout the life span, it is not surprising that after brief exposure to systematic durational variation in non-native vowels, Spanish listeners become very sensitive to short and long vowels.

The present study attempted to uncover the thus far uncertain status of vowel length within Dutch phonology. A cross-linguistic comparison of listeners' pre-attentive processing of duration in a native and in a non-native vowel quality suggests that Dutch is neither like Spanish (a non-quantity language) nor like Czech (a quantity language). Unlike Spanish listeners, Dutch listeners exploit the phonetic dimension of vowel duration in their native language, as reflected by their large MMN in the native vowel condition. And unlike Czech who use the phonetic dimension of vowel duration to cue phonological length categories 'short' and 'long,' Dutch appears to use phonetic vowel duration as a cue to the phonological identity of specific vowel phonemes, as reflected by their smallest MMN in the non-native condition.

Further, Spanish listeners demonstrated large sensitivity to duration changes when exposed to a novel vowel quality. It is plausible that the widely reported Spanish speakers' reliance on vowel duration in a second language is caused by the complete lack of this dimension in their native language. However, further research should show whether this sensitivity is the result of automatic psychoacoustic salience, as proposed by Bohn (1995), or is the result of the learning of frequency distributions, as proposed by Escudero and Boersma (2004).

Acknowledgements

We thank Paul Boersma for valuable comments on experiment design, data and statistical analyses, and previous versions of this paper, Karin Wanrooij for comments on experiment design and for sharing her testing and data analysis experience, Silke Hamann for comments on a previous version of this paper, Dirk Jan Vet for technical assistance, Nele Salveste for recording the stimuli, the Department of Czech Studies at Palacký University and Václav Jonáš Podlipský for providing and customizing their speech lab to our needs, and Barbora Chládková, Clara Martín Sánchez, Vanina Grippo, Ana Díaz Dasi, Michelle van Bokhorst, Sascha Couvee and Gisela Govaart for participant recruitment and assistance in data collection. This study was funded by NWO (The Netherlands Organization for Scientific Research) grant 277-70-008 awarded to Paul Boersma.

References

Adank, P., van Hout, R., & Smits, R. (2007a). An acoustic description of the vowels of Northern and Southern standard Dutch. *J Acoust Soc Am,* 116, 1729-1738.

Adank, P., van Hout, R., & van de Velde, H. (2007b). An acoustic description of the vowels of Northern and Southern standard Dutch II: Regional varieties. *J Acoust Soc Am*, 121, 1130-1141.

Boersma, P., & Weenink, D. (2011). Praat: Doing phonetics by computer. Computer program: Version 5.2.40, retrieved 12 September 2011 from http://www.praat.org/.

Bogacka, A. (2004). On the perception of English high vowels by Polish learners of English. *CamLing 2004: Proceedings of the University of Cambridge second postgraduate conference in language research*, Cambridge, pp. 43-50.

Bohn, O.-S. (1995). Cross-language speech perception in adults: First language transfer doesn't tell it all. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 279-300). Timmonium, MD: York Press.

Booij, G. (1995). *The Phonology of Dutch.* Oxford: Oxford University Press.

Botma, B., & van Oostendorp, M. (2012). A propos of the Dutch vowel system 21 years on, 22 years on. In B. Botma, & R. Noske (Eds.), *Phonological Explorations: Empirical, Theoretical and Diachronic Issues* (pp. 1–16). Berlin: Mouton de Gruyter.

Broersma, M. (2005). Perception of familiar contrasts in unfamiliar positions. *J Acoust Soc Am*, 117, 3890-3901.

Broersma, M. (2010). Perception of final fricative voicing: Native and non-native listeners' use of vowel duration. *J Acoust Soc Am*, 127, 1636-1644.

Cebrian, J. (2006). Experience and the use of duration in the categorization of L2 vowels. *J Phon*, 34, 372-387.

Chládková, K., Escudero, P., & Boersma, P. (2011). Context-specific acoustic differences between Peruvian and Iberian Spanish vowels. *J Acoust Soc Am*, 130, 416-428.

