Auditory cortical tuning to statistical regularities in phonology

Milene L. Bonte[a,*], Holger Mitterer[b], Nadia Zellagui[a], Hanne Poelmans[a], Leo Blomert[a]

[a]*Department of Cognitive Neuroscience, Faculty of Psychology, University of Maastricht, P.O. Box 616, 6200 MD Maastricht, The Netherlands*
[b]*Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands*

Abstract

Objective: Ample behavioral evidence suggests that distributional properties of the language environment influence the processing of speech. Yet, how these characteristics are reflected in neural processes remains largely unknown. The present ERP study investigates neurophysiological correlates of phonotactic probability: the distributional frequency of phoneme combinations.

Methods: We employed an ERP measure indicative of experience-dependent auditory memory traces, the mismatch negativity (MMN). We presented pairs of non-words that differed by the degree of phonotactic probability in a modified passive oddball design that minimizes the contribution of acoustic processes.

Results: In Experiment 1 the non-word with high phonotactic probability (*notsel*) elicited a significantly enhanced MMN as compared to the non-word with low phonotactic probability (*notkel*). In Experiment 2 this finding was replicated with a non-word pair with a smaller acoustic difference (*notsel–notfel*). An MMN enhancement was not observed in a third acoustic control experiment with stimuli having comparable phonotactic probability (*so–fo*).

Conclusions: Our data suggest that auditory cortical responses to phoneme clusters are modulated by statistical regularities of phoneme combinations.

Significance: This study indicates that the language environment is relevant in shaping the neural processing of speech. Furthermore, it provides a potentially useful design for investigating implicit phonological processing in children with anomalous language functions like dyslexia.

1. Introduction

Auditory environments play a crucial role in defining the functional organization of the auditory cortex. Recent studies using invasive electrophysiological recordings have demonstrated that statistical regularities in the basic acoustic properties of environmental sounds shape the neural circuitry in the primary auditory cortex of animals (Nelken, 2004; Zhang et al., 2001). Whether and to what extent neural processing in non-primary auditory cortical areas, for example those involved in phonetic-phonological processing of speech (Jacquemot et al., 2003; Jäncke et al., 2002), are similarly tuned to statistical regularities in

abstract properties of complex sounds, is unknown. In humans, speech constitutes the most relevant complex sound which is routinely dealt with. Thus, it can be hypothesized that the neural system underlying speech perception exploits the distributional properties of speech input to facilitate the acquisition, recognition and representation of spoken language.

The present study investigates event-related potential (ERP) correlates of auditory cortical tuning to distributional frequencies of speech sounds, i.e. phonemes, in the language environment. These distributional frequencies are typically referred to as phonotactic probabilities, where phonotactic refers to the sequential arrangement of phonemes in the syllables and words of a given language (Trask, 1996). In behavioral studies, phonotactic probability has been shown to strongly influence language processing

* Corresponding author. Tel.: +31 43 3884036; fax: +31 43 3884125.
  E-mail address: m.bonte@psychology.unimaas.nl (M.L. Bonte).

across the lifespan. The sensitivity to phonotactic probability emerges during the first year of life, most likely between 6 and 9 months of age (Jusczyk, 1999). Infants may use statistical relationships between neighboring speech sounds to infer which sounds can be combined to form words, and how to segment these words from fluent speech (Saffran et al., 1996). During subsequent language development in children, high phonotactic probability leads to a faster acquisition of words (Storkel, 2001), a better recall of non-words (Gathercole et al., 1999) and a higher accuracy of non-word repetition (Coady and Aslin, 2004). In adults, phonotactic probability affects a large number of language processes, ranging from the speed and ease of spoken word recognition (Vitevitch and Luce, 1999) to listeners' metalinguistic judgments of word-likeness (Bailey and Hahn, 2001). Although these and other behavioral effects are relatively well documented (for a review see Auer and Luce, 2003), neural correlates of phonotactic probability remain largely unexplored.

An ERP measure that is particularly suited for the investigation of experience-dependent neurophysiological changes is the so-called mismatch negativity (MMN). The MMN is elicited by an infrequently occurring discriminable change (i.e. deviant stimuli) in a repetitive aspect of auditory stimulation (i.e. standard stimuli). The MMN represents an automatic change detection response indicative of experience-dependent auditory memory traces (Näätänen, 2001; Näätänen et al., 2001; Picton et al., 2000). The properties of these memory traces may be reflected by the size, latency and topographical distribution of the MMN. Interestingly, the MMN appears to be sensitive to language-specific phoneme representations (Mitterer and Blomert, 2003; Näätänen, 2001; Näätänen et al., 1997; Phillips et al., 2000; Winkler et al., 1999) and lexical representations of words (Jacobsen et al., 2004a; Pulvermüller et al., 2001). In particular, both in adults (Dehaene-Lambertz, 1997; Näätänen et al., 1997) and in infants (Cheour et al., 1998; Dehaene-Lambertz and Baillet, 1998), phonemes that are prototypical in the native language elicit larger MMN responses as compared to phonemes that do not occur in that language but are still discriminable. Moreover, the processing of phoneme contrasts has been shown to elicit an MMN when the phoneme contrast occurs in the native language of the listener but not in case of a non-native contrast that cannot be discriminated (Dehaene-Lambertz et al., 2000).

