Cortical representations for phonological quantity

Sari Ylinen

Academic dissertation to be publicly discussed,

by due permission of the Faculty of Behavioural Sciences

at the University of Helsinki in auditorium XII

on the 7th of June, 2006, at 12 o'clock

Cognitive Brain Research Unit

Department of Psychology

University of Helsinki

Finland

Helsinki 2006

Supervised by

Doc. Minna Huotilainen
Cognitive Brain Research Unit
Department of Psychology
University of Helsinki, Finland

Dr. Viola de Silva
Department of Languages
University of Jyväskylä, Finland

Reviewed by

Assistant Professor Katsura Aoyama
Department of Speech, Language, and Hearing Sciences
Texas Tech University Health Sciences Center, USA

Professor Patricia T. Michie
School of Psychology
University of Newcastle, Australia

Contents

Abstract

Different languages use temporal speech cues in different linguistic functions. In Finnish, speech-sound duration is used as the primary cue for the phonological quantity distinction — i.e., a distinction between short and long phonemes. For the second-language (L2) learners of Finnish, quantity is often difficult to master if speech-sound duration plays a less important role in the phonology of their native language (L1). The present studies aimed to investigate the cortical representations for phonological quantity in native speakers and L2 users of Finnish by using behavioral and electrophysiological methods. Since long-term memory representations for different speech units have been previously shown to participate in the elicitation of the mismatch negativity (MMN) brain response, MMN was used to compare the neural representation for quantity between native speakers and L2 users of Finnish.

The results of the studies suggested that native Finnish speakers' MMN response to quantity was determined by the activation of native-language phonetic prototypes rather than by phoneme boundaries. In addition, native speakers seemed to process phoneme quantity and quality independently from each other by separate brain representations. The cross-linguistic MMN studies revealed that, in native speakers of Finnish, the MMN response to duration or quantity-degree changes was enhanced in amplitude selectively in speech sounds, whereas this pattern was not observed in L2 users. Native speakers' MMN enhancement is suggested to be due to the pre-attentive activation of L1 prototypes for quantity. In L2 users, the activation of L2 prototypes or other L2 learning effects were not reflected in the MMN, with one exception. Even though L2 users failed to show native-like brain responses to duration changes in a vowel that was similar in L1 and L2, their duration MMN response was native-like for an L2 vowel with no counterpart in L1. Thus, the pre-attentive activation of L2 users' representations was determined by the degree of similarity of L2 sounds to L1 sounds. In addition, behavioral experiments suggested that the establishment of representations for L2 quantity may require several years of language exposure.

4

Acknowledgements

Abbreviations

Ag/AgCl	silver/silverchloride
ANOVA	analysis of variance
C	consonant
DAT	digital audio tape
EEG	electroencephalography, electroencephalogram
EOG	electro-oculogram
ERP	event-related potential
fMRI	functional magnetic resonance imaging
F0	fundamental frequency
Hz	Hertz
ISI	interstimulus interval
L1	native language
L2	second language
MEG	magnetoencephalography
MMN	mismatch negativity
MMNm	magnetic mismatch negativity
PAM	Perceptual Assimilation Model
PET	positron emission tomography
RT	reaction time
SD	standard deviation
SOA	stimulus onset asynchrony
SLM	Speech Learning Model
2AFC	two-alternative, forced-choice task
V	vowel

List of original publications

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

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

Study III. Ylinen, S., Shestakova, A., Alku, P., & Huotilainen, M. (2005). The perception of phonological quantity based on durational cues by native speakers, second-language users and nonspeakers of Finnish. *Language and Speech, 48,* 313–338.

Study IV. 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 Research, 1072,* 175–185.

Study V. Ylinen, S., Huotilainen, M., & Näätänen, R. (2005). Phoneme quality and quantity are processed independently in the human brain. *NeuroReport, 16,* 1857–1860.

1. Introduction

For decades, there has been a great deal of debate about how speech is perceived. What is generally agreed, however, is that the flow of speech must be mapped onto some kinds of long-term memory representations (i.e., categories), irrespective of whether this occurs at the level of phonetic features, phonetic segments, syllables, or words (for a review, see Miller & Eimas, 1995). Thus, the acquisition of a phonological system involves the establishment of these representations. The native-language (L1) phonetic categories are established during the first year of life, and serve as building blocks for the further language acquisition (see e.g., Kuhl, 2004, for a review). It seems that, later in life, L1 categories affect the learning of a second language (L2) (e.g., Best, 1994; Flege, 1995; Trubetzkoy, 1939/1969). Yet, to become a competent and fluent L2 user, one must be able to effortlessly map the L2 sounds onto their categories. Sometimes this may require the establishment of new categories for L2 phonemes that may even be cued by phonetic features that are not used contrastively in L1. This is the case for many L2 learners of Finnish because the Finnish phonological system includes quantity, that is, phonological distinction between short and long segments that is primarily cued by duration of the segments. In contrast, in the phonological systems of many other languages, the role of duration is less prominent. The present set of studies addressed the cortical representation of speech-sound duration and phonological quantity in native speakers and L2 users of Finnish. Russian speakers were chosen as the L2 group because the Russian language uses duration cues differently from Finnish. Studying this issue on Russian L2 learners also had practical interest, since Russians are one of the largest immigrant groups in Finland and, consequently, one of the largest groups learning Finnish as their L2.

In the present set of studies, both behavioral and electrophysiological methods were used to investigate the representation of quantity. Research with behavioral methods, such as identification, discrimination, and reaction-time (RT) measurements, has laid the foundation on our understanding of speech perception. However, since behavioral responses require decision making that can be based on different cognitive strategies, it is

sometimes difficult to determine to what extent behavioral methods tap the neural organization of the perceptual space underlying the recognition of speech (see e.g., Massaro, 1987; Schouten et al., 2003). The development of brain-research methods during the past decades has shed some more light on the neural mechanisms of speech perception. On one hand, functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) have allowed to localize brain activity elicited by speech stimuli. On the other hand, event-related potentials (ERP) measured with electroencephalography (EEG) and evoked magnetic fields measured with magnetoencephalography (MEG) have enabled one to measure brain activity with a timescale of milliseconds. In addition to the excellent temporal resolution of the electromagnetic brain-research methods, a component of ERP, the mismatch negativity (MMN), provides further benefits for research on the perception of speech sounds. The MMN reflects the accuracy of sensory-memory encoding in auditory change detection. The MMN does not only reflect the processing of the acoustic features, however, but is also modified by the familiarity of speech stimuli, suggesting that the long-term memory representations for phonemes participate in the MMN generation. Therefore, in addition to behavioral methods, the MMN was used in the current thesis to index the representation of duration and quantity in the brain.

The current thesis will proceed to introduce the establishment of the L1 and L2 categories and the perceptual effects related to them in Chapter 2 and the role of speech-sound duration in the Finnish and Russian languages in Chapter 3. Chapter 4 deals with the ERPs and the MMN that are used to determine the brain representation of duration and quantity in the electrophysiological experiments. The results of the experiments are reported and shortly discussed in Chapter 5, which is followed by a general discussion and conclusions in Chapters 6 and 7.

2. Representations for native- and second-language phonetic features and phonemes

Although young infants can discriminate various phonetic contrasts regardless of their ambient language (for reviews, see e.g., Aslin et al., 1998; Jusczyk, 1997), their discrimination sensitivity to non-native contrasts decreases by the end of the first year of life (Cheour et al., 1998; Polka & Werker, 1994; Werker & Lalonde, 1988; Werker & Tees, 1984). On the other hand, the discrimination of some phonetic contrasts seems to improve with L1 experience (Aslin et al., 1981; Polka et al., 2001; Sundara et al., in press). Evidently, this perceptual reorganization is caused by the ambient language. A probable account for the changes in infant speech-sound discrimination is statistical learning. Maye et al. (2002) have demonstrated that 6- and 8-month-old infants are sensitive to the distributional frequencies of the L1 sounds. Their results suggested that infants exposed to speech-sound distribution with two frequent stimuli were able to discriminate the target stimuli. However, when infants were exposed to distribution with only one frequent stimulus, their discrimination sensitivity was reduced. Similarly, Jusczyk et al. (1994) have shown that 9-month-old infants are sensitive to the frequency of phonotactic patterns as suggested by their preference for the patterns occurring with high probability in their L1. Thus, the statistical properties of speech input are likely to cause the tuning of the infant's perceptual space to match the features that are essential for the L1 phonological system; the infant becomes "neurally committed" to the L1 patterns (Kuhl, 2000, p. 11855; 2004, p. 831).

Since infants are sensitive to the statistical properties of speech and they hear typical and good representatives of the L1 phonemes most frequently, the phonetic categories formed in the infant brain may be based on phonetic prototypes. The view that prototypes are essential in the development of speech perception is supported by the observation that in 6-month-old infants, the sensitivity of vowel discrimination is constrained by the L1 prototypes but not by the non-native prototypes (Kuhl et al., 1992). These prototypical representations may be utilized as a basis for word learning and speech production during the second year of life (e.g., Kuhl, 2000).

The phonetic categories established for the L1 enable the fast recognition of speech sounds, which is required for the comprehension of natural fluent speech. Early experimental findings on these categories by Liberman et al. (1957) suggested that the discrimination of speech sounds is more accurate between the phonetic categories when a phonetic boundary is crossed than within a category. This phenomenon known as *categorical perception* or *phoneme-boundary effect*[1] was thought to suggest that categorization is based on phonemic labels. However, further studies indicated that even infants show the same effect regardless of the fact that they lack the knowledge of any particular phonological system. These results were interpreted as providing evidence for an innate processing mechanism specialized for speech (Eimas et al., 1971). These two views were later challenged by Kuhl and colleagues (Kuhl, 1981; Kuhl & Miller, 1975; Kuhl & Padden, 1983) who demonstrated that non-human animals show a phoneme-boundary effect for speech sounds, suggesting that instead of the real phonetic categories, the effect may sometimes be due to nonlinearity in auditory processing. Kuhl and colleagues (Kuhl, 1981; Kuhl & Miller, 1975; Kuhl & Padden, 1983) proposed that phonetic categories have evolved on the basis of such auditory nonlinearities. In addition, an effect resembling categorical perception was found with non-speech sounds in humans (Miller et al., 1976; Pisoni, 1977). Still, not all phoneme-boundary effects can be explained by these nonlinearities, because different languages have different phonetic features, and the effects are language-specific (e.g., Miyawaki et al., 1975). Thus, the category-related perceptual effects must be primarily determined by language experience.

---

[1] According to the definition of categorical perception by Studdert-Kennedy et al. (1970; see also Liberman et al., 1957), perception is categorical if discrimination performance is determined by the ability to categorize stimuli. That is, discrimination sensitivity should be enhanced at the phoneme boundary, but within-category discrimination should be at a chance level. However, the criterion of the chance-level within-category discrimination has usually not been met in perceptual experiments on vowels and some consonant features, suggesting that their discrimination is constrained but not entirely determined by categorization (for a review, see Strange, 1999). Some authors, such as Wood (1976) and Iverson & Kuhl (2000) have used the term phoneme-boundary effect to refer to a discrimination peak at the boundary and reduced, but not chance-level, sensitivity within the category. See Harnad (1987) for a review on categorical perception.

The basic assumption of categorical perception and phoneme-boundary effect, as the latter term reveals, is that the phoneme boundaries determine the perception of speech sounds: a sound represents a phoneme if it occurs within the boundaries of that phoneme in the perceptual space (see e.g., Liberman et al., 1957). However, more recent studies (Iverson & Kuhl, 1996; Kuhl, 1991; Miller et al., 1983; Samuel, 1982) have demonstrated that phonetic categories are internally organized according to the typicality of instances within a category. On this basis, an alternative hypothesis on the mechanism underlying phonetic categorization has emerged, namely, that the mapping of speech sounds onto categories may be based on the phonetic prototypes rather than boundaries. Kuhl (1994) has suggested that the prototypes "warp" the perceptual space by "pulling" input sounds like magnets (p. 812–813) and thus reducing the discernibility between a prototypical stimulus and other category members (*the magnet effect*) (see also Kuhl, 1991). As a result, similar sounds within the category are assimilated to the prototype, which enables the mapping of variable speech stimuli onto long-term memory representations. Conversely, the discernibility between speech sounds representing different phonetic categories is enhanced, because they are "pulled" apart by different magnets. Thus, even though the two hypotheses on speech-sound mapping approach the issue from the opposite sides ― one addresses phoneme boundaries and the other their prototypical centers ― they predict similar perceptual phenomena at the phoneme boundary.

Whatever the underlying mechanism may be, once the representations for the L1 phonemes are established, they affect the perception of not only the L1 sounds, but also the L2 sounds (e.g., Trubetzkoy, 1939/1969). In order to become a competent L2 user, an L2 learner must modify the phonological system to account for the new features and possibly establish new categories for those sounds that cannot be mapped onto the L1 categories. Several models on the perception of the L2 speech sounds and their relation to the L1 categories have been proposed. For instance, Flege's (1995) Speech Learning Model (SLM) makes predictions about the establishment of new phonetic categories based on similarities between the L1 and L2 phonemes. SLM suggests that new phonetic categories are more likely to be established for those L2 sounds that are discernibly dissimilar from any L1 phoneme than for L2 sounds that have similar L1 counterparts.

The reason for this is that there may be no need to establish a new category for an L2 sound that resembles an L1 sound closely enough for categorization via this L1 category. Thus, SLM predicts that the perceived similarity between the L2 and L1 sounds determines the probability of the category establishment: new phonetic categories are more likely established for "*new*" or "*dissimilar*" L2 sounds than for "*similar*" L2 sounds (Flege, 1992; 1995).