Chládková, K., & Podlipský, V. J. (2012). Native dialect matters: Perceptual assimilation of Dutch vowels by Czech listeners. *J Acoust Soc Am*, 130, EL186-EL192.

Crothers, J. (1978). Typology and universals of vowel systems. In J. Greenberg (Ed.) *Universals of Human Language: Vol 2, Phonology* (pp. 93-152). Stanford, CA: Stanford University Press.

Declaration of Helsinki (2008). World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects. Available from: http://www.who.int/en/.

Dietrich, C., Swingley, D., & Werker, J. (2007). Native language governs interpretation of salient speech sound differences at 18 months. *Proc Natl Acad Sci USA*, 104, 16027-16031.

Escudero, P., Benders, T., & Lipski, S. (2009). Native, non-native and L2 perceptual cue weighting for Dutch vowels: The case of Dutch, German, and Spanish listeners. *J Phon*, 37, 452-465.

Escudero, P., Benders, T., & Wanrooij, K. (2011). Enhanced bimodal distributions facilitate the learning of second-language vowels. *J Acoust Soc Am*, 130, EL206-EL2012.

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., & Chládková, K. (2010). Spanish listeners' perception of American and Southern British English vowels. *J Acoust Soc Am*, 128, EL254-EL260.

Escudero, P., & Vasiliev, P. (2011). Cross-language acoustic similarity predicts perceptual assimilation of Canadian English and Canadian French vowels. *J Acoust Soc Am*, 130, EL277-EL283.

Escudero, P., & Williams, D. (2011). Spanish listeners' perception of Dutch vowels. *J Acoust Soc Am*, 129, EL1-EL7.

Escudero, P., & Williams, D. (2012). Native dialect influences second-language vowel perception: Peruvian versus Iberian Spanish learners of Dutch. *J Acoust Soc Am*, 131, EL406-EL412.

Escudero, P., Simon, E., & Mitterer, H. (2012). The perception of English front vowels by North Holland and Flemish listeners: Acoustic similarity predicts and explains cross-linguistic and L2 perception. *J Phon*, 40, 280-288.

Elsendoorn, B. A. G. (1985). Production and perception of Dutch foreign vowel duration in English monosyllabic words. *Lang Speech*, 28, 231-254.

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

Gulian, M., Escudero, P., & Boersma, P. (2007). Supervision hampers distributional learning of vowel contrasts. *Proceedings of the Internation Congress of Phonetic Sciences*, Saarbrücken, pp. 1893-1896.

Hayes-Harb, R. (2007). Lexical and statistical evidence in the acquisition of second language phonemes. *Second Lang Res*, 23, 1-31.

Hisagi, M., Shafer, V. L., Strange, W., & Sussman, E. S. (2010). Perception of a Japanese vowel length contrast by Japanese and American English listeners. *Brain Res*, 1360, 89-105.

Kazanina, N., Phillips, C., & Idsardi, W. (2006). The influence of meaning on the perception of speech sounds. *Proc Natl Acad Sci USA,* 103, 11381-11386.

Kirmse, U., Ylinen, S., Tervaniemi, M., Vainio, M., Schröger, E., & Jacobsen, T. (2008). Modulation of the mismatch negativity (MMN) to vowel duration changes in native speakers of Finnish and German as a result of language experience. *Int J Psychophysiol,* 67, 131-143.

Kondaurova, M. V., & Francis, A. L. (2008). The relationship between native allophonic experience with vowel duration and perception of the English tense/lax vowel contrast by Spanish and Russian listeners. *J Acoust Soc Am,* 124, 3959-3971.

Lehiste, I. (1970). *Suprasegmentals*. Cambridge, MA: MIT Press.

Lipski, S., Escudero, P., & Benders, T. (2012). Language experience modulates weighting of acoustic cues for vowel perception: an event-related potential study. *Psychophysiology,* 49, 638-650.