Here we employ the MMN to investigate whether the phonotactic probability of phoneme clusters influences the processing of meaningless speech. In two passive auditory oddball experiments we presented non-words with medial consonant clusters that either had a high (HPP) or a low (LPP) phonotactic probability (see Fig. 1; Table 1). To minimize the effects of acoustic differences between stimuli, we used a reversed oddball design that has been shown to be suitable for studying mismatch detection at a relatively abstract level of representation (Eulitz and Lahiri,
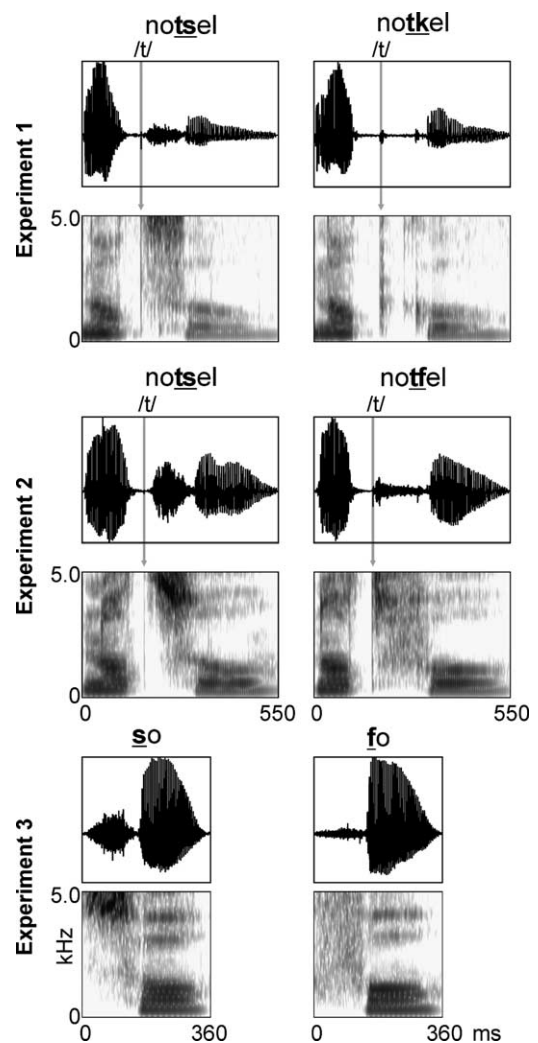

Fig. 1. Waveforms (upper rows) and spectrograms (lower rows) of one exemplar stimulus for each of the non-word stimuli used in Experiment 1: *notsel* [nɔtsəl] and *notkel* [nɔtkəl], Experiment 2: *notsel* [nɔtsəl] and *notfel* [nɔtfəl], and Experiment 3: *so* [sɔ] and *fo* [fɔ]. In Experiments 1 and 2, auditory stimulus deviation occurred around 160 ms at the onset of the stimulus medial consonant clusters /ts/, /tk/, /tf/, and in Experiment 3 at stimulus onset.

2004). In Experiment 1, stimuli were the HPP non-word *notsel* and the LPP non-word *notkel*. In Experiment 2 we used non-words with a comparable difference in phonotactic probability but which are acoustically more similar (see Fig. 1 and Section 2), i.e. the HPP non-word *notsel* and the LPP non-word *notfel*. To further control for purely acoustic effects in Experiments 1 and 2 we conducted a third experiment (Experiment 3) in which we analyzed mismatch effects for /s/ vs. /f/ changes (as in Experiment 2) but in a neutral context. This was achieved by presenting the consonant-vowel syllables *so* and *fo*, which have a comparable phonotactic probability. MMN difference waves were derived from the ERP activity elicited by the same non-word presented as standard and deviant in separate experimental blocks. This enabled us to estimate

Table 1
Log-values of the frequency counts of phonemes and phoneme sequences in our non-word stimuli

cvccvc	cvc[c]vc	cv[cc]vc	c[vcc] vc	cv[ccv]c	c[vccv]c	[cvccvc]
notsel	6.50	5.83	4.32	4.59	3.99	–
notfel	5.69	4.72	3.02	3.82	2.91	–
notkel	6.45	4.51	2.53	0.95	–	–
cv	[c]v	[cv]				
so	6.50	4.60				
fo	5.69	4.72				

Frequency counts were weighted for word frequency and based on the Celex corpus with 42 million Dutch words. [...] indicates the phoneme(s) for which counts are given; c = consonant; v = vowel. For example, the 3rd column: cv[cc]vc, gives frequency counts for /ts/, /tf/ and /tk/, respectively. The frequency counts of the single phonemes /s/, /f/ and /k/ in the 2nd column only include syllable-initial occurrences. – indicates a frequency of 0.

MMN correlates of our non-words independent of variation in general ERP morphology resulting from physical stimulus characteristics. Such variation in ERP morphology is not necessarily related to change detection and may result from the partial overlap of the MMN with other ERP components, e.g. the N1, which may show different amplitude and/or latency characteristics for different consonants (Gage et al., 1998; Obleser et al., 2003).

Enhanced MMNs in response to prototypical phonemes have been suggested to reflect language-specific phoneme traces in the auditory cortex (Näätänen, 2001). Accordingly, if distributional frequencies at the level of phoneme clusters would influence auditory cortical processing of speech sounds, we would expect enhanced MMN responses for the HPP as compared to LPP non-words in Experiments 1 and 2. Importantly, a similar enhancement is not expected for stimuli with comparable phonotactic probability (Experiment 3).