SLM does not, however, make explicit predictions about the perception of non-native phonetic contrasts, as does Best's (1994) Perceptual Assimilation Model (PAM). PAM addresses how L2 sounds are perceptually assimilated into L1 categories. It suggests that if two L2 sounds are perceived as instances of two L1 categories, then the perception is excellent. If they are perceived as instances of a single category but with a different goodness of fit to the prototype, the discrimination is relatively good, but not as good as with the two-category case. A poor accuracy of discrimination is suggested, however, if L2 sounds are perceived as instances of a single category with an equal goodness of fit to the prototype. PAM is based on the assumption that the speech sounds are recognized by mapping them onto the representations of articulatory gestures. In contrast, Kuhl (e.g., 2000) has suggested that speech perception is based on general auditory mechanisms. According to her Native Language Magnet model, the perception of the L2 speech sounds is determined by the L1 phonetic prototypes that warp the perceptual space. In addition, in her Native Language Neural Commitment model, Kuhl (2004) has proposed that "language learning produces dedicated neural networks that code the patterns of native language speech" that interfere with the processing of the L2 sounds and patterns (p. 838). Even though the emphasis of these models is different, they agree with each other on the fact that the extent of the assimilation of the L2 sounds to the L1 categories affects the perception of the L2 sounds.

Since most phonological distinctions are cued by the spectral features, duration processing is explicitly addressed only in a few models on L2 acquisition. Bohn (1995) has discussed the accessibility of the duration cues compared with that of the spectral cues in his Desensitization Hypothesis. He proposed that if listeners are not sensitized to

the spectral differences of the non-native vowels due to the lack of that particular contrast in their L1 and if, at the same time, there are duration differences between the vowels, the listeners tend to utilize the duration cues. According to this hypothesis, duration cues are easily accessible regardless of the role of the duration in the L1. Findings by McAllister et al. (2002) did not, however, support this hypothesis. These authors recruited three groups of L2 users of Swedish with L1s (Spanish, English, and Estonian) that used the duration cue differently. The results suggested that the L2 users' performance in the Swedish quantity contrast varied as a function of the relevance of the duration in the L1. Based on these results, the authors formulated their Feature Prominence Hypothesis, suggesting that the L2 phonological contrasts involving those L2 features that are not used in a phonological distinction in the L1 are more difficult to acquire than those involving features relevant for the L1.

3. Speech-sound duration in Finnish and Russian

In Finnish, quantity degrees can separate both the lexical meanings of words (e.g., /puro/ 'brook' *vs.* /puːro/ 'porridge'; /mɑto/ 'worm' *vs.* /mɑtːo/ 'carpet') and their grammatical functions (e.g., /tɑlon/ 'of a house' *vs.* /tɑloːn/ 'into a house'). All vowels[2] can be either short or long in all positions of the word. In addition, all consonants other than /d, h, j, ʋ/ can appear as short or long within a word, excluding the consonant sequences (however, /p, t, k, s/ can appear as long after a nasal consonant or /r, l/). According to the speech-sound statistics, the probability of occurrence of the long segments is about 10% of all sounds (Aoyama, 2001; Vainio, 1996). This illustrates the importance of the correct perception and production of quantity for the comprehension and intelligibility in Finnish.

---

[2] The Finnish vowel system has eight phonemes: /ɑ, o, i, e, u, y, æ, ø/. According to the St-Petersburg school of phonetics, the Russian vowel system has six phonemes: /ɑ, o, i, e, u, ɨ/ (/o, e/ occuring in a stressed position only) (Bondarko, 1998). According to the Moscow school of phonetics, [i] and [ɨ] are considered allophones of /i/ (Avanesov, 1984).

At the acoustical level, the primary cue of the Finnish quantity is sound duration. However, speech-sound durations do not contribute to quantity distinctions only. As in all languages, various factors, such as the speaking rate, the intrinsic durations of the sounds, the adjacent sounds, the word structure, the word length, the position of a sound in a word and utterance, the prominence and stress relations as well as the speaker-dependent factors, can affect sound durations in Finnish (Iivonen 1974a; 1974b; Lehiste, 1970; Lehtonen, 1970; Marjomaa, 1982; Suomi et al., 2003; Suomi & Ylitalo, 2004; Wiik, 1965; for a review, see e.g., Wiik, 1981). Therefore, it is apparent that quantity can be determined by no absolute values of duration, but rather by the relative duration of the sounds in the speech context (e.g., Lehtonen, 1970). In fact, short sounds produced at a slow speaking rate may even be longer in the absolute duration than long sounds produced at a very fast rate (Marjomaa, 1982).

In addition to the duration cues, slight spectral differences between the Finnish short and long vowels have been found in acoustic measurements. The formant values of the short vowels tend to be more neutralized than those of the long ones, probably because articulators have more time to reach the peripheral target position during the long vowels than during the short ones (Wiik, 1965). To determine the significance of the spectral cues in quantity categorization, O'Dell (2003) presented Finnish listeners with stimuli from two stimulus continua that were modified in duration. The stimuli were of different duration, but in one continuum, the stimuli carried spectral cues of originally short vowels and in the other continuum, spectral cues of originally long vowels. When the duration cue was ambiguous in the middle of the continua, the phoneme boundary was shifted to the direction predicted on the basis of the spectral cues. However, at the endpoints of the continua, the duration cue appeared to determine the categorization response. In addition to the vowel quality, O'Dell's (2003) data suggested that the movement of the fundamental frequency (F0) may serve as a secondary cue for quantity and, thus, slightly shift the phoneme boundary, but this was observed in the middle of the stimulus continua only, that is, when the duration cue was ambiguous. Thus, even though other cues than duration may affect the quantity categorization in ambiguous cases that

are rare in the native speakers' speech, duration is the primary cue of the Finnish quantity and has an important role in the Finnish phonological system (e.g., Lehtonen, 1970).

The recognition of the non-native phonological contrasts may be difficult if they involve features that are not used phonologically in the listener's L1 (McAllister et al., 2002). Cross-linguistic studies on the categorization of the Finnish quantity by non-native listeners support this view. According to Aoyama (2001), the native speakers of another language that has quantity distinctions, Japanese, were able to perceive Finnish short and long consonants correctly with no prior knowledge of Finnish (see Isei-Jaakkola, 2004, for contrastive comparison of Finnish and Japanese). However, non-native speakers were less successful in the studies of Vihanta (1987; 1990), who investigated the perception and production of the quantity distinctions in native French foreign-language learners of Finnish. The main findings were that the French have difficulties especially in recognizing long vowels in word-final position and in producing short vowels in that position. This was suggested to be due to the French transfer: word-final vowels may be perceived in terms of stressed *vs*. unstressed instead of short *vs*. long distinction, since in French, word stress that is cued by longer duration is associated with word-final vowel. In addition, French speakers tended to hear Finnish short consonants as long and produce long consonants that were not long enough in duration.

Similarly to French, Russian is a language that uses speech-sound duration as a cue for word stress and thus uses duration differently than the languages that have quantity contrasts. In Russian, the stressed vowel is longer in duration than the other, unstressed vowels of a word. For example, vowel /ɑ/ has three allophones, durations of which are typically different: the stressed allophone is longer and the two unstressed allophones (the 1st and 2nd degrees of reduction, occurring in different positions of a word) are shorter in duration in ratio of 1:0.5:0.25, respectively (Bondarko, 1998). However, due to their shorter intrinsic duration, the unstressed allophones [u, i, ɨ] are characterized by the 1st degree of reduction only. That is, for the allophones of /u, i, ɨ/ the ratio is 1:0.5 (Bondarko, 1998). The unstressed vowels, especially /ɑ/, are also reduced qualitatively

due to timing constraints (Bondarko, 1998; Verbickaya, 1976). Thus, the word stress affects both the durational and qualitative characteristics of the Russian vowels, but a longer sound duration is considered the most reliable cue of stress (Bondarko, 1998; Verbickaya, 1976). In Russian, stress is lexically determined, and it has a phonologically distinctive role (e.g., /'mukʌ/ 'suffering' *vs.* /mu'kɑ/ 'flour'), whereas in Finnish, stress is associated with the first syllable of the word. Given that, in Finnish, quantity is independent of word stress, the different use of the duration cue may cause confusion for Russians listening to Finnish. Non-Finnish-speaking Russians may perceive the unstressed long vowels in Finnish as stressed based on the vowel duration (de Silva, 1999).

4. Auditory ERPs and MMN as tools of investigating the perception of speech features

4.1. ERPs reflecting acoustic features

Auditory ERPs are electrophysiological responses caused by, and time-locked to, acoustic events. They can be non-invasively recorded from the scalp using the EEG. The long-latency auditory ERPs start with the obligatory (exogenous) components that reflect the transient detection of the physical features of stimulus. In adults, these components are the P1 (P50), N1, P2, and N2. The P1 and N1 peak at about 50 and 100 ms from stimulus onset, respectively. The P2 peaks at 175–200 ms and, depending on stimulus duration, may be followed by the N2 and the sustained potential (for a review, see Näätänen, 1992; see also Kushnerenko et al., 2001).

The N1 is suggested to reflect the transient detection of change in the level of sound energy, such as stimulus onsets (and offsets for stimuli with over 0.5 s of duration) (Näätänen & Picton, 1987). Thus, the N1 amplitude decreases with decreasing stimulus intensity. In addition, the N1 reflects the tonotopical organization of the auditory cortex (e.g., Pantev et al., 1988; 1995; Romani et al., 1982). The N1 amplitude is also modulated by the auditory environment. When a stimulus is repeated, the N1 elicited by the

repetitions is diminished in amplitude in comparison with that elicited by the first stimulus in the stimulus train due to the refractoriness of the neural populations.[3] Depending on the similarity of the stimulus features, a deviant stimulus presented in the stimulus train may again elicit a larger N1 owing to the activation of fresh, non-refracted neural populations in addition to (or instead of) those refracted populations that are tuned to the repeating stimulus (Näätänen & Picton, 1987; Näätänen et al., 1988). The refractoriness is affected by the stimulation rate: the N1 is larger for rarely occurring sounds than for frequently occurring sounds. According to Imada et al. (1997), both onset-to-onset (SOA) and offset-to-onset (ISI) time periods affect the N1 amplitude, but the silent period before the stimulus has a stronger effect.

Some studies have suggested that obligatory responses can be modified by experience on sounds. Pantev et al. (1998) disclosed that the N1 elicited by piano tones was enhanced in amplitude in musicians in comparison with that elicited in control subjects, whereas no enhancement was found for pure tones. Tremblay et al. (2001) reported that the amplitude of the N1-P2 complex reflects the improvement of speech-sound recognition as a result of training. However, more recent data by Sheehan et al. (2005) challenged this view. Sheehan et al. (2005) suggested that the P2 enhancement may be due to an inhibitory process towards responses that are elicited by repetitive stimulation and have no relevance to the individual. Thus, so far, it is controversial whether other obligatory components than the N1 can be modified by experience. Even though Pantev et al. (1998) consider their findings on N1 enhancement as reflecting the cortical representation of auditory stimuli, Näätänen and Winkler (1999) have argued that the process underlying the N1 elicitation reflects feature encoding that does not fulfill the criteria of the neural substrate of stimulus representation, because this feature encoding does not result in a complete, functionally integrated representation that corresponds to conscious perception. Rather, in their view, the integrated representations of auditory events are reflected by the MMN component of ERP.

---

[3] Here the term refractoriness does not refer to the refractory period of single neurons following the generation of an action potential, but rather to stimulus-specific refractoriness of complex neuronal circuits (Näätänen & Picton, 1987).

4.2. MMN as index of change detection

The MMN (Näätänen et al., 1978) is an ERP component elicited by a discriminable change (e.g., change of frequency, duration, intensity, location, or pattern) in a regular stimulus stream of speech or non-speech sounds (for reviews, see Näätänen, 2001; Picton et al., 2000). It peaks at 100–250 ms after change onset and is characterized by a fronto-central scalp distribution and inverted polarity at the mastoids (with nose reference) as a result of the orientation of its generators, namely, a bilateral temporal source at the superior temporal gyri and a frontal source in or near the right inferior frontal gyrus (Giard et al. 1990; Opitz et al., 2002; Rinne et al., 2000; 2005). Typically, the MMN is elicited in an oddball paradigm that contains a repetitive, *standard* stimulus and occasional *deviant* stimuli randomly occurring at a low probability. According to Näätänen (1990), the repetitive standard stimuli form and maintain a sensory-memory representation of the stimulus features. The MMN is elicited by the difference between the perceived features of a deviant stimulus and the features stored in the representation of the standard stimuli (see also Näätänen et al., 2005; Schröger, 1997). However, Winkler et al. (1996; 2001) have demonstrated that the existence of a sensory-memory representation is a necessary but not sufficient condition of MMN elicitation. Rather, it requires the detection of regularities in auditory input and a violation of extrapolations based on these regularities. The MMN amplitude and latency correlate with perception in behavioral discrimination tasks (e.g., Lang et al., 1990; Näätänen et al., 1993; Tiitinen et al., 1994). Generally, the MMN amplitude increases when the acoustic discrepancy between the deviant and standard stimuli increases, which also facilitates behavioral discrimination. However, behavioral tasks require attentive processing and decision making, whereas the MMN is elicited even in the absence of attention. Thus, the MMN reflects automatic, pre-attentive change detection in auditory stimulation.

Since the first studies by Kaukoranta et al. (1989), Joutsiniemi et al. (1998), and Näätänen et al. (1989), it has been consistently found that the MMN and its magnetic equivalent (MMNm) are elicited by sound-duration changes. A large range of duration changes from ten milliseconds up to ca 1 second can elicit an MMN (Amenedo & Escera,

2000; Näätänen et al., 2004). Equally to other MMNs, also the duration MMN correlates with behavioral duration discrimination (Amenedo & Escera, 2000; Jaramillo et al., 2000). Since it has been suggested that the amplitude and latency of the duration MMN are relatively well-replicable in a test–retest setting, the duration MMN is a useful tool in the study of the cognitive functions of different clinical populations (Tervaniemi et al., 1999). For example, the duration MMN has been used to determine the accuracy of sensory-memory representation in individuals with schizophrenia (see Michie, 2001, for a review), and dyslexia (for reviews, see Kujala & Näätänen, 2001; Lyytinen et al., 2004). The duration MMN has been utilized also in studies addressing speech *vs*. non-speech processing in isolated vowels or syllables (e.g., Jaramillo et al., 1999; 2001; Takegata et al., 2004) as well as in words and pseudowords (e.g., Inouchi et al., 2003; Korpilahti et al., 2001; Sussman et al., 2004). In addition, Menning et al. (2002) have used the duration MMN as an index of neural changes occurring as a result of speech-perception training.

