Modulation of the mismatch negativity (MMN) to vowel duration changes in native speakers of Finnish and German as a result of language experience

Ursula Kirmse, Sari Ylienen, Mari Tervaniemi, Martti Vainio, Erich Schröger, Thomas Jacobsen

Abstract

While crucial for phoneme distinctions in the Finnish language, mere vowel duration is of lower relevance as a phonetically distinctive cue in the German language. To investigate the pre-attentive processing of vowel duration between these two languages, the mismatch negativity (MMN), a component of the auditory event-related potential (ERP) that is an index of automatic auditory change detection, was measured in Finnish and German native speakers for vowel duration changes embedded in the pseudoword sasa. Vowel duration changes thereby were presented as a shortening or a lengthening of either the first- or second-syllable vowel. An additional non-speech condition measured the MMN to duration and frequency changes in tones.

In both language groups, diminished MMN amplitudes for the shortening of vowel duration in the word-final syllable suggested a generally more difficult discrimination of vowel duration in a word-final position. Further, shorter MMN latencies for the Finns than the Germans for vowel duration as well as tone duration deviants suggested a generally higher sensitivity to duration contrasts in the Finnish language group. No latency difference between the groups was found for tone frequency processing. Moreover, the Finns, but not the Germans, showed a leftward shift of the MMN scalp distribution for changes in vowel duration, whereas the MMN topography was highly similar between both groups in the tone condition. An enhanced phonetic processing of vowel duration changes and possibly an enhanced processing of sound duration in general is thus indicated for the Finns as a result of their extensive linguistic experience with phonetically distinctive vowel duration contrasts in the native language.

Keywords: Language experience; Cross-linguistic; Vowel quantity; MMN; Vowel duration; Lateralization; ERP; Word stress

1. Introduction

Cross-linguistic studies demonstrated a substantial influence of language experience on the perception and the processing of speech sounds. Listeners usually show increased discrimination sensitivity to those acoustic cues in speech sounds that signal phoneme distinctions in the native language, i.e. that distinguish the language-specific smallest units of speech that can change the meaning of a word. On the other hand, speech-sound differences which are not relevant for phoneme distinctions in the native language, are typically difficult to discriminate (e.g., Iverson et al., 2003; Kuhl et al., 1992; Zhang et al., 2005; see also Trubetzkoy, 1939/1969). This effect can already be observed at the end of the first year of life (Kuhl et al., 1992; Cheour et al., 1998).

An electrophysiological measure that reflects differences in sound discrimination sensitivity is the mismatch negativity (MMN), which is a component of the auditory event-related potential (ERP) and which indexes auditory change detection. The MMN (Näätänen et al., 1978) is elicited by an irregular sound (the deviant) occurring in an otherwise regular sequence
of so-called standard sounds (for review see e.g., Schröger, 1998). This negative deflection is observed in the ERP to the deviant sound compared to the ERP to the standard sound and is best visible in the deviant-minus-standard difference wave. The MMN displays a frontocentral scalp distribution and a latency between 100 and 250 ms after deviance onset. Main neural generators of the MMN have been localized bilaterally in the auditory cortex with additional bilateral generators in the frontal lobe (e.g., Alho 1995; Rinne et al., 2000). Whereas especially the frontal generator usually shows a right-hemispheric predominance when non-speech sounds are presented (e.g., Levänen et al., 1996; Paavilainen et al., 1991), speech stimuli often elicit a more symmetrical or even left-hemispherically lateralized MMN or magnetic MMNm (e.g., Alho et al., 1998; Rinne et al., 1999; Takegata et al., 2004; Tervaniemi et al., 2000), associated with enhanced activity of generators in the left-hemispheric auditory cortex (Nääätänen et al., 1997; Pulvermüller et al., 2001).

Importantly, the MMN is elicited irrespectively of voluntary sound discrimination and even for an auditory stimulation presented outside the focus of voluntary attention. It thus indexes auditory change detection at an automatic, pre-attentive level. Yet, the parameters of the MMN closely correlate with the behavioral discrimination performance measured in additional active discrimination tasks: higher discrimination sensitivity here has been found to be associated with larger MMN amplitudes and/or shorter MMN latencies (e.g., Kujala et al., 1998; Pulvermüller et al., 2001). This negative deflection is observed in the ERP to the standard sound (e.g., Pulvermüller et al., 2001).

In a recent study, enhanced MMN amplitudes and shorter MMN latencies have been found as a consequence of short-term auditory discrimination training for non-speech sounds (e.g., Menning et al., 2000; Nääätänen et al., 1993) as well as speech sounds (e.g., Kraus et al., 1995; Menning et al., 2002). For the processing of speech-sound changes, enhanced amplitudes and shorter latencies of the electric and magnetic MMN response were frequently found for native, in contrast to non-native phoneme contrasts, reflecting the effects of native-language-specific phoneme perception (e.g., Cheour et al., 1998; Dehaene-Lambertz, 1997; Nääätänen et al., 1997; Sharma and Dorman, 2000; Winkler et al., 1999; see Nääätänen 2001 for review).