2. Methods

2.1. Subjects

Thirteen undergraduate students (10 female; 2 left handed) participated in Experiment 1, 14 students (13 female; 2 left handed) in Experiment 2, and 14 students (12 female; 1 left handed) in Experiment 3. None of the subjects participated in more than one of the experiments. Subjects gave their informed consent and received course credits or payment for participation. Approval for the study was granted by the Ethical Committee of the Faculty of Psychology at the University of Maastricht. All subjects were native speakers of Dutch.

2.2. Stimuli and task

In Experiments 1 and 2, stimuli were pairs of bisyllabic non-words, one with high and the other with low phonotactic probability (HPP and LPP). The non-words were phonotactically legal in Dutch and the stress was on the first syllable, the default stress in Dutch. Table 1 gives the phonotactic probabilities of our stimuli as determined by counts of phoneme sequences weighted for word frequency in a Dutch word form database (CELEX corpus; Baayen et al., 1995). In Experiments 1 and 2 the same HPP non-word was used, i.e. *notsel* [nɔtsəl]. The LPP non-words were *notkel* [nɔtkəl] in Experiment 1 and *notfel* [nɔtfəl] in Experiment 2. As can be seen in Table 1, the probability of the individual phonemes, that is /s/, /f/ and /k/ is similar for /s/ and /k/ but much lower for /f/. However, the consonant cluster /ts/ has a higher phonotactic probability than /tf/ and /tk/ (/ts/ > /tf/ > /tk/). The relative difference in phonotactic probability between these stimuli is further increased by the vowels /o/ and /e/ that precede and follow the consonant clusters. As illustrated by the waveforms and spectrograms in Fig. 1, the acoustic difference between the non-words was smaller in Experiment 2 than in Experiment 1, i.e. whereas /s/ and /f/ are both fricatives with a difference in place of articulation (alveolar vs. labial), /k/ is a plosive with a velar place of articulation. No Dutch word starts with the phoneme sequence /nots/, /notk/ or /notf/, or contains the phoneme sequences /notsel/, /notkel/ or /notfel/, thus making top-down lexical effects unlikely.

In Experiment 3, stimuli were the consonant-vowel syllables *so* [sɔ] and *fo* [fɔ] which, like *notsel* and *notfel* in Experiment 2, involve a syllable-initial /s/ vs. /f/ contrast, but with a comparable phonotactic probability (see Fig. 1 and Table 1).

Stimuli were spoken by two female native Dutch speakers and recorded at a sampling rate of 44.01 kHz on a DAT recorder. We used a different female speaker for the stimuli in Experiment 1, and the same female speaker for Experiments 2 and 3. The digitized stimuli were D/A converted with a 16 bit resolution, bandpass filtered (100 Hz to 10.5 kHz), resampled at 22.05 kHz, and edited with Praat software (Boersma and Weenink, 2002). Editing included matching for loudness, by equating the maximal amplitude to 95% of the dynamic range, which resulted in equal rms amplitudes for the stimuli in each of the experiments. In Experiments 1 and 2, stimulus length was equated to 550 ms (original range: 480–600 ms), and in Experiment 3 to 360 ms (original range: 345–371) using PSOLA (140–280 Hz as extrema of the F0 contour).

We carefully checked our stimuli for possible alterations in F0 after length equation and did not find any detectable changes. To minimize the likelihood that the MMN would be determined by only one or a few particular acoustic features (Eulitz and Lahiri, 2004; Jacobsen et al., 2004b), we used four utterances of each non-word stimulus. Fig. 2 illustrates that within each experiment, each set of non-word utterances showed comparable variation in pitch and intensity. In Experiments 1 and 2, auditory stimulus deviation occurred around the onset of the medial consonant clusters /ts/, /tk/, /tf/, about 160 ms after stimulus onset. The onset of stimulus deviation is approximate because we used multiple utterances of natural speech stimuli. The preceding vowel /o/ did not show systematic differences between our non-word stimuli. In Experiment 3, auditory stimulus deviation occurred at stimulus onset.

During the experiments, subjects watched a silent movie while the stimuli were presented binaurally through loudspeakers at 65 dB SPL. Stimuli were presented with an inter-stimulus interval of 650 ms. All three experiments consisted of four experimental blocks with 600 stimuli each. In two of these blocks, the HPP non-word (or neutral /so/ in

Expt. 3) served as standard (84%) and the LPP non-word (or neutral /fo/ in Expt. 3) as deviant (16%). In the other two blocks standards and deviants were reversed, i.e. the LPP non-word (or neutral /fo/) was the standard and the HPP non-word (or neutral /so/) the deviant. The order of experimental blocks was counterbalanced between subjects. In total, 1008 standards and 192 deviants were presented for each non-word.

2.3. EEG recording and analysis

EEG data were recorded (0.01–50 Hz, sampling rate 250 Hz) in a sound-attenuating and electrically shielded room from 30 electrode positions (10–20 International System; Nuwer et al., 1998) relative to a nose reference. Eye movements and blinks were measured with bipolar VEOG/HEOG channels. All electrode impedance levels (EEG and EOG) were kept below 5 kΩ. The raw EEG data were corrected for vertical eye movements (i.e. blink artifacts; Semlitsch et al., 1986). EEG data were epoched from −50 to 850 ms relative to stimulus onset, baseline corrected (50 ms pre-stimulus interval), and 1–30 Hz bandpass filtered. Epochs containing data exceeding a maximum voltage criterion of 75 μV were rejected. Standards immediately following deviants were not included in the analysis.