When measuring the duration MMN, one should note some methodological aspects specific to duration changes. The MMN peak is usually measured from a difference signal obtained by subtracting the response to standard stimulus from that to deviant stimulus. With the duration MMN, however, the subtraction of the ERP responses to standard and deviant stimuli with different physical timing may distort the difference signal, because sound continuation or termination considerably affects the obligatory responses (e.g., Jacobsen & Schröger, 2003; Kushnerenko et al., 2001). This may result in the underestimation of the MMN with duration decrements and the overestimation of the MMN with duration increments (Jacobsen & Schröger, 2003). One possible solution to this problem is to reverse the roles of the standards and deviants in a separate block and to use these reversed-condition responses (or responses from some other control design) for subtraction (Jacobsen & Schröger, 2003). An inevitable consequence of the use of the reversed-condition blocks is that either the SOA or the ISI is different between the oddball and reversed-condition blocks. The presentation rate does not significantly affect the MMN amplitude (Näätänen et al., 1987; Schröger & Winkler, 1995) (with the exception of the intensity MMN; see Schröger, 1996), but both the SOA and the ISI affect the refractoriness state of obligatory responses (Imada et al., 1997). This, in turn, is

reflected in difference signals. However, if the duration change eliciting the MMN occurs at around or after 100 ms from stimulus onset, it is unlikely that the refractoriness effects caused by the presentation rate could distort the MMN, since they are observed only in the P1, N1, and P2 components (Imada et al., 1997; Jacobsen & Schröger, 2003) that thus would not overlap with the MMN peak.

Another possible confound for the interpretation of duration MMN is the perceived loudness, since shorter sounds may sound softer than longer sounds with equal intensities due to their smaller amount of physical energy (Munson, 1947). This loudness integration is one example of temporal integration that applies to sounds falling within the time window of ~200 ms (e.g., Näätänen, 1992). As a result, the duration MMN elicited with stimuli falling within this window could partly be contaminated by the intensity MMN. However, in a study addressing this issue, Todd and Michie (2000) found no significant contribution of the intensity cues to the duration MMN, even though sound durations falling within the temporal window of integration (50 ms *vs.* 125 ms) were used.

4.3. MMN as index of the long-term memory representations for phonemes

The MMN is a useful tool for language studies because the MMN elicited by speech stimuli does not only reflect acoustical discrepancy between the standard and the deviant stimuli as with most non-speech sounds but is modified by the long-term memory representations for speech features. This was first demonstrated by Näätänen et al. (1997) with vowels and Dehaene-Lambertz (1997) with consonants. In Näätänen et al. (1997), the Finnish vowel [ø] elicited a larger-amplitude MMN than the Estonian vowel [ɤ] in native speakers of Finnish regardless of the fact that [ɤ] was acoustically more deviant to the standard than [ø]. Since a reversed pattern was found in the native speakers of Estonian, the effect observed in the native speakers of Finnish was clearly language-specific. With the MEG, the source of the phonetic enhancement of the MMN was localized on the left auditory cortex (see also Alho et al., 1998; Shtyrov et al. 1998; 2000). The language-specific MMN enhancement caused by native-language phoneme

changes was also found by Dehaene-Lambertz (1997) who studied MMNs elicited by native- and non-native across- and within-category phonetic contrasts in consonants. It was found that the native across-category change elicited a larger-amplitude MMN than did the native within-category change. In contrast, for non-native changes, no MMN was elicited.

Several studies have corroborated the first evidence presented by Näätänen et al. (1997) and Dehaene-Lambertz (1997) that the MMN is affected by language experience and long-term memory. For instance, Winkler et al. (1999b) presented orthogonal across- and within-category phonetic contrasts to two language groups and found that the MMN was larger in amplitude for the native-language across-category change in both language groups. Further, MMNs elicited by across- and within-category changes in voice onset time (VOT) were compared with each other by Sharma and Dorman (1999). Their results suggested that an across-category change elicited an MMN, whereas a within-category change did not. In another related study, Sharma and Dorman (2000) used syllables starting with pre-voiced stop consonants (-10 and -50 ms VOT). Behaviorally, native English listeners categorized both stimuli as /ba/, whereas Hindi listeners categorized the former as /pa/ and the latter as /ba/. A significant MMN was only observed in the Hindi listeners, whose category boundary was crossed. Another important demonstration of the long-term memory contribution to the MMN was provided by Phillips et al. (2000) and Shestakova et al. (2002) who suggested that the MMNm is also elicited by acoustically varying speech stimuli that are unlikely to form a memory representation needed in MMN(m) elicitation unless they are processed categorically. In addition to phonemes, an MMN enhancement has recently been found for the L1 syllable structure (Dehaene-Lambertz et al., 2000) and for L1 words relative to pseudowords (Pulvermüller et al., 2001; 2004; Shtyrov & Pulvermüller, 2002).

The language-specific MMN enhancement has also been observed for L2 features. Winkler et al. (1999a) demonstrated that adult second-language users can establish long-term memory representations that are pre-attentively activated and determine the MMN amplitude similarly as those of the native speakers (cf. Peltola et al., 2003, for reduction

instead of enhancement of the MMN amplitude in L2 learners). Furthermore, Cheour et al. (2002) have reported that in 3–6-year-old children, foreign-language learning is reflected in the MMN elicited by foreign-language speech sounds as rapidly as within two months of exposure to the target language (see also Shestakova et al., 2003). Thus, the MMN can be used as a tool to probe the effects of L2 learning at the neural level.

Not all attempts to demonstrate language-specific effects with the MMN have been successful, however. The MMN studies (Aaltonen et al., 1992; Maiste et al., 1995; Sams et al., 1990; Sharma et al., 1993) preceding Dehaene-Lambertz (1997) had repeatedly aimed at revealing the neural correlates of categorical perception, but failed to do so. The MMN response patterns were not significantly different between across-category and within-category changes. This raises the question as to whether crossing the boundary between two phonetic categories is, after all, the critical factor in eliciting the effect at the neural level. Since the phonetic categories are internally organized according to the typicality of instances within a category (Iverson & Kuhl, 1996; Kuhl, 1991; Miller et al., 1983; Samuel, 1982), the extent of typicality may sometimes account for the category-related effects even in those studies that were interpreted to reflect categorical perception or phoneme-boundary effect. Unfortunately, it is difficult to assess the effect of stimulus typicality on some of the above-mentioned results, because no data on this issue were reported. Some more recent MMN studies have, however, emphasized the role of the phonetic prototypes in MMN elicitation (Dehaene-Lambertz et al., 2000; Huotilainen et al., 2001; Näätänen et al., 1997). Even though the interpretations of the determinants of phonetic representations are controversial, the body of evidence suggests that extensive exposure to a certain language facilitates the processing of the acoustic changes that are linguistically relevant in that language, which, in turn, is reflected as an enhanced MMN response.

5. Experiments

5.1. Aims of the studies

Study I aimed at determining whether the relevance of the duration feature in the L1 affects the accuracy of the duration processing in the brain, as indexed by the MMN brain response. Native speakers of Finnish and advanced Russian L2 users of Finnish, whose native language does not have a phonological quantity distinction, were compared with each other. The goal of **Study II** was to determine the L2 users' accuracy in processing duration in different L2 sounds. On the basis of a hypothesis proposed by Flege (1995), the duration processing of the L2 sounds was expected to be affected by the L1 to a different extent depending on whether a new L2 vowel category has been established (with *dissimilar* sounds) or the L1 and L2 sounds are processed as belonging to the same category (with *similar* sounds).

Study III aimed to determine whether, as a result of exposure to Finnish, Russian L2 users of Finnish have been able to establish quantity categories that can be accessed on the basis of the duration cues, even though their L1 uses duration cues differently. In order to extend the findings of Study III, **Study IV** addressed the questions of whether the phoneme-boundary effect is reflected in the pre-attentive processing of the Finnish quantity categories and, further, whether the effect is, at the neural level, indeed induced by crossing a phoneme boundary. In addition, Study IV aimed at determining whether language learning is reflected in Russian L2 users' behavioral and brain responses to quantity and, further, whether L2 users show similar category-related effects as do the native speakers. Finally, the purpose of **Study V** was to determine whether phoneme quality and quantity have a common representation or separate representations in the brain, as indicated by the additivity of MMNs to phoneme quality and quantity.

5.2. Methods

Subjects

In Studies I and **II**, the subjects were 10–14-year-old native speakers of Finnish and Russian L2 users of Finnish. They were the same in both studies, with the exception of Study II having one Finnish subject less. In Study I, the native-speaker group consisted of 14 monolingual native speakers of Finnish and, in Study II, it consisted of 13 monolingual native speakers of Finnish. The L2-user group included 11 advanced second-language users of Finnish, speaking Russian as their L1. From both subject groups, one subject was excluded due to poor signal-to-noise ratio.

Studies III and IV included three adult subject groups with different language backgrounds, namely, native speakers of Finnish, Russian L2 users of Finnish, and non-Finnish-speaking Russians (hereafter referred to as *naïve* Russians). In Experiment 1 of **Study III**, 229 native speakers, 57 Russian L2 users of Finnish, and 60 naïve Russians who reported not having been exposed to Finnish were included in the analysis. The L2-user group was further divided into the short-exposure L2 group and the long-exposure L2 group according to the length of residence in Finland. In Experiment 2 of Study III, 20 native speakers of Finnish and 20 native speakers of Russian participated. Ten subjects in the Russian group reported speaking no Finnish. The other ten Russians spoke Finnish as a second language at a basic level of proficiency. **In Study IV**, the native-speaker group consisted of 13 native speakers of Finnish, while the group of the naïve Russians consisted of 12 non-Finnish-speaking Russians, and the L2-user group of 14 Russian subjects speaking Finnish as their L2 at an intermediate level of proficiency. Two L2 subjects were excluded from the MMN experiment on the basis of their behavioral results. In **Study V**, 14 adult native speakers of Finnish participated in the experiment. Two subjects were excluded from the analysis due to very frequent eye-blink artifacts.

Experimental conditions and stimuli

Studies I and **II** were MMN experiments, where stimuli were presented in a passive oddball paradigm. In both studies, duration change occurred in two stimulus types. They included a complex tone (components 500, 1000, and 1500 Hz) and syllable /kɑ/ in Study I and syllables /kɑ/ with a *similar* sound and /kæ/ with a *dissimilar* sound in Study II.[4] In both studies, the duration of the standard stimulus was 200 ms and that of the deviant stimulus 150 ms, while the SOA was 650 ms.

Study III included two behavioral experiments. Experiment 1 was a two-alternative, forced-choice (2AFC) categorization task with two conditions, where the categorization of the vowel quantity was investigated in the first- and second-syllable positions. Two pseudoword continua with seven steps differing from each other in the vowel duration were used as stimuli ([tuku] *vs.* [tuːku] in the first-syllable position and [tuku] *vs.* [tukuː] in the second-syllable position). Each stimulus was presented ten times in random order. Trains of five stimuli were presented with a 2-s ISI and a 4-s inter-train interval. In Experiment 2, categorization and discrimination performances were compared with each other. The categorization task was a corresponding 2AFC as in the first-syllable position of Experiment 1. Discrimination was studied with an AX ("same–different") task, where the adjacent stimuli of the same stimulus continuum were randomly presented in pairs. The ISI within the stimulus pairs was 1 s while the inter-pair interval was 2.5 s.

Study IV included categorization, word production, and MMN experiments. Behavioral 2AFC tests of quantity categorization were carried out with words and isolated vowels to determine the individual boundaries between the categories for each subject. In the word condition, a stimulus sequence identical to that in the first-syllable position of Study III (i.e., [tuku] *vs.* [tuːku] word continuum) was used. The isolated-vowel condition was otherwise identical, but the [u] *vs.* [uː] vowels were presented without the word context. With a reading task, word-production data on the same pseudowords were gathered from

---

[4] For Russian L2 users, the realizations of the Finnish and Russian /ɑ/ vowels following /k/ are hardly distinguishable, and can thus be considered similar. In contrast, there is no phoneme /æ/ in Russian and, therefore, the realizations of /æ/ can be regarded as dissimilar from any L1 phoneme.

the native speakers, whereas the L2 users read Russian sentences that included a similar Russian word that had stress on the first syllable. Moreover, an MMN experiment with two conditions, word and isolated vowel, was conducted using a passive oddball paradigm. In the word condition, [tuːku] *vs*. [tuːku] represented a within-category pair and [tuːku] *vs*. [tuku] represented an across-category pair. In the isolated-vowel condition, the corresponding vowel stimuli were presented without the word context. In the word and isolated-vowel conditions, the SOAs were 1000 ms and 500 ms, respectively. In addition, the oddball deviants were presented with a 100% probability in separate blocks.

Study V, too, used MMN methodology. The MMNs to consecutive phoneme-quality and -quantity changes as well as their sum were compared with the MMN elicited by a simultaneous change in both quality and quantity. Stimuli were pseudowords [itːi] (standard), [ipːi] (quality deviant), [iti] (quantity deviant), and [ipi] (double deviant, differing from the standard in quality and quantity). In addition, 3 blocks of reversed conditions, in which each of the deviants were used as standards and the rest of the stimuli as deviants, were presented. The stimuli were delivered in a passive oddball paradigm with a 1000-ms SOA.

Data acquisition and analysis

In **Studies I**, **II**, **IV**, and **V**, the EEG was recorded with NeuroScan system and SYNAMPS amplifier. In Studies I and II, Ag/AgCl electrodes were placed at the scalp sites F3, F4, C3, C4, T3, T4, P3, P4, and the two mastoids, and in Studies IV and V at F3, Fz, F4, C3, Cz, C4, P3, Pz, P4, and the two mastoids according to the international 10-20 system (Jasper, 1958). Eye movements were monitored with the electro-oculogram (EOG) attached below the eye and the canthus of the eye. The EEG was recorded while auditory stimuli were presented via headphones in an acoustically and electrically shielded room. Subjects were instructed not to pay attention to sound stimuli, but rather to a self-selected, muted movie.

Table 1. Details on the data acquisition and analysis in the MMN studies.