German and Finnish, the two here investigated languages, differ in the extent to which they apply vowel duration changes to signal phonetic distinctions. In so-called quantity languages, such as Finnish, Japanese or Thai, “long” and “short” vowels (and/or consonants) are distinguished as different phonemes based on the relative physical durations of the speech sounds in a sequence. In Finnish, for example, vowel duration contrasts (so-called “quantity oppositions”) are used to distinguish words (e.g., /tuli/ [short] = “fire” vs. /tuuli/ [long] = “wind”) as well as to mark grammatical inflections and cases (e.g., /talon/ [short] = ‘of the house’ vs. /italon/ [long] = ‘into the house’). Non-quantitative languages as e.g., Russian, on the other hand, do not differentiate between different vowel quantities. Although German conventionally differentiates between long and short vowel phonemes, it can neither be regarded as a genuine quantity nor a pure non-quantity language. German quantity oppositions are concurrently accompanied by large spectral differences for most vowels 1, and it is not satisfactorily clarified which cues, durational or spectral, German listeners actually use to distinguish between both instances (cf. Becker, 1998; Ramers, 1988). Furthermore, compared to Finnish, German shows a rather restricted use of these vowel oppositions: these occur predominantly in stressed syllables (and thus not in the typically unstressed, word-final positions) and do not indicate grammatical functions. On the other hand, vowel duration in German is an important cue for word stress (see Dogil and Williams, 1999). The role of vowel duration as a phonologically distinctive cue thus is more restricted and controversial in German (cf. Becker, 1998), whereas it is highly relevant for phonological distinctions in Finnish (see e.g., Lehtonen, 1970; Suomi and Ylitalo, 2004; Suomi, 2005).

Recent studies showed that linguistic experience with vowel duration as a phonetically distinctive cue is reflected in the pre-attentive MMN response: Native speakers of Finnish (quantity language), compared to native speakers of Russian (non-quantity language), were found to show an enhanced MMN to deviants in vowel duration (Nenonen et al., 2003, 2005; Ylinen et al., 2006). This difference was apparent even though the investigated children (aged 10 to 14 years) were fluent second-language (L2) users of the quantity language Finnish. Together with studies applying fMRI or NIRS to investigate the processing of vowel duration (e.g., Gandour et al., 2002; Minagawa-Kawai et al., 2005), these MMN results suggest that vowel duration processing is crucially influenced by the linguistic experience with vowel duration in the phonological system of the native language. Furthermore, they indicate an influence of the linguistic status of the native language as a quantity or non-quantity language already at the pre-attentive level of speech-sound processing.

The intermediate status of the German language regarding the use of distinctive vowel duration differences therefore stimulates the question to which extent vowel duration processing, especially on the pre-attentive level, might differ when native speakers of German are contrasted to speakers of a genuine quantity language such as Finnish. Previously, Menning et al. (2002) found larger magnetic MMN responses with shorter latencies for Japanese subjects when comparing the processing of duration contrasts embedded in Japanese words between native speakers of German and Japanese. However, since the stimuli were meaningful words in Japanese but not German, this might have at least partly been due to lexical processing in the Japanese subjects, which has been found to foster the MMN response (e.g., Pulvermüller et al., 2001). In a recent study,

1 This difference is similar to the “tense” versus “lax” opposition in English and is due to a different constellation of the speech-organs (see e.g., Hertrich and Ackermann, 1997). Generally spoken, German “short” vowels are articulated with a lower, more lax tongue position compared to their long counterparts [as in the English contrast of ‘pull’ (lax) vs. ‘pool’ (tense); cf. Becker, 1998]. This affects F1 and F2 formant frequencies which in particular determine the timbre of the vowel. An exception is the vowel /a/ (see Materials and methods). Since the vowel classes in German nevertheless are usually referred to as “short” and “long”, we will keep this terminology in the proceeding of this article.
Tervaniemi et al. (2006) found enhanced MMN responses for Finns in contrast to Germans for deviants in tone duration, but not for deviants in tone frequency. The authors concluded that the different long-term experience with vowel duration as a linguistic cue in Finnish and German has been transferred to the processing of sound duration on a non-speech level. Since the processing of duration, however, may differ between speech and non-speech sounds (e.g., Nenonen et al., 2003), it is fundamentally interesting to compare duration processing between the two language groups by using phonetic stimuli.

The present study measured the electroencephalographic MMN response elicited by changes in vowel duration in adult native speakers of Finnish and German. Vowel duration changes were presented in the context of a naturally sounding pseudoword, the pseudoword sasa. On the one hand, the choice of pseudoword stimuli avoided the contribution of possible effects of lexical processing. On the other hand, it allowed a direct comparison of vowel duration processing in the first- and second-syllable position of a word within as well as between the language groups. In an additional tone condition, tone duration and tone frequency changes were presented similarly as in Tervaniemi et al. (2006), to assess language-specific differences in the duration processing of non-speech sounds (duration deviant), as well as possible non-language related differences between the groups (frequency deviant).

2. Materials and methods

2.1. Participants

A group of Finnish and a group of German native speakers were tested. Finnish participants were recorded in Helsinki (Cognitive Brain Research Unit, University of Helsinki, Finland); German participants were recorded in Leipzig (Research Group of Cognitive and Biological Psychology, University of Leipzig, Germany) by the same experimenter (first author). For each language group, 13 participants who were matched between the groups in sex and in approximate age were included in the data analysis (mean age (±SD) for Finns vs. Germans 25.46 (±5.2) vs. 26.1 (±4.1) years; 5 females per group). Additionally recorded participants were excluded from the data analysis due to heavy blinking (3; two Finns) or due to technical failure during the recording of the electroencephalographic data (3, three Finns). In order to obtain two equally large groups in the data analysis, two further German participants, who had the poorest fit with the remaining Finns concerning the sociodemographic data, were also excluded from the analysis.

All participants grew up in monolingual families and were native speakers of the respective language only. German participants reported no knowledge of or extended passive exposure to the Finnish language. Participants were healthy, had no history of neurological or hearing disorders, and apart from one Finnish participant who was a compensated stammerer, had no language-related disorders. Participants gave their informed consent before the experiment and received payment or course credits for their participation.

