The effects of long-term native language exposure on event-related potential (ERP) components N1 and MMN in typically reading adults

Annika Eräkannas
Master's thesis
University of Jyväskylä
Department of Psychology
Autumn 2009

Supervisor of the graduate thesis	Paavo H.T. Leppänen, Docent, Ph.D., Acting Professor, Department of Psychology University of Jyväskylä, Finland
Research project	NeuroDys, Dyslexia genes and Neurobiological pathways European Dyslexia Research Project
Research site	Centre of Excellence for Human Development and Risk factors, Department of Psychology, University of Jyväskylä, Finland
Supported by	European Union Academy of Finland Department of Psychology, University of Jyväskylä, Finland

Neurodys
Dyslexia genes and Neurobiological pathways
European Dyslexia Research Project

CONTENTS:

ABSTRACT

Author: Annika Eräkannas
Master's Thesis in Psychology
Autumn 2009
University of Jyväskylä
70 pages

Speech perception is a complex process, which has been studied with different methods. In this Master's thesis, the effects of long-term native language exposure on auditory brain responses were studied by exploiting the brain event-related potential (ERP) technique. The main interest was in the N1 and MMN (mismatch negativity) responses from which the latter has been found to be affected by memory traces and top-down processes. In the ERP measurement, synthetic /i/ and /y/ vowels and equivalent sinusoidal versions of them were used as stimuli. The non-speech sounds were also studied to estimate the effects of acoustical distances on brain responses. The brain responses were studied using a passive MMN paradigm. The Finnish-Hungarian (native) /y/ vowel served as the prototype of the /y/ phoneme category whereas the French and German /y/ vowels were chosen as the non-prototype members of the /y/ category. The participants were typically reading Finnish adults without any exposure to French, German or Hungarian. The results revealed minor differences between the responses to different languages. By visual observation, the prototype elicited smaller responses than the non-prototypes. This difference was statistically significant in the fronto-central-parietal channels in a late time window of the MMN response and revealed language specific processing. In further studies, to unquestionably show the language specificity, the effects of long-term native language exposure must be assessed in a comparison with behavioural tests. In order to get reliable information of the processing of prototypes and non-prototypes, it could be useful to use within-category stimuli in one oddball condition.

Keywords: Speech perception, categorical perception, language exposure, ERP, MMN, N1, top-down effects, crosslinguistic studies

---

TIIVISTELMÄ

Tekijä: Annika Eräkannas
Psykologian Pro gradu -tutkielma
Syksy 2009
Jyväskylän Yliopisto
70 sivua

Puheen havaitseminen on monimutkainen prosessi, jota on tutkittu erilaisista lähtökohdista. Tässä Pro gradu -tutkielmassa tarkasteltiin pitkäaikaisen äidinkielelle altistumisen vaikutuksia auditiivisiin aivovasteisiin. Herätevasteita (ERP) käytettiin altistumisen vaikutusten arviointiin. Erityisesti tarkastelun kohteena olivat N1 ja MMN (mismatch negatiivisuus) vasteet, joista jälkimmäiseen on aikaisemmissa tutkimuksissa havaittu pitkäaikaisten muistijälkien ja ylempien kognitiivisten prosessien vaikuttavan. Ärsykkeinä käytettiin synteettisesti tuotettuja /i/ ja /y/ vokaaleja sekä siniäänistä muodostettuja vastaavia muunnoksia. Ei-puheäänien tarkastelulla haluttiin arvioida ärsykkeiden välisten akustisten erojen vaikutusta vasteisiin. Herätevasteet mitattiin passiivisessa MMN paradigmassa. Suomi-unkarin /y/ vokaalia käytettiin tutkimuksessa prototyyppinä /y/ kategorialle ja ranskalainen ja saksalainen /y/ vokaali valittiin edustamaan /y/ kategorian ei-prototyyppisiä edustajia. Tutkittavat olivat tyypillisesti lukevia suomalaisia aikuisia, jotka eivät olleet altistuneet ranskan, saksan tai unkarin kielelle. Kielten välillä havaittiin pieniä tilastollisia eroja. Näönvaraisesti tarkasteltuna prototyyppiärsyke tuotti pienemmän aivovasteen verrattuna ei-prototyyppien tuottamiin vasteisiin. Tämä ero oli tilastollisesti havaittavissa MMN vasteen myöhäisellä aikaikkunalla etu-, keski-, ja päälaenkanavilla (frontaaliset, sentraaliset ja parietaaliset kanavat) viitaten kielispesifiseen prosessiin. Jatkossa pitkäaikaisen äidinkielelle altistumisen vaikutusta on syytä, herätevasteiden ohella, tutkia arvioimalla koehenkilöiden kykyä erotella ja arvioida käytettyjä ärsykkeitä. Jotta prototyyppien ja ei-prototyyppien prosessoinnista saataisiin luotettavaa tietoa, voisi olla hyödyllistä lisätä ERP kokeeseen kategorian sisäisiä ärsykkeitä sisältävä koetilanne.

Avainsanat: Puheen havaitseminen, kategorinen havaitseminen, altistuminen kielelle, ERP, MMN, N1, ylemmät kognitiiviset prosessit, kieltenväliset koeasetelmat

INTRODUCTION

Researches have found out that large changes occur in auditory perception during the infancy and childhood. Several studies have discovered the infant ability to distinguish any foreign language phonemes (Kuhl, 1991, 1993), whereas adults' auditory processing abilities are diminished for foreign language phonemes which are not relevant in their native language (Yamada & Tohkura, 1992). These observations have aroused interest towards language learning and the effects of long-term exposure on native language (Halsband, 2006). The long-term exposure to native language starts at birth and the infant starts to prefer his native language phonemes and loses the ability to distinguish foreign language phonemes at around the age of six months (Cheour et al., 1998, however see Polka, Colantonio, & Sundara, 2001; Sundara, Polka, & Genesee, 2006 for the results of remained ability to discriminate foreign language phonemes). This language-specific perception is based on category prototypes (Kuhl, 1993) or in other terms; the memory traces of phonemes i.e. phonological representations, which develop during early childhood as a result of the influence of native language (Best, McRoberts, Lafleur, & Silver-Isenstadt, 1995; Best, McRoberts, & Sithole, 1988).

Several fields of science have focused on examination of speech perception with different methods. In neuropsychology event-related potentials (ERPs) are used for investigating brain processes elicited by stimuli (Eggermont & Ponton, 2002). The objective of this Master's thesis was to explore whether the native language prototype affects the basic auditory processing of speech sounds within a phoneme category as well as non-speech sounds without phoneme categories. This was done by using auditory ERPs in a classic passive oddball paradigm and focusing on basic auditory processing component N1 and automatic change detection component mismatch negativity (MMN; Näätänen, 1992, 2001) in typically reading adults. In the following sections, first an overview to speech perception/processing and general concepts of speech perception will be presented and the following paragraphs of introduction will concentrate on the models of speech perception/processing. The last paragraphs of introduction will discuss the issues of ERPs and the memory trace effects on ERPs examined mainly in crosslinguistic studies.

Speech perception and processing

Speech perception is a complex process where acoustic waveforms are coded into phonological units and into words. Speech has been considered as a unique characteristic of human communication. Studies have shown, however, that also animals show same kind of patterns in speech perception and processing (from now on referred to speech perception) as humans (Ramus, Hauser, Miller, Morris, & Mehler, 2000) as measured by EEG or other brain imaging techniques (Fitch, Miller, & Tallal, 1997). These studies reveal that processing speech until some stage seems to be a process which does not require language or cognitive processing; rather it is an obligatory (automatic) process of elements of language (i.e. sounds and phonemes) without the knowledge of language. On the contrary, Liberman (1998) and Liberman and Whalen (2000) suggested in their essay that speech indeed is special and speech stimuli from the very start is processed as speech and not just a sound with different acoustical features. In this thesis speech and non-speech stimuli were used to compare whether the speech perception and ERPs elicited by stimuli are induced by the acoustical differences between the stimuli (non-speech stimuli) or are the responses affected by long term memory traces also (i.e phonological representations, speech stimuli) of native language.

There are multiple theories concerning speech perception; nevertheless, the approach in this thesis is related to auditory and neural theories and models closely connected to them. Evaluating all the multiple theories is beyond this thesis, thus only the models concerning closely the objective of this thesis are described. The neural theories of speech perception examine perception in the brain level. Various terms are used in speech perception research and following are important in this thesis.

Formants of speech and vowels. Speech is an acoustic signal consisting of multiple co-occurring frequency bands, called formants. A formant is a peak in the frequency spectrum of a sound caused by acoustic resonance (Ladefoged, 2006). Vowel sounds consist of specific combinations of static frequencies whereas consonants contain more variable frequencies (Fitch et al., 1997; Ladefoged, 2006). Related to vowel perception, Nearey (1989) points out that four types of information is implicated in the perception of vowels. These are: static properties (formants and fundamental properties), dynamic properties (context), intrinsic relational properties (the relations between formants within vowels) and extrinsic properties (the relations to other vowels produced by same speaker or vowel duration changes). Speakers vary widely in their fundamental frequency (F0) of speech. Still a normal listener identifies phonemes despite this variation and it seems that no specific frequency is important in identifying phonemes. Rather, recognition depends on the relative combination of frequencies (Phillips, 2001). The effect of fundamental frequency (F0), context and dynamic characteristics of formants in perceiving vowels have been, nevertheless, shown in some studies (Aaltonen, Eerola, Lang, Uusipaikka, & Tuomainen 1994; Assmann & Nearey, 2003, Nearey 1989; Rosner & Pickering, 1994).

By definition, the information that human requires to distinguish between vowels can be represented purely quantitatively by the frequency content of the vowel sounds. The interest in this thesis was the processing of vowels. Considering the perception of vowels, the significant aspect is the presence of the formant frequencies whose manipulation in this study created the differences between stimuli. The formant with the lowest frequency is called F1, the second F2, which describes the F1-F2 ratio and the third is F3 (Ladefoged, 2006). Most often the two first formants, F1 and F2, are enough to disambiguate a vowel (Rosner & Pickering, 1994). These two formants determine the quality of vowels in terms of the open/close and front/back dimensions (which have traditionally, though not entirely accurately, been associated with the position of the tongue). In the Methods section the formants used in this thesis are described in detail (in Figure 2 vowel map of stimuli is presented).

Categorical perception, allophonic perception and phonological representations. The perception of speech requires categorical perception. Perception is considered to be categorical when discrimination is better for stimuli belonging to different categories than for stimuli belonging to the same category, even though physical differences are the same in both the cases (Liberman, Harris, Hoffman, & Griffith, 1957). Categorical perception thus points to the situation where a continuum is divided into categories and the units in the continuum are considered to be members of a particular subgroup (Aaltonen, Eerola, Hellström, Uusipaikka, & Lang, 1997). Aaltonen and colleagues (1997) for example examined the categorical perception of the neighbouring Finnish /i/ and /y/ vowels illustrating the known fact that these phonemes belong to separate categories in Finnish. Members of a category can be seen as allophonic exemplars of a certain phoneme (Ladefoged, 2006). For example, when producing vowel /i/, listener identifies the exemplar (or allophone) to be /i/ vowel (phoneme), whether it is produced by a person with different F0 frequency or different dialect. Processing speech and phonemes requires categorical perception, skill to make a difference between phonemes (Liberman et al., 1957). The enhancement of between-category differences and the reduction of within category differences are crucial for neglecting irrelevant variations in the stimuli and representing distinctions that are meaningful for speech perception (Serniclaes, Ventura, Morais, & Kolinsky, 2005). Categorical perception of speech could be seen as an innate property of humans (and animals) but also as the result of learning native language. According to Kuhl (Kuhl, 1991, 1993) the exposure to native language affects the formation of categories (described in detail later). Categorical processing thus seems to be a biological necessity for many species, including humans.

Discrimination of between-category differences is an important skill in understanding speech and learning to speak. Deficits found in dyslexic people show the importance of adequate discrimination skills. Dyslexic children have been found to be better in discrimination of within-category differences than typical readers (Serniclaes, Van Heghe, Mousty, Carré, Sprenger-Charolles, 2004). These categorical perception anomalies might be one causal factor of dyslexia. Serniclaes et al. (2005) investigated categorical perception of literate and illiterate adults to examine whether these anomalies are the consequence of the poor reading skills. They found that the illiterate and literate adults did not display the categorical perception deficit thus suggesting that deficits found in categorical perception are not a result of poor reading skills. In general, categorical perception was more accurate to between- than within-category pairs (Serniclaes et al., 2004; 2005). In their behavioral and PET study, Dufor and colleagues (Dufor, Serniclaes, Sprenger-Charolles, & Démonet, 2006) examined phoneme learning in dyslexic and typically reading adults. At first acoustic mode (sine wave synthesis) stimulus pairs were presented either between or within pairs of /ba/ and /da/ syllables. Second, the same stimulus pairs were presented as speech mode, i.e., either the same sine waves were perceived as syllables after training or pitch-modulated (F0 modulation) sine waves resembled synthesized speech. Categorical discrimination responses appeared after training in both conditions of speech mode. In addition they found differences between the dyslexic and normal adults in learning phoneme categorization: the typically reading adults were more accurate in categorizing between stimulus pairs. Finally, discrimination of between-category stimulus pairs (/ba/ vs. /da/) was more accurate in all conditions compared to within-category stimulus pairs (/da/ vs. /da/) in both groups.

In general, phonological representations can be seen as long-term memory traces formed by long-term exposure of native language phonemes during the development. However, phonological representations seem to be hard to detect and the existence of them might be abstract (Phillips, 2001). Furthermore, invalid or weak phonological representations in dyslexic readers have been considered as one of the major causes for dyslexia. Access to phonological representations, as stable entities stored in perceptual long-term memory, is a prerequisite to the categorical perception of speech. Moreover, it has been observed that the increase of processing time of speech might relate to accessing the phonological representations with which incoming stimuli had to be matched before a participant could make the decision; this top-down processing has an additional time cost effect (Dufor et al., 2006).

Categorical perception has been traditionally studied in behavioral settings but also electro-physical measures have been conducted (Aaltonen, Paavilainen, Sams, & Näätänen, 1992). Following paragraphs thus cover mainly ERP studies.

Models of speech assimilation and learning. Several models of speech perception have been formulated to explain speech assimilation and language learning. Perceptual assimilation model (Best et al., 1988) and Speech learning model (Flege, Schirru, & McKay, 2003) aim to explain speech perception in a comprehensive way, whereas the Native language magnet model (NLM) by Kuhl (1991, 1993) concentrates to explain the assimilation of certain phoneme categories. These models are described in more detail in the following paragraphs.

The Perceptual assimilation model and the Speech learning model. The Perceptual assimilation model (PAM) developed by Catherine Best (Best, Halle, Bohn, & Faber, 2003; Best, McRoberts, & Goodell, 2001; Best et al., 1995; Best et al., 1988) is one of the important models explaining how non-native segments are perceived. The PAM proposes three explanations for processing non-native language: non-native sounds can be perceived as exemplars of an existing native language phonetic category (either good or bad exemplars of it, “goodness of fit”), as exemplars of a sound category non-existing in native language or even as non-speech sounds. Also, Best et al. (1988) suggest that when the non-native phoneme contrasts (e.g. /r-l/ for Japanese) are perceived as exemplars of one native phoneme category, the discrimination between them is difficult. Predictions of PAM are mainly based on the (behavioral) studies of phoneme contrasts of languages either for consonants (Best et al., 2001; 1995) or vowels (Best et al., 2003) in adults (Best et al., 2003; 2001; 1988) or in infants (Best et al., 1995; 1988). The phonetic characteristics of language (i.e. position of tongue in producing speech, lip-rounding of phonemes in different languages) have also been considered in the assumptions of PAM. One study of Best et al. (2001) examined the perception of non-native vowels and the effects of listeners' native language. Discrimination ability of Norwegian vowels among American, Danish and French adults was measured. Syllables with naturally produced Norwegian phonemes were chosen: /si/, /sy/, /su/ and /sū/. It was supposed that participants of different countries will assimilate these syllables differently: American participants (English speaking) should assimilate Norwegian /i-u/ to their native /i-u/ with less than perfect fit. Danish and French should assimilate /i-u/ fairly good exemplars of their native /i-u/ continuum. Danish should discriminate /i-y/, /y-u/ and /u-ū/ excellently (as native contrasts /i-y/, /y-u/ and /u-y/, respectively). They should discriminate /y-ū/ well (goodness difference with native /y/) but less well than /i-y/, /y-u/ and /u-ū/. French should discriminate /y-u/, /y-ū/ and /u-ū/ excellently (as native /i-u/, /i-y/ and /u-y/), however, /i-y/ somewhat less well (goodness difference in native /i/). The PAM was able to predict these discrimination abilities correctly. One surprising result was that Danish discriminated the /y-ū/ contrasts well, given that they detected only modest goodness difference of the contrast (i.e. they did not perceive the exemplars of contrasts to be good exemplars of these vowels). Furthermore, Polka and colleagues (Polka et al., 2001) results assume that some part of speech perception is not influenced only by the native language but also by the acoustical characteristics of sounds, specifically the extreme F1/ F2 values in vowel space. Therefore some discrimination ability to non-native sounds, even tough, not relevant in native language, remains intact.

The Speech learning model (SLM) (Flege, Bohn, & Jang, 1997; Flege & McKay, 2004; Flege et al., 2003) shares the implications with the PAM but it proposes that the exemplars of PAM are most effective in initial stages of non-native (second language learning) acquisition. The SLM addresses primarily the issues of learning a new language. The SLM have been studied with natural tokens of languages (Flege et al., 1997). Similarly to the PAM, the SLM assumes that non-native phonemes can be assessed as members of existing native category, either as good or bad representatives of it. Learners of a new language can also formulate new categories for new phonemes which do not exist in their native language.

The Native language magnet and the Perceptual magnet effect. The Native language magnet model (NLM) developed by Patricia Kuhl (1991, 1993) considers the perception to be based on general auditory properties which are common to man and other species i.e. the auditory perception in general is not unique to humans. The main emphasis is on the language-specificity of speech perception which originally relies on general perceptual ability to categorize and it is elaborated by the experience of the native language (Iverson et al., 2003; Kuhl, 1991, 1993). This language-specific perception is based on the formulation of perceptual magnets, namely the category prototypes (i.e. phonological representations), which develop during early childhood as a result of the influence of mother tongue (Cheour, 1998). According to this view of perception, the acoustic input is compared with the existing prototypes of categories, and thus it takes the position that there is a hierarchical structure within the perceptual categories. In a nutshell, the prototypes work as a basis for perception. The prototypes or the centres of the categories are developed when the most often perceived category members (allophones of phonemes) form them.

From the neuro-cognitive point of view this formulation can be explained as follows: memory traces are formed for perceived phonemes of speech, and the more frequent they are heard, the stronger the synaptic connections becomes. The NLM and Perceptual magnet effect (PME) accentuate the role of category centres or prototypes, which are compared to magnets as they are pulling the sounds in the immediate vicinity towards the centre, thus inhibiting the acoustic difference between the prototype and the neighboring sounds from being discriminable. As a consequence, category coherence is enhanced. When the category members become more distant from the centre, the magnet loses its grip and finally another magnet for another category takes hold of the speech sound. Discrimination of the phonemes near the prototype is more difficult than phonemes further from centre (Kuhl, 1991).