Lipski, S. C., Lahiri, A., & Eulitz, C. (2007). Differential height specification in front vowels for German speakers and Turkish-German bilinguals: an electroencephalographic study. *Proceedings of the International Congress of Phonetic Sciences*, Saarbrücken, pp. 809-812.

Lipski, S. C., & Mathiak, K. (2007). A magnetoencephalographic study on auditory processing of native and nonnative fricative contrasts in Polish and German. *Neuroscience Letters*, 415, 90-95.

Maddieson, I. (1984). *Patterns of Sounds.* Cambridge: Cambridge University Press.

Maye, J., & Gerken, L. A. (2001). Learning phonemes: how far can input take us? *BUCLD 25 Proceedings*, Somerville, MA, pp. 480-490.

Maye, J., Weiss, D., & Aslin, R. (2008). Statistical phonetic learning in infants: facilitation and feature generalization. *Dev Sci,* 11, 122-134.

Maye, J., Werker, J. F., & Gerken, L. A. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition,* 82, B101-B111.

McGee, T. J., King, C., Tremblay, K., Nicol, T. G., Cunningham, J., & Kraus, N. (2001). Long-term habituation of the speech-elicited mismatch negativity. *Psychophysiology,* 38, 653-658.

Meister, E., Werner, S., & Meister, L. (2011). Short vs. long category perception affected by vowel quality. In Proceedings of the 17th International Congress of Phonetic Sciences, Hong Kong, pp. 1362–1365.

Menning, H., Imaizumi, S., Zwitserlood, P., & Pantev, C. (2002). Plasticity of the human auditory cortex induced by discrimination learning of non-native, mora-timed contrasts of the Japanese language. *Learn Mem,* 9, 253-267.

Moulines, E., & Charpentier, F. (1990). Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones. *Speech Commun,* 9, 453-467.

Moulton, W. (1962). The Vowels of Dutch: Phonetic and Distributional Classes. *Lingua,* 11, 294-312.

Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R., Luuk, A., Allik, J., Sinkkonen, J., & Alho, K. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature,* 385, 432-434.

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin Neurophysiol,* 118, 2544-2590.

Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P., & Winkler, I. (2001). Primitive intelligence in auditory cortex. *Trends in Neuroscience,* 25, 283-288.

Nenonen, S., Shestakova, A., Huotilainen, M., & Näätänen, R. (2003). Linguistic relevance of duration within the native language determines the accuracy of speech-duration processing. *Cogn Brain Res,* 16, 492-495.

Nenonen, S., Shestakova, A., Huotilainen, M., & Näätänen, R. (2005). Speech-sound duration processing in a second language is specific to phonetic categories. *Brain Lang,* 92, 26-32.

Nooteboom, S. G., & Doodeman, G. J. N. (1980). Production and perception of vowel length in spoken sentences. *J Acoust Soc Am,* 67, 276-287.

Obleser, J., Lahiri, A., & Eulitz, C. (2004). Magnetic brain response mirrors extraction of phonological features from spoken vowels. *J Cogn Neurosci*, 16, 31-39.

Rauber, A. S., Escudero, P., Bion, R., & Baptista, B. O. (2005). The interrelation between the perception and production of English vowels by native speakers of Brazilian Portuguese. *Proceedings of Interspeech*, Lisbon, pp. 2913-2916.

Scharinger, M., Idsardi, W., & Poe, S. (2011a). A comprehensive three-dimensional cortical map of vowel space. *J Cogn Neurosci,* 23, 3972-3982.

Scharinger, M., Merickel, J., Riley, J., & Idsardi, W. J. (2011b). Neuromagnetic evidence for a featural distinction of English consonants: Sensor- and source-space data. *Brain Lang,* 116, 71-82.

Scharinger, M., Monahan, P. J., & Idsardi, W. J. (2011c). You had me at "Hello": Rapid extraction of dialect information from spoken words. *NeuroImage,* 56, 2329-2338.

Sharma, A., & Dorman, M. F. (2000) Neurophysiologic correlates of cross-language phonetic perception. *J Acoust Soc Am,* 107, 2697–2703.