2.2. Design

Vowel duration changes were presented in either the first or second syllable of the pseudoword sasa in four separate experimental conditions (see Table 1). For each of the two syllable positions, deviants had a shorter vowel duration in one condition (condition names First-Syllable Duration-Decrement and Second-Syllable Duration-Decrement), but a longer vowel duration in a second condition, in which the roles of deviant and standard stimuli were respectively inverted (condition names First-Syllable Duration-Increment and Second-Syllable Duration-Increment). As a result of this reversed design, MMN difference waves for the Pseudoword conditions could subsequently be obtained by subtracting responses to physically identical stimuli (see Jacobsen and Schröger, 2003). Pseudoword deviants occurred pseudo-randomly with a probability of 15% and at least two standards following each deviant. In total, 1400 stimuli were presented for each Pseudoword condition, respectively divided to two experimental blocks. The inter-stimulus interval (ISI) was randomly varied between 350 ms and 650 ms, so that no rhythmical pattern could be established by the temporal characteristics of the acoustic stimulation per se. Thus, the obtained MMN responses can only be explained as responses to the changes in vowel duration.

For the Tone stimuli, two different deviants were presented in one experimental condition (see Table 1): one deviant differed from the standard in tone duration (Tone-Duration deviant), the other in tone frequency (Tone-Frequency deviant). In the Tone condition, each deviant type occurred randomly with a probability of 10% among a total of 2000 stimuli presented with a steady stimulus-onset asynchrony (SOA) of 300 ms. Altogether, nine experimental blocks were presented to the participants in randomized order.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Stimuli presented as standards and deviants in the six different deviant conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deviant condition</td>
<td>Standard stimulus</td>
</tr>
<tr>
<td>Pseudowords</td>
<td>First-Syllable</td>
</tr>
<tr>
<td></td>
<td>Second-Syllable</td>
</tr>
<tr>
<td>Tones</td>
<td>Tone-Duration</td>
</tr>
<tr>
<td></td>
<td>Tone-Frequency</td>
</tr>
</tbody>
</table>

The bold font signifies what aspect of the stimulus changed between standard and deviant stimuli in each condition.

Note: Roles of standard and deviant stimuli were respectively inversed in the decrement vs. increment Pseudoword conditions.
Subsequent to the EEG recordings, data on the subjective perception of the pseudoword stimuli were collected. To this end, participants were presented with two short oddball sequences taken from the pseudoword conditions, both either representing the decrement or increment conditions for both syllable positions. Participants were asked to listen attentively and to identify the feature that distinguished the deviant from the standard stimulus. The presentation of decrement/increment sequences and the order of first-/second-syllable position presentation thereby were counterbalanced over the participants in each language group. In a second task, each pseudoword stimulus was presented in isolation and participants were asked to write it down (to test the correct identification) and further to rate the typicality of the stimulus with respect to the native language on a seven-step rating scale (1—very typical; 7—very untypical).

2.3. Stimuli

Three different pseudoword stimuli, differing only in vowel duration, were used in the experiment (see Table 1 / Fig. 1). One of the three stimuli contained two short vowels (pseudoword /sasa/; \(a_1\)=78 ms; \(a_2\)=110 ms; see Fig. 1a). The two further stimuli respectively contained a longer vowel duration in either the pseudowords' first syllable (pseudoword /s\(a\)\(\ast\)\(a\)/, \(a_1\)=135 ms; \(a_2\)=110 ms; see Fig. 1b) or second syllable (pseudoword /s\(a\)\(a\)/, \(a_1\)=78 ms; \(a_2\)=190 ms; see Fig. 1c). None of these pseudowords was a lexical item in Finnish or German. The three stimuli were produced from one primary stimulus of the pseudoword /sasa/ recorded from a female native speaker of Finnish [first-syllable vowel: F0=212 Hz, F1=785 Hz, F2=1380 Hz; second-syllable vowel: F0=194 Hz, F1=745 Hz, F2=1280 Hz (for details on stimulus construction see below)]. For each syllable position, shorter vowels had a duration of 58% of the duration of the longer vowel (/s/-consonants had a constant duration of 95 ms \([s_1]\) and 71 ms \([s_2]\) in all three pseudoword stimuli).

Vowel duration contrasts were presented in the vowel /a/ because long and short realizations of this vowel have a very similar formant structure within each of the two languages, as well as between both languages (F1 ca. 700 Hz, F2 ca. 1250 Hz, cf. also Jacobsen et al., 2004). Note that the vowel /a/ in this respect is an exception in the German phonology system (see Becker, 1998; Hertrich and Ackermann, 1997). Using this particular vowel, however, long and short vowels were similarly valid for both groups in our stimuli. In principle, the Germans thus could have attributed phonetic relevance to the vowel duration changes in our stimuli as well. Differences in the duration processing between the groups thus purely reflect the influence of the general linguistic experience.

In the Tone condition, three harmonically complex tones (fundamental frequency plus two upper partials) with a rise and fall time of 5 ms each were presented (identical to the stimuli used by Tervaniemi et al., 2006). These Tone stimuli had (1) a frequency of 500 Hz and a duration of 75 ms (Standard), (2) a frequency of 500 Hz and a duration of 25 ms (Tone-Duration deviant) and (3) a frequency of 550 Hz and a duration of 75 ms (Tone-Frequency deviant, see Table 1).

2.4. Construction of the pseudoword stimuli

In constructing the pseudoword stimuli, special attention was paid to generate non-artificial, naturally sounding speech stimuli. In a first step, a large set of utterances of the three different pseudowords (/sasa/, /s\(a\)\(\ast\)\(a\)/and /s\(a\)\(a\)/) was recorded in a sound proof studio at the Department of Speech Sciences of the University of Helsinki from three different female native speakers of Finnish. These recordings were made directly into a computer hard-drive, using a 44.1 kHz sampling frequency and a 16-bit quantization. The succession of the pseudowords thereby was repeatedly alternated between the utterances in all possible orders and the pseudowords were spoken with word stress given to the first syllable as prototypical for disyllabic words in both languages. Typical durations of the speech sounds in the respective pseudowords then were determined from these recordings and a representative exemplar of the pseudoword /sasa/ was chosen as the primary stimulus for further stimulus construction.