Relationship with the NLM and PAM emerge from theoretical assumptions of models: the PAM presents that non-native sounds can be perceived as exemplars of an existing native language phonetic category. According to the NLM these non-native sounds are processed as non-prototypes or allophones of a certain phoneme category wherein native prototype serves as a magnet.

Even though, the findings of native language magnet seem convincing there are diverging results, as well. For example, Frieda and colleagues (Frieda, Walley, Flege, & Sloane, 1999) showed in their study that personal prototypes exist and the dialect used by participants also affects to the prototype chosen. Rosner and Pickering (1994) also proposed that personal prototypes exist for speech vowels and prototypes formulated by natural speech might not be working when stimuli used in the studies are semisynthesized speech. Lotto, Kluender and Holt (1998) observed that vowels are affected by contextual sounds thus the prototypes shift and change in different contexts (i.e represented in isolation or in pairs, making category goodness judgments or detecting differences). Kazanina, Phillips and Ibsardi (2006) in their crosslinguistic study of the speakers of Russian and Korean showed instead that the meaning of phonemes affects categories: when the phoneme had a distinctive role in the meaning, it was separated to two different phoneme categories. In a study of Aaltonen et al. (1997) two different groups of categorizers were found: a group which prototype for /i/ had a lower F2 value (Good categorizers) and a group with a high F2 value (Poor categorizers). The group with the lower value was more accurate in categorizing stimuli behaviorally and the results of this group have been interpreted as a perceptual magnet effect. The high value group failed to show the magnet effect. Best et al. (2001) also showed that the discrimination of the sounds away from the prototype was actually worse than discrimination of the sounds near the prototype (i.e. discrimination of prototypical phoneme contrasts) as opposite to the NLM. The results of Sharma and Dorman (1998) indicated that discrimination of prototype-like stimulus pair (/i/ vowels) was not worse than discrimination of non-prototype-like pair, which, according to NLM should be easier to detect.

Auditory event-related potentials as a method for studying speech perception

In neuropsychology one of the most popular ways of examining brain processes is measuring auditory event-related potentials (ERPs) Temporally, ERPs are very accurate; one can see the changes in ERP waveform within milliseconds. However, the ERP generators are difficult to disentangle from ERP data because the electric activation spreads over the scalp (for a reviews see Dehaene-Lambertz & Gliga, 2004; Kraus & Cheour, 2000) and thus the spatial resolution is not very good.

ERPs are changes in electroencephalogram (EEG) produced by stimuli. Generally ERP components are divided into two groups: exogenous and endogenous (Coles, Gratton, &

Fabiani, 1990). Exogenous components develop because of the physical characteristics of stimuli and they are also called obligatory responses to stimuli. It has also been proposed that (obligatory) auditory ERPs of latencies earlier than approximately 100 ms cannot reflect language specific differences and perception is auditory as general (Eggermont & Ponton 2002). These components occurring 50 ms after a stimulus presentation are also called long-latency components. Endogenous components are related to cognitive processes and are determined by task or situation. They are quite independent of physical characteristics of stimuli. This division between exogenous and endogenous components is partly outdated because both components depend on each other (Näätänen, Paavilainen, Alho, Reinikainen, & Sams, 1989). ERPs are series of negative and positive waves and these components are typically named after their polarity (P for positivity; N for negativity), latency or the order of appearance (used in this thesis). Perception of stimulus differences or stimulus change, in both humans and animals, is strongly correlated with the amplitude, latency and other parameters of various surface-recorded evoked potentials (Eggermont & Ponton, 2002). Related to this thesis, the processing of speech and non-speech differences for example in amplitude, latency and locations of the ERP components are considered.

Auditory stimuli produce variety of long-latency components. The first peak in ERP is called P1 and the first negative deflection occurring is N1 (Čeponienė, Rinne, & Näätänen, 2002). The N1 is followed by mismatch negativity (MMN) when there is a change in the repetitive stimuli. Other components elicited by auditory stimulus are for example P2 and N2 (Luck, 2005). Attention specific components (and thus quantified as endogenous components) found are P3a, N4 and late discriminative negativity (LDN) (Luck, 2005). The main emphasis in this thesis was on the N1 and MMN components and therefore they are described in more detail in following paragraphs.

N1 in speech perception. N1 is the first major negative peak in ERP data, occurring about 100 ms after the stimulus onset. The N1 is an obligatory component and it is presumed to be linked with general conscious auditory sound level change detection and orienting of attention (Näätänen & Picton, 1987). The N1 is elicited by simple repetitive auditory stimuli such as tones or syllables (Näätänen & Picton, 1987). The N1 has several distinct subcomponents. The supratemporal subcomponent is generally observed as negativity at frontocentral site reversing polarity over the Sylvian fissure and being positive at temporomastoid areas (Näätänen & Picton, 1987). Other identified components are T-complex; generated in the association cortex in the superior temporal gyrus and non-specific N1 component; negativity at the vertex (Luck, 2005). In some studies (Čeponiené, Torki, Alku, Koyama, & Townsend, 2008) the amplitude of N1 has been seen larger to non-speech syllables than speech syllables (sinusoidal tones with the corresponding formant frequencies, semisynthetic syllables and consonant-vowel transitions, respectively).

Mismatch negativity (MMN). MMN is a frontal negative deflection in the human event-related potential that typically appears when repeated auditory stimulus changes in some manner (Näätänen, 1992; Näätänen & Picton 1987). It was first found by Näätänen, Gaillard & Mäntysalo (1978). The MMN usually peaks around 100-250 ms after stimulus-change onset and its duration varies between 150-300 ms. Stimulus changes can vary from simple changes in a single stimulus feature to abstract changes in the relationship between stimuli (Picton, Alain, Otten, Ritter, & Achim, 2000). The MMN amplitude generally increases when there is an acoustical difference gain between standards and deviants (Aaltonen et al., 1994). Lower amplitudes in the MMN elicited by vowels have been observed in a comparison to pure tones (Aaltonen et al., 1994). The MMN is classically measured in the passive oddball paradigm where the participant pays no attention to the stimuli. The MMN response is detectable also in infants although the polarity of the response can differ from adults (for reviews see Cheour, Leppänen, & Kraus, 2000 and Dehaene-Lambertz & Gliga, 2004, see also Maurer, Bucher, Brem, & Brandeis 2003).

Relation of the N1 and MMN. The N1 and MMN are linked to each other and distinction between them can be challenging. The N1 is a clear peak in the ERP data whereas the MMN is seen as a slope after the N1 between standard and deviant stimulus response in adults. The MMN has been studied extensively since after its discovery. The MMN exists separately without N1. In their review Picton et al. (2000) portrayed elegantly the five main differences of the N1 and MMN. First, decreasing intensity of stimulus diminish the N1 response but the MMN remains intact, furthermore the duration of the stimuli affects the size of MMN response more than N1. Second, the N1 responses tend to be more susceptible for changes in the inter-stimulus-interval (ISI) than MMN. Third, the N1 component reacts to physical changes of stimuli. The magnitude of N1 response increases when the physical characteristics of stimulus increase (e.g. intensity). On the other hand, the MMN can be larger for smaller changes, if the change is relevant (Näätänen, 1992). This is supposed to occur because of short-term memory or memory trace effects. Fourth, the difference between standard and deviant stimulus responses affects the MMN latency while the latency of N1 response remains the same. Fifth, the generators of MMN and N1 seem to be quite distinct: scalp distribution for these components is slightly different, the MMN being more anterior than the N1 (Tiitinen et al., 1993).

Generators and spatial distribution of the MMN. The MMN is seen as negativity in the frontocentral areas and as positivity below the supratemporal auditory cortex (Alho, 1995). It has been discussed and researched if the frontal negativity and the temporal positivity seen in the MMN represent different functional separation processes. The change detection process indexed by the MMN involves a network of brain areas with distinct roles: a superior temporal generator associated with processing the sensory input against a formed memory trace, and a frontal generator related to triggering of involuntary attention (Dual generator model; Näätänen & Mitchie, 1979; Paavilainen et al., 2003). The positivity in the supratemporal cortex emerges before the frontal negativity (Rinne, Alho, Ilmoniemi, Virtanen, & Näätänen, 2000). Shalgi and Deouell (2007) found support for the Dual generator model in their ERP research. They made an effort to functionally distinguish the frontal and temporal MMN components from each other by looking at their modulation by specific task conditions: ignore, attend-spatial and attend-pitch. Temporal activity of MMN was affected in different conditions but frontal activity remained unaffected (same) between conditions. Their results also postulated that frontal generator might not be dependent on temporal generator.

Response to the auditory non-speech stimulus is shown usually more in the right auditory cortex (Tervaniemi, Radil, Radilova, Kujala, & Näätänen, 1999) and phoneme/language related stimuli more in the left auditory cortex (Möttönen et al., 2006). In some studies such differences have not been found (Aaltonen et al., 1994). Results of laterality effects of MMN come mainly from MEG and hemodynamic studies (see for example Dehaene-Lambertz et al., 2005). In fMRI condition Dehaene-Lambertz and colleagues (2005) found asymmetries of activation in speech versus non-speech stimuli: despite of the same physical characteristics of the stimuli the activation to the speech stimuli was more asymmetric favoring the left thalamus and the left supramarginal gyrus. This result suggests that general auditory characteristics are not sufficient to explain hemispheric asymmetry on phoneme perception. The left hemisphere presents a specialization for speech stimuli that goes beyond stimulus-driven characteristics.

Top-down processing affecting the MMN. While the MMN is hold as an automatic change detection process, which arises when sensory memory trace formed by standard stimulus is interrupted with rare deviant stimulus, another point of view to understand the speech perception of native language or specially vowel perception is to study the effects of higher (top-down) processes on MMN (see Näätänen, 2001 for review). Phonological top-down processes are of interest. Dehaene-Lambertz and colleagues (2005) found some supporting results for top-down processes in their ERP-fMRI study. They studied the changes in brain activity in a discrimination paradigm. The aim of their study was to compare the activation in speech and non-speech stimuli i.e., when the same auditory stimuli (sine waves) were perceived as speech (phonemic stimuli). They found out in a ERP condition that a mismatch response (MMR) for the phonemic contrast (across-category stimuli group) was fast and extended over 100 ms whereas in the acoustic contrasts (equivalent version of phonemic stimuli), the onset of significant MMR was delayed and the MMR duration shorter. These results are in line with those previous studies that have shown that a longer duration and an earlier onset of mismatch responses are correlated with better discrimination performances (Kujala, Kallio, Tervaniemi & Näätänen, 2001). The study of Kujala et al. (2001) also revealed that the MMN was significant only when the standard and deviant stimuli were also behaviorally discriminated.

The MMN has been shown to change as a result of learning (Winkler et al., 1999). When identifying stimulus pair of Finnish /e/ and /ae/ synthesized vowels Finnish participants and Hungarians, who were fluent in Finnish, accurately categorized the stimuli to two different categories. Furthermore, the native Hungarians failed to categorize the stimulus pair; rather it was proposed that stimulus pair of two distinct Finnish phonemes were categorized to one Hungarian phoneme category of vowel /ɛ/ (Winkler et al., 1999). In addition, Winkler and colleagues found that the MMN for a contrast between two Finnish phonemes (/e/ and /ae/) was elicited in the fluent Hungarians and not in the native Hungarians (without the knowledge of Finnish). As a conclusion, they suggested that the Hungarian participants who have not learned to discriminate the two Finnish vowels do not show MMN in the ERPs (Winkler et al., 1999). MMN has shown to appear when the participant learned to discriminate between two similar-sounding stimuli (synthetic speech stimuli of /da/ phoneme) after training (Kraus et al., 1995). Moreover, training in the discrimination of stimuli resulted in changes in the duration and magnitude of MMN response (earlier onset and longer duration) with participant initially showing the MMN between the standard and deviant stimulus /da/ (Kraus et al., 1995).

Finally, Aaltonen et al. (1997) showed that the MMN response was larger for prototype /i/ than for non-prototype /i/ in Finnish participants. Equal differences between the standard and deviant stimuli elicited lower MMN responses when the standard was the prototype /i/ and deviant being non-prototype /i/, whereas in a condition where the standard stimulus was the non-prototype /i/ and prototype /i/ served as deviant a clear MMN response was elicited to the deviant stimulus. This phenomenon was seen only in Good categorizers. Poor categorizers failed to show similar differences between prototype and non-prototype condition, rather the responses seemed opposite.

Crosslinguistic studies as a method for investigating speech perception

Comparing speakers from different languages enables detection of language specific memory traces (i.e. effects of native language) and it enables the exploration of the influence of language to automatic processes which do not demand attention toward the stimuli. Most of the studies have concentrated on between-category stimuli or very specific aspects of language: duration in Finnish (Tervaniemi et al., 2006) or specific phonemes and phoneme categories in certain languages (Näätänen et al., 1997; Sharma & Dorman, 2000).

General findings of the long-term native language exposure effects to speech perception.

Zhang, Kuhl, Imada, Kotani and Tohkura (2005) compared native speakers of Japanese and native speakers of American English. The stimuli were two synthesized continua, /ba-wa/ and /ra-la/. In Japanese /r/ and /l/ consonants are not separate phonemes and thus do not exist as a phoneme categories in adult native Japanese speakers. In the study, all participants showed similar performance for the /ba-wa/ continuum and striking differences for the /ra-la/ continuum. Identification curve for the Japanese participants was continuous without separate categories whereas Americans showed distinct categories for these phonemes. Zhang, Kuhl, Imada, Kotani and Stevens (2001) and Iverson et al. (2003) found similar results in their /r-l/ continuum study (with Japanese and American adults; Japanese, German and American adults respectively).

The MMN as an index of memory traces for phonemes. As clarified earlier, the MMN provides sensitive measure of timing and cortical utilization of different stages of cognitive processing (for a review see Näätänen, 2001). The following studies explain the utilization of MMN in crosslinguistic studies finding the effects of native language exposure.

Finnish is a so-called quantity language, there is a phonological distinction between short and long phonemes and the quantity degrees thus distinguish words with different meanings. Ylinen, Shestakova, Huotilainen, Alku and Näätänen (2006) examined MMN in crosslinguistic setting where they compared native Finnish speakers, Russian speakers with Finnish as a second language and native Russians without any exposure to Finnish. The stimuli were represented in word condition (/tuku/ or /tuuku/) or as isolated vowels (/u/ or /uu/) as within- or between-categories. The main hypothesis in their study was that MMN is enhanced in the Finnish speakers because their knowledge of Finnish phonemes and phonological categories. The MMN response was larger among the native Finnish speakers than in two other groups, however they did not find a solid support for their hypotheses: explicit differences between the groups were not found in the response pattern of across- and within-category changes in the MMN response. In addition, duration differences were more detectable in word condition than in isolated vowel condition.

Language-specific categories of phonemes affecting the MMN. The study of Näätänen and his colleagues (1997) examined the between category differences. They examined the MMNm (the magnetic equivalent of the electric mismatch negativity) and the MMN (ERP) in Finnish and Estonian adults in a paradigm consisting of phonemes which occurred in both languages (/o/, /e/ and /ö/) or only in the other (/õ/). Phonemes /o/, /e/ and /ö/ are common in Finnish and Estonian whereas /õ/ is part of the Estonian phoneme system and does not exist in Finnish. The standard stimuli in this passive oddball paradigm were phonemes /e/ and /o/ and deviants were /ö/ and /õ/. Results showed that the MMN response to the deviant stimulus /õ/ vowel (/e/ as standard stimulus) was larger in the Estonian participants compared to the Finnish participants. Researches explained the smaller amplitude of MMN response in the Finnish participants by the lack of long-term memory trace for /õ/ phoneme. Studies of MMN in crosslinguistic settings showed the MMN in the Finnish participants but not in the native Hungarians, who did not speak Finnish at all (Winkler et al., 1999; Winkler, Kujala, Alku, & Näätänen, 2003). Similarly, Rivera-Gaxiola, Csibra, Johnson and Karmiloff-Smith (2000) observed that native English speakers lacked the MMN response to phoneme contrasts (between-category syllables in oddball paradigm) relevant in Hindi.

Peltola et al. (2003) also examined the effects of native language vowel perception in a comparison between native Finnish speakers, advanced Finnish students of English and native English speakers without the knowledge of Finnish. Stimuli used were vowels rated as good category representatives by the native speakers and the oddball paradigm used consisted of contrasts (between-category stimuli) of native language (e.g. Finnish /i-e/) and foreign language (English) contrasts (/i-e/, /i-I/ and /e-I/). The study showed that the brain responses to speech stimuli were larger for the native language contrasts in a comparison with foreign language ones with the same acoustic distance between the stimuli. This was most evident in the MMN responses of the native Finnish speakers, but it was also seen in the native English speakers. Also in a study of Dehaene-Lambertz (1997) significant MMN response was found to the change which crossed a phonemic boundary of phonetically relevant category. Studies above offer convincing evidence that lacking of MMN response or having a decreased response, reveals that the brain does not process foreign language phonemes same as native language phonemes.

Support for the Native language magnet model, Perceptual magnet effect (PME), Perceptual assimilation model (PAM) and Speech learning model (SLM) have been puzzling and the evidences supporting them are mainly based on behavioural measures. Thus, the speech perception measured at the brain function level needs further investigation. ERP studies described above have mainly examined certain specific phonemes (e.g. quantity in Finnish) existing only in other examined language; contrary to that stimuli used in this thesis were vowels naturally existing in Finnish phoneme system.

The objective of this Master’s thesis. The objective of this thesis was to use exemplars of different language (Finnish, Hungarian, French and German) phonemes in one paradigm and investigate speech perception and the effects of native language exposure on the MMN and N1 in typically reading Finnish adults with the long-term exposure only to Finnish. This master’s thesis is a part of the ongoing NeuroDys research and it is described in detail in the Methods section.

The hypothesis of this Master's thesis was that all the selected deviant stimuli (/y/) were expected to fall to one phoneme category where one deviant stimulus served as a native magnet/prototype for Finnish participants and other deviant stimuli worked as exemplars of non-prototype, as suggested by NLM. Similarly, implications of PAM and SLM foreshadowed that all deviant stimuli were exemplars of one category in Finnish and the members of that should have been processed either good or bad representatives of it, whereas the selected standard stimulus (/i/) belonged to different phoneme category than the deviant stimuli. Comparison of the responses to non-speech stimuli enabled to evaluate the effect of physical distance and acoustical cues. This within- and between-category paradigm thus deviated partly from some earlier studies where within- or between-category paradigm has been used (Aaltonen et al., 1997; Näätänen et al., 1997). The paradigm thus enabled the novel perspective for evaluation of existing models of speech perception.

To summarize, according to the NLM, it was hypothesized that the responses elicited by deviant stimuli differ and the prototype stimulus elicits distinguishable response in a comparison with non-prototypes. Whereas, the PAM and SLM lead to propose that the deviant stimuli can be processed very similarly. Predictions of PAM and SLM concerns only behaviorally examined speech sounds and thus the direct assumptions concerning ERPs or non-speech sounds are not straightforward.