Statistical analyses were performed on electrodes covering the frontocentral (Fz, Cz, FC3, FC4) and centroparietal regions (Pz, CP3 and CP4). At frontocentral sites, MMN amplitude and latency measures were calculated from the ERP signal rereferenced to the average signal of the left and right mastoids, which provides an integrated measure of the total neural activity underlying the auditory MMN (Eulitz and Lahiri, 2004; Schröger, 1998). This integrated measure of MMN activity led to identical experimental results as obtained with the plain ERP signal, but to an improved detection of MMN responses (at frontocentral electrodes) on an individual subject level.

For each experiment, we first analyzed the strength and timing of ERP activity elicited by standard and deviant stimuli using a repeated-measures ANOVA with *stimulus type* (standard vs. deviant), *phonotactic probability* (HPP vs. LPP; Expts. 1, 2) / *syllable* (so vs. fo; Expt. 3), and *electrode sites* as within-subjects factors. Timing of ERP responses in the MMN window was measured by determining individual peak latencies between 250 and 360 ms after stimulus onset (about 90–200 ms after auditory stimulus deviation) in Experiments 1 and 2, and between 90–250 ms after stimulus onset (=auditory stimulus deviation) in Experiment 3. Amplitude measures included both individual peak amplitude and mean amplitude taken from the 50 ms window around the individual peak latency (separately for each subject, condition and electrode).

Mismatch effects were further examined with deviant-standard difference waveforms that were calculated
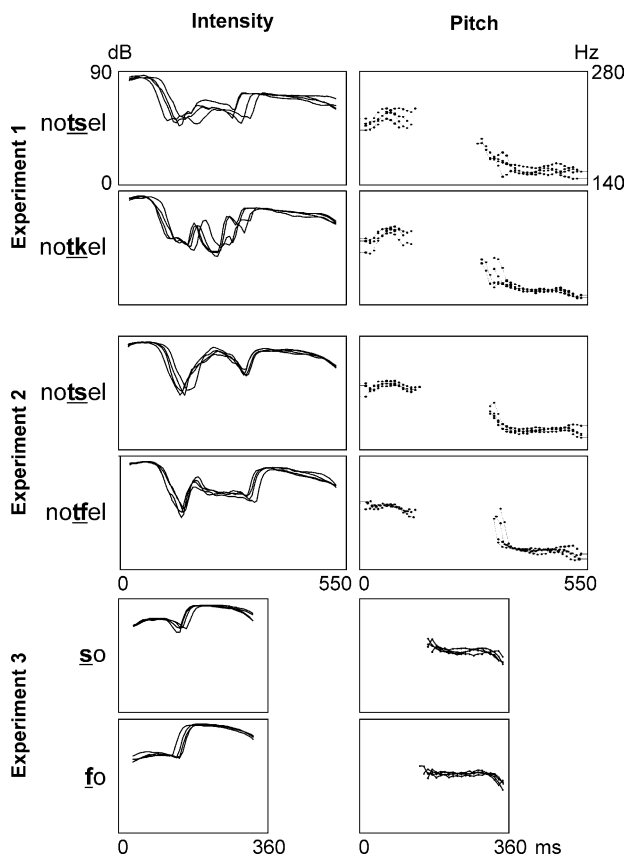