	Study I	Study II	Study IV	Study V
Sampling rate	250 Hz	250 Hz	500 Hz	500 Hz
Reference during the recording	Right mastoid	Right mastoid	Nose	Nose
Re-referencing	The average of the mastoids	The average of the mastoids	The average of the mastoids	No
Filter	1–15 Hz	2–15 Hz	1–20 Hz	1–20 Hz
Artefact rejection	±75 μV	±75 μV	±75 μV	±50 μV
Epoch duration (ms)	-50–650	-50–650	WC*: -50–1000 IC**: -50–500	-50–1000
Window of baseline correction (ms)	-50–0	-50–0	-50–0	50 ms pre-change
Amplitude measurement	20-ms windows centered at the grand-average peaks	20-ms windows centered at the grand-average peaks	20-ms windows centered at the grand-average peaks	100–150 ms and 150–200 ms from change onset

* WC = word condition; ** IC = isolated-vowel condition

The details of data acquisition and analysis are presented in Table 1. In Studies I and II, the averaged standard-stimulus responses were subtracted from those for the deviants.[5] In Studies IV and V, the difference signals were created by using the ERP signals elicited by the same stimulus in the low- and high-probability positions (oddball deviants and responses from 100% blocks or reversed-condition standards, respectively). For Study V,

---

[5] The subtraction of the ERP to a shorter standard from that to a longer deviant may result in the underestimation of the MMN amplitude (Jacobsen & Schröger, 2003). However, the way of subtraction did not distort the comparisons between the subject groups for the different stimulus types that were of interest, since the effect of subtraction was the same for all groups and stimulus types used in the studies.

a modeled double deviant was created by adding the two single-deviant responses after adjusting the change-onset points. In all studies, the mean-amplitude data were submitted to t-test to assess the significance of the MMN component. Further, analyses of variance (ANOVA) with the MMN mean amplitude as a dependent variable were performed.

In Experiment 1 of **Study III**, subjects responded to the stimuli in an answer sheet. After the data collection, difference curves for each subject's categorization functions were calculated by subtracting the contiguous data points. The mean of the difference curve, representing the category-boundary location (the 50% cross-over point of the functions), and the standard deviation (SD), representing the consistency of the categorization, were calculated. The data on the consistency of the categorization and the boundary location were subjected to separate one-way ANOVAs for the two syllable positions. The overall categorization functions were tested in separate two-way ANOVAs. In the categorization task of Experiment 2 of Study III, the subjects categorized the stimuli by pressing keys of a response pad. The RT was measured from the offset of the stimulus. The data on the category boundary and the consistency of the categorization were analyzed as in Experiment 1 and corresponding statistical tests were used. In addition, the RT data were subjected to a two-way ANOVA. The procedure of the discrimination task was the same as that of the categorization task with the exception that subjects responded according to whether the stimuli within each pair were the same or different. As an index of the discrimination sensitivity, d' scores were used. They were calculated according to the signal detection theory: $d' = Z_N - Z_{SN}$, where Z_N was obtained by converting $1 - p$(false alarms) and Z_{SN} by converting $1 - p$(hits) to z scores (Gescheider, 1985). The RT was measured from the offset of the second stimulus of the pair. For statistical testing, d' data and RT data for the same pairs and the different pairs were submitted to separate two-way ANOVAs. In addition, t-tests were performed on the d' data to assess whether discrimination sensitivity was above the chance level.

In the categorization task of Study IV, subjects responded to the stimuli in an answer sheet. After the data collection, the normal distribution was fitted to each subject's categorization functions. The mean value of the normal distribution represented the

boundary location and its SD indicated the consistency of the categorization. The mean and SD values were used as dependent variables in two-way ANOVAs. In the production experiment of Study IV, the vocalizations were recorded with a digital audio tape (DAT) recorder. Then the sound durations were measured, and the relative durations were calculated by dividing the absolute durations of the vowels of interest by the mean duration of a sound in the utterance. The data were submitted to one-way ANOVAs to compare the responses of the Finnish speakers and the Russian speakers with each other.
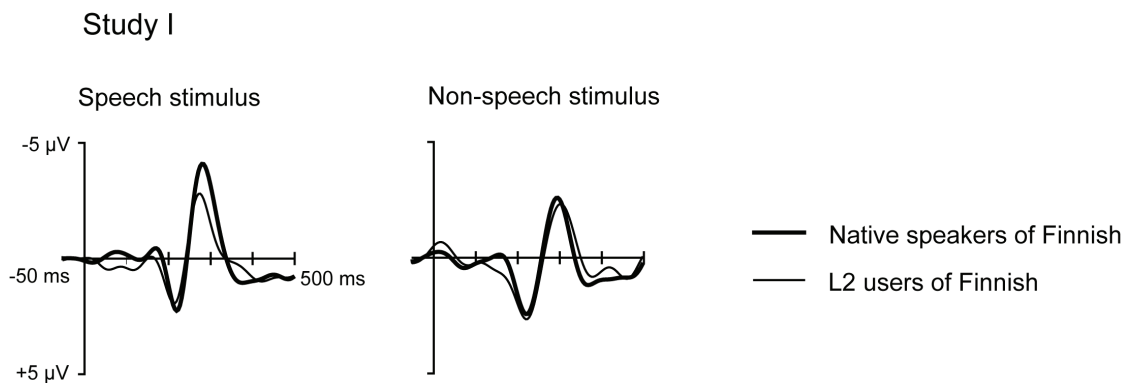
5.3. Results and discussion

Study I

Figure 1. Study I: MMN elicited by duration change in speech and non-speech sounds in native speakers and second-language (L2) users of Finnish. Grand-average difference signals.

The results of **Study I** revealed that the amplitude of the MMN brain response was similar in the native speakers and the L2 users of Finnish with a non-speech stimulus. With a speech sound, however, the MMN amplitude was larger than with a non-speech sound in the native speakers, but not in the L2 users (Group x Stimulus type interaction [$F(1, 21) = 7.7$, $p < 0.05$], see Fig. 1). In **Study II**, the MMNs amplitude for duration changes in two Finnish vowels did not differ significantly from each other in the native speakers. In the L2 users, however, duration change in a *similar* [ɑ] vowel that could be categorized through the L1 phonological system elicited a smaller MMN than that in a *dissimilar* [æ] vowel that could not be categorized through L1 categories (Group x

31

Stimulus type interaction [$F(1, 20) = 8.75$, $p < 0.01$], see Fig. 2). The results were interpreted to suggest that, in the native speakers of Finnish, duration processing is tuned or facilitated with speech sounds in comparison with non-speech sounds (see also Jaramillo et al., 2001). In contrast, in the L2 users, no such tuning was observed in Study I. However, Study II suggested that the lack of tuning in L2 users holds for L2 sounds that can be categorized through L1 phonological system, whereas duration processing in such sounds, for which new phonetic categories are established (*dissimilar* sounds), may achieve native-like facilitation.
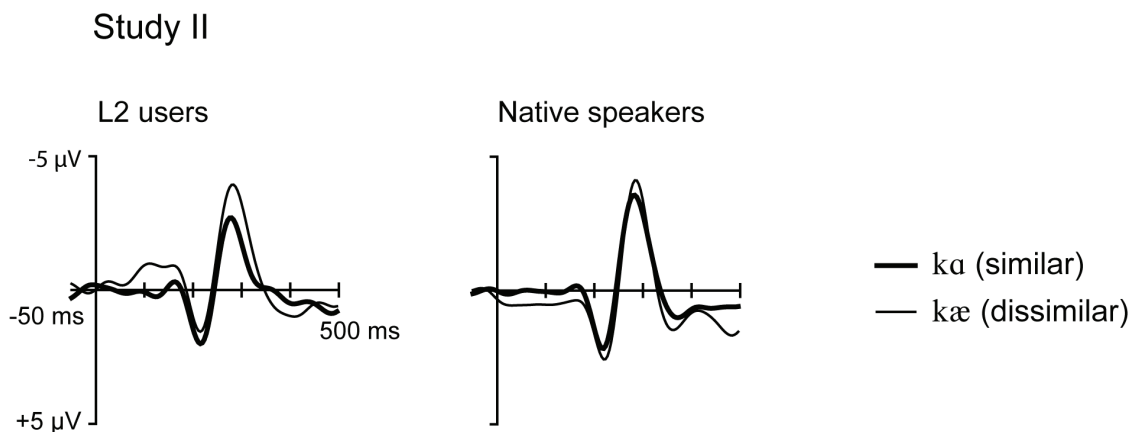
Study II

Figure 2. Study II: MMN elicited by duration changes in syllables [kɑ] and [kæ] in native speakers and second-language (L2) users of Finnish. For the L2 users, /ɑ/ was similar in their native and second languages, whereas /æ/ was dissimilar to any native-language phoneme. Grand-average difference signals.

In **Study III**, the main finding of Experiment 1 was that the consistency of categorization differed between the groups with the different language backgrounds in both syllable positions (first-syllable [$F(3, 342) = 40.17$, $p < 0.001$], see Fig. 3, top; second-syllable [$F(3, 342) = 17.88$, $p < 0.001$]). The native speakers of Finnish were more consistent in the categorization than the short-exposure L2 group and the naïve Russians. Moreover, the long-exposure L2 group and the naïve Russians were more consistent than the short-exposure L2 group. Significant differences between the groups were also found in the category-boundary location. However, the relative consistency of categorization rather than the boundary location or the shape of the categorization functions seemed to reflect

the access to the quantity categories. Thus, the results suggested that the native speakers and some of the long-exposure L2 users had access to the quantity categories, whereas the short-exposure L2 users and the naïve Russians did not.
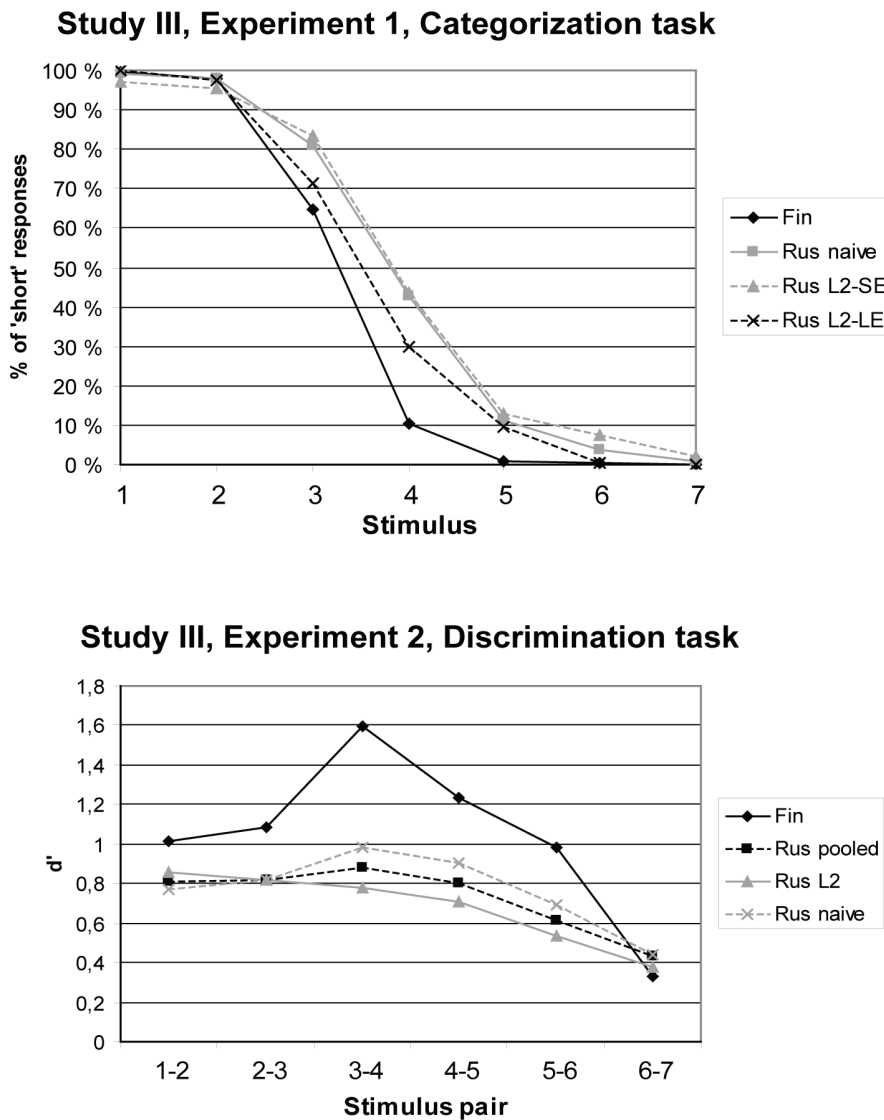
Study III, Experiment 1, Categorization task

Study III, Experiment 2, Discrimination task

Figure 3. The main findings of Study III. *Top:* The categorization functions of the native speakers of Finnish (Fin), the Russian second-language users of Finnish with short exposure (Rus L2-SE) and long exposure (Rus L2-LE), and the non-Finnish-speaking Russians (Rus naive) in the first-syllable position of Experiment 1. *Bottom:* The *d'* scores of the native speakers of Finnish (Fin), the Russian second-language users of Finnish (Rus L2), and the non-Finnish-speaking Russians (Rus naive), and the average of the two Russian groups (Rus pooled) in the discrimination task of Experiment 2.

In Experiment 2 of Study III, the main finding was that the d' data of the discrimination task showed that the native speakers had a significant peak in their discrimination sensitivity at the category boundary, reflecting the phoneme-boundary effect (cf. Bastian & Abramson, 1962), whereas no such effect occurred in the Russians (Language background x Stimulus pair interaction [$F(5, 190) = 2.75$, $p < 0.05$], see Fig. 3, bottom). Thus, the results suggested that the quantity categories facilitated the discrimination at the category boundary in the native Finnish subjects, but not in the non-native subjects. Moreover, the RT data supported this interpretation. At the same time, the native speakers' categorization was significantly more consistent than that of the Russians [$F(1, 38) = 9.14$, $p < 0.01$]. Thus, the results of Experiment 2 supported the interpretation of the results of Experiment 1, namely, that the relative consistency of quantity categorization may reveal the extent to which the quantity categories are accessed.