METHODS

This study is a part of the ongoing research project NeuroDys (Dyslexia genes and neurobiological pathways; European Dyslexia Research, www.neurodys.com) in which ten European countries have participated. Aim of the NeuroDys is to examine etiology and the biological basis of dyslexia. NeuroDys project is divided into three subfields: genetics, environment and neuroscience. EEG studies of dyslexia belong to neuroscience subfield. There are four universities taking part in EEG studies: The University of Jyväskylä, Finland (the coordinator of ERP studies); The University of Toulouse, France; Ludwig-Maximilians-Universität München, Germany and Institute for Psychology of the Hungarian Academy of Sciences, Hungary. As a result of the collaboration of these universities and some other researchers, all the EEG labs are using the same cross-linguistic paradigm and research protocol. The exactly same paradigm and protocol give opportunity to inspect dyslexia (and speech processing in general) in the crosslinguistic setting. The main participants in NeuroDys research were children. Pilot studies with dyslexic and typically reading adults were carried out to examine the functionality of the paradigm.

Participants

The participants of this Master's thesis were adults who were pilot participants for the NeuroDys study in Finland. The piloting experiment was carried out to ensure that chosen crosslinguistic paradigm works in general in typically reading adults. For this thesis, 14 typically reading adults participated (9 females and 5 males; age range: 33.7-64.0 years, mean age: 44.6 years and SD: 8.36 years). All the participants belonged to a subgroup (control group) of Jyväskylä Longitudinal Study of Dyslexia (JLD) (Lyytinen et al., 2004) and they were recruited for the project between years 1993 and 1996 in the Province of Central Finland (for detailed information of participant selection criteria of JLD, see Leinonen et al., 2001). The participants did not have any reading problems or familial background of reading difficulties (dyslexia in their first degree relatives). None of the participants reported neurological problems or hearing deficit.

Hearing ability. Participants' hearing ability for both ears was tested individually before the ERP experiment. Frequencies tested in hearing ability test followed the recommendations of

The American Speech-Language-Hearing Association (ASHA) and BSA Hearing and Speech centre (www.asha.org; www.bsahearing.com). The Hearing ability for following frequencies was tested: 250, 500, 1000, 2000, 4000, 6000 and 8000 Hz. If the participant failed to hear the sounds below 20 dB, it was estimated whether this occurred in both ears in several frequencies important in speech perception (250, 500, 1000, 2000 and 4000 Hz). The average of mentioned frequencies was then calculated for both ears and the threshold of average was set to 20 dB. The average for all 14 participants was 4.8 dB (range: -3.75-12.50 dB). Participants above this threshold were excluded (initially 15 participants were selected but one was excluded after hearing ability testing).

Participants' background information was gathered by a questionnaire. Exclusion criteria for the ERP experiment were an exposure to French or German either living in French or German speaking countries (as an adult over six months or over one year in childhood), one parent speaking mentioned languages to subject in childhood or majoring these languages (exposure to Hungarian was not an exclusive criterion because the Hungarian /y/ used in this study were same in Finnish). Neurological problems, psychological disorders or medication for mental problems were also reasons for exclusion.

Stimuli

Stimuli used in the ERP paradigm of this thesis were speech vowels of Finnish-Hungarian, French and German and equivalent non-speech versions of them. The stimulus set consisted of four synthesized vowels (three /y/ vowels and one /i/ vowel), and four complex non-speech stimuli consisting of five sine tones located at frequencies corresponding to the lowest five (F1-F5) formants of the synthesized vowels. Distances between stimuli were defined and calculated by Euclidean distance. The distinctive features between stimuli were formant frequencies and detailed information of them is provided later. Quality of the stimuli was also examined by the stimulus related questions which were asked after the ERP experiment. In this section, the creation of the stimuli is described in detail; Euclidean distance is described later in its own paragraph.

Creation of the stimuli. Natural vowels of several women were recorded in Finland, France, Germany and Hungary. The formant frequencies of vowel samples were then utilized in the creation of the synthetic speech and non-speech stimuli. Stimuli were synthesized by Jyrki Tuomainen, PhD (University of Turku) by Praat software (v. 4.5; Boersma & Weenink, 2006; www.praat.org). In an initial stage, a set of 135 synthetic vowels were created covering the formant space of the /y/ vowel and the surrounding vowels of Finnish, French, German and Hungarian (120 stimuli categorized as /y/, and 15 qualified as /i/, /e/, or /oe/). These vowels were then used to evaluate phoneme boundaries and goodness (acceptability) of the vowel stimuli by listeners (ca. 20 from each participating country) of the four languages. From this set, all stimuli that were identified as /y/ with 95% accuracy were further inspected for their goodness values. For each four country groups, one of the vowels receiving the highest goodness ratings was selected to be used in the ERP paradigm as a representative of the country. The results indicated that the Finnish and Hungarian listeners preferred the same vowels, thus only three /y/-vowels were chosen: Finnish-Hungarian /y/ (y-fh), French (y-fr), and German (y-ge). Furthermore, one Euro-i vowel was synthesized representing the average formant frequencies of a typical /i/ vowel over the four languages. Stimuli were created using several formant frequencies but the description in this thesis focuses on formants distinguishing the /y/ stimuli (for the Euro-i vowel 4 additional formants were used: F6, 6500; F7, 7500; F8, 8500, and F9, 9500 Hz. For all /y/ vowels, 5 additional formants were used, F6, 5500; F7, 6500; F8, 7500; F9, 8500; F10, 9500 Hz).

The speech stimuli were synthesized using the Praat software. The sounds were synthesized to resemble natural speech sound in following actions: The glottal source was created by converting the pitch and timing information to a glottal source signal (0.1% noise was added to make the signal sound more natural). The duration of the source signal was 150 ms and the pitch fell linearly from 230 Hz at the onset to 200 Hz at the offset (mean pitch 215 Hz). Female pitch characteristics were used.

The non-speech stimuli were created using Praat software by synthesizing five separate sine wave tones at the frequencies corresponding to the first five formant peaks used in the synthesis parameters of speech stimuli. The amplitudes of the sine tones were matched according to values obtained by directly measuring the formant amplitudes of the selected synthesized vowels with Praat. Finally, all five sine tones were then combined to create one complex tone (i.e. one stimulus).

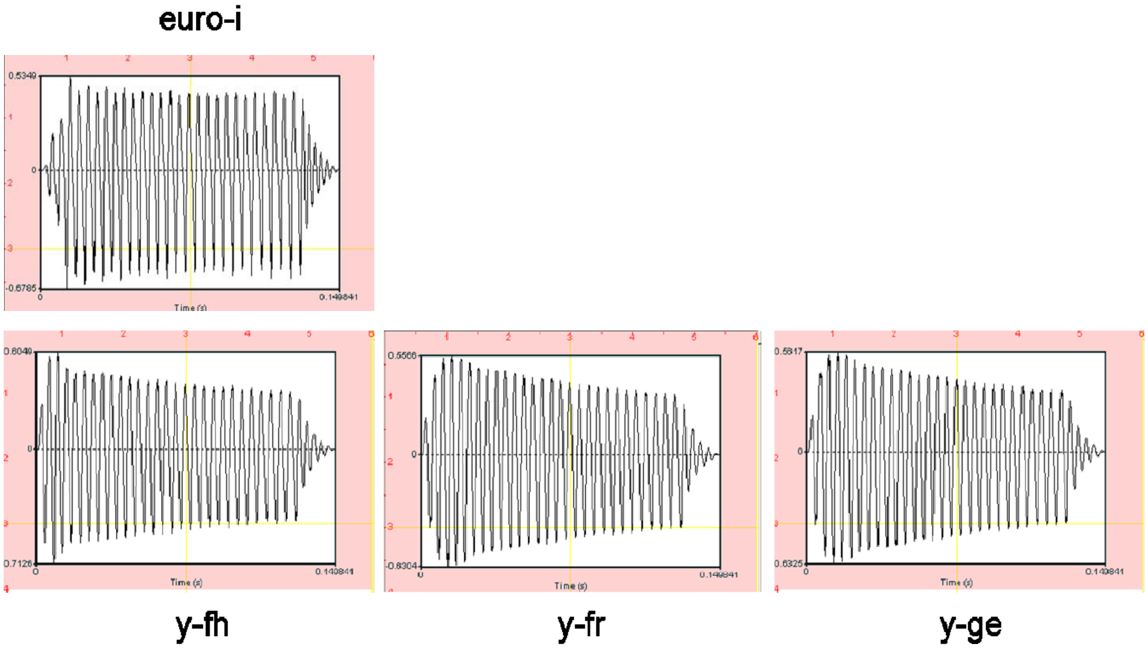
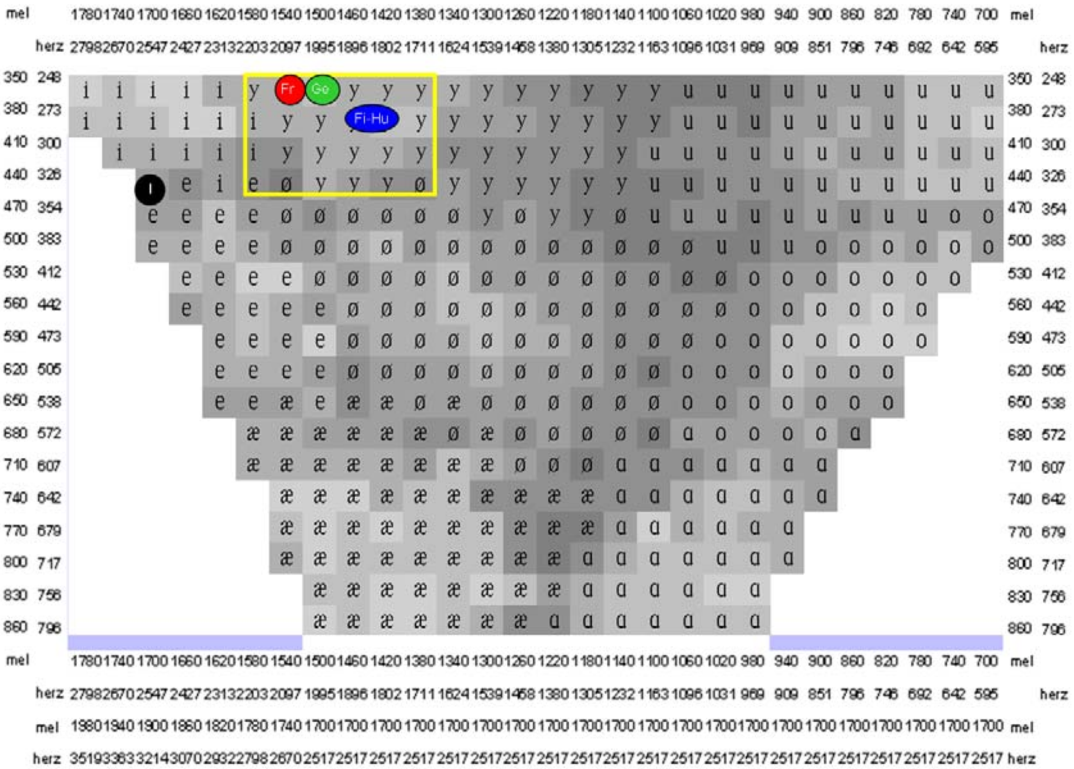


Figure 1. The stimulus waveforms showing all formant frequencies for the 4 stimuli. The duration of all stimuli is 150 ms.

Formant frequencies and the differentiation of the stimuli. F1 and F2 formant frequencies (the peak values of formants) were used to distinguish stimuli from each other, whereas F3, F4, F5 had same frequencies in all /y/ stimuli (though differing from Euro-i). Finnish-Hungarian deviant differed from French and German deviants in both frequencies. See the values of the lowest formant frequencies on the table below.

Table 1. The first five formant frequencies (as Hz) for the stimuli used. Euclidean distance shows the distance between Euro-i and each deviant stimulus as Hz.

Formant frequency	F1	F2	F3	F4	F5	Euclidean distance
Euro-i	335	2638	3500	4500	5500	
y-fh	274	1886	2400	3500	4500	755
y-fr	250	2086	2400	3500	4500	559
y-ge	250	2018	2400	3500	4500	626



mel		1780	1740	1700	1660	1620	1580	1540	1500	1460	1420	1380	1340	1300	1260	1220	1180	1140	1100	1060	1020	980	940	900	860	820	780	740	700	mel	
	herz	2798	2670	2547	2427	2313	2203	2097	1995	1896	1802	1711	1624	1539	1468	1380	1305	1232	1163	1096	1031	969	909	851	796	746	692	642	595		herz
350	248	i	i	i	i	i	y	Fr	Ge	y	y	y	y	y	y	y	y	y	y	u	u	u	u	u	u	u	u	u	u	350	248
380	273	i	i	i	i	i	i	y	y	Fi-Hu		y	y	y	y	y	y	y	y	u	u	u	u	u	u	u	u	u	u	380	273
410	300		i	i	i	i	i	y	y	y	y	y	y	y	y	y	y	y	u	u	u	u	u	u	u	u	u	u	u	410	300
440	326			I	e	i	e	ø	y	y	y	ø	y	y	y	y	y	y	u	u	u	u	u	u	u	u	u	u	u	440	326
470	354			e	e	e	e	ø	ø	ø	ø	ø	ø	y	ø	y	y	ø	u	u	u	u	u	u	u	u	u	o	o	470	354
500	383			e	e	e	e	ø	ø	ø	ø	ø	ø	ø	ø	ø	ø	ø	ø	ø	u	u	u	o	o	o	o	o	o	500	383
530	412				e	e	e	e	ø	ø	ø	ø	ø	ø	ø	ø	ø	ø	ø	ø	ø	o	o	o	o	o	o	o		530	412
560	442				e	e	e	e	e	ø	ø	ø	ø	ø	ø	ø	ø	ø	ø	ø	ø	o	o	o	o	o	o			560	442
590	473					e	e	e	e	ø	ø	ø	ø	ø	ø	ø	ø	ø	ø	ø	o	o	o	o	o	o	o			590	473
620	505					e	e	e	e	ø	ø	ø	ø	ø	ø	ø	ø	ø	ø	o	o	o	o	o	o	o				620	505
650	538					e	e	æ	e	æ	æ	ø	æ	ø	ø	ø	ø	ø	ø	o	o	o	o	o	o	o				650	538
680	572						æ	æ	æ	æ	æ	æ	ø	æ	ø	ø	ø	ø	ø	ɑ	o	o	o	o	ɑ					680	572
710	607						æ	æ	æ	æ	æ	æ	æ	æ	ø	ø	ø	ɑ	ɑ	ɑ	ɑ	ɑ	ɑ	ɑ						710	607
740	642							æ	æ	æ	æ	æ	æ	æ	æ	æ	æ	ɑ	ɑ	ɑ	ɑ	ɑ	ɑ	ɑ						740	642
770	679							æ	æ	æ	æ	æ	æ	æ	æ	æ	æ	ɑ	ɑ	ɑ	ɑ	ɑ	ɑ							770	679
800	717							æ	æ	æ	æ	æ	æ	æ	æ	æ	ɑ	ɑ	ɑ	ɑ	ɑ	ɑ	ɑ							800	717
830	756								æ	æ	æ	æ	æ	æ	æ	æ	ɑ	ɑ	ɑ	ɑ	ɑ	ɑ								830	756
860	796								æ	æ	æ	æ	æ	æ	æ	ɑ	ɑ	ɑ	ɑ	ɑ	ɑ	ɑ								860	796
mel		1780	1740	1700	1660	1620	1580	1540	1500	1460	1420	1380	1340	1300	1260	1220	1180	1140	1100	1060	1020	980	940	900	860	820	780	740	700	mel	
	herz	2798	2670	2547	2427	2313	2203	2097	1995	1896	1802	1711	1624	1539	1468	1380	1305	1232	1163	1096	1031	969	909	851	796	746	692	642	595		herz
	mel	1980	1940	1900	1860	1820	1780	1740	1700	1700	1700	1700	1700	1700	1700	1700	1700	1700	1700	1700	1700	1700	1700	1700	1700	1700	1700	1700	1700	mel	
	herz	3519	3383	3214	3070	2932	2798	2670	2517	2517	2517	2517	2517	2517	2517	2517	2517	2517	2517	2517	2517	2517	2517	2517	2517	2517	2517	2517	2517	herz	

Figure 2. The Figure shows the stimuli in F1 and F2 formant frequencies. The column on the left shows F1 values in hertz and the left row shows F2 values. The vowels inside the (yellow) box are the vowels used in stimulus selection. Black /i/ stands for Euro-i standard, red Fr for French /y/, green Ge for German /y/ and blue Fi-Hu for Finnish-Hungarian /y/ deviant. See also the Turku vowel test; http://fon.utu.fi/ for Finnish vowel map).

Euclidean distance. The distances between Euro-i and /y/ vowels were calculated by Euclidean distance. Euclidean distance is a measure which counts the distances between factors in multidimensional space (the formula below is for two-dimensional distances):

$$d_e = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}.$$

In phonetics the values of formant frequencies can be seen as coordinates in multidimensional space and in this study as coordinates of two-dimensional space (See also the Finnish vowel map in Figure 2). The Euclidean distance is counted as follows: F1 formant values of two stimuli are used as x-values and F2 values in y-values in the formula above. When measured with the Euclidean distance the distance between Euro-i and y-fh was the longest, y-fr and y-ge vowel were closer to Euro-i. The distances between Euro-i and /y/ stimuli are given in Table 1.

Paradigm

The stimuli were presented in a classic passive oddball paradigm. The whole ERP experiment consisted of Speech, Non-speech, and Equal probability condition. A standard stimulus was Euro-i and /y/ vowels were used as deviant stimuli.

The stimulus representation was divided into separate blocks of each language. Non-speech bocks were always presented first in a randomized order (y-fh, y-fr or y-ge first), then speech blocks followed in the order of non-speech blocks. Presenting the speech stimuli first might have induced the participants to process non-speech stimuli as vowels. Extra blocks were collected if the quality of the data were not good enough (for two participants in this study). Each language (either speech or non-speech) condition consisted of a total of 880 stimuli: 748 Euro-i standards (85%) and 132 (15%) of one of the three deviant stimuli (132 of y-fh, y-fr and y-ge). Each language was presented separately in its own condition (both Speech and Non-speech condition yielded the total of 5280 stimuli: 4488 Euro-i standards and 792 /y/ deviants). Equal probability control condition consisted of eight stimuli (speech and non-speech stimuli), all with 12.5 % probability. The offset-to-onset interstimulus interval (ISI) varied randomly between 350-450 ms. Onset-to-onset (stimulus onset asynchrony, SOA) changed randomly between 500-600 ms. Stimulus duration was 150 ms. Figure 3 below illustrates the presentation of stimuli in the oddball paradigm and shows ISI, SOA, and stimulus duration.