Fig. 2. Pitch and intensity contours of the non-word stimuli used in Experiment 1: *notsel* and *notkel*, Experiment 2: *notsel* and *notfel*, and Experiment 3: *so* and *fo*. We used four utterances of each non-word. Within each experiment, each set of non-word utterances showed comparable variation in pitch and intensity.
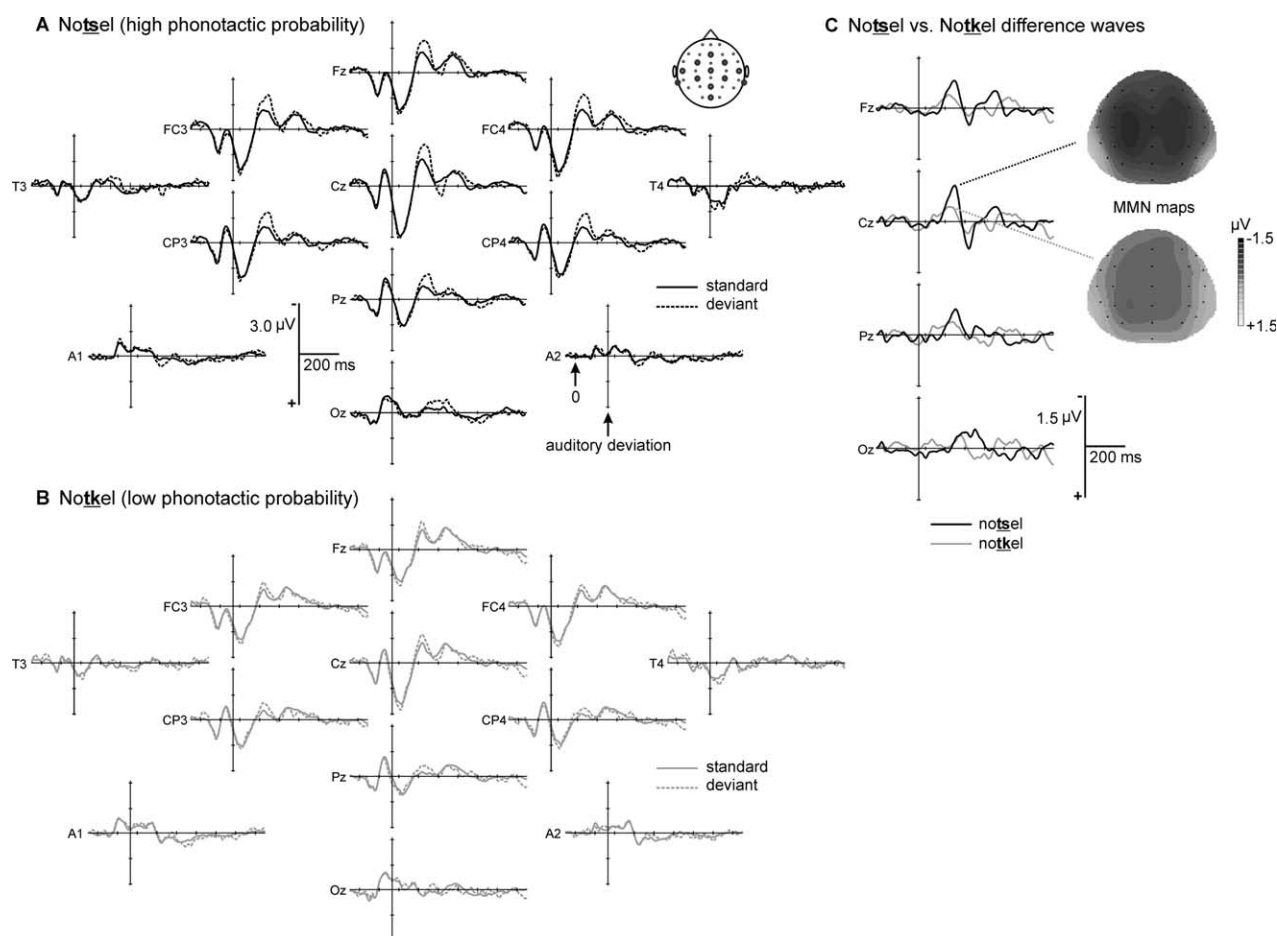
Fig. 3. Grand average standard and deviant waveforms for the non-words (A) *notsel* (high phonotactic probability) and (B) *notkel* (low phonotactic probability) in Experiment 1. (C) Deviant-standard difference waves, and topographic maps of the mismatch negativity (MMN-maps) for *notsel* and *notkel*. Note that auditory stimulus deviation occurred around 160 ms after stimulus onset.

3. Results

3.1. Experiments 1 and 2: effects of phonotactic probability

Figures 3 and 4 show the grand average ERPs (A,B) and difference waveforms (deviant-standard stimuli) (C) for non-words with a high vs. low phonotactic probability in Experiments 1 (Fig. 3) and 2 (Fig. 4). All non-word

separately for each non-word (across blocks), e.g. *notsel* as deviant minus *notsel* as standard. MMN peak latency, peak amplitude and mean amplitude (50 ms window around the peak) were again determined individually for each subject, condition and electrode. MMN parameters were tested with a repeated-measures ANOVA with *phonotactic probability* (HPP vs. LPP; Expts. 1, 2) / *syllable* (*so* vs. *fo*; Expt. 3) and *electrode sites* as within-subjects factors, followed by post hoc *t* tests. We only report results reaching significance at $P < 0.05$ after Geisser–Greenhouse correction.

contrasts elicited an MMN between 100–200 ms after the onset[1] of auditory stimulus deviation (Figs. 3C and 4C—difference waves), with a topographical distribution that is typically reported (Figs. 3C and 4C—MMN maps; Picton et al., 2000; Schröger, 1998). Crucially, in both experiments, the strength of the MMN responses followed the relative phonotactic probability of our non-word stimuli.

Amplitude and latency characteristics of ERP activity in the MMN time-window were first analyzed using a 2 (stimulus type) by 2 (phonotactic probability) by 7 (electrode sites) repeated-measures ANOVA. In Experiment 1, there were general morphological differences in ERP responses to the different non-words. Indeed, *notsel* elicited significantly stronger ERP responses than *notkel* as indicated by main effects of *phonotactic probability* for peak amplitude [$F(1,12) = 38.1$; $P = 0.000$] and mean amplitude in the 50 ms window around the individual