Šimáčková, Š., Podlipský, V. J., & Chládková, K. (2012). Czech spoken in Bohemia and Moravia. *J Int Phon Assoc,* 42, 225-232.

Tervaniemi, M., Jacobsen, T., Röttger, S., Kujala, T., Widmann, A., Vainio, M., Näätänen, R., & Schröger, E. (2006). Selective tuning of cortical sound-feature processing by language experience. *Eur J Neurosci,* 23, 2538-2541.

van der Feest, S., & Swingley, D. (2011). Dutch and English listeners' interpretation of vowel duration. *J Acoust Soc Am,* 129, EL57-EL63.

van Leussen, J. W., Escudero, P., & Williams, D. (2011). Acoustic properties of Dutch steady-state vowels: Contextual effects and a comparison with previous studies. *Proceedings of the 17th International Congress of Phonetic Sciences*, Hong Kong, pp. 1194-1197.

van Oostendorp, M. (1995). *Vowel quality and phonological projection*. PhD dissertation, Tilburg University.

Ylinen, S., Shestakova, A., Huotilainen, M., Alku, P., & Näätänen, R. (2006). Mismatch negativity (MMN) elicited by changes in phoneme length: A cross-linguistic study. *Brain Res,* 1072, 175-185.

Zimmerman, S. A., & Sapon, S. M. (1958). Note on vowel duration seen cross-linguistically. *J Acoust Soc Am,* 30, 152-153.

Zonneveld, W. (1993). Schwa, superheavies, stress and syllables in Dutch. *The Linguistic Review*, 10, 61-110.

Tables & Figures

Table 1. MMN amplitude (in μV) averaged across 9 sites (Fz, FCz, Cz, F3, F4, FC3, FC4, C3, C4). The table shows means and 95% confidence intervals (c.i.) per vowel-quality, language, and duration-type, and the number of subjects (n) in each group.

language	n	duration type	n	native vowel		non-native vowel	
				mean	95% c.i.	mean	95% c.i.
Czech	23	long	11	-1.782	-2.465..-1.099	-1.771	-2.284..-1.258
		short	12	-0.884	-1.378..-0.389	-1.344	-1.824..-0.863
		average across long and short		-1.313	-1.742..-0.885	-1.548	-1.883..-1.213
Spanish	22	long	11	-0.795	-1.502..-0.088	-1.783	-2.408..-1.159
		short	11	-0.706	-1.173..-0.238	-1.278	-2.085..-0.470
		average across long and short		-0.750	-1.137..-0.364	-1.531	-2.010..-1.051
Dutch	24	long	12	-1.830	-2.715..-0.944	-0.972	-1.541..-0.404
		short	12	-0.655	-1.121..-0.188	-0.926	-1.333..-0.518
		average across long and short		-1.242	-1.768..-0.717	-0.949	-1.271..-0.627

Fig. 1. F1 and F2 plot of the two vowels produced by a female Estonian speaker that served as stimuli (native vowel quality = blue filled circle; non-native vowel quality = orange filled circle), and the female vowel inventories of the participants' native languages, specifically their native dialects: Moravian Czech (Šimáčková, Podlipský & Chládková, 2012), Iberian Spanish (Chládková et al., 2011), and Randstad Dutch (van Leussen et al., 2011). Symbols represent the mean value of the population, ellipses show 2 standard deviations. Marks are in Hz, axes are scaled in Erb. The quality of the native stimulus resembles the native (long or short) phoneme /a(ː)/ in all three participant languages, while the non-native stimulus does not resemble any phoneme in any of the three participant languages.