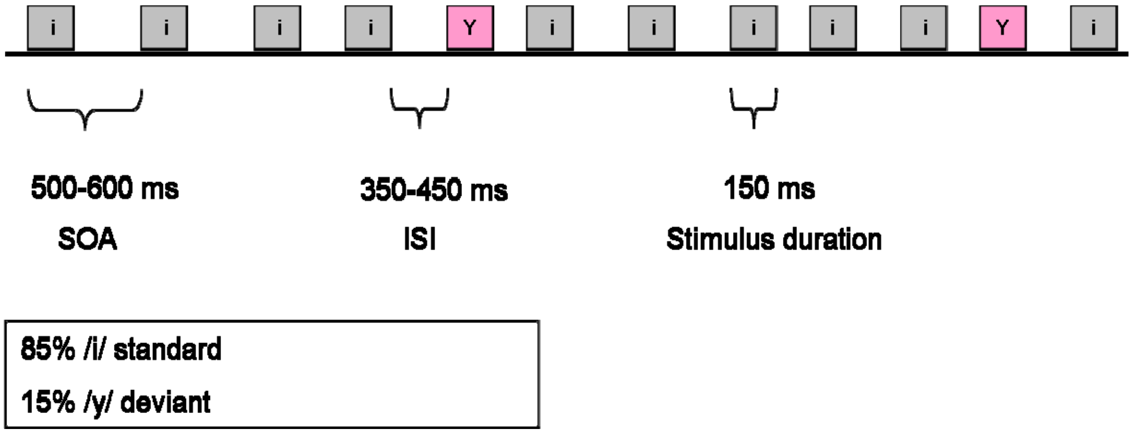


Figure 3. The Figure illustrates the presentation of stimuli in a classic oddball paradigm showing the changing ISI (interstimulus interval) and SOA (stimulus onset asynchrony), and the imagined order of stimuli presented.

ERP Procedure and recording

The experiment took place in the EEG lab in the Department of Psychology, University of Jyväskylä. Its total duration per participant was about 2.5 hours including the net application and breaks. The experiments were carried out from April 2007 to March 2008. The stimuli were presented via Phillips headphones model SHS4700/00 with intensity of 70 dBA. During the experiment the participants watched a self-selected silent movie without subtitles. They were instructed not to pay any attention to the sounds.

Preparation of the experiment. Calibration of the sound intensity was made using Brüel and Kjaer dB meter (type 2235) and Styrofoam head set with the headphones. Sound intensity checking was done before ERP experiment to ensure that the sound volume was same for all participants. Net impedances were measured before and after non-speech blocks and after the whole experiment. Location of the electrodes was determined with the Polhemus Fastrak System before experiment to make sure that the net was in the right position.

EEG recording. The EEG data were recorded with the Netstation program 4.2.1 and the Electric Geodesics Inc, (EGI) EEG-system using Ag-AgCl -electrodes attached to the EGI 128-channel sensor net. Sampling rate was 500 Hz and the electrodes were referred to the vertex electrode. All electrodes (except EOG electrodes) were located with the same distance of each other. Eye movements (electro-oculogram, EOG) were recorded by the electrodes below and above the eyes. The EEG data were filtered online with the high pass filter of 0.1 Hz and low pass 200 Hz. An effort was made to keep the electrode impedances under 50 kΩ before and during the recording. During the experiment data were monitored and if necessary the bad channels´ impedances (e.g. channels with electrical noise) were fixed.

Pre-processing and averaging of ERP data

The averaged ERPs were re-referred off-line to the averaged reference with the BESA 5.1.8 analysis program. The ERP data of each subject were inspected visually to check the quality of the data. The following filter settings were used: notch filter of 50 Hz (width 2 Hz), low cut off 0.32 Hz (12 dB, filter type zero phase), and high cut off 30 Hz (12 dB, filter type zero phase). Channels with a lot of artefacts throughout the data (dried electrodes or initially not touching the skin properly during ERP experiment) were considered as bad and were omitted from the averaging. After visual inspection, blink correction was carried out basing on the principal component analysis (PCA) method of BESA. The filters used in the blink correction were: low cut off 0.53 Hz (6 dB, filter type forward), high cut off 8.0 Hz (12 dB, filter type zero phase). Eye blinks in the data were chosen manually (20-200 self-selected blinks). EEG epochs containing peak-to-peak deflections over 150 µV or transient peaks over 115 µV at any channels were excluded from averaging.

ERPs were calculated by averaging EEG epochs from -50 to 600 ms. Only epochs with each deviant stimuli and pre-deviant standard stimuli were averaged to equal the signal-to-noise ratios. For the stimuli of interest in this study, the number of accepted EEG epochs was required to be more than 70 in order to be accepted for further analysis. The general quality of the data was good; hence all the participants had the required amount of trials. Mean number of accepted trials for the pre-deviant standards for the 12 participants was 124 (range 106-132) and for the deviants was 124 (range 100-132). The mean number of accepted trials for two participants with extra blocks collected was 118 (range 89-165) for pre-deviant standards and 118 for deviants (range 89-164). Bad channels were interpolated after averaging by using spherical spline interpolation in BESA (Perrin, Pernier, Bertrand, & Echalier, 1989).

After averaging each individual data were visually inspected and more channels were interpolated afterwards in some cases if the general quality of the data cleared up from this. Difference waves were calculated from averaged ERP waveforms by calculating the difference between the standard and deviant response (the response to the deviant stimulus minus the response to the standard stimulus, see Figure 5 as illustration of the ERP waveforms elicited by the standard and deviant stimulus) in the Speech and Non-speech condition. The function of the difference waves is to show the processing differences of standard and deviant stimuli. The residual indicates processes which are related to deviant stimuli, although difference wave might show the residual of N1 wave when the smaller N1 response to the standard stimulus is subtracted from the larger N1 response to the deviant stimulus (for a review see Kujala, Tervaniemi, & Schröger, 2007). Grand average of averaged waveforms and difference waves were created and visually inspected (inspection also for each individual difference waves). Visual inspection concentrated on N1 and MMN responses. The visual inspection of the data is described in detail in the Results section.

Statistical analyses of the data

The average-referenced difference waves were exported and imported to SPSS 15.0. First, temporal principal component analysis (PCA) for the difference waves were applied and temporal principal components (PCs) were chosen. Statistical analyses were performed for the factor scores of chosen components and the selected channels including outlier detection, t-tests and multivariate analysis of variance (MANOVA).

Principal Component Analysis. Principal Component Analysis (PCA) is one of the factor analytic procedures and it has been used in the decomposition of event-related potentials. PCA takes into account all the information in the ERP data and constitutes it from a second-order covariance or correlation matrix (Kayser & Tenke, 2003) into combinations (i.e. PCA components). A covariance matrix was chosen based on the recommendations made by Kayser and Tenke (2003) and Dien, Beal and Berg (2005). PCA reduces the huge amount of information into a small number of components that represent the underlying ERP components. From the initial matrix, a certain number of components are retained or rotated. The rotation method used in this study was an oblique rotation (Promax) which accepts factors to correlate with each other. The reason for using Promax rotation in this study arises from the assumption that in ERP datasets the latent variables can be significantly correlated (Dien, Spencer, & Donchin, 2003; Dien et al., 2005). PCA extracts the components in a hierarchical manner: the first factor accounts for the largest proportion of the total variance in the data and the following components can be oblique, i.e., related to the preceding and account for the largest residual variance (Dien et al., 2005).

Temporal PCA was conducted to identify the time points of ERPs that varied systemically across experimental conditions. PCA was performed for the difference waves, 325 time points served as the variables (original data was averaged from -50 to 600 ms epoch and sampling rate being 500 Hz creates 325 time points). 14 participants, 12 conditions[1] and 129 electrodes generated 21672 averaged ERP waveforms which were the cases for the PCA. The number of temporal principal components (PCs) was determined by the value of variance explained. 33 components explained 99% of cumulative variance in the data of averaged waveforms (Kayser & Tenke, 2003). PCA assigned temporal factor loadings and scores. The factor loadings over the participants and cases represent the systematic contribution of each PCA component to the voltage at each time point. Factor scores represent the contribution of each component to each individual ERP difference wave.

The selection of components. At first the factor loading values of components were imported to BESA. This was performed to formulate the loading values to comparable form (i.e. topographic distribution of PCA) with the original difference waves. The selection of the components was done by the visual inspection of the difference waves and the factor loading pictures (from now on referred as factor score maps) with BESA. The components were compared to the original ERP responses and if they were related to each other through their spatial configurations (i.e. the topographic distribution of the factor scores) and temporal characteristics (i.e. the peak latencies of the components), they were chosen for further analyses. The selected three components are reported in detail in the Results section.

The selection of channels for further statistical analyses. In this study, the channel selection was based on earlier literature and widely used channels in ERP studies (Näätänen, 1992). One-sample t-tests (two-tailed) were carried out on the factor scores in each of the channels of interest to test whether the responses differed from the zero. The selected channels were 25 (F3, on the left hemisphere) and 124 (F4, on the right) as frontal channels; central channels were 37 (C3) and 105 (C4). These channels indicated the fronto-central negativity. Left mastoid 57 (LM) and right mastoid 101 (MR) were chosen to indicate the temporal positivity. Parietal channels 60 (P3) and 86 (P4) responses were rather small thus not differing from the zero. However, the visual inspection of factor score pictures showed scalp distribution differences in these locations elicited by /y/ stimuli. See the positions of the channels in Figure 4 below.

---

[1] PCA was done for all 12 all conditions (speech, non-speech and equal probability) to keep all conditions comparable with each other for further analyses.

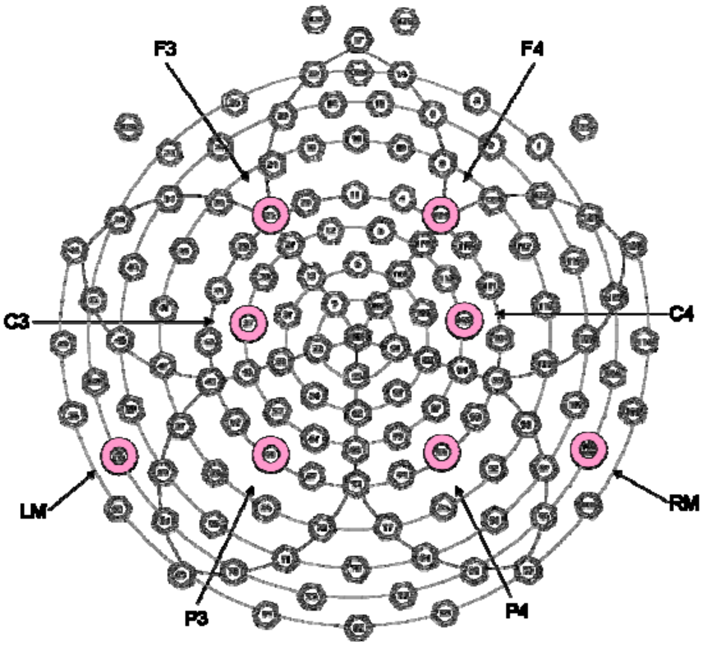


Figure 4. The eight channels chosen to further statistical analyses: F3/25, F4/124, C3/37, C4/105, P3/60, P4/86, LM/57, and RM/101.

Outlier detection, means and standard deviations. Outliers were detected from the factor scores of chosen components and channels using box-plots of SPSS. A value was considered as extreme outlier if it laid more than three interquartile range (IQR) from the closed end of the box (box containing the 50% of values). A mild outlier was any value that laid more than one and a half times the IQR from the closed end of the box. Some outliers were found but in the visual inspection of the original difference waves and the averaged waveforms, the outliers did not originate from the artifact like responses; rather the variation stem from the individual variation. When considering all the channels in each component and condition, the outliers remained within factor score variation of all channels i.e., in some channels the variation was small-scaled compared to others. Therefore the original values were not corrected and they were used in the analyses as the dependent variables.

Assumptions. The normal distribution of factor scores was assumed because of the physiological nature of data. The alpha level for all analyses was 0.05.

Statistical comparisons. The Multivariate analyses of variance (MANOVA) for factor scores were carried out to see whether any significant changes with Language (3; Finnish-Hungarian, French and German), Hemisphere (2; left and right) and Location (3; frontal, central and parietal) being the within subject factors. The main interest in this thesis was to detect whether the processing of native language prototype (Finnish-Hungarian) differed from two non-prototypes (French and German) in the Speech condition. Within-subjects contrasts were used to reveal whether the Finnish-Hungarian condition differed from the French and German conditions. The Non-speech condition comparisons were analyzed to see whether the processing of the non-speech stimuli was driven by the distance, not the language specific characteristics of the stimuli. The Multivariate analyses of variance were carried out to reveal the processing differences and the time behaviour differences between the Speech and Non-speech conditions by Condition (2; speech and non-speech) by Language (3; Finnish-Hungarian, French, German) by Hemisphere (2; left and right) by Location (3; frontal, central and parietal) as the within subject factors. When the main effects or interactions of Hemisphere, Location or Language were found, post hoc paired t-tests were carried out to detect the origin of differences.

RESULTS

Behavioural measures

Stimulus related questions were asked after the ERP experiment to collect the information of the quality of the stimuli. Seven participants (7/12, 58%) reported to hear other than speech sounds during the Speech condition and seven participants reported to hear the non-speech sounds as pips. The data were missing from two participants.

ERP waveforms

The responses for all conditions (the original grand averaged waveforms for the standard and deviant stimuli, the difference waves and the principal component's (PCs) topographies) were visually inspected using BESA. The focus was on the grand average difference waves and the frontal, central, parietal and mastoid channels (F3/25, F4/124, C3/37, C4/105, P3/60, P4/86, LM/57, and RM/101). The comparison of the standard and deviant stimuli in the Speech and Non-speech condition showed that the responses to the deviant stimulus elicited MMN component. The presence of MMN was also confirmed by visually comparing the oddball and the equal probability conditions.

Original ERP waveforms. The latencies of the averaged ERP components (N1 and MMN) appeared quite early (as observed at the frontal and central channels). The visualization of the original ERP waveforms and voltage maps can be seen in Figure 5. The N1 peaked approximately at 100 ms latency to the standard stimulus; the onset of N1 to the deviant stimulus appeared slightly later. The N1 was most prominent in the fronto-central areas but present also in the temporal and parietal areas, and it was larger in response to deviant stimuli. The MMN followed the N1 at the latency of 100-200 ms (the deviant response's onset to offset) and it was more frontally activated than the N1. Generally the non-speech stimuli elicited larger responses at the central and frontal channels than the speech stimuli (see Figure 5). In the Non-speech condition Finnish-Hungarian /y/ deviant stimulus elicited the largest N1 response compared to French and German. In the Speech condition the differences were more complex; at several latencies the response (N1/MMN) to Finnish-Hungarian /y/ was smallest.

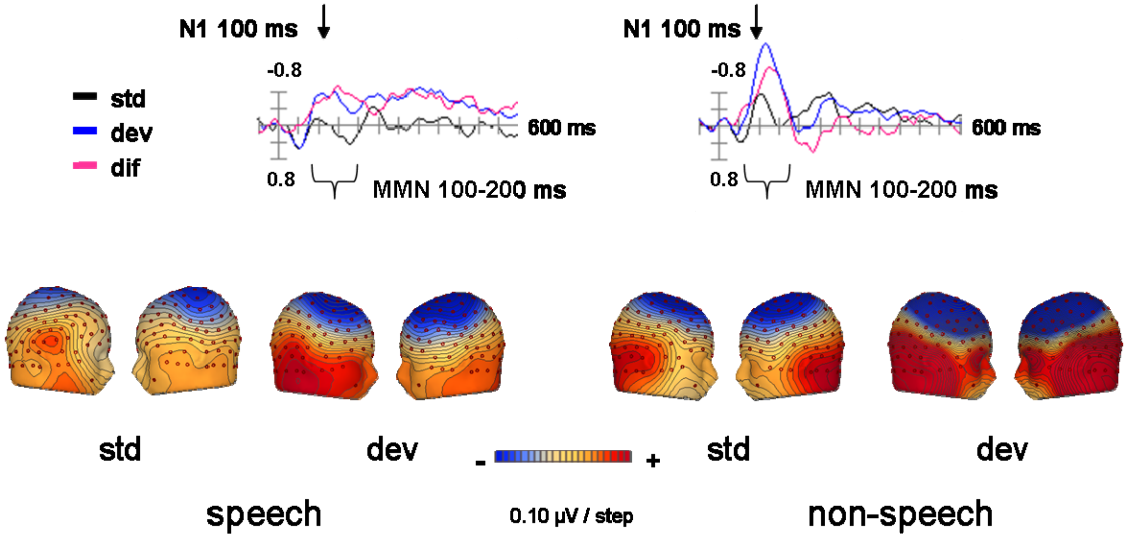


Figure 5. An illustration of the original averaged waveforms and the difference waves at the channel F3 of 14 typically reading adults. Waveforms show the responses to Euro-i standard, Finnish-Hungarian /y/ deviant and Finnish-Hungarian difference wave (dif) in the Speech (left-hand) and Non-speech (right-hand) condition at 100 ms latency (peak of N1 to standard stimulus). Topographic maps show the response to Euro-i standard and Finnish-Hungarian /y/ deviant.

Speech difference waves. The speech difference waves lacked a clear single peak and the responses were broad and the latencies of responses were later for speech (in a comparison with non-speech) (see Figure 5 above). The negative deflection was seen between 90-186 ms (early onset to offset). The highest peak in the difference waves reflects the onset of MMN response and possible residual of N1 peak of the averaged waveforms. It peaked approximately at 158 ms latency. Some latency differences between the languages were observed: the peak of the response to Finnish-Hungarian condition (dif-fh) was earlier than to French (dif-fr) or in German (dif-ge) conditions at some channels (C3 and F4) whereas for other channels latency differences between the languages were not observed (F3, C4, P3, P4, LM, and RM). Figure 6 below presents the speech difference waves of all languages. The response elicited in Speech condition had negativity over the frontal and central areas being largest at the frontal area. The positivity was seen at both temporal areas constructing dipole structure in temporal areas between positivity and negativity. Towards the end of the response, the negativity expanded broader to the central areas and the response was slightly more frontally activated. The temporal positivity shifted backwards in both sides. The response elicited by Finnish-Hungarian emphasized frontally leftward (reflected in PCs 2-122 and 5-186) compared to the responses elicited by French and German whose activation was more central. The topographic distribution of PC 2-122 for Speech condition showed that the responses (fronto-central negativity) elicited by French and German were actually slightly towards the right hemisphere (Figure 9).