---

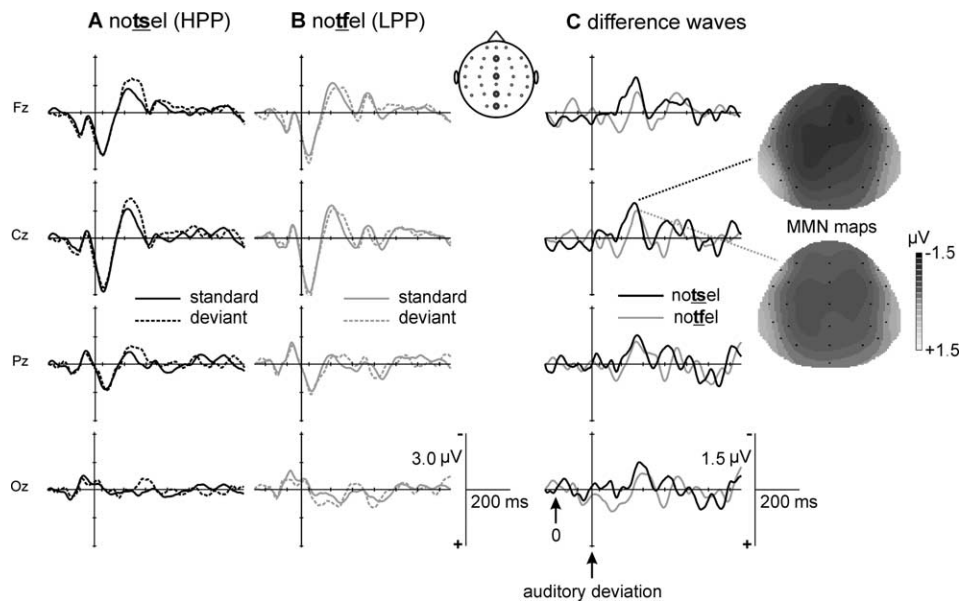[1] This onset is approximate because we used multiple tokens of natural stimuli.

Fig. 4. Grand average standard and deviant waveforms for the non-words (A) *notsel* (high phonotactic probability, HPP) and (B) *notfel* (low phonotactic probability, LPP) in Experiment 2. (C) Deviant-standard difference waves, and topographic maps of the mismatch negativity (MMN-maps) for *notsel* and *notfel*. Note that auditory stimulus deviation occurred around 160 ms after stimulus onset.

peaks $[F(1,12)=42.8; P=0.000]$. Conversely, in Experiment 2 ERP activity for the non-words *notsel* and *notfel* did not show any significant difference in response strength. The timing of ERP responses was very similar for both *notsel* and *notkel* (Expt. 1), and *notsel* and *notfel* (Expt. 2), as indicated by the absence of significant effects of *phonotactic probability* for peak latency.

The main goal of the present study was to investigate whether ERP mismatch responses would be sensitive to manipulation of phonotactic probability. Thus, we further analyzed the responses reflecting stimulus mismatch. As expected, deviant stimuli elicited a significantly stronger negative ERP response than standard stimuli in both experiments (main effects of *stimulus type* for peak amplitude in Expt. 1: $F(1,12)=24.1; P=0.000$, Expt. 2: $F(1,13)=14.9; P<0.005$); and for mean amplitude in Expt. 1: $F(1,12)=15.9; P<0.005$, Expt. 2: $F(1,13)=5.8; P<0.05$). Importantly, this mismatch effect showed a significant interaction with phonotactic probability in both experiments [*stimulus type*-by-*phonotactic probability* interaction for peak amplitude in Expt. 1: $F(1,12)=9.6; P<0.01$, Expt. 2: $F(1,13)=4.6; P=0.05$; and for mean amplitude in Expt. 1: $F(1,12)=9.4; P=0.01$, Expt. 2: $F(1,13)=6.3; P<0.05$]. As illustrated in Figs. 3 and 4, this interaction resulted from stronger mismatch responses to the HPP non-word *notsel* as compared to both the LPP non-word *notkel* (Expt. 1) and the LPP non-word *notfel* (Expt. 2). Peak latency did not show any differences in Experiment 1, but showed a general delay for deviants as compared to standards in Experiment 2 [main effect of *stimulus type* $F(1,13)=14.0; P<0.005$].

To further test the mismatch effects, we examined within non-word differences, by subtracting the activity elicited by standards and deviants across blocks; e.g. *notsel* deviant minus *notsel* standard. Based on the observation that in all conditions, the MMN was prominent at frontocentral sites (Figs. 3 and 4), we tested MMN effects using a 2 (phonotactic probability) by 4 (frontocentral electrode sites) repeated-measures ANOVA. The analysis led to main effects of *phonotactic probability* for MMN peak amplitude [Expt. 1: $F(1,12)=8.9; P<0.025$, Expt. 2: $F(1,13)=10.2; P<0.01$], and mean MMN amplitude [Expt. 1: $F(1,12)=8.2; P<0.025$ and Expt. 2: $F(1,13)=5.3; P<0.05$]. MMN peak latency did not show any significant effect of phonotactic probability.