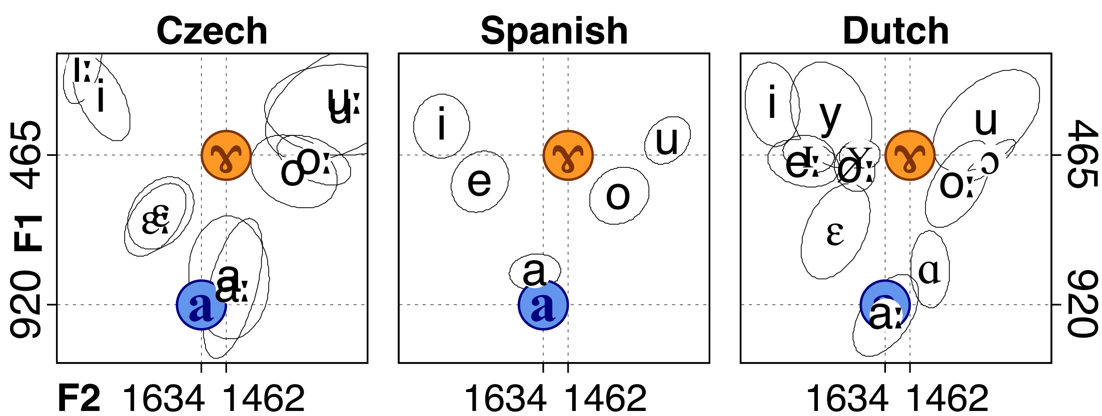


Fig. 2. Grand-average difference waveforms at FCz of the three language groups in the native vowel quality (blue scale; upper graphs) and the non-native vowel quality (orange scale; bottom graphs), for the short stimuli (dashed dark line) and the long stimuli (solid light line). Topographic MMN distributions at average MMN peak latencies at FCz per language, duration-type, and vowel quality (upper plots = native vowel, lower plots = non-native vowel). Peak latencies at FCz are given above the respective scalps.