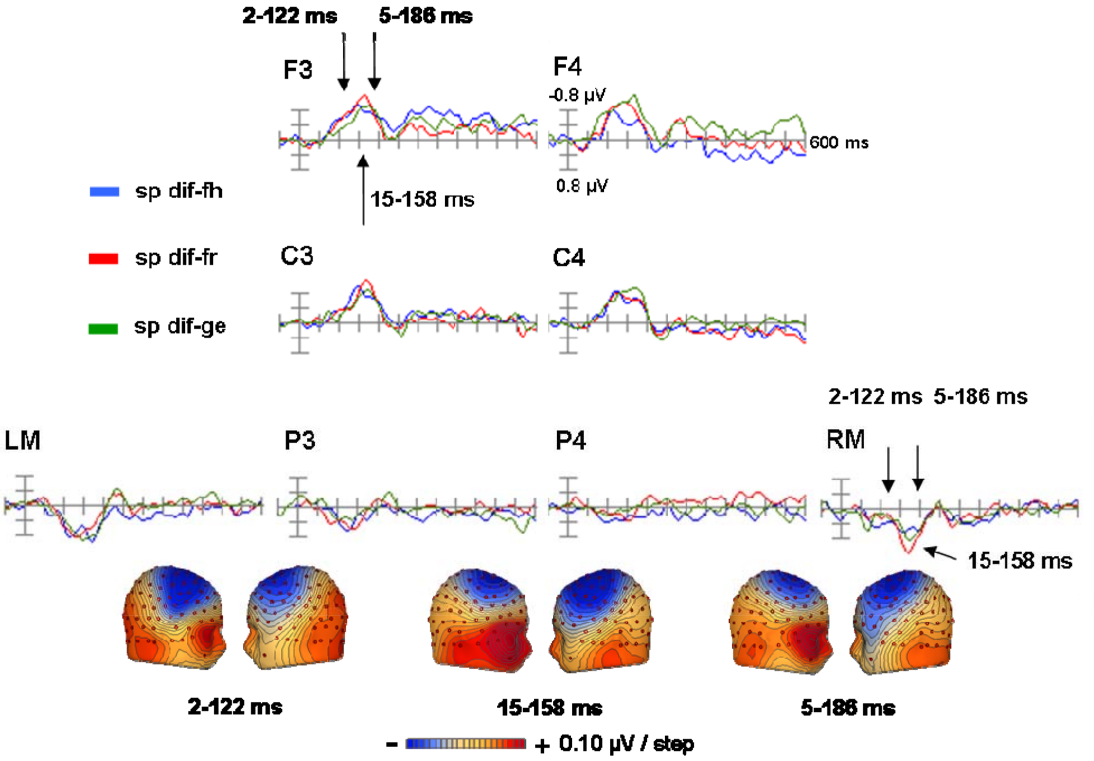


Figure 6. Difference waves for Finnish-Hungarian (dif-fh; blue), French (dif-fr; red), and German (dif-ge; green) in Speech condition of 14 typically reading adults. The arrows show the positions (in the timeline) of the three chosen temporal principal components. At C3 and F4 the Finnish-Hungarian-response is earlier than the responses to French or German. Topographic maps of the original voltage scores show the response to dif-fh at the peak loading latencies of PCA components.

Non-speech difference waves. The negative deflection was seen at the latency between 80-160 ms. The highest peak appeared at 122 ms in the Non-speech condition (Figure 7). Only small differences in amplitudes between the responses within different languages were observed. Non-speech difference waves followed the expected responses of Finnish-Hungarian condition eliciting largest responses at the frontal site (F3), French smallest, adhering to the largest (Euclidean) distance between Euro-i standard and Finnish-Hungarian /y/ deviant. However, also in the Non-speech condition Finnish-Hungarian (i.e. Finnish-Hungarian /y/ minus Euro-i) elicited smaller response at the central scalp areas (C3, C4). The fronto-central negativity in the Non-speech condition tended to be larger at left side while the temporal positivity did not show prominent differences in the left-right scalp distribution.

Non-speech responses, however, were larger at the left mastoid, compared to the right mastoid in all languages (see Figure 7 below). The Non-speech responses were narrower than the speech responses i.e., processing of the stimuli continued longer for the speech stimuli. Topographic distribution of the responses showed differences between languages.

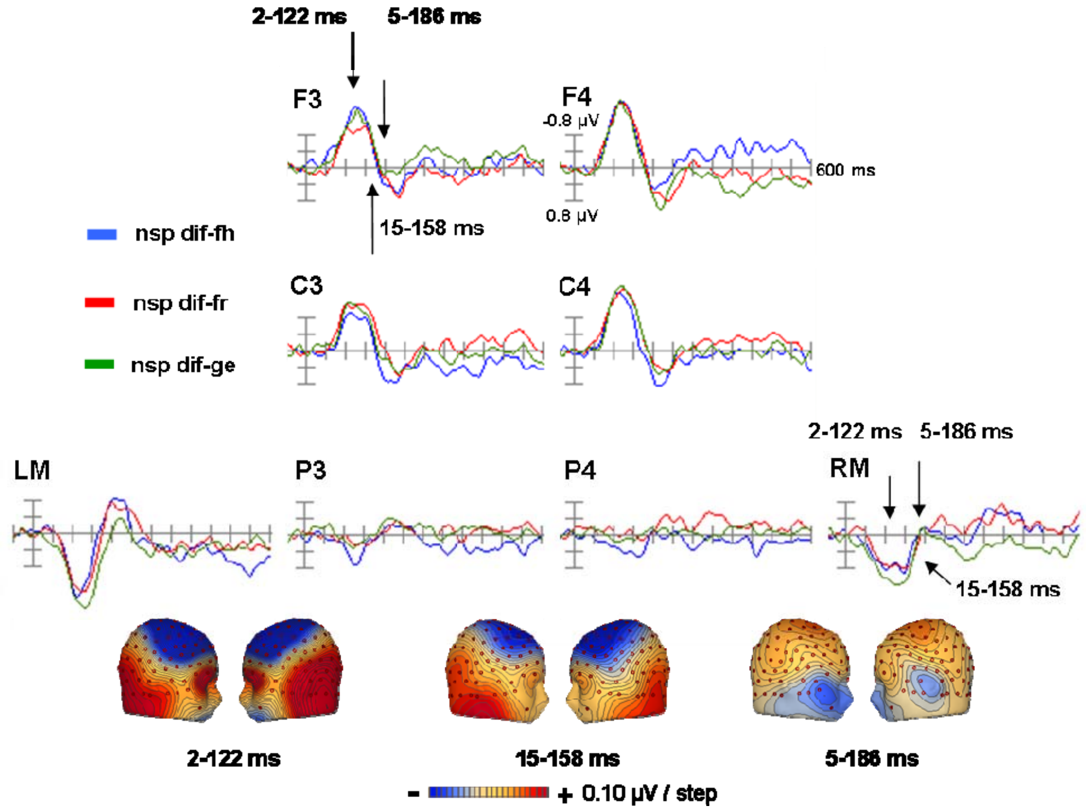


Figure 7. Difference waves for Finnish-Hungarian (dif-fh; blue), French (dif-fr; red), and German (dif-ge; green) in Non-speech condition of 14 typically reading adults. The arrows show the positions (in the timeline) of the three chosen temporal principal components. The latencies of the responses are very similar for all languages. Topographic maps of the original voltage scores show the response to dif-fh at the peak loading latencies of PCA components.

Principal components of ERP

Temporal principal components of interest. For further analyses three representative PCs were chosen from difference wave data (Figures 6 and 7 show the original difference waves and ERP components and the arrows point the peaks of PCs). The peak latencies of chosen PCs and latencies of ERP components in the grand average difference waves corresponded well to each other. The components are named after their loading order and latency.

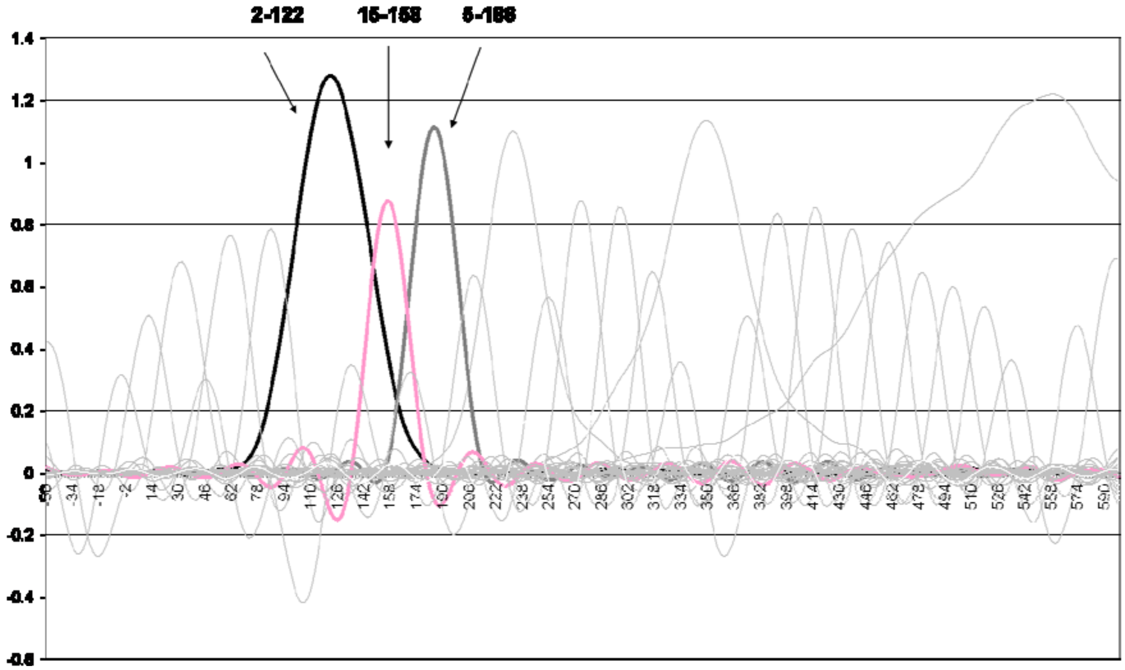


Figure 8. The factor loadings of temporal PCA components. Three selected components are highlighted and marked with the number of the temporal component and the peak latency of the loading. PCs are named after their order and latency (2-122, 15-158, 5-186, respectively). The X-axis shows the timeline (-50-598 ms).

The second largest component (10 % of the variance explained) peaked at 122 ms latency and it corresponded to the latency of the N1 response and the onset of MMN to the Non-speech condition (early time window). The component number 15 (0.6%) seemed to reflect the N1 and the onset of MMN response to the speech stimuli (middle time window), peaking at 158 ms, corresponding to the latency observed in the speech difference waves. A separate component (number 5; 2.7%) was found to MMN at the latency of 186 ms (late time window) and it settled to the end of MMN response for the Speech condition.

Signal strength of the responses. One sample t-tests (two-tailed) for the factor scores were carried out on each of the selected channels to examine whether the responses differed from zero, thus affirming the existence of the components. In the Speech condition the factor scores (responses) of PC 2-122 differed significantly from zero, except for the right mastoid and the parietal channel P4/86. In the Speech condition, the factor scores of PCs 15-158 and PC 5-186 at the parietal channels P3 and P4 did not differ from zero. In the Non-speech condition, the response in the parietal channels did not differ from zero in PCs 2-122 and 15-158. In the Non-speech condition the factor scores of PC 5-186 did not differ from zero at any channel, the result indicating the fact that PC 5-186 loaded only for speech.

Language effects on PCs 2-122, 15-158 and 5-186

PC 2-122 structure and topography. The early negative component at 122 ms was greater for the non-speech than for speech stimuli (see Figure 7) at the latency of 122 ms. PC 2-122 showed fronto-central negativity with positive amplitudes over temporal and occipital areas. In the speech condition, the negativity was more frontally activated and smaller in amplitude, at the latency of 122 ms the difference waves in the Speech condition did not reach its maximum amplitude yet. In the Speech condition, the responses showed amplitude (Figures 6 & 7) and topographic differences (Figure 9) between languages, whereas responses to the Non-speech condition were very similar (Figure 9).

Language and topographical effects on PC 2-122 in the Speech condition. For the fronto-central and parietal channels, a Language (3; Finnish-Hungarian, French and German) by Hemisphere (2; left and right) by Location (3; frontal, central and parietal channels) MANOVA and within- subjects contrasts were carried out to detect language differences. In the Speech condition processing between hemispheres differed significantly (see Table 2 below); all languages were processed more negatively at the right hemisphere (Figure 13; F4 and C4). Also locations showed significant main effect but the difference was due to fronto-central negativity (F3, C3, F4, and C4) differing from parietal positivity (P3, P4) ($F(1,13) = 26.178$, $p<.001$) as shown by the within-subjects contrasts. No main effects for language or any language related interactions were found. See Table 2 for hemisphere and location effects of MANOVA. Post hoc paired t-tests were carried out to explore the origin of the hemisphere effects. The t-tests for hemisphere effects did not reveal any significant differences between separate channels.

For the mastoids, a Language (3; Finnish-Hungarian, French, and German) by Hemisphere (2; left and right) MANOVA and within-subjects contrasts were carried out to test language, hemisphere and location effects. Nearly significant hemisphere main effect was found (p = 0.051; see Table 2), whereas language related effects were not found. The factor scores were larger at the left mastoid than at the right for thus indicating that all the speech stimuli were processed more at the left. However, post hoc t-tests did not show differences between the mastoid channels in tests for each language separately.

Language and topographical effects on PC 2-122 in the Non-speech condition. A corresponding Language by Hemisphere by Location MANOVA and within-subjects contrasts did not reveal any language main effects or interactions. The MANOVA with the fronto-central-parietal channels showed Location main effect (Table 2) which was confirmed by within-subjects contrasts. Differences between the frontal and central negativity were not found but the parietal positivity differed from both significantly (F(1,13) = 53.834, p<.001). For the mastoid channels, A MANOVA (Language by Hemisphere) showed that all languages were processed significantly more at the left hemisphere (LM, Table 2). The paired t-tests showed the processing differences at Finnish-Hungarian condition between channels LM and RM (t(13) = 2.438, p<.05). Paired t-tests for French or German did not show significant differences.

Table 2
Manova tests (F-values) for PC 2-122

Comparison	df	Hemisphere F-value		df	Location F-value	
SP F-C-P	1, 13	4.819*	W = 0.730	1.160, 12	16.022***	W = 0.272
NSP F-C-P	1, 13	4.138	W = 0.759	2, 12	26.277***	W = 0.186
SP LM/RM	1, 13	4.603	W = 0.739			
NSP LM/RM	1, 13	14.704**	W= 0.469			

Note: df = degrees of freedom; W = Wilk's Lambda. F-C-P means comparison including six channels (frontal, central, and parietal), SP = Speech condition; NSP = Non-speech condition.
* p<0.05
** p<0.01
*** p<0.001

PC 2-122: the Speech vs. Non-speech conditions. A Condition (2; speech and non-speech) by language (3; Finnish-Hungarian, French, German) by Hemisphere (2; left and right) and by Location (frontal central, and parietal channels) MANOVA and a Condition (2; speech and non-speech) by language (3; Finnish-Hungarian, French, German) by Hemisphere (2; left and right, mastoid channels) MANOVA were carried out to see the differences between the speech and the non-speech responses. Within-subjects contrasts were also carried out. Main effect for Condition (F(1,13) = 36.523, p<.001, Λ = 0.263) was found. Also the MANOVA with the mastoid channels showed main effect for condition (F(1,13) = 27.957, p<.001, Λ = 0.317). Paired t-tests were carried out to explore the origin of differences. The difference between the Speech and the Non-speech conditions arose from the larger responses to the non-speech stimuli at F4, C3, C4, LM, RM in French and German condition (See table 3 below for results).

Table 3
Paired t-tests (t-values) for PC 2-122

Channel	df	SP/NSP FH	df	SP/NSP FR	df	SP/NSP GE
F4			13	4.047***		
C3			13	2.748*	13	2.455*
C4	13	2.669*	13	4.271***	13	2.572*
LM			13	-2.407*	13	-3.452**
RM					13	-2.470*

Note: df = degrees of freedom; FH = Finnish-Hungarian; FR = French; GE = German; SP = Speech condition; NSP = Non-speech condition.
* p<0.05
** p<0.01
*** p<0.001

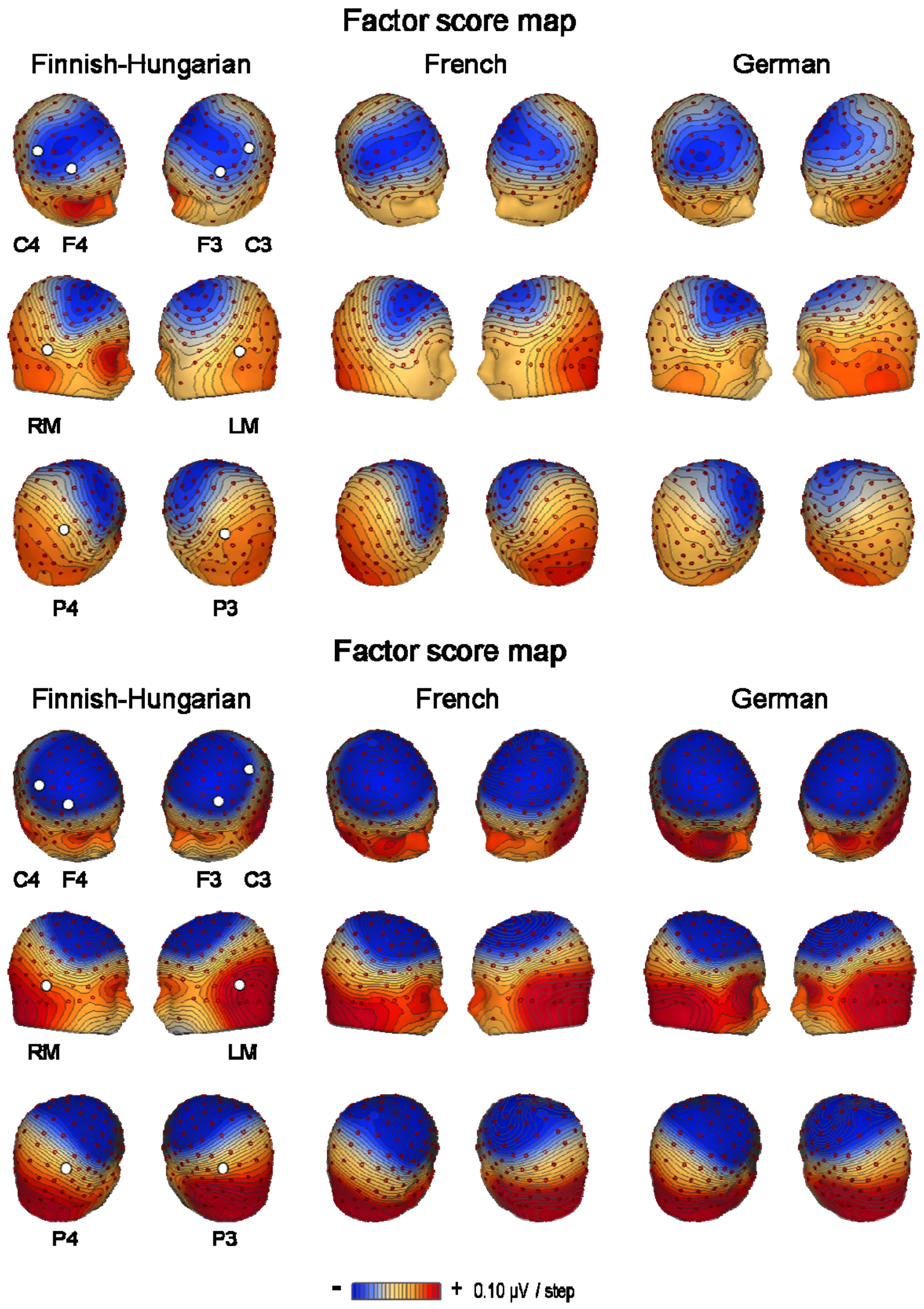


Figure 9. Topographic maps of the factor scores (difference waves) of PC 2-122 in the Speech (above) and the Non-speech condition (below) of 14 typically reading adults. The white spots indicate the channels selected for MANOVA. The responses to the Speech condition have not reached the peak at 122 ms latency. In the Non-speech condition the responses are very intense and similar between languages.