Figure 5 illustrates MMN amplitude and latency measures at frontal, central and parietal midline electrodes. In Experiment 1, *notsel* (HPP) elicited a significantly stronger MMN than *notkel* (LPP) at Fz ($P<0.01$), at Cz ($P<0.05$), but not at Pz ($P=0.16$). In Experiment 2, *notsel* (HPP) elicited a significantly stronger MMN than *notfel* (LPP) at Fz ($P<0.05$), almost at Cz ($P<0.06$), but not at Pz ($P=0.97$). As for the timing of the MMN responses at these midline electrodes, peak latencies occurred around 325 ms in Experiment 1 and 338 ms in Experiment 2. Relative to auditory stimulus deviation (~ 160 ms after stimulus onset), these latencies correspond to ~ 165 and ~ 178 ms, respectively. With the exception of a significantly longer latency for *notsel* (HPP) vs. *notkel* (LPP) at electrode Cz ($P<0.05$), MMN latencies did not show clear differences between non-words with a high vs. low phonotactic probability.
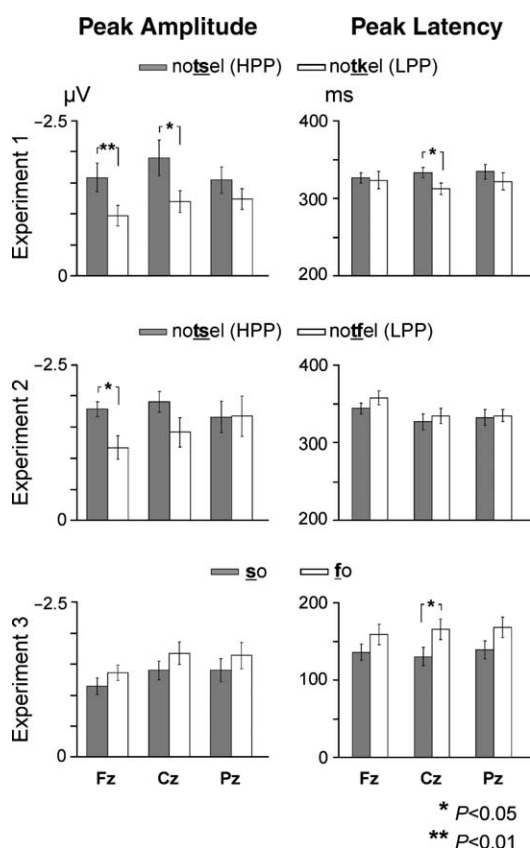
Fig. 5. Mean (SEM) peak latencies and peak amplitudes of mismatch responses elicited by non-words in Experiments 1, 2 and 3. Asterisks indicate significant differences between conditions (post hoc *t* comparisons). Note that latency values are given relative to stimulus onset. Auditory stimulus deviation occurred at ∼ 160 ms after stimulus onset in Experiments 1 and 2 and at stimulus onset in Experiment 3.

3.2. Experiment 3: /s/ vs. /f/ in a neutral context

To further investigate the nature of our findings in Experiments 1 and 2, we performed a third experiment in which we compared mismatch effects for a /s/–/f/ change (as in Expt. 2) with stimuli of comparable phonotactic probability: *so* and *fo*. An interpretation of the enhanced MMN response for the HPP non-word *notsel* vs. the LPP non-word *notfel* observed in Experiment 2 (and *notsel* vs. *notkel* in Expt. 1) in terms of phonotactic probability, would predict such enhancement for /s/ to be disrupted by its presentation in the neutral context of Experiment 3 (i.e. no difference in phonotactic probability). Vice versa, an interpretation of the modulation of MMN responses observed in Experiments 1 and 2 in terms of acoustic differences between the stimuli would predict a similar MMN modulation in Experiment 3. The analysis of the ERP activity elicited by deviants vs. standards of *so* (Fig. 6A,C) and *fo* (Fig. 6B,C) indicates that there is no significant difference between mismatch responses for /s/ and /f/ when phonotactic probability is controlled for, thus lending support to the first interpretation.

As in Experiment 1 and 2, both *so* and *fo* elicited a significant MMN between 100–200 ms after stimulus mismatch (which corresponded to stimulus onset), with a typical topographical distribution (Fig. 6C, compare to MMN maps in Figs. 3C and 4C). A 2 (stimulus type) by 2 (syllable) by 7 (electrode sites) repeated-measures ANOVA of amplitude and latency of ERP activity in the MMN window showed that there were no significant differences in response strength or timing of ERP responses to *so* vs. *fo* stimuli. Importantly, however, deviant stimuli elicited
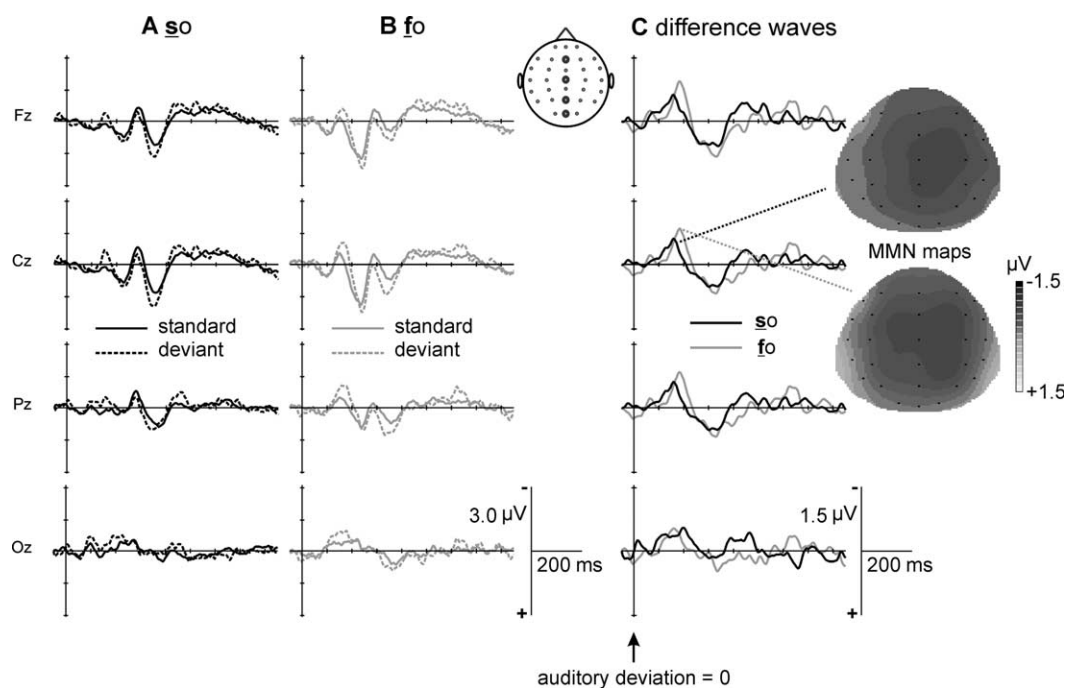