In **Study IV**, the results of the categorization task were in accordance with the findings of Study III: across the two conditions, the native speakers' categorization was significantly more consistent than that of the naïve Russians (Group main effect [$F(2, 36) = 3.40$, $p < 0.05$]). A similar but only marginally significant trend was found between the native speakers and the L2 users. The consistency of the categorization did not significantly differ between the Russian L2 users of Finnish and the naïve Russians, suggesting that the L2 users probably did not have access to the quantity categories. Thus, the L2 users' categorization might be determined by the L1 rather than the L2. In the MMN experiment of Study IV, the main finding was that across the conditions and stimulus types, the MMN amplitude was significantly larger in the native speakers than in each of the two Russian groups (Group main effect [$F(2, 34) = 5.38$, $p < 0.01$], see Fig. 4). In contrast, the Russian L2 users and the naïve Russians did not significantly differ from each other. Since the amplitude pattern of the MMN elicited by the across- and within-category changes was the same for the Finnish-speaking and non-Finnish-speaking subjects who lack the categories for Finnish quantity, it appears unlikely that the crossing of the phoneme boundary affected the MMN amplitude in the Finnish-speaking groups. Rather, the degree of the match of the stimuli to the L1 prototypes may have determined the MMN amplitude: the two deviant stimuli used in the experiment matched the L1

prototypes of the native speakers but did not match those of the Russian groups. The results also suggested that Finnish prototypes were not pre-attentively activated in the L2 learners as in the native speakers, since the L2 users did not differ from the naïve Russians in the MMN amplitude.
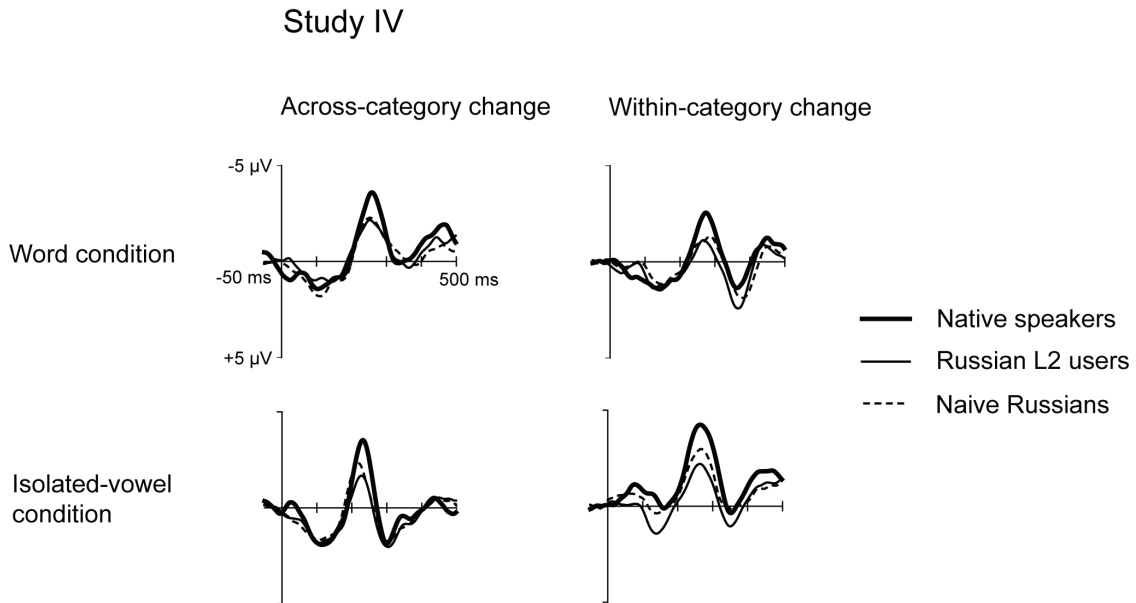
Study IV

Fig. 4. Study IV: The MMN responses elicited by the across- and within-category quantity changes in the native speakers of Finnish, the Russian second-language (L2) users of Finnish, and the non-Finnish-speaking (naive) Russians in word and isolated-vowel conditions. Grand-average difference signals.

Study V indicated that in comparison with the two single deviants (quality and quantity), the double deviant that differed from the standard in both quality and quantity elicited an MMN with a significantly larger amplitude in one of the two time windows of the amplitude measurement, whereas the MMN amplitude for the double deviant closely resembled the sum of the quantity and quality MMNs (Deviant type x Time window interaction [$F(3, 33) = 7.94$, $p < 0.001$], see Fig. 5). These results suggest that the MMNs elicited by changes in phoneme quality and quantity were additive and, consequently, were independently processed using separate representations, which implies that phoneme quality and quantity represent different levels in the phonological system of Finnish.
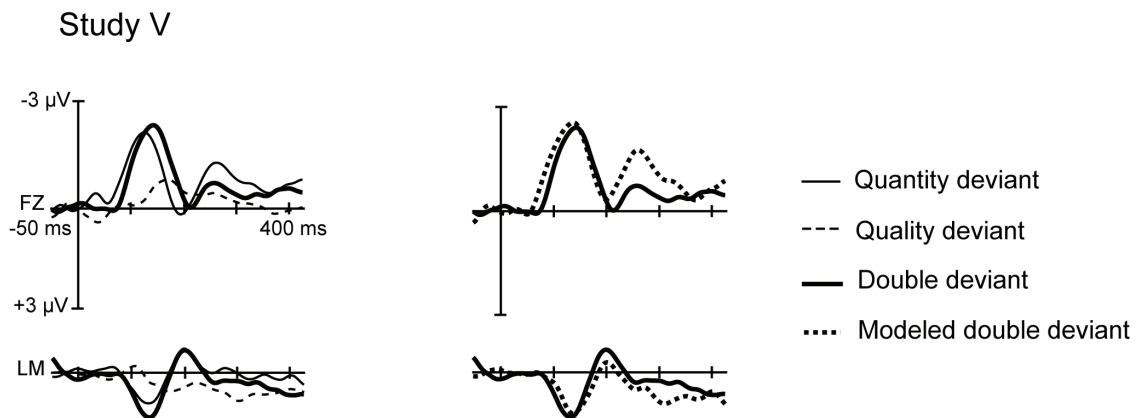
35

Study V

Fig. 5. Study V. *Left:* The MMN responses elicited by quality deviant (consonant change), quantity deviant (duration change), and double deviant (consonant and duration change). *Right:* the comparison of the MMN response to the double deviant with the modeled double deviant (the sum of the MMNs elicited by the quality and quantity deviants). The time scale is in relation to change onset. Grand-average difference signals at Fz and left mastoid (LM).

6. General discussion

With behavioral and electrophysiological methods, the present studies investigated the cortical representations for phonological quantity cued by speech-sound duration in native speakers and L2 users of Finnish. Possible differences between the groups in the mapping of stimuli onto the L1 and L2 categories as well as the establishment of the L2 phonetic categories were of interest. Below, three topics relevant to the studies will be discussed in more detail: (1) whether the experiments tapped the processing of duration or quantity; (2) whether the categorization was based on the category boundaries or the phoneme prototypes in the studies and how the quantity was represented in the phonological system; and (3) what have the studies revealed on the L2 learning.

6.1. Processing of duration or quantity?

Studies I and II and the isolated-vowel condition of Study IV were intended to address duration processing, whereas Studies III, the word condition of Study IV, and Study V

36

were intended to determine quantity at a phonological level. In the latter studies, word stimuli were used on the basis of the assumption that the activation of the representations for quantity would require context information at no less than the word level, so that the speaking rate and the relative durations were available as cues for the quantity categorization. However, on the basis of Study IV, it seems unlikely that the division to duration processing in syllables and quantity processing in words took place; the words and the isolated vowels elicited a similar MMN pattern across the subject groups that either did or did not have categories for quantity. Thus, even though the results of Studies I and II were originally interpreted to reflect the processing of duration and those of Studies IV and V were interpreted to reflect the processing of quantity, it is possible that all studies, in fact, tapped the same processes. Below, the duration processing and category activation accounts for the result will be discussed.

The duration-processing interpretation can be considered more conservative than the category-activation account. Since changes between phonetic categories must be accomplished by changes in acoustic features, to conclude that the MMN reflects category activation one should be able to rule out the possibility that the acoustic or non-categorical phonetic features caused the pattern of results. For example, in Study I, a larger-amplitude MMN was elicited by duration changes in speech than in non-speech sounds in the native speakers of Finnish, which *could* be interpreted to reflect category activation in speech stimulation. However, this result is not *necessarily* due to the category activation, since the result could have been caused by speech-specific sensitivity to duration changes in the native speakers of a quantity language. This sensitivity could also account for the results of Studies II and IV. In Study II, the results were interpreted to reflect speech-sound duration processing that is specific to the phonetic categories, because the duration-change MMN for a Finnish sound that could be categorized through Russian was reduced in comparison with that for a Finnish sound that has no Russian equivalent. In Study IV, the duration MMN elicited by the speech stimuli was larger in amplitude in native speakers than in Russian L2 users or non-Finnish-speaking Russians throughout the conditions regardless of whether the vowels were presented in word

context or not. Alternatively to the explanation proposed in Study IV, this could be due to the native Finnish speakers' general sensitivity to duration changes in speech sounds.

In the original interpretation of the results of Studies I and II, a conservative approach not suggesting category activation was preferred partly because the activation of the representations for quantity was assumed to require context information at no lower than the word level. However, this assumption may have been incorrect, since it is possible that an experimental design with constant stimulation, such as the oddball paradigm used in MMN studies, provides a sufficient context for the activation of the prototypes or categories for quantity. A conservative approach was favored also because the subjects were advanced L2 users of Finnish, and their pronunciation of Finnish suggested that they had categories for the Finnish quantity. However, even though no category activation was observed in the L2 users' MMN data, this does not exclude the possibility that the native speakers' results reflected the activation of the L1 categories or prototypes. Even though the quantity categories utilized in both production and perception of Finnish would have been established by the L2 users, it is nevertheless possible that the mapping of stimuli onto these categories is not automatic to such extent that the categories would be pre-attentively activated; the categorization of the L2 quantity may require attention and cognitive effort. At the pre-attentive level, stimuli may be processed via the L1 rather than the L2 if suitable L1 categories or prototypes are available.

Support for the category-activation account for the present results rises from previous cross-linguistic studies (e.g., Dehaene-Lambertz, 1997; Näätänen et al., 1997). These studies have suggested that larger-amplitude MMN responses are elicited by phonetic stimuli in native than in non-native listeners and, further, that stimuli representing L1 phonemes elicit larger-amplitude MMN responses than those not representing any phoneme in the L1. These phenomena are suggested to be due to the activation of L1 phonetic categories, and the present data are compatible with these patterns of results. In addition, Dehaene-Lambertz et al. (2000) demonstrated that the MMN reflects the parsing of the pseudoword syllable structure into L1 prototypes. Since in Finnish, there is a phonological distinction consisting of (C)V *vs*. (C)VV syllables (e.g., /u/ *vs*. /uː/ in uni

/uni/ 'dream' *vs*. uuni /uːni/ 'oven'; tuli /tuli/ 'wind' *vs*. tuuli /tuːli/ 'fire') and there are CV *vs*. CVV and VCV *vs*. VCCV minimal-pair words [e.g., te /te/ 'you (plural)' *vs*. tee /teː/ 'tea'; ala /ɑlɑ/ 'area' *vs*. alla /ɑlːɑ/ 'under'], the stimuli used in Studies I, II, IV, and V match syllable structures that have representations in the brains of native speakers of Finnish. Thus, regardless of whether the quantity is recognized via phoneme or syllable representations, the activation of the L1 prototypes, rather than the tuning of duration processing, may account for the results of the present studies. Also Sussman (2005) and Sussman et al. (2004) seem to have abandoned the original tuning account of the results reported in Study I, since they refer to it as evidence for categorical perception (see Footnote 1 for the definition of categorical perception).

Since there is no unequivocal agreement as to how sensory and long-term memory are linked (see Huotilainen et al. 2001; Phillips et al. 2000), one way to approach the problem of the substance of processing is to regard category formation as equivalent to the facilitation in the processing of the phonetic features. In this case, for example, there would be no difference between the interpretations that the results of Studies I, II, IV, and V reflected the activation of L1 prototypes for quantity or that they reflected the tuning of speech-sound duration processing, because the two interpretations would mean the same. This proposal could be falsified if, for example, the sensitivity to duration changes was found in native Finnish speakers for such speech sounds that fall out of the range of durations that could be categorized as short or long Finnish phonemes (e.g., clearly over-long speech sounds). This result would support the general-sensitivity account for the duration processing in the native speakers of Finnish. Since the present studies did not include such a stimulus condition (all the stimuli were clearly identifiable to one or the other quantity category), it would be important to include such a condition in the future studies of this topic.

6.2. Categorization

Studies III and IV addressed the quantity categorization, that is, the recognition of the quantity degree by mapping stimuli onto the representations for the quantity degrees. In Study III, the phoneme-boundary effect was found in behavioral measurements in the native speakers of Finnish: when stimulus pairs from a continuum with step-wise changes in vowel duration were presented to the native Finnish listeners, they perceived the duration changes most accurately, when a phoneme boundary was crossed by the change. This may be interpreted to imply that the phoneme boundaries are the determinants of the categorization. However, even though the phoneme-boundary effect was found in behavioral measurements, the MMN data of Study IV reflected no boundary effect at the neural level, as indicated by the same MMN pattern in the across- and within-category changes in the Finnish-speaking subjects who had the categories and in the non-Finnish-speaking subjects who did not have the categories. Instead of the category boundary, the MMN amplitude appeared to be determined by the degree of the typicality of the stimuli in subjects' L1. The results support the view that at the neural level, the phonemes or phonetic features are represented in long-term memory in terms of prototypes that can be pre-attentively activated. In Study IV, this was observed in the native speakers only.