To obtain the stimuli /s\(a\)\(\ast\)\(a\)/ and /s\(a\)\(a\)/, this primary stimulus was modified in the length of the respective vowel by duplicating and/or removing one to two glottal period sized chunks of the vowel signal. These manipulations were limited to the static part of the vowel (in order to avoid discontinuities) and occurred outside of the extent of the formant transitions and segmentally conditioned, microprosodic perturbations in fundamental frequency. The choice of manipulation points thus
minimized distortions of fundamental frequency, spectral characteristics (as formant frequencies and spectral slopes) or intensity. Thus, the only perceivable difference between the stimuli was vowel duration. To remove excess sibilance from the fricatives, the stimuli were finally low-pass filtered using a filter cut-off frequency of 8700 Hz and a very mild slope, ensuring that both, vowels and fricatives, retained their spectral characteristics.

For the /sasa/ vs. /sasəsə/ as well as the /sasa/ vs. /sasa/ contrast, the respective stimuli were virtually identical to a point approximately 20 ms after vowel onset. The onset of deviance was determined as the time point at which the intensity of the vowel signal became below .02 dB relative intensity, assuming that lower intensity parts would not contribute significantly to the sound perception (171 vs. 320 ms for the first, respectively second vowel). Due to linguistic phenomena as utterance-final lengthening and, specifically, the phenomenon of the Finnish “half-long” vowel, the word-final short vowel in the primary stimulus /sasa/ had a longer duration than the first-syllable short vowel. This difference was kept in order to preserve the naturalness of the stimuli. To account for this primary difference in vowel duration, an equal percentage of duration change was applied for the two syllable positions, instead of an equal absolute amount.

2.5. Experimental procedure and data acquisition

Participants were comfortably seated in an acoustically and electrically shielded cabin and were instructed to ignore the presented sounds, while watching a muted, self-selected movie. The movie was presented on a computer monitor with subtitles given in the participant’s native language. Sound stimuli were presented binaurally via headphones at 40 dB (SPL) above the individually determined hearing threshold (BrainStim (Helsinki)/ERTS (Leipzig) presentation software).

EEG data were collected from 128 active Ag/AgCl electrodes with technically identical recording systems in both laboratories (BioSemi system). Electrodes were embedded in a prefabricated cap and were arranged equidistantly (5% distance) in concentric circles around the vertex plus extending to inferior-posterior regions. Further electrodes were attached to the left and right mastoid site and the tip of the nose. The vertical and horizontal electrooculogram (VEOG/HEOG) were measured using electrodes placed above and below the left eye (VEOG) and lateral to the outer canthi of both eyes (HEOG). EEG data were recorded continuously (sampling rate 512 Hz; 128 Hz low-pass filter) with a reference electrode left-posterior to the vertex and were offline referenced against the nose. Data analysis was carried out using the EEP analysis software (Max-Planck-Institute of Cognitive Neuroscience/A.N.T. Software).

2.6. Data analysis

The continuous EEG data were filtered with a 1 to 15 Hz band-pass filter and ERPs epochs were calculated with a 100 ms pre-stimulus interval and a further length of 820 ms (Pseudoword stimuli) or 300 ms (Tone stimuli), respectively. For the Pseudoword stimuli, baseline correction was applied over a 100 ms interval previous to the deviance onset to avoid potential influences of language-specific processing differences for the preceding speech sounds or sustained shifts over the rather long stimuli. Importantly, an identical baseline interval thereby was used for standards and deviants belonging to one condition. For first-syllable vowel duration changes, the deviance onset was at 171 ms, for second-syllable vowel duration changes, at 320 ms (cf Fig. 2). For the less complex Tone stimuli, the baseline was generally calculated for a 100 ms interval before stimulus onset. Trials containing voltage changes exceeding a range of 120 μV, the first ten trials of each block, as well as each two standards following a deviant were excluded and responses to standard and deviant stimuli were then averaged separately for each participant and condition. Deviant-minus-standard difference waves were calculated for each participant and condition (in the Pseudoword conditions subtracting responses to physically identical stimuli) and used as the basis of further statistical analysis. Statistical analysis only included those participants with at least 80 accepted deviant trials in each condition.

For MMN quantification, MMN amplitudes were determined for each participant and condition as the mean amplitude within an 80 ms time window centered to the average MMN peak latency of both groups measured in the grand-average waveforms at Fz. For each group and condition, statistical presence of the MMN was confirmed at Fz using one-tailed t-tests against zero for the nose-referenced data (with Bonferroni-corrected significance levels according to the number of calculated t-tests). In order to compare the hemispheric distribution of MMN activity between the frontal and posterior regions, as well as to include the polarity inverted MMN activity over inferior-posterior regions into the analysis, a laterality index was computed based on the nose-referenced data. This index was calculated as the root mean square (RMS) of all electrodes respectively located over the left-frontal, right-frontal, left-posterior and right-posterior scalp regions, as separated by and excluding the central coronal and sagittal lines. Calculating the RMS includes the absolute values, independently from negative or positive polarity, in the average and thus provides a joint index of positive and negative activations. Since statistical comparisons of the four sectors showed significant effects of hemispheric lateralization at frontal scalp regions only, this supported a re-referencing of the signal to the average voltage at both mastoids. Further analyses of MMN amplitudes and topography accordingly were restricted to a grid of twelve frontocentral electrodes, approximately corresponding to the positions AF3, AFz, AF4, F3, Fz, F4, FC3, FCz, F4, C3, Cz, C4 in the international ten–twenty system. Scalp current density maps (SCDs, see Fig. 6) are given to provide a display of the possible MMN generator structures (Perrin et al., 1989; SCDs were calculated using a free Matlab

2 In a structure CVCV (a disyllabic word with two short vowels), the second short vowel has a distinctly longer duration than the first short vowel in Finnish and is therefore characterized as “half-long”. The degree of this prolongation is subject to situational and dialectal variations (see e.g., Suomi and Ylitalo, 2004).