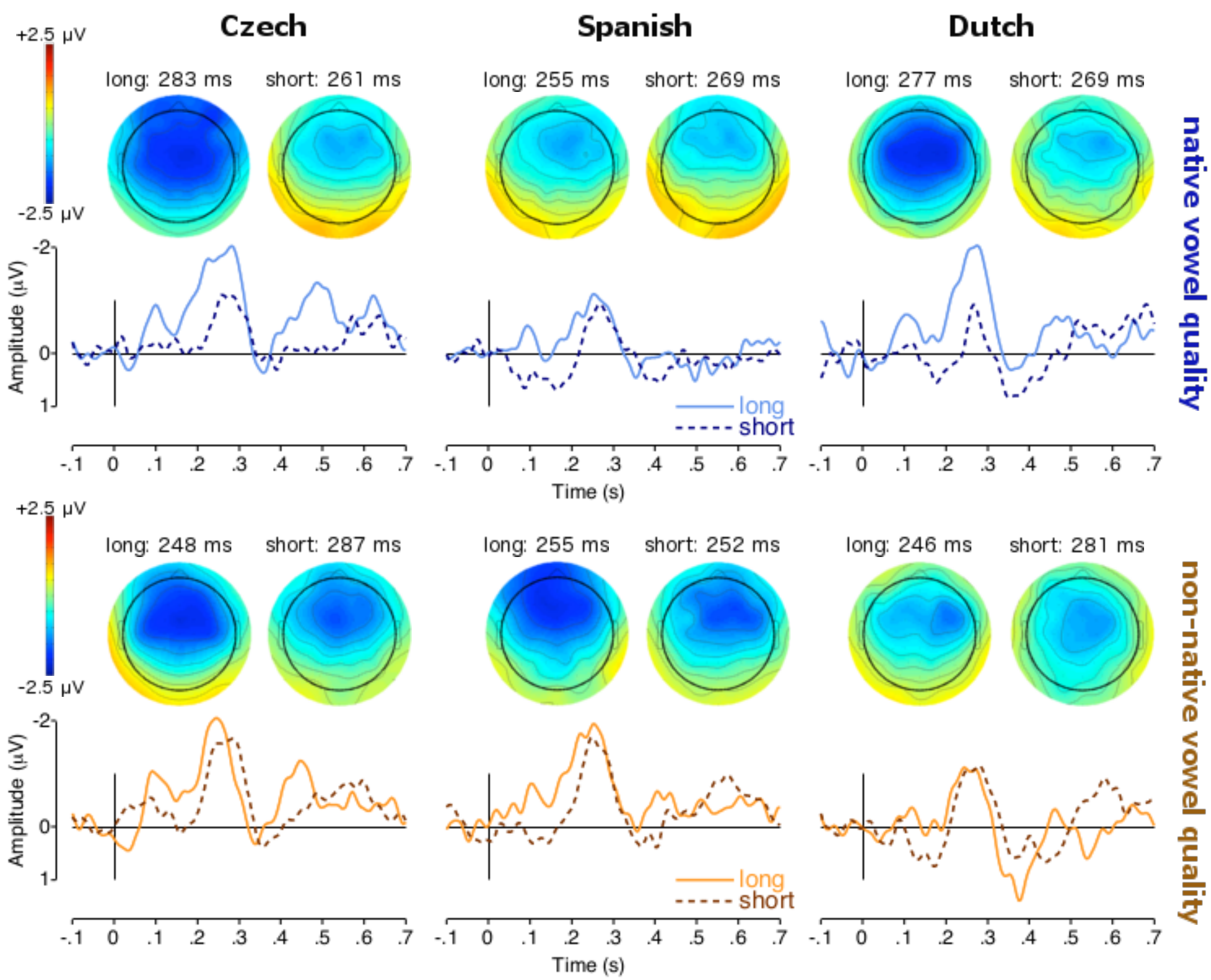


Fig. 3. Grand-average standard (dashed darker lines) and deviant (solid lighter lines) waveforms for long stimuli (top 2 rows) and short stimuli (bottom 2 rows) at FCz of the three language groups (per column) in the native condition (blue scale) and the non-native condition (orange scale).