Unlike Study IV, some previous studies comparing the MMNs elicited by across- and within-category changes have suggested that the MMN is enhanced by crossing the phoneme boundary (Dehaene-Lambertz, 1997; Sharma & Dorman, 1999; Winkler et al., 1999b; cf. Aaltonen et al., 1992; Maiste et al., 1995; Sams et al., 1990; Sharma et al., 1993). However, since it is well-established that the phonetic categories have an internal structure in terms of the degree of typicality (Iverson & Kuhl, 1996; Kuhl, 1991; Miller et al., 1983; Samuel, 1982), the results on the across- and within-category changes may be affected by the typicality of the stimuli used in these studies. For example, Sharma and Dorman (2000) observed a significant MMN in Hindi listeners whose phoneme boundary was crossed, but not in English listeners who had no boundary between the stimuli. At the same time, however, the deviant stimulus was more prototypical to Hindi than to English

speakers (see Lisker & Abramson, 1964). Similarly in Dehaene-Lambertz's (1997) study, the across-category /ba/ vs. /da/ distinction elicited an MMN in French listeners, whereas the dental vs. retroflex distinction, which is not phonemic in French, did not. Again, the Hindi retroflex was less prototypical to the French listeners than the stimuli used in the native across-category change, suggesting that the results may be affected by typicality. The typicality is not, however, as likely account for the difference found between the MMNs to the native across- and within-category changes in the studies of Dehaene-Lambertz (1997) and Sharma and Dorman (1999), where the stimuli appeared to be within the typical range for the target languages. Since previous studies (e.g., Fry et al., 1962; Pisoni, 1973) have shown that perceptual phenomena are different between the vowels and consonants, the processing of some consonant features might be less affected by typicality than that of the vowels.

The phoneme-boundary effect was observed in the behavioral experiment of Study III but not in the MMN experiment of Study IV. A possible account for this discrepancy is differences in the tasks and their requirements with regard to the sensory and categorical representations. According to Winkler et al. (1996) and Winkler & Cowan (2005), in the oddball paradigm a sensory-memory representation as well as extrapolations on future events are formed on the basis of the repetitive standard stimulus. An incoming stimulus, if compatible with the representation and the extrapolations, strengthens them or, if incompatible, updates them automatically, eliciting an MMN. With speech sounds and larger speech units, the MMN is affected by both sensory and long-term memory representations (e.g., Huotilainen et al., 2001; Näätänen et al., 1997; Winkler et al., 1999b). However, no MMN is elicited without a sensory-memory representation. In contrast, sensory-memory representations for acoustic detail may not always be strong enough to be utilized in an AX task, which was used in Study III, because they are not strengthened by repetition as is the case with the passive oddball paradigm. An AX task requires active decision making after hearing a particular stimulus pair only once in a roving stimulus environment, where A and X are constantly changing. With increasing memory load, sensory memory may not be able to represent the fine acoustic detail, but if stimulus A is recognized as belonging to a long-term memory category, then this

information, or possibly the activated prototype of the category, may be used for comparison. In another words, stimulus A of an AX pair may be more prone to be categorically processed, for example in terms of a prototype, rather than in terms of both acoustic detail and phonetic features of the stimulus, as is suggested to be the case with MMN (Näätänen et al., 1997; Winkler et al., 1999b). Consequently, if stimuli A and X represented the same category, then X would be compared with the prototype, which would diminish the difference between them (Kuhl, 1991; Samuel, 1982). However, if A and X represented different categories, then the discrepancy between them would be enhanced, resulting in a peak in the behavioral discrimination sensitivity at the category boundary, as was observed in Study III.

The view that the strength of the memory representation exerts an influence on the phoneme-boundary effect in an AX task is supported by previous observations (Crowder, 1982; Iverson & Kuhl, 2000; Macmillan et al., 1988; Pisoni, 1973), according to which the effect may be modified by the manipulation of the experimental conditions, such as the task type, the ISI, or the level of context variation. Additional support comes from the MMN literature, since MMN elicited by speech stimuli is also affected by the manipulation of the strength of the memory representation. Huotilainen et al. (2001), measuring the MMN elicited by deviants following prototypical and non-prototypical standard stimuli in a roving-standard paradigm, found that the MMN was larger in amplitude for deviants following two or three repetitions of the prototypical standards than after the same number of the non-prototypical standards. However, this effect disappeared with the increasing number of repetitions strengthening the representations. Thus, even though the MMN and the behavioral responses to phonetic stimuli reflect different processes in the sense that the MMN is based on pre-attentive, automatic memory processes and the behavioral measures require active decision making, both of them seem to be modified by the strength of the memory representation and the typicality of the phonetic stimuli.

In addition to the determinants of the quantity categorization, the organization of the phonological system with regard to the quantity was addressed. Theoretically, long

segments can be simply interpreted as a sequence of two short segments (Karlsson, 1969; cf. Harrikari, 2000), the short and long phonemes can be considered equal to the other phonemic distinctions in the phonological system, or the phonemes may be regarded as being comprised of a qualitative and a quantitative component that represent separate levels in the phonological system. Study V suggested that the MMN responses to changes in the phoneme quality and quantity were additive, indicating that the analysis of these features was independent of each other. This suggests that the phoneme quality and quantity were processed by separate neural processes that may thus be regarded as different levels in the phonological system.

6.3. L2 learning

An important aspect of the current studies was L2 learning. In particular, the establishment of the quantity categories in the L2 users was investigated with behavioral methods in Study III, while Studies I, II, and IV addressed the MMN brain response to L2 sounds in non-native speakers of Finnish. Previous research has suggested that the L2 quantity distinction is often difficult to non-native speakers and listeners (Han, 1992; Hayes-Harb, 2005; Hirata, 2004; McAllister et al., 2002). Similar conclusions can be drawn from the results of Studies III and IV, where the Russian L2 users and the non-Finnish-speaking, naïve Russians were compared with each other. Study III suggested that some of the L2 users who have been exposed to Finnish for a longer time (5 years or more) had access to the phonological quantity categories, whereas the L2 users with a shorter exposure and the naïve subjects did not have such categories. In addition, it was found that similarly to the naïve subjects, the non-native subjects with low proficiency in Finnish did not show a phoneme-boundary effect for the quantity as the native speakers did, corroborating the lack of the categories in this group. In Study IV which addressed the representation of the quantity in the brain, no difference was found between the L2 users' and naïve subjects' MMN responses to quantity. Thus, the L2 learning had not affected the processing of the quantity at the pre-attentive level: Finnish prototypes were

not pre-attentively activated in the L2 learners, as was suggested to be the case in the native speakers. As discussed earlier, Study I may reflect the same phenomenon.

The results of Study IV indicating no effect of the L2 learning on the MMN are in contrast with those of previous studies (e.g., Cheour et al., 2002; Peltola et al., 2003; Shestakova et al., 2003; Winkler et al., 1999a) that showed such an effect when segmental-level phonetic stimuli were used. Cheour et al. (2002) and Shestakova et al. (2003) observed an enhanced MMN to L2 sounds after exposure to L2. However, since the age of the L2 learning is one important factor having an influence on the L2 performance (e.g., Flege & MacKay, 2004; Piske et al., 2001), these studies are not fully comparable to Study IV: their subjects were 3–6-year-old children and thus exposed to L2 markedly earlier than were the adult subjects of Study IV. Subjects in Peltola et al. (2003) were adults, but had started to study English at school at around 9 years of age. While the other studies suggested an enhanced MMN response to L2 speech sounds after L2 exposure, in foreign-language learners of Peltola et al. (2003), a diminished response to L1 sounds was observed. In Study IV, neither increase nor decrease of the MMN amplitude in the L2 users as compared with that in the naïve subjects was found, however. Considering the age of the onset of the L2 exposure and the length of the residence in an L2-speaking country, the vowel study by Winkler et al. (1999a) is the most comparable with Study IV. In Study IV, the L2 users' age at the onset of learning Finnish was 14–22 years and the length of the residence in Finland was 3–12 years, while, in Winkler et al. (1999a), the corresponding values were 13–32 years (with one exception) and 2–13 years, respectively. Unlike the results of Study IV, those of Winkler et al. (1999a) suggested that the Hungarian L2 users of Finnish have a native-like MMN response to a Finnish vowel.

The lack of the language-learning effect in Study IV as opposed to the results of Winkler et al. (1999a) might be due to the fact that these studies involved L2 contrasts that were cued by different features of speech sounds. In their Feature Prominence Hypothesis, McAllister et al. (2002) proposed that re-attunement to duration can be difficult for those L2 learners with an L1 that does not exploit this feature phonologically. Moreover, these

L2 learners may be better able to attune their phonological systems to the spectral rather than to the durational cues of the L2, since the spectral cues are used in the L1 phonological contrasts, whereas the durational cues are not. As noted by Mueller (2005), the learning of different features may also be differently constrained by the age of acquisition. Studies I and IV, when compared with Winkler et al. (1999a) vowel study, suggest that in the studies involving duration feature, a less native-like MMN was elicited than in the vowel study. This holds regardless of the fact that the subjects of Study I can be considered early L2 learners, whereas subjects in Winkler et al. (1999a), as well as those in Study IV, were adolescents or adults at the L2-acquisition onset. Therefore, the discordance between our results and those of Winkler et al. (1999a) might suggest that plastic changes in the brain resulting in L2-category establishment are less likely to occur for the L2 contrasts cued by the duration than for those cued by the spectral features, if the L2 learner's L1 does not have phonological quantity contrasts.

A possible account for the variable difficulty in the establishment of the categories for the L2 contrasts that are cued by different kinds of features is the different optimal periods for acquisition (Werker & Tees, 2005). Infants are sensitive to the rhythmic properties of speech, and the perception of these features is attuned to the L1 earlier than that of the sound segments (see Nazzi & Ramus, 2003, for a review). Werker and Tees (2005) have suggested a cascading model of the optimal periods of language acquisition, where each level of processing reinforces and constrains the preceding and following sensitivities. If the rhythmic features are acquired at an earlier phase than the sound segments in a cascading system, then these language-specific sensitivities might be more difficult to change after the higher levels of the L1 processing have already been established.

In line with McAllister et al. (2002), Studies III and IV suggested that the establishment of the L2 quantity categories is difficult and may require several years of exposure to the target language, since a significant language-learning effect was only observed in the long-exposure L2 group of Study III. Even in this group, the category establishment probably had not occurred in all individuals. The non-native listeners' problem in the establishment of the quantity categories does not seem to be primarily in the perception

of the duration cues, however, since all subject groups discriminated duration changes above the chance level in Studies I-IV. One possible account for the category establishment in some of the L2 users of Study III but not in the others, regardless of their access to the relevant cues, is selective attention. Pisoni et al. (1994) have suggested that the linguistic experience modifies the attentional processes underlying perception. In addition, Iverson et al. (2003) have demonstrated that the Japanese L2 learners of English have difficulties in the recognition of the /r/ *vs*. /l/ distinction due to sensitivity to irrelevant cues that make it harder to focus attention on those cues that the native speakers use. Regardless of the fact that the duration cues were acoustically salient enough to be discerned by the non-native listeners in Studies I–IV, the L2 users may not always be able to utilize the cues in everyday communication in such a way that stable categories would be established to long-term memory and that they would be automatically activated by speech input. Especially in the early stages of the L2 acquisition, the comprehension of L2 speech in a natural communication situation requires the processing of many unfamiliar features in a limited time and with a limited processing capacity. Since phonological distinctions based on spectral cues are generally much more frequent and thus potentially more important than those based on duration cues, a language learner with L1 not using duration cues in phoneme distinction may attend to the spectral cues at the expense of the durational cues when listening to L2 speech.

Even though the most of the data in Studies I–IV showed that the L2 learners were not native-like in their processing of the Finnish quantity distinctions, there was one exception: Study II compared the advanced L2 users' MMNs to duration changes in two kinds of L2 sounds; namely, in an L2 vowel *similar* to an L1 vowel and in an L2 vowel *dissimilar* to any L1 vowel. The results showed a native-like duration MMN in a *dissimilar* vowel, but a smaller duration MMN in a *similar* vowel. The result was probably due to the establishment of a new vowel category for a *dissimilar* L2 vowel, whereas a *similar* L2 vowel could be pre-attentively mapped via the L1 phonological system. Apparently, the L1 that does not use the duration phonologically had a smaller influence on the processing of a duration change in a *dissimilar* vowel than in a *similar*

46

vowel. Thus, despite L1 constraints, the L2 users could learn to utilize the duration cues in a native-like fashion. These L2 users were, however, early learners and had achieved an advanced level of proficiency. Additional studies are needed to determine whether this data pattern could be replicated in late learners. Recently, a near-infrared spectroscopy study of Minagawa-Kawai et al. (2005) suggested that the late but nevertheless proficient L2 learners' brain activation for the Japanese quantity is different from that of the native speakers; even though the L2 learners behaviorally categorized quantity similarly to native speakers, they did not show such left-hemisphere brain responses for across-category quantity-degree changes as did the native speakers. The similarities between the L1 and L2 phonological systems were not discussed by Minagawa-Kawai et al. (2005), however.

It has been shown that the MMN is a useful tool to study the brain correlates of speech perception and L2 learning. It is noteworthy, however, that if the MMN data indicate no pre-attentive activation of L2 prototypes in L2 learners and there is no corroborating behavioral evidence, one should be cautious in concluding that L2 users have no categories for the L2 contrast in question. As suggested earlier, depending on the L2 competence level, the mapping of the stimuli onto the established L2 categories may sometimes require active effort and not be automatic as the mapping of stimuli onto the L1 sounds. This would suggest that the MMN can be used as a tool to reveal whether the possible categories are automatically accessed in a native-like fashion by comparing the MMN responses of native speakers with those of L2 users [as was the case in the Winkler et al. (1999a) study]. However, since access to categories may require active effort from the L2 users, MMN results may not always reveal whether such long-term memory representations are established or not. On the other hand, behavioral methods may not reveal the differences in the neural representation of the speech features underlying the overt behavioral responses (e.g., Minagawa-Kawai et al., 2005). If a study is intended to address the L2-category establishment in the brain, behavioral and brain-research methods should be combined.