MMN latencies were determined for the mastoid-referenced data by an automatic software procedure retrieving the time point of the most negative peak within a 120 ms wide analysis window centered to the grand-average peak latency at Fz. Latencies were analyzed relative to the time point of deviance onset. MMN latencies were analyzed at the electrodes Fz and FCz, at which the MMN amplitude was found to be maximal.

For statistical comparison, MMN responses were firstly analyzed comparing Pseudoword and Tone stimuli in overall repeated-measures analyses of variance (ANOVAs) including the between-subjects factor group (Finnish vs. German) and the within-subject factor stimulus class (Pseudoword stimuli vs. Tone stimuli). For these comparisons, data were averaged over the respective Pseudoword and Tone conditions (four Pseudoword vs. two Tone conditions). For the MMN latencies, however, the factor stimulus class included the averaged Pseudoword conditions in contrast to the Tone-Frequency condition only, since here a comparison to a non-durational control was needed rather than a comparison between Pseudoword and Tone stimuli per se. Comparisons further included topographical within-subject factors depending on the analyzed parameter: for the analysis of the laterality index the factors hemisphere (left/right) and scalp region (anterior/posterior), for MMN amplitudes the factors gradient (AF-, F-, FC- and C-line) and laterality (left, midline, right) and for MMN latencies the factor electrode (Fz, FCz). To compare between the individual conditions, Pseudoword and Tone stimuli were further analyzed in separate repeated-measures ANOVAs including the within-subject factor condition (Pseudoword stimuli: First-Syllable Duration-Decrement, Second-Syllable Duration-Decrement, First-Syllable Duration-Increment and Second-Syllable Duration-Increment; Tone stimuli: Tone-Duration and Tone-Frequency). Greenhouse–Geisser corrected degrees of freedom are reported when the assumption of sphericity was not met for the within-subject factors. For post-hoc analyses, subsequent ANOVAs were run separately for each language group or stimulus type. Analysis of stimulus rating and identification included only the data of those participants included in the ERP data analysis (n=13 for each group).

3. Results

3.1. Presence of MMN

Negative peaks were observed in the deviant-minus-standard difference waves for both language groups in all deviant conditions (see Fig. 3; see Fig. 2 for examples of standard and deviant ERPs). These negative deflections were statistically significant for all groups and conditions except the Second-Syllable Duration-Decrement condition for the Germans (see Fig. 5). They were characterized by typical MMN latencies (120 to 180 ms after deviance onset) and scalp distributions (a frontocentral maximum with a polarity inversion at the mastoid leads). The overall analysis of the laterality index showed stronger MMN activations over frontal than posterior scalp regions (main effect scalp region: $F_{1,24}=45.15; p<.001$). Additionally, a main effect of hemisphere ($F_{1,24}=4.31; p<.05$) indicated a right-hemispheric lateralization of the MMN response, however, an

Fig. 2. Standard and deviant ERPs for vowel duration decrements and the Tone stimuli. a) First-Syllable Duration-Decrement: deviant /sasa/ (standard /sa*sa/) and standard /sasa/ from the reversed condition (First-Syllable Duration-Increment). b) Second-Syllable Duration-Decrement: deviant /sasa/ (standard /sa*sa/) and standard /sasa/ from the reversed condition (Second-Syllable Duration-Increment). c) ERPs for the Tone-Duration, the Tone-Frequency deviant and the tone standard.
additional interaction of hemisphere × stimulus class ($F_{1,24}=6.92; p<.05$) indicated differences in the MMN lateralization between the stimulus types. In post-hoc analyses, an effect of hemisphere was only found for the Tone stimuli (main effect hemisphere: $F_{1,24}=7.69; p<.05$).

For the MMN amplitudes, the overall ANOVA reflected the characteristic frontocentral maximum of the MMN (main effect gradient: $F_{3,72}=18.06; p<.001$, main effect laterality: $F_{2,48}=13.61; p<.001$ interaction gradient × laterality: $F_{6,144}=4.26; p<.05$). Again, an additional interaction of laterality × stimulus class ($F_{2,23}=8.34; p<.01$) indicated the hemispheric distribution of MMN activity to differ between Pseudoword and Tone stimuli. Post-hoc analyses showed a significant effect of laterality for the Tone stimuli only, here again reflecting a right-hemispheric lateralization of the MMN response (main effect laterality: $F_{2,48}=21.37; p<.001$).

3.2. Between group effects

3.2.1. MMN latencies and amplitudes

Between the language groups, no differences were found in MMN amplitude. MMN latencies, however, showed a significant interaction of stimulus class × group ($F_{1,23}=6.22; p<.05$). Separate analyses here showed significantly shorter MMN latencies for the Finnish than the German participants for the Pseudoword stimuli (see Fig. 4; main effect group for Pseudoword stimuli: $F_{1,23}=24.37; p<.001$). On the other hand, MMN latencies did not differ between the Finns and the Germans for the Tone-Frequency condition. Analysis of MMN latencies for the Tone conditions yielded no main effect of group, but an interaction of group × condition ($F_{1,23}=5.11; p<.05$). Post-hoc analyses revealed shorter MMN latencies for the Finnish compared to the German participants in the Tone-Duration condition (main effect group: $F_{1,24}=25.58; p<.001$), but not the Tone-Frequency condition (see Fig. 4). The results thus showed duration-specific differences in MMN latency between the groups.