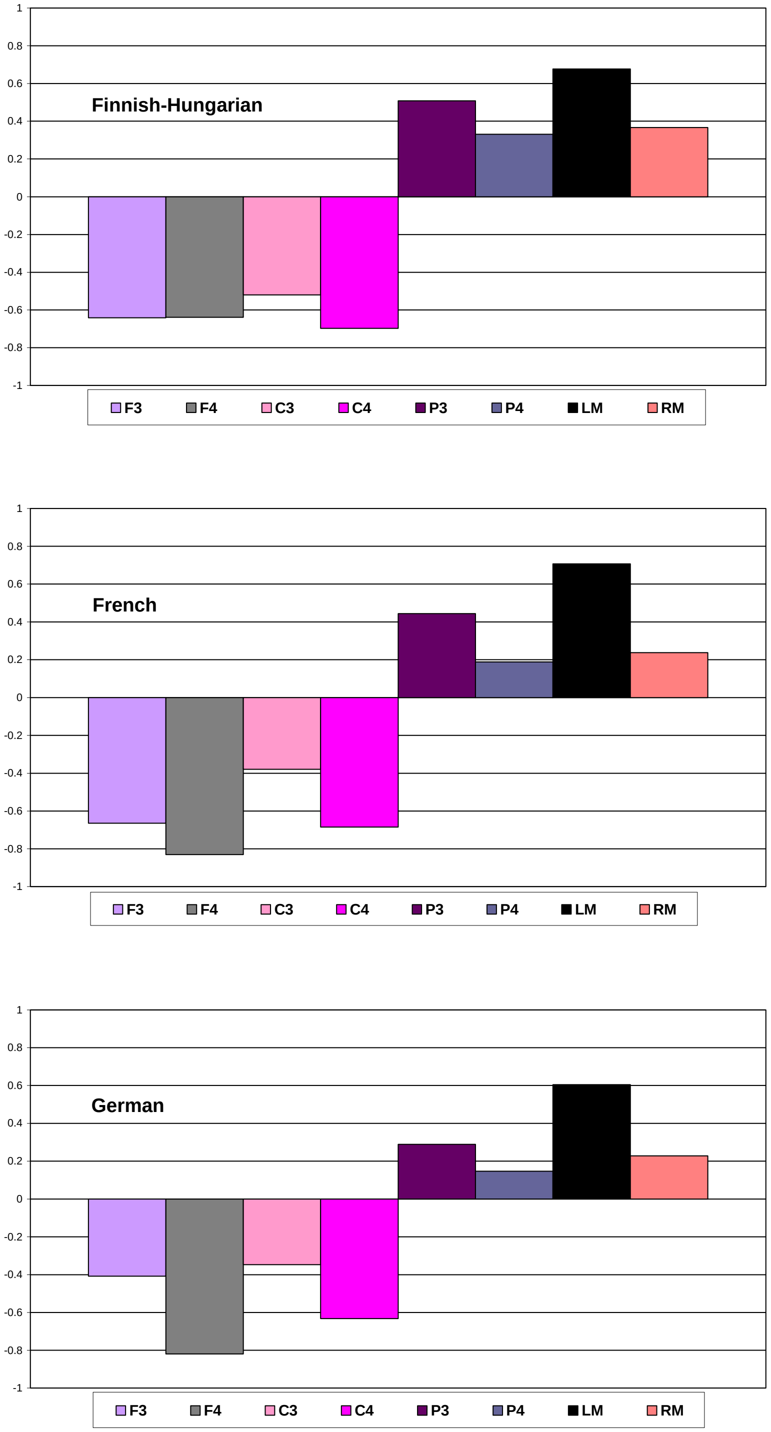


Figure 10. The PCA factor scores of the selected channels (from the left F3, F4, C3, C4, P3, P4, LM, and RM) of PC 2-122 in the Speech condition. The response at the left mastoid is displaying more positive activation compared to the right mastoid in all languages.

PC 15-158 structure and topography. PC 15-158 showed fronto-central negativity with positive amplitudes over temporal areas (Figure 11) and it indicated the onset of MMN response to the speech stimuli. At 158 ms latency, responses to the Speech condition peaked, whereas responses to the Non-speech condition were smaller and obscure thus reflecting the fact that the responses to the non-speech stimuli were already diminishing at this latency. Topographically observed the magnitudes of responses were similar between the Speech and Non-speech conditions. In the Speech condition, it was visually observed that the response to Finnish-Hungarian condition was smaller in the fronto-central areas, whereas the responses to French and German conditions showed more similar pattern fronto-centrally. The response to German condition was stronger at the mastoids than in Finnish-Hungarian or French conditions (see Figure 11).

Language and topographical effects on PC 15-158 in the Speech condition. The language effects for fronto-central and parietal channels, were analyzed by using a Language (3; Finnish-Hungarian, French and German) by Hemisphere (2; left and right) by Location (3; frontal, central and parietal channels) MANOVA and within-subjects contrasts showing one location main effect (Table 4 below). The location effect was due to difference between fronto-central negativity and parietal positivity within all languages (F(1,13) = 13.743, p=.003) (Figure 12). MANOVA comparisons for mastoid channels did not show any language or hemisphere main effects or interactions in speech. Topographic distributions of factor scores are given in Figure 11.

Language and topographical effects on PC 15-158 in the Non-speech condition. Corresponding Language (3; Finnish-Hungarian, French and German) by Hemisphere (2; left and right) by Location (3; frontal, central and parietal channels) MANOVA and within-subject contrasts showed Location main effect (Table 4 below). The MANOVA (Language by Hemisphere) and within-subjects contrasts for the mastoid channels did not show any Language or Hemisphere main effects or interactions for the Non-speech condition.

PC 15-158: the Speech vs. Non-speech conditions. A Condition (2; speech and non-speech) by Language (3; Finnish-Hungarian, French and German) by Hemisphere (2; left and right) by Location (3; fronto-central channels including the parietal channels (F3, F4, C3, C4, P3, and P4) MANOVA and a MANOVA with mastoids channels; LM, and RM were carried out to compare the responses to the Speech and Non-speech conditions. Within-subjects contrasts were used as well. No main effect or interactions for Condition was found thus indicating the similar amplitude of all responses at 158 ms latency.

Table 4
Manova tests (F-values) for PC 15-158

Comparison	df	Hemisphere F-value	df	Location F-value
SP F-C-P	1, 13	2.158 W = 0.858	1.084, 12	17.634*** W = 0.254
NSP F-C-P	1, 13	2.316 W = 0.849	1.379, 12	6.625* W = 0.489

Note: df = degrees of freedom; W = Wilk’s Lambda. F-C-P means comparison including six channels (frontal, central, and parietal channels), SP = Speech condition, NSP = Non-speech condition
* $p<0.05$
*** $p<0.001$

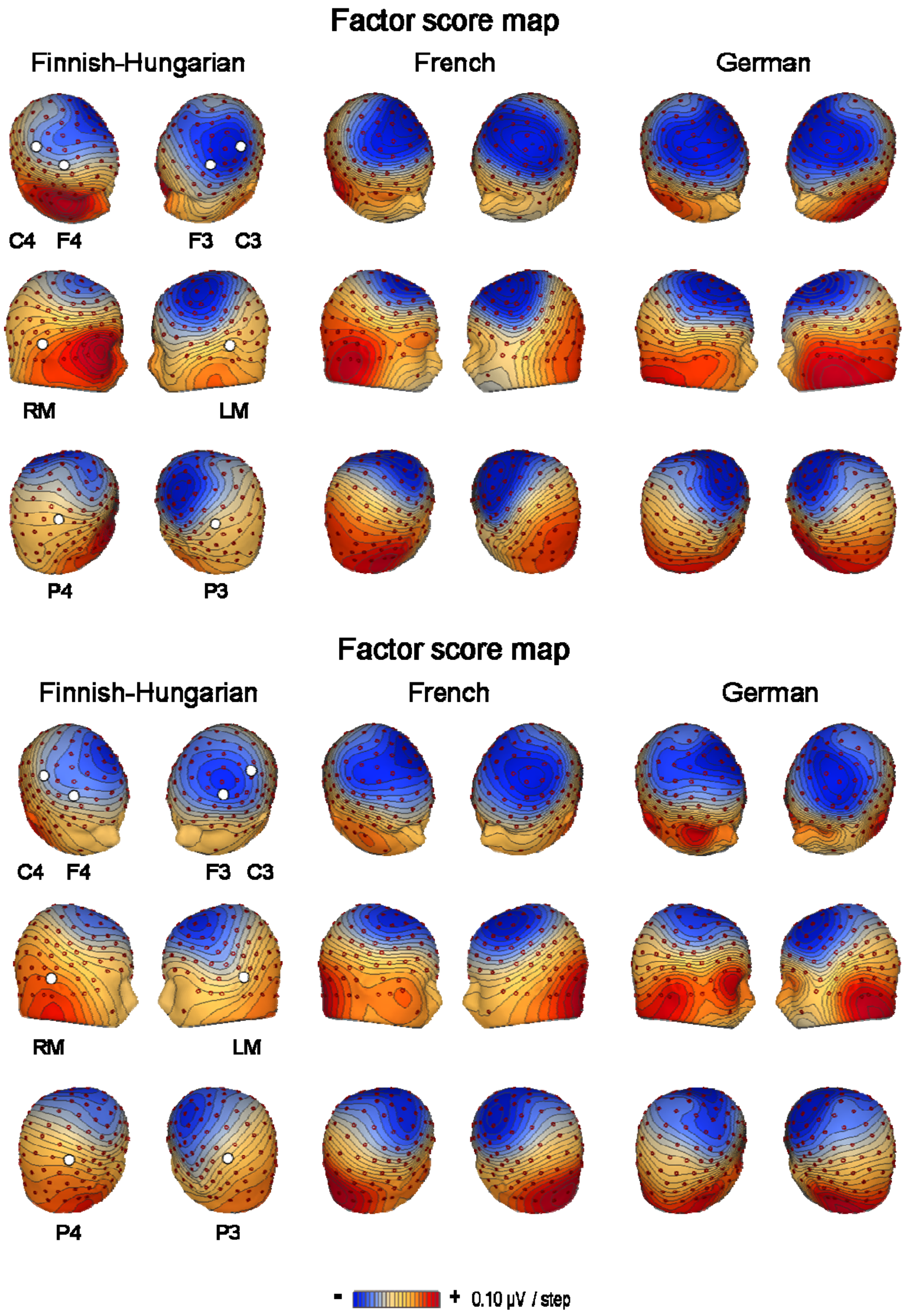


Figure 11. Topographic maps of the factor scores (difference waves) of PC 15-158 in the Speech (above) and Non-speech condition (below) of 14 typically reading adults. The white spots indicate the channels selected for MANOVA.

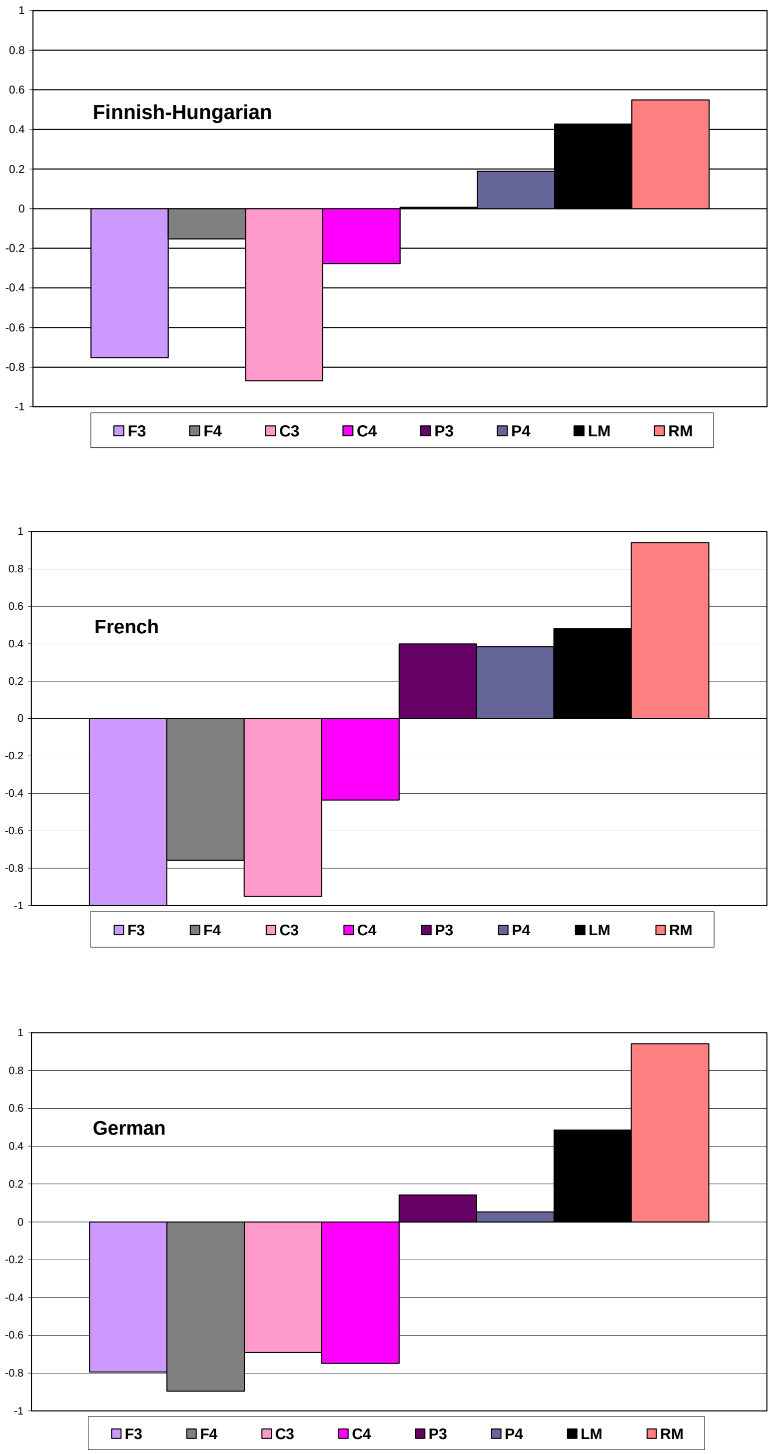


Figure 12. The PCA factor scores of selected channels (from the left F3, F4, C3, C4, P3, P4, LM, and RM) of PC 15-158 in the Speech condition.

PC 5-186 structure and topography. PC 5-186 was mainly activated by the speech stimuli and it showed fronto-central negativity with positive amplitudes over temporal areas (Figure 13). The latency of component 5-186 did not indicate a clear peak in the grand average difference waves, on the contrary it settled to a slope of the responses at all channels. At 186 ms latency, the responses to the Non-speech conditions had faded away. In the Speech condition, the responses to French and German conditions showed more similar fronto-central negativity than the response to Finnish-Hungarian condition, which was more frontally and leftwards activated.

Language and topographical effects on PC 5-186 in the Speech condition. For the fronto-central and parietal channels, a Language (Finnish-Hungarian, French and German) by Hemisphere (left and right) by Location (3; frontal, central and parietal channels) MANOVA and within- subjects contrasts to detect language effects was carried out. The language main effect was found in speech comparison ($F(2, 12) = 4.751$, $p<.04$, $\Lambda = 0.558$) indicating the smaller and more frontal response to Finnish-Hungarian (Figures 13 and 14). Within-subjects contrasts revealed that Finnish-Hungarian response differed significantly from French ($F(1,13) = 9.454$, $p<.010$) and German ($F(1,13) = 5.197$, $p<.05$). Post hoc paired t-tests for each channel, however, did not show statistically significant results between languages. Hence the language effect found in MANOVA reflects the sum of fronto-central-parietal channels and not any particular difference between any (pair of) channels. In addition, main effect for Location was found ($F(2, 12) = 13.627$, $p<.01$, $\Lambda = 0.306$). Within-subjects contrasts indicated that the main effect arouse from the difference between the fronto-central negativity and the parietal positivity across all languages ($F(1,13) = 23.135$, $p<.001$).

A Language by Hemisphere MANOVA with the mastoid channels did not show any main effects or interactions.

Language and topographical effects on PC 5-186 in the Non-speech condition. No main effects or interactions for the Non-speech condition were found with corresponding MANOVA and within-subjects contrasts comparisons.

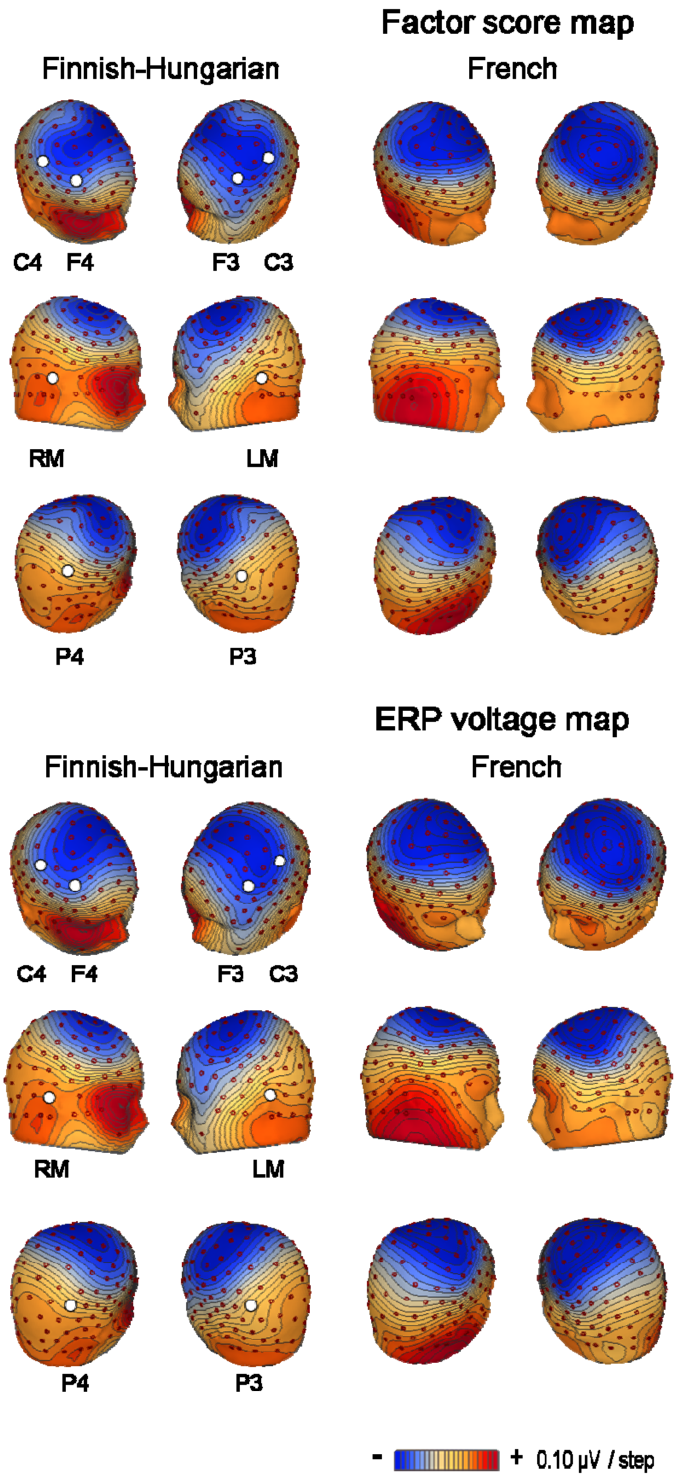


Figure 13. Topographic maps of the factor scores (difference waves) of the principal component analysis (above) of PC 5-186 in the Speech condition and the corresponding original voltage map at the same latency (below) of 14 typically reading adults. The white spots indicate the channels selected for MANOVA. The responses to French and German conditions show more similar fronto-central negativity than the response to Finnish-Hungarian condition, which is more frontally and leftwards activated.