7. Conclusions

This thesis suggests that the native speakers process the phoneme quantity and quality independently by using separate brain representations. It also shows that, in the native speakers of Finnish, the pre-attentive mapping of speech sounds onto the quantity categories does not occur on the basis of the phoneme boundaries. Rather, the L1 phonetic prototypes may be involved. The native Finnish-speakers' MMN brain response to quantity or duration changes was enhanced in amplitude in comparison with that of the L2 users of Finnish. This may be due to the pre-attentive and automatic activation of the L1 prototypes in the native, but not in the non-native listeners, who did not seem to have an automatic access to the L2 prototypes. Similarly, other signs of L2-category establishment, such as the phoneme-boundary effect, were not found in the inexperienced non-native listeners' behavioral responses. Some L2 users with over five years of exposure to Finnish, however, appeared to have established categories for the L2 quantity distinctions. In the present set of studies, the L2 users showed a native-like duration MMN only for those speech sounds that could not be processed via the L1 phonetic categories. The L2 users' difficulty to establish categories for quantity may be due to their inability to utilize duration cues linguistically even though they can be perceived non-linguistically.

References

Aaltonen, O., Paavilainen, P., Sams, M., & Näätänen, R. (1992). Event-related potentials and discrimination of steady-state vowels within and between phoneme categories: A preliminary study. *Scandinavian Journal of Logopedics and Phoniatrics, 17,* 107-112.

Alho, K., Connolly, J. F., Cheour, M., Lehtokoski, A., Huotilainen, M., Virtanen, J., Aulanko, R., & Ilmoniemi, R.J. (1998). Hemispheric lateralization in preattentive processing of speech sounds. *Neuroscience Letters, 258*, 9-12.

Amenedo, E., & Escera, C. (2000). The accuracy of sound duration representation in the human brain determines the accuracy of behavioural perception. *The European Journal of Neuroscience, 12*, 2570-2574.

Aoyama, K. (2001). *A psycholinguistic perspective on Finnish and Japanese prosody: Perception, production and child acquisition of consonantal quantity distinctions.* Boston: Kluwer Academic Publishers.

Aslin, R. N., Jusczyk, P. W., & Pisoni, D. B. (1998). Speech and auditory processing during infancy: Constraints on and precursors to language. In D. Kuhn, & R. S. Siegler (Eds.), *Handbook of child psychology: Cognition, perception, and language* (pp. 147-198). New York: John Wiley & Sons.

Aslin, R. N., Pisoni, D. B., Hennessy, B. L., & Perey, A. J. (1981). Discrimination of voice onset time by human infants: New findings and implications for the effects of early experience. *Child Development, 52*, 1135-1145.

Avanesov, R. I. (1984). *Russkoye literaturnoye proiznoshenie.* Moscow: Prosveschchenie.

Bastian, J., & Abramson, A. S. (1962). Identification and discrimination of phonemic vowel duration. *The Journal of the Acoustical Society of America, 34*, 743-744.

Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J. C. Goodman, & H. C. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 167-224). Cambridge, MA: The MIT Press.

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. 279-304). Baltimore: York Press.

Bondarko, L. V. (1998). *Fonetika sovremennogo russkogo jazyka*. St.-Petersburg: University of St.-Petersburg.

Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., & Näätänen, R. (1998). Development of language-specific phoneme representations in the infant brain. *Nature Neuroscience, 1*, 351-353.

Cheour, M., Shestakova, A., Alku, P., Ceponiene, R., & Näätänen, R. (2002). Mismatch negativity shows that 3-6-year-old children can learn to discriminate non-native speech sounds within two months. *Neuroscience Letters, 325*, 187-190.

Crowder, R. G. (1982). Decay of auditory memory in vowel discrimination. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 8*, 153-162.

de Silva, V. (1999). *Quantity and quality as universal and specific features of sound systems. Experimental phonetic research on interaction of Russian and Finnish sound systems*. Jyväskylä: University of Jyväskylä.

Dehaene-Lambertz, G. (1997). Electrophysiological correlates of categorical phoneme perception in adults. *Neuroreport, 8*, 919-924.

Dehaene-Lambertz, G., Dupoux, E., & Gout, A. (2000). Electrophysiological correlates of phonological processing: A cross-linguistic study. *Journal of Cognitive Neuroscience, 12*, 635-647.

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science, 171*, 303-306.

Flege, J. E. (1992). Speech learning in a second language. In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development. Models, research, implications* (pp. 565-604). Timonium, MD: York Press.

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

Flege, J. E., & MacKay, I. R. A. (2004). Perceiving vowels in a second language. *Studies in Second Language Acquisition, 26,* 1-34.

Fry, D. B., Abramson, A. S., Eimas, P. D., & Liberman, A. M. (1962). The identification and discrimination of synthetic vowels. *Language and Speech, 5*, 171-189.

Gescheider, G. A. (1985). *Psychophysics. Method, theory, and application*. Hillsdale, NJ: Lawrence Erlbaum associates.

Giard, M. H., Perrin, F., Pernier, J., & Bouchet, P. (1990). Brain generators implicated in the processing of auditory stimulus deviance: A topographic event-related potential study. *Psychophysiology, 27*, 627-640.

Han, M. S. (1992). The timing control of geminate and single stop consonants in Japanese: A challenge for nonnative speakers. *Phonetica, 49*, 102-127.

Harnad, S. (Ed.). (1987). *Categorical perception. The groundwork of cognition.* Cambridge: Cambridge University Press.

Harrikari, H. (2000). *Segmental length in Finnish – studies within a constraint-based approach*. Helsinki: University of Helsinki.

Hayes-Harb, R. (2005). Optimal L2 speech perception: Native speakers of English and Japanese consonant length contrasts. *Journal of Language and Linguistics, 4*, 1-29.

Hirata, Y. (2004). Training native English speakers to perceive Japanese length contrasts in word versus sentence contexts. *The Journal of the Acoustical Society of America, 116*, 2384-2394.

Huotilainen, M., Kujala, A., & Alku, P. (2001). Long-term memory traces facilitate short-term memory trace formation in audition in humans. *Neuroscience Letters, 310*, 133-136.

Iivonen, A. (1974a). Äännekeston riippuvuus ilmauksen pituudesta. *Virittäjä, 78*, 399-402.

Iivonen, A. (1974b). Äännekeston riippuvuus sanan pituudesta irrallaan äännetyissä sanoissa. *Virittäjä, 78*, 134-151.

Imada, T., Watanabe, M., Mashiko, T., Kawakatsu, M., & Kotani, M. (1997). The silent period between sounds has a stronger effect than the interstimulus interval on auditory evoked magnetic fields. *Electroencephalography and Clinical Neurophysiology, 102*, 37-45.

Inouchi, M., Kubota, M., Ferrari, P., & Roberts, T. P. (2003). Magnetic mismatch fields elicited by vowel duration and pitch changes in Japanese words in humans: Comparison between native- and non-speakers of Japanese. *Neuroscience Letters, 353*, 165-168.

Isei-Jaakola, T. (2004). *Lexical quantity in Japanese and Finnish*. Helsinki: University of Helsinki.

Iverson, P., & Kuhl, P. K. (1996). Influences of phonetic identification and category goodness on American listeners' perception of /r/ and /l/. *The Journal of the Acoustical Society of America, 99*, 1130-1140.

Iverson, P., & Kuhl, P. K. (2000). Perceptual magnet and phoneme boundary effects in speech perception: Do they arise from a common mechanism? *Perception & Psychophysics, 62*, 874-886.

Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A., & Siebert, C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition, 87*, B47-57.

Jacobsen, T., & Schröger, E. (2003). Measuring duration mismatch negativity. *Clinical Neurophysiology, 114*, 1133.

Jaramillo, M., Alku, P., & Paavilainen, P. (1999). An event-related potential (ERP) study of duration changes in speech and non-speech sounds. *Neuroreport, 10*, 3301-3305.

Jaramillo, M., Ilvonen, T., Kujala, T., Alku, P., Tervaniemi, M., & Alho, K. (2001). Are different kinds of acoustic features processed differently for speech and non-speech sounds? *Brain Research. Cognitive Brain Research, 12*, 459-466.

Jaramillo, M., Paavilainen, P., & Näätänen, R. (2000). Mismatch negativity and behavioural discrimination in humans as a function of the magnitude of change in sound duration. *Neuroscience Letters, 290*, 101-104.

Jasper, H. H. (1958). The ten twenty electrode system of the International Federation. Appendix to the report of the committee on methods of clinical examination in electroencephalography. *Electroencephalography and Clinical Neurophysiology, 10*, 371-375.

Joutsiniemi, S. L., Ilvonen, T., Sinkkonen, J., Huotilainen, M., Tervaniemi, M., Lehtokoski, A., Rinne, T., & Näätänen, R. (1998). The mismatch negativity for duration decrement of auditory stimuli in healthy subjects. *Electroencephalography and Clinical Neurophysiology, 108*, 154-159.

Jusczyk, P. W. (1997). *The discovery of spoken language*. Cambridge, MA: The MIT Press.

Jusczyk, P. W., Luce, P. A., & Charles-Luce, J. (1994). Infants' sensitivity to phonotactic patterns in native language. *Journal of Memory and Language, 33*, 630-645.

Karlsson, F. (1969). Suomen yleiskielen segmentaalifoneemien paradigma. *Virittäjä, 73*, 351-362.

Kaukoranta, E., Sams, M., Hari, R., Hämäläinen, M., & Näätänen, R. (1989). Reactions of human auditory cortex to a change in tone duration. *Hearing Research, 41*, 15-21.

Korpilahti, P., Krause, C. M., Holopainen, I., & Lang, A. H. (2001). Early and late mismatch negativity elicited by words and speech-like stimuli in children. *Brain and Language, 76*, 332-339.

Kuhl, P. K. (1981). Discrimination of speech by nonhuman animals: Basic auditory sensitivities conducive to the perception of speech-sound categories. *The Journal of the Acoustical Society of America, 70*, 340-349.

Kuhl, P. K. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics, 50*, 93-107.

Kuhl, P. K. (1994). Learning and representation in speech and language. *Current Opinion in Neurobiology, 4*, 812-822.

Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences of the United States of America, 97*, 11850.

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience, 5*, 831-843.

Kuhl, P. K., & Miller, J. D. (1975). Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants. *Science, 190*, 69-72.

Kuhl, P. K., & Padden, D. M. (1983). Enhanced discriminability at the phonetic boundaries for the place feature in macaques. *The Journal of the Acoustical Society of America, 73*, 1003-1010.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science, 255*, 606-608.

Kujala, T., & Näätänen, R. (2001). The mismatch negativity in evaluating central auditory dysfunction in dyslexia. *Neuroscience and Behavioral Reviews, 25*,535-543.

Kushnerenko, E., Ceponiene, R., Fellman, V., Huotilainen, M., & Winkler, I. (2001). Event-related potential correlates of sound duration: Similar pattern from birth to adulthood. *Neuroreport, 12*, 3777-3781.

Lang, H., Nyrke, T., Ek, M., Aaltonen, O., Raimo, I., & Näätänen, R. (1990). Pitch discrimination performance and auditory event-related potentials. In C. H. M. Brunia, A. W. K. Gaillard, A. Kok, G. Mulder & M. N. Verbaten (Eds.), *Psychophysiological brain research, vol. 1,* (pp. 294-298). Tilburg: Tilburg University Press.

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

Lehtonen, J. (1970). *Aspects of quantity in standard Finnish*. Jyväskylä: University of Jyväskylä.

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology, 54*, 358-368.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word, 20*, 384-422.

Lyytinen, H., Ahonen, T., Eklund, K., Guttorm, T., Kulju, P., Laakso, M.-L., Leiwo, M., Leppänen, P., Lyytinen P., Poikkeus, A.-M., Richardson, U., Torppa, M., & Viholainen, H. (2004). Early development of children at familial risk for dyslexia-follow-up from birth to school age. *Dyslexia, 10*, 146-178.

Macmillan, N. A., Goldberg, R. F., & Braida, L. D. (1988). Resolution for speech sounds: Basic sensitivity and context memory on vowel and consonant continua. *The Journal of the Acoustical Society of America, 84*, 1262-1280.

Maiste, A. C., Wiens, A. S., Hunt, M. J., Scherg, M., & Picton, T. W. (1995). Event-related potentials and the categorical perception of speech sounds. *Ear and Hearing, 16*, 68-90.

Marjomaa, I. (1982). Englannin ja suomen vokaalien kestoista puhenopeuden vaihdellessa. In A. Iivonen, & H. Kaskinen (Eds.), *XI fonetiikan päivät - Helsinki 1982* (pp. 119-137). Helsinki: University of Helsinki.

Massaro, D. W. (1987). Categorical partition: A fuzzy-logical model of categorization behavior. In S. Harnad (Ed.), *Categorical perception. The groundwork of cognition* (pp. 254-283). Cambridge: Cambridge University Press.

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

McAllister, R., Flege, J. E., & Piske, T. (2002). The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of Phonetics, 30*, 229-258.

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. *Learning & Memory, 9*, 253-267.

Michie, P. T. (2001). What has MMN revealed about the auditory system in schizophrenia? *International Journal of Psychophysiology, 42*, 177-194.

Miller, J. D., Wier, C. C., Pastore, R. E., Kelly, W. J., & Dooling, R. J. (1976). Discrimination and labeling of noise-buzz sequences with varying noise-lead times: An example of categorical perception. *The Journal of the Acoustical Society of America, 60*, 410-417.

Miller, J. L., Connine, C. M., Schermer, T. M., & Kluender, K. R. (1983). A possible auditory basis for internal structure of phonetic categories. *The Journal of the Acoustical Society of America, 73*, 2124-2133.

Miller, J. L., & Eimas, P. D. (1995). Speech perception: From signal to word. *Annual Review of Psychology, 46*, 467-492.

Minagawa-Kawai, Y., Mori, K., & Sato, Y. (2005). Different brain strategies underlie the categorical perception of foreign and native phonemes. *Journal of Cognitive Neuroscience, 17*, 1376-1385.

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J. J., & Fujimura, O. (1975). An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception & Psychophysics, 18*, 340.

Mueller, J. L. (2005). Electrophysiological correlates of second language processing. *Second Language Research, 21*, 152-174.