3.2.2. MMN topography

Besides the overall characteristic MMN scalp distribution, statistical analysis revealed a stimulus-class specific group difference in MMN lateralization at frontal electrodes. For the Pseudoword stimuli, i.e. the vowel duration contrasts, Finnish participants showed a rather left-hemispheric frontal MMN
lateralization, whereas a clear right-hemispherically lateralized MMN was observed for the German participants (see Fig. 6a/Table 2). However, there was no group difference in MMN lateralization for the Tone stimuli, for which a right-hemispheric lateralization of the MMN was observed for the Finns as well as the Germans (see Fig. 6b/Table 2).

In the analysis of the laterality index, this dissociation was reflected by a 4-way interaction of stimulus class × hemisphere × scalp region × group (F_{1,24}^2 = 7.83; p < .01). Post-hoc ANOVAs showed, firstly, an interaction of stimulus class × hemisphere × group for the frontal (F_{1,24}^1 = 8.49; p < .01), but not the posterior scalp regions. Secondly, an interaction of group × hemisphere × scalp region (F_{1,24}^1 = 11.81; p < .05) was observed for the Pseudoword stimuli only. Thirdly, a significant interaction of stimulus class × hemisphere × scalp region was only found for the Finns (F_{1,12} = 8.90; p < .05). Differences in MMN lateralization thus were found only over frontal scalp regions, where the MMN lateralization differed between the Finns and the Germans for the Pseudoword stimuli, and between the stimulus types only for the Finns. Post-hoc test at frontal scalp regions showed that for the Pseudoword stimuli larger values were observed for the laterality index over the right than left hemisphere for the Germans (F_{1,13}^3 = 2.44; p < .05), whereas no difference between the hemispheres was found for the Finns. This confirmed a more symmetric MMN for the Finnish group, which numerically was even larger over the left hemisphere (cf. Fig. 6, middle).

For the MMN amplitudes, this stimulus-type specific group difference in MMN lateralization in the overall ANOVA was reflected in an interaction of stimulus class × laterality × group (F_{2,23}^3 = 9.02; p < .01). Similarly as the results of the laterality index, this interaction reflected differences in MMN lateralization between the stimulus types only for the Finns (interaction of stimulus class × laterality for the Finns: F_{2,14}^3 = 16.84; p < .001). On the other hand, it again indicated different group effects on MMN lateralization for Pseudoword and Tone stimuli (see Fig. 6). For the Pseudoword stimuli, a group × laterality interaction (F_{2,48}^2 = 3.61; p < .05) indicated a leftward shift of the MMN for the Finns [post-hoc tests: no difference between left and midline electrode positions, but a tendency (p = .1) for smaller amplitudes at right vs. midline electrodes], whereas Germans showed a clear right-hemispheric lateralization of the MMN [post-hoc tests: significantly smaller MMN amplitudes at left compared to midline electrode positions (p < .01), but not right than midline positions, see Fig. 6a/Table 2]. For the Tone stimuli, an interaction of group × laterality (F_{2,48}^2 = 3.68; p < .048) contrariwise reflected that Finns here showed an even more right-lateralized MMN [post-hoc tests: smaller MMN amplitudes at left compared to midline and left compared to right electrode positions (both p < .01)] compared to the Germans [post-hoc tests: significantly smaller MMN amplitudes over left compared to midline electrodes (p < .05), but no difference between right and left electrodes].

3.3. General effects of stimulus class and deviant condition

Analysis of the MMN amplitudes for the four Pseudoword conditions yielded a significant main effect of condition (F_{3,72}^1 = 4.91; p < .007), which was due to a significantly smaller MMN amplitude in the Second-Syllable Duration-Decrement condition compared to the other Pseudoword conditions [see Fig. 5; post-hoc comparisons: p = .006 (vs. First-Syllable Duration-Decrement), p = .013 (vs. First-Syllable Duration-Increment), p = .009 (vs. Second-Syllable Duration-Increment)]. This effect was descriptively larger for the Germans, but no interaction with the factor group was observed. MMN latencies were generally shorter for increments compared to decrements of vowel duration (main effect condition: F_{3,72}^1 = 8.2; p < .001; post-hoc tests: all p < .02, except p = .07 for the comparison of Second-Syllable Duration-Increment vs. Second-Syllable Duration-Decrement), probably due to a contribution of an enhanced N1 response in the increment conditions, because of the increased stimulus energy for longer deviants.

Overall ANOVAs further showed general effects of stimulus class and the analysis of the Tone conditions showed further effects of condition on MMN amplitudes and latencies (cf. Figs. 4/5). However, these effects are not described here, since they were due to different amounts of deviance between these conditions and are not related to our cross-linguistic research question. The overall ANOVA for the MMN amplitudes further showed an interaction of stimulus class × gradient.
directly identified the stimulus German. Writing down the stimuli, German subjects correlated functional role of vowel duration as a word-stress cue in stimuli as changes in the word-stress pattern, reflecting the Finnish participants, described the differences between the standard and deviant stimuli as changes between long and short German participants (cal

\( F_{3,22} = 19.89; p < .001 \)), reflecting a more spatially constrained maximum of the MMN amplitude in the Tone condition. No further significant effects were found for the ERP data.