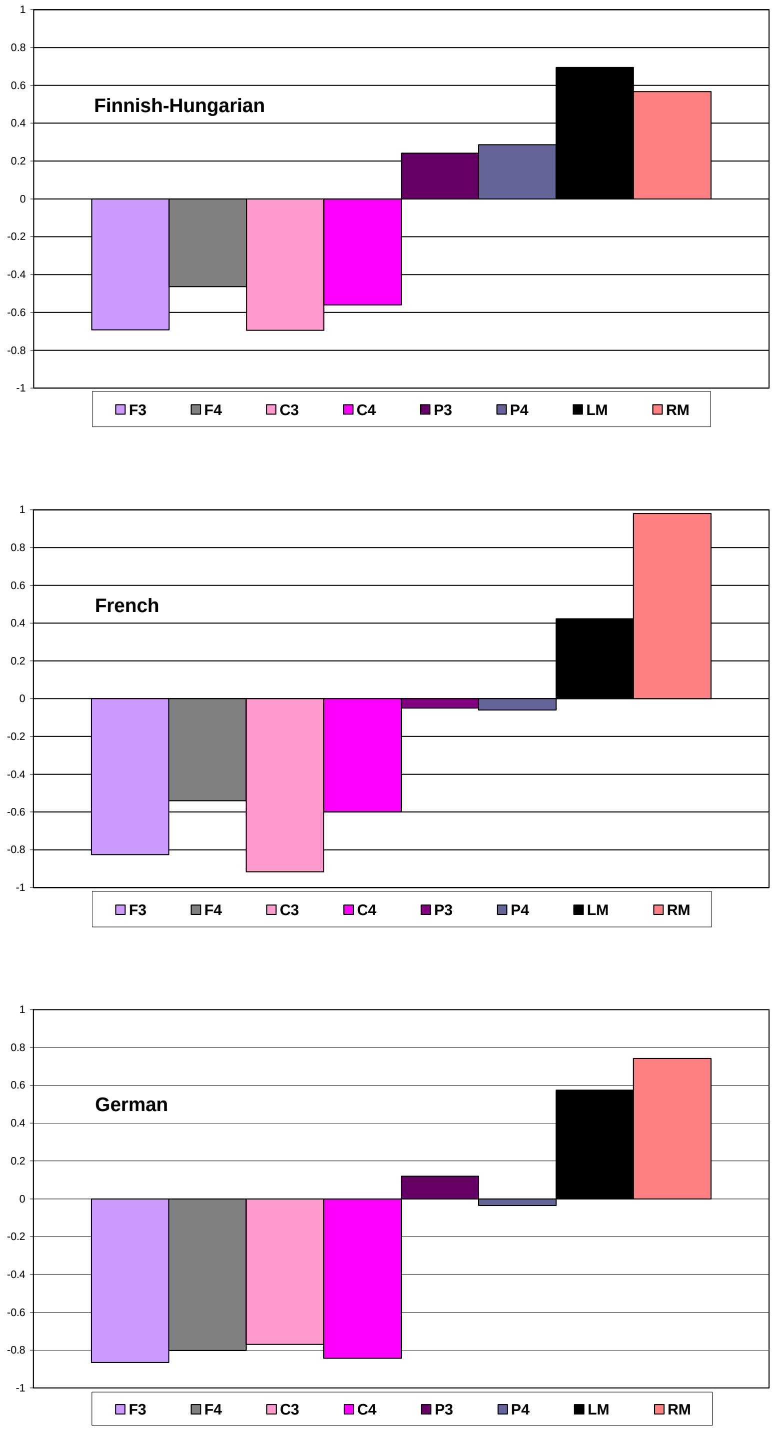


Figure 14. The PCA factor scores of selected channels (from the left F3, F4, C3, C4, P3, P4, LM and RM) of PC 5-186 in the Speech condition. Factor scores of channels show that response to Finnish-Hungarian is more frontally activated than French or German hence the parietal channels are positively activated. Response in fronto-central channels is also slightly smaller to Finnish-Hungarian.

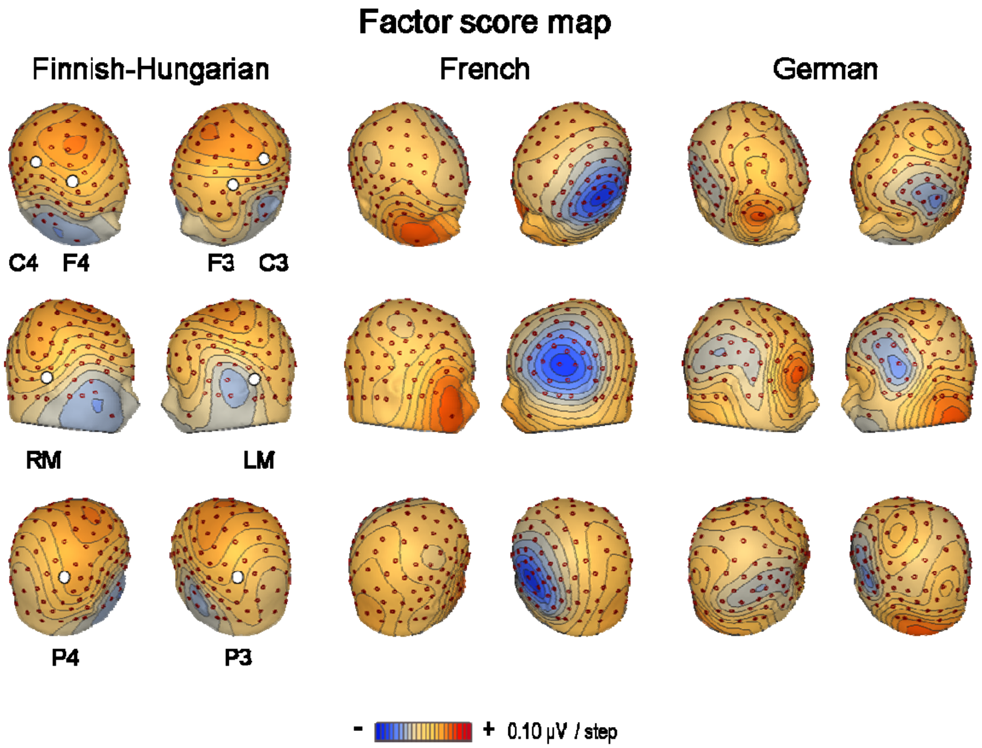


Figure 15. Topographic maps of the factor scores (difference waves) of PC 5-186 in the Non-speech condition of 14 typically reading adults. The white spots indicate the channels selected for MANOVA. The responses to the non-speech stimuli have been faded at 186 ms latency.

PC 5-186: the Speech vs. Non-speech conditions. A Condition (2; speech and non-speech) by Language (3; Finnish-Hungarian, French and German) by Hemisphere (2; left and right) by Location (3; fronto-central channels including the parietal channels (F3, F4, C3, C4, P3, and P4) MANOVA and a MANOVA with the mastoids channels; LM, and RM were carried out. In addition, within-subjects contrasts were carried out as well. The MANOVA with frontal, central and parietal channels (F3, F4, C3, C4, P3, and P4) showed main effect for Condition (F(1, 13) = 26.180, p<.001, Λ = 0.434) thus confirming the processing differences of speech and non-speech stimuli. An interaction between Condition and Location was also found (F(5,9) = 4.521 p<.05, Λ = 0.269). Paired t-tests showed that the differences between Speech and Non-speech condition arose mainly from fronto-central channels of right hemisphere indicating the extended response to the speech stimuli (see Figures 13 and 15) The mastoid channel comparison between the Speech and Non-speech condition showed significant difference (F(1,13) = 17.925, p<.001, Λ = 0.420). The main effect arose from the stronger positive response to the Speech conditions on both mastoids and on the contrary slightly negative response at the left mastoid for all the Non-speech conditions in all languages. The t-tests confirmed the difference of both mastoids between the Speech and Non-speech condition. The results of all the t-tests are shown in Table 5 below.

Table 5
Paired t-tests (t-values) for PC 5-186

Channel	df	SP/NSP FH		df	SP/NSP FR		df	SP/NSP GE	
F3				13	-2.824	-2.824**			
F4				13	-2.367	-2.367*	13	-4.585	-4.585***
C3	13	-4.165	-4.165***						
C4	13	-3.176	-3.176**	13	-3.176	-3.176**	13	-2.161	-2.161*
LM	13	2.280	2.280*	13	2.493	2.493*			
RM				13	3.152	3.152**	13	2.649	2.649*

Note: df, degrees of freedom; FH = Finnish-Hungarian; FR = French; GE = German; SP = Speech; NSP = Non-speech.
* $p<0.05$
** $p<0.01$
*** $p<0.001$

DISCUSSION

In this Master's thesis, the effects of long-term exposure of native language on ERPs, especially on the N1 and the mismatch negativity (MMN) responses were evaluated. The aim was to examine the N1/MMN response to Finnish /y/ prototype in a comparison with the non-prototype /y/ vowels (which were prototypes for French and German) in a crosslinguistic MMN paradigm. Euro-i served as the standard stimulus and was a member of /i/ vowel category and /y/ vowel of each language served as the deviant stimulus and was member of /y/ vowel category. The participants in this study were 14 typically reading Finnish adults without the knowledge of French or German. The approach and theoretical background for this thesis arose from the earlier findings of memory traces and top-down effects affecting the MMN and Kuhl's (1991, 1993) findings of the NLM and PME (the native language magnet model and the perceptual magnet effect, respectively).

It has been suggested that the exposure to native language formulates language specific representations of phonemes, which can be detected from ERPs (Cheour et al., 1998; Halsband, 2006). Several studies have successfully discovered differences in processing native and non-native speech sounds (Näätänen et al., 1997; Rivera-Gaxiola et al., 2000; Winkler et al., 1999) and convincing evidence for the NLM and PME have been found (Iverson et al., 2003; Kuhl, 1991, 1993), whereas some studies have found that native and non-native sounds are discriminated equally well (Polka et al., 2001) or the findings have been inconsistent with current theories: for example in relation to the NLM or PME (Best et al., 2001, Lotto et al., 1998). The effects of native language exposure have been observed by the larger MMN responses to native stimulus and the lack of MMN or smaller MMN response to non-native stimulus (Dehaene-Lambertz, 1997). However, this was not observed in the present Master's thesis. On the contrary, the response to Finnish-Hungarian prototype /y/ elicited smaller magnitude of N1 and MMN responses compared to French and German non-prototypes and it was statistically compared in the late time window of MMN response in the Speech condition (PC 5-186), wherein the responses showed language differences in the fronto-central negativity. Significant differences between languages in statistical comparisons for early (PC 2-122) and middle (PC 15-158) time windows of the responses were not found. All in all, the results suggested that processing the stimuli was very similar between the languages in the early and middle fronto-central negativity and the temporal positivity of the responses (all principal components).

ERP waveforms and the processing differences between the speech and non-speech stimuli. The speech stimuli elicited responses in the N1 and MMN time windows. Amplitudes of responses to the speech stimuli were rather small as observed also in the studies of Aaltonen et al. (1994) and Čeponiené et al. (2008). The stimuli of the Speech condition elicited broad responses when compared to the responses to the Non-speech condition. Similarly, Kujala et al. (2001) observed broad responses to speech stimuli. In a line with studies of Möttönen et al. (2006) and Tervaniemi et al. (1999) the speech stimuli were processed more at the left auditory cortex (i.e. the left mastoid). Visual observation of the non-speech stimuli showed larger amplitudes (Aaltonen et al., 1994; Čeponiené et al., 2008) and more narrow response (Dehaene-Lambertz et al., 2005) when compared to the waveforms elicited by the speech stimuli. The non-speech stimuli were processed more on the RM site. The latencies of N1 and MMN responses appeared earlier for the Non-speech condition (similarly in Dufor et al., 2006) than for the Speech condition.

The Speech and Non-speech conditions were compared to examine these visually observed processing differences. MANOVA comparisons showed that in the early and late time windows (PCs 2-122 and 5-186) the responses in the Speech and Non-speech conditions differed significantly. At the latency of 122 ms, the speech responses were smaller than the non speech responses; on the contrary, at 186 ms latency the situation was inversed, indicating the longer duration of the speech responses.

Aaltonen et al. (1994) suggested that processing of speech is more demanding and therefore the magnitude of response remains smaller for phonemic stimuli than for the more simple acoustic stimuli. In their study, it was proposed that the processing of F0 and F2 formant frequency changes in speech stimuli were more demanding than the changes in the pure tone stimuli. Also Dufor et al. (2006) proposed that the increase of latency arose from the processing demands of more complex speech stimuli. In several earlier studies participants have been taking part in behavioral tests before ERP measurements (Aaltonen et al., 1997; Sharma & Dorman, 1998; 2000). Kraus et al. (1995) have shown the enhancement in the MMN magnitude as a result of training. These former observations may explain the generally small amplitudes of responses in this Master's thesis, wherein a testing of the stimuli was not used before the ERP experiment.

Native language exposure effects on the N1 and MMN responses. The MANOVA revealed no significant changes or language main effects in the beforehand selected channels at the frontal, central, parietal, and temporal areas for the early and middle time window of responses (PCs 2-122 and 15-158) in the Speech condition. Frontal negativity with left hemisphere predominance was elicited to Finnish-Hungarian at 186 ms and the responses were statistically confirmed to be smaller for Finnish-Hungarian condition in a comparison to French or German condition. The effect did not arise from single channels; rather it illustrated the sum of fronto-central-parietal channels. The response to Finnish-Hungarian was more frontal and smaller in amplitude and that was seen in greater parietal positivity whereas, the responses to French and German were larger and the negativity spread more to parietal channels as well. In MEG study with mismatch field (MMF), Zhang et al. (2005) compared the MMF responses of American and Japanese participants in /ra-la/ and /ba-wa/ continuum (non-native pair and native pair, for Japanese participants, respectively). They proposed that the processing of non-native sounds demands more activation than the native sounds and therefore the non-native responses in Japanese participants were broader (longer duration) and greater in amplitude when compared to processing of native sounds in them.

The non-speech stimuli (equivalent formant values with the speech stimuli) were used to study the effects of distance only. However, despite the largest distance between Euro-I standard stimulus and Finnish-Hungarian deviant stimulus, the response to this condition did not differ significantly from the responses to French and German conditions (with shorter distance to Euro-I standard).

Furthermore, it is possible that the participants experienced Finnish-Hungarian /y/ vowel to be more familiar than the others and therefore processing of two other /y/ vowels demanded more involuntary attention (the fronto-central negativity) and change detection (the temporal positivity) towards the stimuli (Paavilainen et al., 2003). Thus, it could be possible that processing of familiar speech sound did not demand as much attention as foreign stimuli. However, the attention shift detector component P3a was not observed in this study. Also, in the earlier studies the frontal negativity of MMN has been seen quite unchangeable and not related to the stimuli (Paavilainen et al., 2003; Shalgi & Deouell, 2007).

In this Master's thesis Finnish-Hungarian /y/ vowel had the lowest F2 value compared to French or German (see Table 1 for formant values). Results of Aaltonen et al. (1997) might provide another explanation for the surprising results of the small response to Finnish-Hungarian in the Speech condition. They discovered that the speech stimuli (/i/ vowels) with high F2 frequency values elicited larger MMN responses than the stimuli with lower F2 values in both participant groups, despite the equal distances between the standard and deviant stimuli used in the study. In addition, they observed that different pattern for Good and Poor categorizers emerged: Poor categorizers selected the stimulus high F2 value, whereas Good categorizers qualified the stimulus with low F2 values to present a best exemplar of /i/ vowel (synthetic speech stimulus). Sharma and Dorman (1998) found similar F2 effect on the magnitude of MMN response. In their MMN experiment, three different stimulus pairs were used, the pair (the standard and deviant stimulus were both exemplars of /i/ vowel) with highest F2 frequency values elicited the greatest MMN response compared to the responses elicited by the two other pairs. The difference between all the standard and deviant stimuli was always same.

Polka et al. (2001) studied the discrimination ability of /d-ð/ contrasts of French and English (American) speaking adults and infants (French and American) and observed that French speaking adults surprisingly discriminated non-native /d-ð/ (e.g. doze-those) pair comparable with two groups of infants, even though this consonant difference is not relevant in French and therefore it was assumed that for French adults the discrimination is difficult. The F2 value of used /ð/ phoneme was high. Possibly the French adults used the high F2 value as a cue for discriminating /d-ð/ pair. Similarly, Iverson et al. (2003) proposed that behavioral discrimination is affected irrelevantly by the F2 value when Japanese participants categorize the /r-l/ continuum (/r/ and /l/ are members of the same phoneme category in Japanese). In the same study, German participants categorized the /r-l/ continuum with the sensitivity for more critical acoustic cues like F3 formant values.

To summarize, the lower F2 value of Finnish-Hungarian /y/ vowel may have resulted in the smaller magnitude for Finnish-Hungarian and the larger F2 values for French and German /y/ vowels increased the magnitude for them. In the Non-speech condition, the largest N1/MMN response was elicited by Finnish-Hungarian deviant stimulus, however, it did not differ statistically from French and German deviant stimulus response, despite the largest Euclidean distance between Euro-i standard stimulus and Finnish-Hungarian deviant stimulus. It has been observed the larger the difference between standard and deviant stimulus, the stronger the MMN response elicited by deviant stimulus (see for example Aaltonen et al., 1994). This is in contrast to the findings of the present Master's thesis.

The Native language magnet effect considered. The original theory of Kuhl's Native language magnet and the Perceptual magnet effect (NLM and PME, respectively) does not directly provide an explanation for the results of present study. Originally, Kuhl (1991, 1993) proposed that discrimination between category member (near to prototype) and category prototype is difficult. This assumption arose from the behavioral measures of discriminating vowel pairs and it was observed that discrimination was easier for pairs which were near to category boundary. In this Master's thesis the ERP paradigm consisted of the /i/ standard stimulus and /y/ deviant stimuli, hence the condition crossed the category boundary. Furthermore, one of the deviant /y/s was qualified as prototype for the Finnish and two other served as non-prototypes. These differences prevent the direct evaluation of the NLM and PME. However, it was hypothesised that the responses to these /y/ vowels should differ and the results of the present study did not confirmed this without question. The late time window of the fronto-central negativity of MMN response showed differences between languages. Individually chosen prototypes were not used in this study and according to earlier studies (Aaltonen et al., 1997; Frieda et al., 1999) this might have affected the lack of differences between languages in temporal positivity and in the early and middle time windows of the responses. Moreover, it is possible that the participants perceived the Finnish-Hungarian prototype as non-prototype and thus all the stimuli used were perceived as non-prototypes of /y/ (i.e allophones of Finnish /y/ category). Kuhl (1991, 1993) proposed that discrimination near the category prototype is difficult. Perhaps the stimuli in this thesis were all perceived near to the /y/ prototype and thus not differentiated. From that point of view, the findings of this Master's thesis are in a line with the assumptions of NLM.