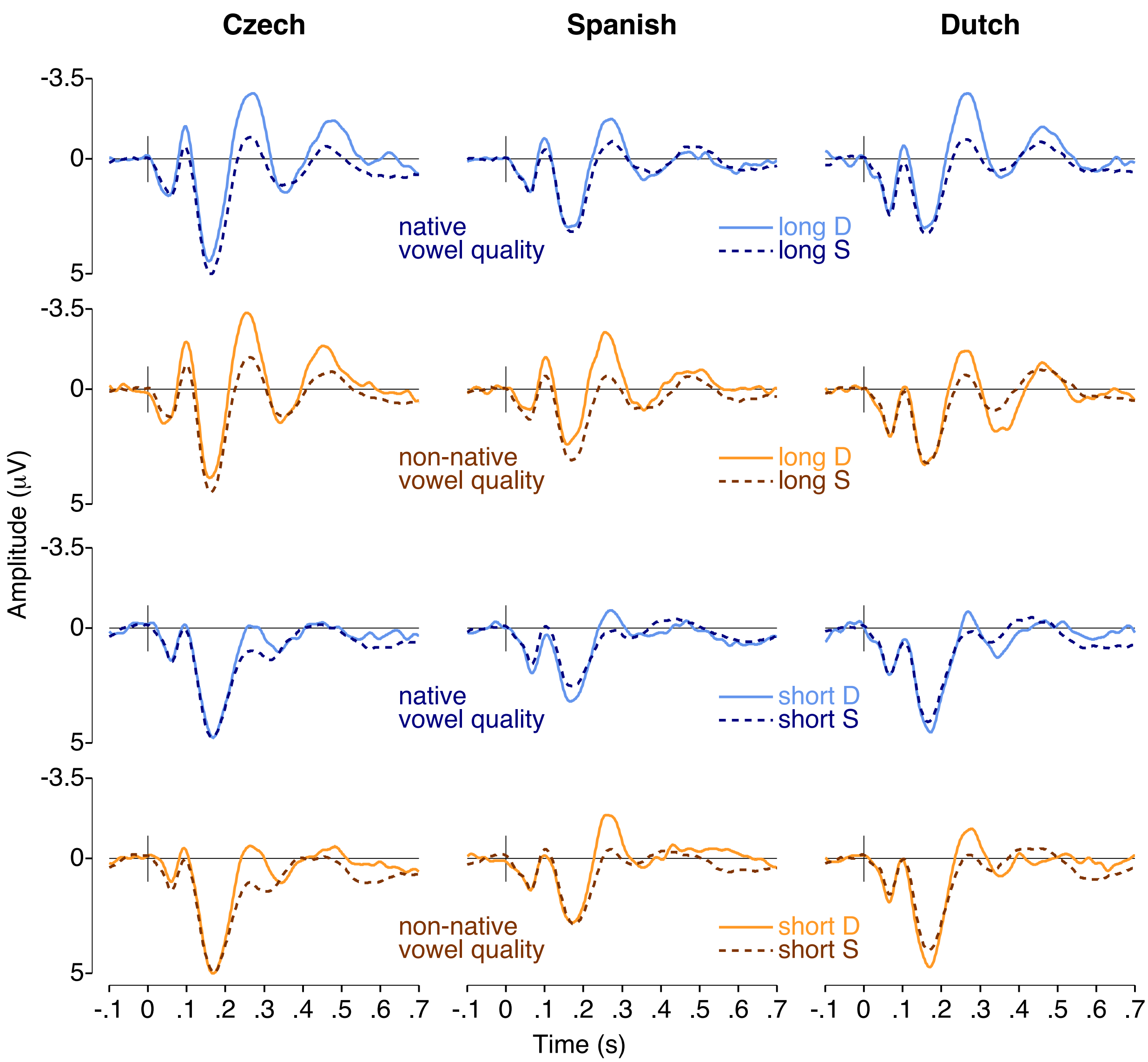


Fig. 4. Mean MMN amplitude per language (Cs = Czech, Sp = Spanish, Du = Dutch) and vowel quality (native = [a], non-native = [ɤ]) pooled across 9 sites (Fz, FCz, Cz, F3, F4, FC3, FC4, C3, C4). The bars indicate 95% confidence intervals, the thick lines mark the mean. Asterisks mark significant differences between groups; ***p*<0.05, **p*=0.065.

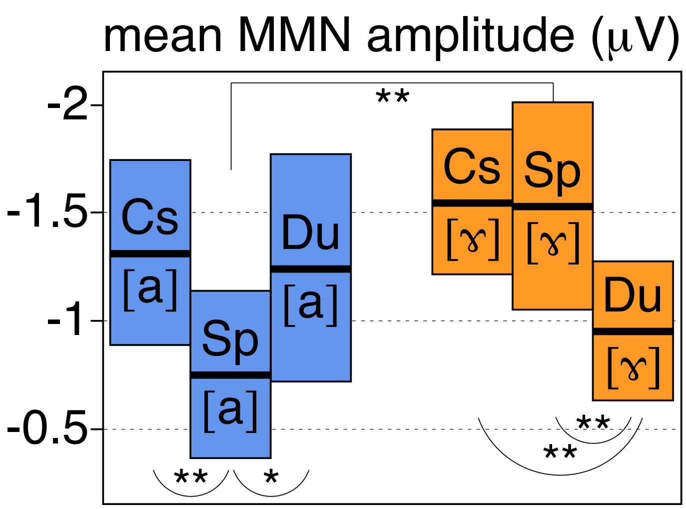


Fig. 5. Model of Czech, Spanish, and Dutch perception based on our findings. Top graph: Czech and Dutch have some linguistic categories defined by duration. In Czech they are abstract length categories independent of vowel quality, while in Dutch these categories are specific vowels (the vowel /a:/ in this case). Bottom graph: Czech but not Dutch can apply their native linguistic categories for duration in perception of non-native vowel qualities. In Spanish, vowel duration does not contribute to native linguistic categories (top graph), and therefore, Spanish listeners can use this dimension to create new non-native categories in a second language (bottom graph).