### 3.4. Stimulus rating and identification

After listening to the samples of the oddball sequences, all participants consistently described the stimuli as non-lexical speech. Twelve of the Finnish (\( n = 13 \)) but only eight of the German participants (\( n = 13 \)) described the differences between standard and deviant stimuli as changes between long and short vowels. Additionally, seven of the German, but only one of the Finnish participants, described the differences between the stimuli as changes in the word-stress pattern, reflecting the functional role of vowel duration as a word-stress cue in German. Writing down the stimuli, German subjects correctly identified the stimulus /sasa'/ less often than the other two stimuli (correct identifications: Germans: 4 (/sasa'/) vs. 9 (/sasa') and 10 (/sasa/) subjects, \( n = 13 \); Finns: 11 (/sasa'/) vs. 13 (/sasa') and 12 (/sasa/) subjects, \( n = 13 \)). The typicality in respect to the native language was rated slightly higher by Finns than Germans \( ^3 \) (rating 1 = very typical, 7 = very untypical; median ratings Finns: 3 (/sasa'), 3 (/sasa'), 2 (/sasa'); median ratings Germans: 5 (/sasa'), 5 (/sasa'), 3 (/sasa')).

### 4. Discussion

We investigated whether Finnish and German native speakers differ in the pre-attentive processing of vowel duration. Whereas this feature is highly phonetically relevant in Finnish, it is less important in the German language. Both groups thus differ in their linguistic experience with vowel duration contrasts. The MMN elicited by decrements and increments of vowel duration in the first or second syllable of the pseudoword sasa, as well as the MMN elicited by duration and frequency deviants in complex tones, was measured in separate passive oddball paradigms. The obtained group effects regarding MMN amplitudes and latencies on the one hand, and MMN topography on the other hand, will be separately discussed in the following.

#### 4.1. Language group effects on MMN amplitudes and latencies

The presented vowel duration contrasts elicited MMNs in all Pseudoword conditions, indicating that the vowel duration changes have been pre-attentively discriminated by both language groups. As indicated by smaller MMN amplitudes, word-final duration decrements thereby were generally more difficult to discriminate. Although the stimulus identification data only suggested a less clear vowel quantity categorization for the word-final vowel for the Germans, and the Germans indeed numerically showed a larger decrease of MMN amplitude here, the effect was observed for the Finnish participants as well, even though quantity oppositions frequently occur in word-final positions in the Finnish language. This might be related to findings that deviances in the initial part of a sound elicit larger MMNs than deviances in later parts (Grimm and Schröger, 2005; Sussman et al., 2004).

In contrast to previous studies, which compared vowel duration processing between speakers of genuine quantity and non-quantity languages (Nenonen et al., 2003, 2005; Ylinen et al., 2006), we did not find group effects on MMN amplitude. Group differences, however, were found for the MMN latency. Finns, for Pseudoword as well as Tone stimuli, showed shorter MMN latencies than Germans in all duration conditions, but notably not for deviants in Tone-Frequency. Shorter MMN latencies thereby might indicate faster or more automated pre-attentive processing of an auditory change. The MMN latency has been reported to decrease as a function of increased auditory discrimination performance (e.g., Näätänen et al., 1993; Novitski et al., 2004; Tiitinen et al., 1994) and, crucially, also in consequence of discrimination training for speech-sound duration (Menning et al., 2002). Importantly, it seems to be an especially sensitive indicator for differences in discrimination sensitivity at a high level of discrimination accuracy: In a study of Näätänen et al. (1993), improvements of behavioral discrimination performance after auditory discrimination training were in general accompanied by an increase in MMN amplitude and a decrease in MMN latency. Participants, however, who were already able to discriminate the changes before the training, showed no increase of MMN amplitude, but a decrease of MMN latency only. Thus, although the Finnish

---

\( ^3 \) The less typicality attributed to the stimuli by the Germans is most likely due to the fact that the recording of the primary stimulus from a Finnish speaker involved a for the German language untypical realization of the consonant /s/ in the word-initial position. However, this should be independent from the processing of the duration of subsequent vowels.
and the German phonological system feature a similar, duration-based, short–long phoneme distinction for the vowel /a/, a processing advantage for vowel duration changes in this vowel was observed for the Finns. Generally higher vowel duration discrimination sensitivity thus is suggested for the Finns, which can be explained by the increased long-term experience with vowel duration as a phonologically distinctive speech cue.

Shorter MMN latencies, however, could concurrently reflect a contribution – or an enhanced contribution – of phonetic processing to the analysis of vowel duration in the Finnish group. Although, by using pseudoword stimuli the influence of semantics is excluded, two different pseudowords are still distinguished based on their contained phonemes and their analysis therefore involves phonetic processing. Winkler et al. (1999) showed that phonetically distinctive, spectral vowel changes can elicit shorter MMN latencies than acoustically matched, non-distinctive, spectral vowel changes (for similar results see also Huotilainen et al., 2001; Peltola et al., 2005). The shorter latencies shown by the Finns therefore might secondly indicate increased phonetic processing elicited in the Finns compared to the Germans. Germans, on the other hand, might rather have processed the vowel duration changes in terms of word stress, as it is also indicated by the German participants’ subjective reports. Even though duration can be a distinctive cue for the vowel /a/ in German, vowel duration in the German language is in general more closely linked to signaling word-stress patterns. As processing of word stress should be expected to be a more time-consuming process, because it involves a more complex analysis of the word as a whole, the MMN elicited by word-stress deviants was correspondingly found to show a rather long latency (Honbo-lygo et al., 2004). Using a vowel phoneme for which mere duration changes in German would probably only be associated with word stress (as e.g., the vowel /i/), presumably would have resulted in more pronounced differences in vowel duration processing between the language groups, however, this would have introduced a confound in our design.