Munson, W. A. (1947). The growth of auditory sensation. *The Journal of the Acoustical Society of America, 19*, 584-591.

Näätänen, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral and Brain Sciences, 13*, 201-288.

Näätänen, R. (1992). *Attention and brain function*. Hillsdale, NJ: Lawrence Erlbaum Associates.

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*, 1-21.

Näätänen, R., Gaillard, A. W., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica, 42*, 313-329.

Näätänen, R., Jacobsen, T., & Winkler, I. (2005). Memory-based or afferent processes in mismatch negativity (MMN): A review of the evidence. *Psychophysiology, 42,* 25-32.

Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R.J., 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., Alho, K., Reinikainen, K., & Sams, M. (1987). Interstimulus interval and the mismatch negativity. In C. Barber, & T. Blum (Eds.), *Evoked potentials III. The third international evoked potentials symposium* (pp. 392-397). London: Butterworths.

Näätänen, R., Paavilainen, P., & Reinikainen, K. (1989). Do event-related potentials to infrequent decrements in duration of auditory stimuli demonstrate a memory trace in man? *Neuroscience Letters, 107*, 347-352.

Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology, 24*, 375-425.

Näätänen, R., Sams, M., Alho, K., Paavilainen, P., Reinikainen, K., & Sokolov, E. N. (1988). Frequency and location specificity of the human vertex N1 wave. *Electroencephalography and Clinical Neurophysiology, 69*, 523-531.

Näätänen, R., Schröger, E., Karakas, S., Tervaniemi, M., & Paavilainen, P. (1993). Development of a memory trace for a complex sound in the human brain. *Neuroreport, 4*, 503-506.

Näätänen, R., Syssoeva, O., & Takegata, R. (2004). Automatic time perception in the human brain for intervals ranging from milliseconds to seconds. *Psychophysiology, 41*, 660-663.

Näätänen, R., & Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychological Bulletin, 125*, 826-859.

Nazzi, T., & Ramus, F. (2003). Perception and acquisition of linguistic rhythm by infants. *Speech Communication, 41*, 233-243.

O'Dell, M. (2003). *Intrinsic timing and quantity in Finnish*. Tampere: University of Tampere.

Opitz, B., Rinne, T., Mecklinger, A., von Cramon, D. Y., & Schröger, E. (2002). Differential contribution of frontal and temporal cortices to auditory change detection: fMRI and ERP results. *NeuroImage, 15*, 167-174.

Pantev, C., Bertrand, O., Eulitz, C., Verkindt, C., Hampson, S., Schuierer, G., & Elbert, T. (1995). Specific tonotopic organizations of different areas of the human auditory cortex revealed by simultaneous magnetic and electric recordings. *Electroencephalography and Clinical Neurophysiology, 94*, 26-40.

Pantev, C., Hoke, M., Lehnertz, K., Lutkenhoner, B., Anogianakis, G., & Wittkowski, W. (1988). Tonotopic organization of the human auditory cortex revealed by transient auditory evoked magnetic fields. *Electroencephalography and Clinical Neurophysiology, 69*, 160-170.

Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L. E., & Hoke, M. (1998). Increased auditory cortical representation in musicians. *Nature, 392*, 811-814.

Peltola, M. S., Kujala, T., Tuomainen, J., Ek, M., Aaltonen, O., & Näätänen, R. (2003). Native and foreign vowel discrimination as indexed by the mismatch negativity (MMN) response. *Neuroscience Letters, 352*, 25-28.

Phillips, C., Pellathy, T., Marantz, A., Yellin, E., Wexler, K., Poeppel, D., McGinnis, M., & Roberts, T. (2000). Auditory cortex accesses phonological categories: An MEG mismatch study. *Journal of Cognitive Neuroscience, 12*, 1038-1055.

Picton, T. W., Alain, C., Otten, L., Ritter, W., & Achim, A. (2000). Mismatch negativity: Different water in the same river. *Audiology & Neuro-Otology, 5*, 111-139.

Piske, T., MacKay, I. R. A., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics, 29*, 191-215.

Pisoni, D. B. (1973). Auditory and phonetic memory codes in the discrimination of consonant and vowels. *Perception & Psychophysics, 13*, 253-260.

Pisoni, D. B. (1977). Identification and discrimination of the relative onset time of two component tones: Implications for voicing perception in stops. *The Journal of the Acoustical Society of America, 61*, 1352-1361.

Pisoni, D. B., Lively, S. E., & Logan, J. S. (1994). Perceptual learning of nonnative speech contrasts: Implications for theories of speech perception. In J. C. Goodman, & H. C. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 121-166). Cambridge, MA: The MIT Press.

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

Polka, L., & Werker, J. F. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 421-435.

Pulvermüller, F., Kujala, T., Shtyrov, Y., Simola, J., Tiitinen, H., Alku, P., Alho, K., Martinkauppi, S., Ilmoniemi, R.J., & Näätänen, R. (2001). Memory traces for words as revealed by the mismatch negativity. *NeuroImage, 14*, 607-616.

Pulvermüller, F., Shtyrov, Y., Kujala, T., & Näätänen, R. (2004). Word-specific cortical activity as revealed by the mismatch negativity. *Psychophysiology, 41*, 106-112.

Rinne, T., Alho, K., Ilmoniemi, R. J., Virtanen, J., & Näätänen, R. (2000). Separate time behaviors of the temporal and frontal mismatch negativity sources. *NeuroImage, 12*, 14-19.

Rinne, T., Degerman, A., & Alho, K. (2005). Superior temporal and inferior frontal cortices are activated by infrequent sound duration decrements: An fMRI study. *NeuroImage, 26*, 66-72.

Romani, G. L., Williamson, S. J., & Kaufman, L. (1982). Tonotopic organization of the human auditory cortex. *Science, 216*, 1339-1340.

Sams, M., Aulanko, R., Aaltonen, O., & Näätänen, R. (1990). Event-related potentials to infrequent changes in synthesized phonetic stimuli. *Journal of Cognitive Neuroscience, 2*, 344-357.

Samuel, A. G. (1982). Phonetic prototypes. *Perception & Psychophysics, 31*, 307-314.

Schouten, B., Gerrits, E., & van Hessen, A. (2003). The end of categorical perception as we know it. *Speech Communication, 41*, 71-80.

Schröger, E. (1996). The influence of stimulus intensity and inter-stimulus interval on the detection of pitch and loudness changes. *Electroencephalography and Clinical Neurophysiology, 100*, 517-526.

Schröger, E. (1997). On the detection of auditory deviations: A pre-attentive activation model. *Psychophysiology, 34*, 245-257.

Schröger, E., & Winkler, I. (1995). Presentation rate and magnitude of stimulus deviance effects on human pre-attentive change detection. *Neuroscience Letters, 193*, 185-188.

Sharma, A., & Dorman, M. F. (1999). Cortical auditory evoked potential correlates of categorical perception of voice-onset time. *The Journal of the Acoustical Society of America, 106*, 1078-1083.

Sharma, A., & Dorman, M. F. (2000). Neurophysiologic correlates of cross-language phonetic perception. *The Journal of the Acoustical Society of America, 107*, 2697-2703.

Sharma, A., Kraus, N., McGee, T., Carrell, T., & Nicol, T. (1993). Acoustic versus phonetic representation of speech as reflected by the mismatch negativity event-related potential. *Electroencephalography and Clinical Neurophysiology, 88*, 64-71.

Sheehan, K. A., McArthur, G. M., & Bishop, D. V. (2005). Is discrimination training necessary to cause changes in the P2 auditory event-related brain potential to speech sounds? *Brain Research. Cognitive Brain Research, 25*, 547-553.

Shestakova, A., Brattico, E., Huotilainen, M., Galunov, V., Soloviev, A., Sams, M., Ilmoniemi, R.J., & Näätänen, R. (2002). Abstract phoneme representations in the left temporal cortex: Magnetic mismatch negativity study. *Neuroreport, 13*, 1813-1816.

Shestakova, A., Huotilainen, M., Ceponiene, R., & Cheour, M. (2003). Event-related potentials associated with second language learning in children. *Clinical Neurophysiology, 114*, 1507-1512.

Shtyrov, Y., Kujala, T., Ahveninen, J., Tervaniemi, M., Alku, P., Ilmoniemi, R.J., & Näätänen, R. (1998). Background acoustic noise and the hemispheric lateralization of speech processing in the human brain: Magnetic mismatch negativity study. *Neuroscience Letters, 251*, 141-144.

Shtyrov, Y., Kujala, T., Palva, S., Ilmoniemi, R. J., & Näätänen, R. (2000). Discrimination of speech and of complex nonspeech sounds of different temporal structure in the left and right cerebral hemispheres. *NeuroImage, 12*, 657-663.

Shtyrov, Y., & Pulvermüller, F. (2002). Neurophysiological evidence of memory traces for words in the human brain. *Neuroreport, 13*, 521-525.

Strange, W. (1999). Perception of consonants: From variance to invariance. In J. M. Pickett (Ed.), *The acoustics of speech communication: Fundamentals, speech perception theory, and technology* (pp. 166-182). Boston: Allyn and Bacon.

Studdert-Kennedy, M., Liberman, A. M., Harris, K. S., & Cooper, F. S. (1970). Theoretical notes. Motor theory of speech perception: A reply to Lane's critical review. *Psychological Review, 77*, 234-249.

Sundara, M., Polka, L., & Genesee, F. (in press). Language-experience facilitates discrimination of /d-ð/ in monolingual and bilingual acquisition of English. *Cognition*.

Suomi, K., Toivanen, J., & Ylitalo, R. (2003). Durational and tonal correlates of accent in Finnish. *Journal of Phonetics, 31*, 113-138.

Suomi, K., & Ylitalo, R. (2004). On durational correlates of word stress in Finnish. *Journal of Phonetics, 32*, 35-63.

Sussman, E. S. (2005). Integration and segregation in auditory scene analysis. *The Journal of the Acoustical Society of America, 117*, 1285-1298.

Sussman, E., Kujala, T., Halmetoja, J., Lyytinen, H., Alku, P., & Näätänen, R. (2004). Automatic and controlled processing of acoustic and phonetic contrasts. *Hearing Research, 190*, 128-140.

Takegata, R., Nakagawa, S., Tonoike, M., & Näätänen, R. (2004). Hemispheric processing of duration changes in speech and non-speech sounds. *NeuroReport, 15*, 1683–1686.

Tervaniemi, M., Lehtokoski, A., Sinkkonen, J., Virtanen, J., Ilmoniemi, R. J., & Näätänen, R. (1999). Test-retest reliability of mismatch negativity for duration, frequency and intensity changes. *Clinical Neurophysiology, 110*, 1388-1393.

Tiitinen, H., May, P., Reinikainen, K., & Näätänen, R. (1994). Attentive novelty detection in humans is governed by pre-attentive sensory memory. *Nature, 372*, 90-92.

Todd, J., & Michie, P. T. (2000). Do perceived loudness cues contribute to duration mismatch negativity (MMN)? *Neuroreport, 11*, 3771-3774.

Tremblay, K., Kraus, N., McGee, T., Ponton, C., & Otis, B. (2001). Central auditory plasticity: Changes in the N1-P2 complex after speech-sound training. *Ear and Hearing, 22*, 79-90.

Trubetzkoy, N. S. (1939/1969). *Principles of phonology*. Berkeley, CA: University of California Press.

Vainio, M. (1996). Phoneme frequencies in Finnish text and speech. In A. Iivonen, & A. Klippi (Eds.), *Studies in logopedics and phonetics 5* (pp. 181-192). Helsinki: University of Helsinki.

Verbickaya, L. A. (1976). *Russkaya orfoepia*. Leningrad: University of Leningrad.

Vihanta, V. V. (1987). Suomen äännekestot ranskalaisen suomenoppijan kannalta. In P. Hurme, & H. Dufva (Eds.), *Papers in speech research 7. Papers from the 14th meeting of Finnish phoneticians* (Jyväskylä ed.) (pp. 101-122) Jyväskylä: University of Jyväskylä.

Vihanta, V. V. (1990). Suomi vieraana kielenä foneettiselta kannalta. In J. Tommola (Ed.), *Foreign language comprehension and production: AFinLA yearbook 1990* (pp. 199-225). Turku: AFinLA.

Werker, J. F., & Lalonde, C. E. (1988). Cross-language speech perception: Initial capabilities and developmental change, *Developmental Psychology, 24*, 672-683.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development, 7*, 49-63.

Werker, J. F., & Tees, R. C. (2005). Speech perception as a window for understanding plasticity and commitment in language systems of the brain. *Developmental Psychobiology, 46*, 233-234.

Wiik, K. (1965). *Finnish and English vowels*. Turku: University of Turku.

Wiik, K. (1981). *Fonetiikan perusteet*. Porvoo: WSOY.

Winkler, I., & Cowan, N. (2005). From sensory to long-term memory: Evidence from auditory memory reactivation studies. *Experimental Psychology, 52*, 3-20.

Winkler, I., Karmos, G., & Näätänen, R. (1996). Adaptive modeling of the unattended acoustic environment reflected in the mismatch negativity event-related potential. *Brain Research, 742*, 239-252.

Winkler, I., Kujala, T., Tiitinen, H., Sivonen, P., Alku, P., Lehtokoski, A., Czigler, I., Csepe, V., Ilmoniemi, R.J., & Näätänen, R. (1999a). Brain responses reveal the learning of foreign language phonemes. *Psychophysiology, 36*, 638-642.

Winkler, I., Lehtokoski, A., Alku, P., Vainio, M., Czigler, I., Csepe, V., Aaltonen, O., Raimo, I., Alho, K., Lang, H., Iivonen, A., & Näätänen, R. (1999b). Pre-attentive detection of vowel contrasts utilizes both phonetic and auditory memory representations. *Brain Research. Cognitive Brain Research, 7*, 357-369.

Winkler, I., Schröger, E., & Cowan, N. (2001). The role of large-scale memory organization in the mismatch negativity event-related brain potential. *Journal of Cognitive Neuroscience, 13*, 59-71.

Wood, C. C. (1976). Discriminability, response bias, and phoneme categories in discrimination of voice onset time. *The Journal of the Acoustical Society of America, 60*, 1381-1389.