Several studies have not found support for the NLM and PME. Sharma and Dorman (1998) failed to find behavioural evidence for the magnet effect, the discrimination near to the prototype did not differ from discrimination of non-prototypes in English (British) participants. On the contrary to the NLM, Best et al. (2001) found that (behaviorally measured) discrimination was better near the prototype and actually worse near the phoneme boundary. Best et al. (1988) proposed that adults maintain some discrimination abilities of non-native phonemes to which they have had a little exposure and Polka et al. (2001) also showed some intact discrimination abilities of non-native phonemes in adults. Considering results of these studies, the similarity of the responses between within-category members (/y/ vowels) is not extraordinary and the minor differences between the language responses follow the findings of categorical discrimination in the within-category situation (Aaltonen et al., 1992; Liberman et al., 1957; Serniclaes et al., 2005).

To summarize, it seems plausible that speech specific factors and auditory structure of the stimuli affected the MMN response: the distance between standard and deviant stimulus was not the dominating characteristic of stimuli since the largest distance did not elicit the largest MMN response. This was due to the non-speech as well: statistical differences between language responses were not found. In addition, the inconsistent results concerning the NLM and PME put it into consideration whether there is a stabile phonemic prototype for categories. Rather, it seems that several factors affect the categories, their boundaries and also to the category prototypes (i.e individual prototypes, context, and formant values).

The Perceptual assimilation model and the Speech learning model considered. It is worth considering that primarily predictions of the Perceptual assimilation model (PAM) (Best et al., 2003; 2001; 1995; 1988) and the speech learning model (SLM) (Flege et al., 1997; 2004; 2003) are based on behavioral measures, thus the direct evaluation to ERPs is not straightforward. Concerning the stimuli in this Master's thesis study, the PAM and SLM propose that all /y/ vowels belong to the same category: French and German /y/ prototypes are processed as members of Finnish /y/ category and the similarity of /y/ vowels might complicate the discrimination. These models also lean on the phonetic characteristics of language and context effects. Furthermore, the observations of Liberman et al. (1957) have shown the categorical perception difficulties in within-category situation. It must be taken into account that a minor exposure to French and German vowels (language) is possible in Finland through the use of these languages in the Finnish media. In addition, Kraus et al. (1995) and Winkler et al. (1999) observed that the discrimination ability of stimuli is associated with the strength of MMN response. When stimuli are hard or impossible to discriminate, the MMN response may lack. Thus it is possible that /y/ vowels were hard to discriminate from each other and therefore the differences in languages were not detected. Moreover, the participants were not beforehand trained to discriminate the stimuli or made goodness judgments which may have had enhanced the magnitude of MMN response. Taken together, the small differences in responses are in line with the PAM and SLM.

Limitations of the study. There are some possibly limiting factors concerning this study that should be taken into account. One problem in this Master's thesis data was the small amount of participants. In the visual inspection of the individual data it was seen that averaged waveforms varied between the participants. The response to standard Euro-i varied between blocks although the stimulus was same in all blocks. This observation also tells about the context effect which is discussed above. As described above, the different stimuli could have been affected the results. As noticed earlier, it is possible that there are some small shifts in personal prototypes (Aaltonen et al., 1997; Frieda et al., 1999) which could have affected the results. Thus, detecting the individual prototypes for the participants might help. Luck (2005) highlighted the problem of difference waves. When subtracting the response to standard stimulus from the response to the deviant stimulus, the latency differences of these responses affects the subtraction. However, the methodological consideration and evaluation of difference waves were beyond this thesis. The responses to the Speech condition showed strongly at the right hemisphere's auditory cortex. This might indicate that the speech stimuli resembled non-speech sounds and were not speech like enough. Topographically inspected the responses between languages differed. Hence, the different channel selection or using average of multiple channels should be taken into account in further studies of the present data.

From the phonetics point of view the investigation of speech processing or language specificity by ERPs is not straightforward. Perception and processing of speech (or language) exist in the context of the environment. Speech is not only a signal consisting of different formants; it varies by the speaker and region (Frieda et al., 1999). Each person has their own way to speak and the fundamental frequency (F0) of speech varies by person. It might be that the anchor or prototype of certain phoneme is the phoneme that a person produces oneself. Several studies have also showed that perception of vowels is dependent on the F0 frequency as well as F1 and F2 frequency (Assmann & Nearey, 2003; Nearey, 1989) and other contextual factors (Rosner & Pickering, 1994). Furthermore, Rosner and Pickering (1994) suggested that the vowels have their intrinsic formant values and they are produced differently in different languages. Therefore examining speech perception with the same F0 in all speech stimuli may not attain the whole phenomena of speech perception of different languages.

Aaltonen et al. (1997) discussed in their study that possibly the discrimination differences between the participants arise partly from the difficulty of perceiving synthetic vowels. They considered whether the pronunciation of /y/ vowel with lip-rounding is important and this phenomenon is not possible to achieve by synthesizer. Half of the participants in this Master's thesis reported to hear pips or other sounds during the Speech condition. It is thus possible that the speech stimuli were not speech-like enough. There is a need to considerate natural speech stimuli utilization in the further studies.

Ylinen et al. (2006) observed that vowels presented within syllables elicited stronger MMN responses than presentation of isolated vowels. Lotto et al. (1998) showed in their study that context affects prototypes (vowels presented in pairs or in isolation). Previous studies have been using the stimuli related strongly to characteristics of the language. In these cross-linguistic studies with significant language specific findings, the stimuli used have generally been vowels (Näätänen et al., 1997; Winkler et al., 1999) or other specific features existing only in one language category used in a study (Sharma & Dorman 2000).

Considerations for the further studies. To get reliable information of the processing of prototypes and non-prototypes, it might be worthwhile to use within-category stimuli in one oddball condition, in addition to between-category condition. In this comparison, it is possible to see the effects of different context. Vowels could be assessed also in a context of syllables, with a comparison of presentation in isolation. In further studies, to demonstrate the language specificity and the long-term exposure of language, the ERP data should be compared with the results of behavioral tests, which were carried out, but not used in this Master's thesis. It is worth to consider whether it is more effective to carry out behavioural measures before ERP measurement and thus prepare the participants to perceive the stimuli.

Summary. Taken together, the results of this Master's thesis indicate that in the between-category situation (Euro-i as standard stimulus and /y/ as the deviant stimulus) the processing of the prototype and the non-prototypes was very similar and the differences between /y/ vowels were hard to detect. However, evidence was found for a language specific processing indicated by the smaller fronto-central negativity elicited by Finnish-Hungarian /y/ vowel. The Finnish-Hungarian response was the smallest compared to the responses to French and German /y/ vowels. This language specific effect was contradictory to many earlier studies and thus challenging to interpret. To summarize, the F2 effect seemed to be one of the important interpretive factors affecting responses. The other important factor, proposed by Zhang et al., (2005) was that non-native sounds induce more activation at the brain level. Native sounds are perceived more familiar and the processing requires less activation than the non-speech sounds. The results of this Master's thesis were in a line with the Native language magnet model (NLM), the Perceptual assimilation model (PAM and with the Speech learning model (SLM).

REFERENCES

Aaltonen, O., Eerola, O., Hellstrom, C., Uusipaikka, E., & Lang, A. (1997). Perceptual magnet effect in the light of the behavioral and psychophysiological data. *Journal of the Acoustical Society of America, 101, 1090-1103.*

Aaltonen, O., Eerola, O., Lang, A. H., Uusipaikka, E., Tuomainen, J. (1994). Automatic discrimination of phonetically relevant and irrelevant vowel parameters as reflected by mismatch negativity. *Journal of Acoustical Society of America, 96(3), 1489-1493.*

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

Alho, K. (1995). Cerebral generators of mismatch negativity (MMN) and its magnetic counterparts (MMNm) elicited by sound changes. *Ear and Hearing, 16(1), 38-51.*

Assmann, P. F., Nearey T. M (2003). Frequency shifts vowel identification. *15th ICPhs Barcelona.*

Best, C. T., Halle, P., Bohn, O-S., & Faber, A. (2003). Cross-language perception of non-native vowels: phonological and phonetic effects listeners' native languages, *15th ICPhs Barcelona.*

Best, C. T., McRoberts, G. W., & Goodell, E. (2001) Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of Acoustical Society of America, 109, 775-794.*

Best, C. T., McRoberts, G. W., Lafleur, R., and Silver-Isenstadt, J. (1995). Divergent developmental patterns for infants' perception of two non-native consonant contrasts. *Infant Behavior and Development. 18, 339–350.*

Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance, 14(3), 345–360.*

Čeponienė, R., Rinne, T., & Näätänen, R. (2002). Maturation of cortical sound processing as indexed by event-related potentials. *Clinical neurophysiology*, 113, 870-882.

Čeponienė, R., Torki, M., Alku, P., Koyama, A., Townsend, J. (2008). Event-related potentials reflect spectral differences in speech and non-speech stimuli in children and adults. *Clinical Neurophysiology, 119, 1560-1577.*

Cheour, M., Čeponienė, 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(5), 351–353.*

Cheour, M., Leppänen, P.H.T., Kraus, N. (2000). Mismatch negativity (MMN) as a tool for investigating auditory discrimination and sensory memory in infants and children. *Clinical Neurophysiology, 111, 4–16.*

Coles, M. G. H., Gratton, G., & Fabiani, M. (1990). Event-related brain potentials. *In J. T. Capioppo & L.G. Tassinary (Eds.), Principles of Psychophysiology: Physical, social and inferential elements (pp. 413-455). Cambridge: Cambridge University Press.*

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

Dehaene-Lambertz, G., Gliga, T. (2004). Common neural basis for phoneme processing in infants and adults. *Journal of Cognitive Neuroscience, 16, 1375–87.*

Dehaene-Lambertz, G., Pallier, C., Serniclaes, W., Sprenger-Charolles, L., Jobert, A., & Dehaene, S. (2005). Neural correlates of switching from auditory to speech perception. *NeuroImage, 24, 21-33.*

Dien, J., Spencer, K. M., Donchin, E. (2003). Localization of the event-related novelty response as defined by principal component analysis. *Cognitive Brain Research, 17, 637-650.*

Dien, J., Beal, D. J., Berg, P. (2005). Optimizing principal component analysis of event-related potentials: Matrix type, factor loading weighting, extraction, and rotations. *Clinical Neurophysiology, 116, 1808-1825.*

Dufor, O., Serniclaes, W., Sprenger-Charolles, L., & Démonet, J-F. (2006). Top-down processes during auditory phoneme categorization in dyslexia: A PET study. *NeuroImage, 34, 1692-1707*

Eggermont, J. J., Ponton, C.W. (2002). The neurophysiology of auditory perception: from single units to evoked potentials, *Audiology & Neuro-Otology, 7, 71-99.*

Fitch, R. H., Miller, S., Tallal, P. (1997). Neurobiology of speech perception. *Annual Review of Neuroscience, 20, 331-353*.

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

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

Flege, J. E., Schirru, C., MacKay, I. R. A. (2003). Interaction between native and second language subsystems. *Speech communication, 40, 467-491.*

Frieda, E. M., Walley, A. C., Flege, J. E. & Sloane, M. E. (1999). Adults' perception of native and nonnative vowels: Implications for the perceptual magnet effect. *Perception & Psychophysics, 61(3), 561-577.*

Halsband, U. (2006). Bilingual and multilingual language processing. *Journal of Physiology - Paris, 99, 355-369.*

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

Kazanina, N., Phillips, C., & Ibsardi, W. (2006). The influence of meaning on the perception of speech sounds. *Proceedings of the National Academy of Sciences, 103(30), 11381-11386.*

Kayser, J., & Tenke, C. E. (2003). Optimizing PCA methodology for ERP component identification and measurement: theoretical rationale and empirical evaluation. *Clinical Neurophysiology, 114, 2307-2325.*

Kraus, N., & Cheour, M. (2000). Speech sound representation in the brain. *Audiology and Neuro-Otology, 5, 140-150.*

Kraus, N., McGee, T., Carrel, T.D., King, C., Tremblay, K., & Nicol, T. (1995). Central auditory system plasticity associated with speech discrimination training. *Journal of Cognitive Neuroscience, 7:1, 25-32.*

Kuhl, P. K. (1991). Perception, cognition, and the ontogenetic and phylogenetic emergence of human speech. *In Steven E. Brauth, William S. Hall, Robert J. Dooling (Eds.), Plasticity of development, chapter five, 73-106, University of Maryland, College Park.*

Kuhl, P. K. (1993). Early linguistic experience and phonetic perception: implications for theories of developmental speech perception. *Journal of Phonetics, 21, 125-139.*

Kujala, T., Kallio, J., Tervaniemi, M., Näätänen, R. (2001). The mismatch negativity as an index of temporal processing in audition. *Clinical Neurophysiology, 112, 1712–1719.*

Kujala, T., Tervaniemi, M., & Schröger, E. (2007). The mismatch negativity in cognitive and clinical neuroscience: Theoretical and methodological considerations. *Biological Psychology, 74, 1-19.*

Ladefoged, P. (2006). A Course in Phonetics (Fifth Edition), *Boston, MA: Thomson Wadsworth.*

Leinonen, S., Muller, K., Leppanen, P. H. T., Aro, M., Ahonen, T., & Lyytinen, H. (2001). Heterogeneity in adult dyslexic readers: Relating processing skills to the speed and accuracy of oral text reading. *Reading and Writing: An Interdisciplinary Journal 14,*265-296.

Liberman, A. M. (1998). When theories of speech meet the real world. *Journal of Psycholinguistic Research, 27(2), 111-122.*

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.*

Liberman, A. M., Whalen, D. H. (2000). On the relation of speech to language. *Trends in Cognitive Sciences, 4, 187–196.*

Lotto, A. J., Kluender, K. R., & Holt, L. L. (1998). Depolarizing the perceptual magnet effect. *Journal of Acoustical Society of America, 103(6), 3648-3655.*

Luck, S. J. (2005). An introduction to event-related potential technique. *Cambridge, MA, UK: The MIT press.*

Lyytinen, H., Aro, M., Eklund, K., Erskine, J., Guttorm, T., Laakso, M-L., Leppänen, P. H. T., Lyytinen, P., Poikkeus, A-M., Richardson, U., & Torppa, M. (2004). The development of children at familial risk for dyslexia. Birth to early school age. *Annals of Dyslexia, 54, 184-220.*

Maurer, U., Bucher, K., Brem, S., Brandeis, D. (2003) Development of the automatic mismatch response: from frontal positivity in kindergarten children to the mismatch negativity. *Clinical Neurophysiology, 114, 808–817.*

Möttönen, R., Calvert, G. A., Jääskeläinen, I. P., Matthews, P. M., Thesen, T., Tuomainen, J., & Sams, M. (2006). Perceiving identical sounds as speech or non-speech modulates activity in the left posterior superior temporal sulcus. *NeuroImage, 30, 563-569.*

Nearey, T. M. (1989). Static, dynamic, and relational properties in vowel perception. *Journal of Acoustical Society of America, 85, 2088-2113.*

Näätänen, R. (1992). *Attention and Brain function*. LEA, Hillsdale.

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.K., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychogica*, 42, 313-329.

Näätänen, R., Michie, PT. (1979). Early selective attention effects on the evoked potential: A critical review and reinterpretation. *Biological Psychology, 8, 81–136.*

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., 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. (1989). Do event-related potential reveal the mechanism of auditory sensory memory in the human brain. *Neuroscience letters, 98, 217-221.*

Paavilainen, P., Mikkonen, M., Kilpeläinen, M., Lehtinen, R., Saarela, M., Tapola, L. (2003). Evidence for the different additivity of the temporal and frontal generators of mismatch negativity: A human auditory event-related potential study. *Neuroscience Letters, 349, 79-82.*

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.*

Perrin, F., Pernier, J., Bertrand, O., & Echallier, J.F. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalography and Clinical Neurophysiology*, 72, 184-187.

Phillips, C. (2001). Levels of representation in the electrophysiology of speech perception. *Cognitive Science, 25, 711-731.*

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.*

Polka, L., Colantonio, C., & Sundara, M. (2001). A cross-language comparison of /d/-/ð/ perception: Evidence for a new developmental pattern. *Journal of Acoustical Society of America, 5, 2190–2200.*

Ramus, F., Hauser, M. D., Miller, C., Morris, D., Mehler, J. (2003) Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science, 288, 349-351.*

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.*

Rivera-Gaxiola, M., Csibra, G., Johnson, M. H., & Karmiloff-Smith, A. (2000). Electrophysiological correlates of cross-linguistic speech perception in native English speakers. *Behavioural Brain Research, 111, 13-23.*

Rosner, B. S., Pickering, J. B. (1994) Vowel perception and production. *Oxford: Oxford University Press.*

Serniclaes, W., Van Heghe, S., Mousty, P., Carré, R., & Sprenger-Charolles, L. (2004). Allophonic mode of speech perception in dyslexia. *Journal of Experimental Child Psychology, 87, 336-361.*

Serniclaes, W., Ventura, P., Morais, J., & Kolinsky, R. (2005). Categorical perception of speech sounds in illiterate adults. *Cognition, 98, B35-B44.*

Shalgi, S., & Deouell, L. Y. (2007). Direct evidence for differential roles of temporal and frontal components of auditory change detection. *Neuropsychologia, 45, 1878-1888.*

Sharma, A., & Dorman, M. F. (1998). Exploration of the perceptual magnet effect using the mismatch negativity auditory evoked potential. *Journal of the Acoustical Society of America, 104(1), 511-517.*

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

Sundara, M., Polka, L., Genesee, F. (2006) Language experience facilitates discrimination of /d-ð / in monolingual and bilingual acquisition of English. *Cognition, 100, 369-388.*

Tervaniemi, M., Radil, T., Radilova, J., Kujala, T., & Näätänen, R. (1999). Pre-attentative discriminability of sound order as function of tone duration and interstimulus interval: A mismatch negativity study. *Audiology & Neuro-Otology, 4, 303-310.*

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

Tiitinen, H., Alho, K., Huotilainen, M., Ilmoniemi, R. J., Simola, J., & Näätänen, R. (1993). Tonotopic auditory cortex and the magnetoencephalographic (MEG) equivalent of the mismatch negativity. *Psychophysiology, 30(5), 537-540.*

Yamada, RA., Tohkura, Y. (1992). The effects of experimental variables on the perception of American English /r/ and /l/ by Japanese listeners. *Perception & Psychophysics, 52(4), 376-392.*

Ylinen S, Shestakova A, Huotilainen M, Alku P, Näätänen R. (2006). Mismatch negativity (MMN) elicited by changes in phoneme length: a crosslinguistic study. *Brain Research, 107, 175–85.*

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

Winkler, I., Kujala, T., Alku, P., & Näätänen, R. (2003). Language context and phonetic change detection. *Cognitive Brain Research, 17, 833-844*.

Zhang, Y., Kuhl, PK., Imada, T., Kotani, M., & Stevens, E. (2001). MMFs to native and nonnative syllables show an effect of linguistic experience. *In Nenonen, J., Ilmoniemi, R. J., Katila, T. (eds.), Proceedings of the 12th International Conference on Biomagnetism, Espoo, Finland, Helsinki University of Technology, 379-382.*

Zhang, Y., Kuhl, PK., Imada, T., Kotani, M., Tohkura, Y. (2005). Effects of language experience: neural commitment to language-specific auditory patterns. *NeuroImage, 26, 703–720.*