The results observed in the Tone condition correspond with the results reported by Tervaniemi et al. (2006), although we observed the group dissociation for Tone-Duration and Tone-Frequency processing for MMN latencies instead of MMN amplitudes. Our results thus support the assumption that the
different language experience with durational cues in the two languages is transferred to the non-speech domain. Even if phonetic processing might concurrently contribute to the shorter latencies for Finnish subjects for the processing of vowel duration changes, the shorter latencies for the processing of tone duration indeed indicate a language-specifically enhanced sensitivity for the discrimination of duration in the Finns.

4.2. Language group effects on MMN topography

Besides the differences in MMN latency, long-term language experience in the present study also had an influence on the lateralization of the MMN. The Finns, but not the Germans, showed an enhanced activation over (frontal) left-hemispheric scalp areas for the processing of the vowel duration deviants. Both groups, however, showed a right-lateralized MMN topography for the Tone stimuli, here showing an even stronger right-hemispheric lateralization for the Finnish subjects. In order to investigate the MMN generator structures in our data more closely, we calculated SCD maps that are given in Fig. 6. Although the increased frontal left-hemispheric lateralization of the MMN in our data could in principle also stem from an uncommonly oriented generator in the right hemisphere (as it might be proposed by the rather strong posterior activation in the right hemisphere shown by the Finnish subjects in the Pseudoword conditions), this would be in contrast to the remaining MMN literature and is thus rather unlikely, especially since no statistical differences have been found for the activation at posterior areas. The SCDs indicated typical MMN generator structures with bipolar generators separately in each hemisphere for both groups and stimulus types and also reflected the structural difference between the groups in the generator activity over frontal areas for the Pseudoword stimuli. Interestingly, SCDs for the Germans thereby, compared to those for the Finns, showed a stronger positive MMN generator over left-posterior areas for Pseudoword as well as Tone stimuli, which, however, here reflected a general group difference. The lateralization difference found between the groups for the MMN in the Pseudoword conditions, on the other hand, is specific for vowel duration processing and might reflect the different linguistic functions associated with vowel duration in the two languages.

Several studies suggested that the increased left-hemispheric activation often observed for the processing of speech stimuli might depend on the phonetic distinctiveness of a speech-sound contrast in the native language. Enhanced left-hemispheric activity was found for phonetically distinctive speech-sound contrasts compared to phonetically non-distinctive speech-sound contrasts (Kasai et al., 2001; Minagawa-Kawai et al., 2005; Sittiprapaporn et al., 2005), or compared to equivalent contrasts in non-speech stimuli (also Kasai et al., 2001). Importantly, a left-hemispheric shift has as well been shown for the MMNm elicited by phonetically distinctive in contrast to phonetically non-distinctive vowel duration contrasts (Takegata et al., 2004). However, here the influence of the native language phonological system seems to be crucial: Increased left-hemispheric activation for vowel duration processing was found in native speakers of a quantity language in contrast to speakers of a non-quantity language (Gandour et al., 2002), but also in contrast to high-proficiency L2 speakers of the quantity language (Minagawa-Kawai et al., 2005). Recent results showed that also native infants only showed a left-hemispheric dominance for phonetically distinctive duration contrasts after the age of twelve months (Minagawa-Kawai et al., 2007), i.e. the age in which speech-sound discrimination starts to be language-specific (Kuhl et al., 1992; Cheour et al., 1998). Since enhanced left-hemispheric activation thus can be associated with phonetic processing, we presume the group difference in the lateralization of the MMN for the vowel duration contrasts to reflect increased phonetic processing of vowel duration contrasts for the Finns in contrast to the Germans.

Processing of linguistic prosody, as it possibly occurred in the German group, in contrast has usually been attributed to right-hemispheric brain areas (Gandour et al., 2003, 2004; see also Friederici and Alter, 2004). However, only few MMN studies addressed the processing of word stress so far. Whereas no hemispheric dominance was found for the processing of word-stress deviants in infants (Weber et al., 2004), a right-hemispheric preponderance of the MMN on word stress has been observed in adults (Honbolygo et al., 2004). A right-hemispheric MMN thus might support the notion of a processing of the deviants in terms of word stress for the Germans. However, in general it here indicates that less phonetic processing was activated for the vowel duration contrasts in the German group compared to the Finnish group in our study. Taken together, the lateralization difference found between the groups for the Pseudoword, but not the Tone stimuli, reflects the different functional roles (phonetic vs. prosodic) of vowel duration as a linguistic cue in the two languages as a result of long-term linguistic experience.

5. Conclusion

The present study showed differences in the pre-attentive processing of vowel duration between Finns and Germans as a consequence of the long-term experience with the native language. Generally, the discrimination of vowel duration thereby was more difficult for a decrement of duration in a word-final position. Shorter MMN latencies for vowel duration and tone duration contrasts for the Finns, however, indicated an enhanced sensitivity to duration contrasts for the Finns due to the frequent exposure to duration as a phonologically distinctive cue in the native language. On the other hand, differences in the MMN lateralization between the groups reflected the different role of vowel duration in the two languages: The left-hemispheric shift of MMN activation observed specifically for vowel duration contrasts for the Finns, suggested that these changes elicited increased phonetic processing in the Finns in contrast to the Germans.

Acknowledgements

Parts of these data were acquired during a stay of Ursula Kirmse at the Cognitive Brain Research Unit in Helsinki,
Finland. We wish to thank the members of the CB RU for their support in various matters at this time. We further thank our anonymous reviewers for their very helpful comments on earlier versions of this manuscript.

References
Alho, K., 1995. Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNm) elicited by sound changes. Ear Hear. 16 (1), 38–51.


Näätänen, R., 2001. The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). Psychophysiology 38 (1), 1–21.