Fig. 6. Grand average standard and deviant waveforms for the consonant-vowel syllables (A) *so* and (B) *fo* in Experiment 3. Like *notsel* and *notfel* in Experiment 2, these syllables involve a /s/ vs. /f/ contrast, but they do *not* differ in phonotactic probability. (C) Deviant-standard difference waves, and topographic maps of the mismatch negativity (MMN-maps) for *so* and *fo*. Auditory stimulus deviation occurred at stimulus onset.

significantly stronger negative ERP responses than standard stimuli, similarly for *so* and *fo*. This is indicated by significant main effects of *stimulus type* for peak amplitude [$F(1,13) = 46.1$; $P = 0.000$] and mean amplitude [$F(1,13) = 39.7$; $P = 0.000$] and the absence of significant *stimulus type*-by-*syllable* interactions for peak amplitude [$F(1,13) = 0.77$; n.s.] and mean amplitude [$F(1,13) = 0.75$; n.s.]. Peak latency did not show any significant difference between standard and deviant responses.

As in Experiments 1 and 2, mismatch effects were further tested by calculating a difference wave for each syllable (e.g. *so* deviant minus *so* standard). Amplitude and latency characteristics were tested with a 2 (syllable) by 4 (frontocentral electrode sites) repeated-measures ANOVA. In concordance with the previous analysis, amplitude measures did not show significant main effects of *syllable*. However, unlike standard and deviant ERPs, the peak latency of the difference waves showed a significant effect of *syllable* [$F(1,13) = 5.1$; $P < 0.05$]. As can be seen in Fig. 6C, this is due to a longer MMN peak latency for *fo* as compared to *so*. Fig. 5 (lower panel) shows that at midline electrodes, the MMN peak occurred around 135 ms for *so* stimuli and around 165 ms for *fo* stimuli. At the level of single midline electrodes, the latency difference was significant at Cz ($P < 0.05$), but not at Fz ($P = 0.18$) or Pz ($P = 0.12$). Peak amplitude measures shown in the left panel of Fig. 5 again illustrate that the results of Experiment 3 did not replicate a response enhancement for a /s/ vs. /f/ contrast in a neutral context.

4. Discussion

The present ERP study used a reversed oddball paradigm to investigate how the phonotactic probability of phoneme clusters influences the processing of meaningless speech. Our results revealed an enhanced MMN response to non-words with a high vs. low phonotactic probability in two separate experiments. Furthermore, a similar MMN enhancement was not found in a third acoustic control experiment in which we presented a single phoneme contrast in the context of a stimulus pair that did not differ in phonotactic probability. The present findings may thus reflect auditory cortical tuning to distributional frequencies of phoneme clusters in the language environment.

Acoustic change detection and phoneme-specific processes have been shown to contribute in parallel to the MMN elicited by speech stimuli (Näätänen, 2001). In this study the MMN was used as a neurophysiological index of a relatively high-level phonological process (as opposed to a mere acoustic process). This type of investigation necessarily requires using spectrally complex stimuli, which not only differ in the high-level variable of interest but also in several other low-level acoustic features. These low-level features may influence amplitude and latency characteristics of the MMN (Näätänen, 2001; Picton et al., 2000). Our

design and analysis ensured that the contribution of such acoustic processes was minimized (see Eulitz and Lahiri, 2004, for a similar example). First, MMN difference waves were derived from the ERP activity elicited by the same non-word presented as standard and deviant in separate experimental blocks. This provided a measure of MMN responses to HPP vs. LPP non-words unconfounded by variation in general ERP morphology that may result from physical stimulus characteristics per se. Second, we used four utterances of each non-word stimulus so that both standard and deviant stimuli were acoustically variable (see Fig. 2). Such acoustic variability is typical in natural speech and reduces the likelihood that the MMN is determined by only one or a few particular acoustic features (Eulitz and Lahiri, 2004; Jacobsen et al., 2004b). Third, across experimental blocks, the acoustic difference between standard and deviant stimuli was identical, with the only difference being the directionality of change, e.g. standard→deviant was either *notsel→notkel* (*notsel→notfel*) or *notkel→notsel* (*notfel→notsel*). It may be argued that our results actually reflect such a directionality effect, e.g. a (t)k→(t)s change would elicit a different response than a (t)s→(t)k change. However, this interpretation appears unlikely. Non-word stimuli were acoustically more similar in Experiment 2 (*notsel* and *notfel*) than in Experiment 1 (*notsel* and *notkel*; see Fig. 1 and Section 2). Thus, if directionality of consonant change would have caused the enhanced MMN for *notsel* in the context of *notkel* (Expt. 1), *notsel* would have been expected to elicit a smaller MMN in the context of *notfel* (Expt. 2). In contrast, *notsel* elicited a very similar MMN in both experiments. Moreover, an MMN enhancement for a /s/ vs. /f/ contrast was not shown in the context of the phonotactically comparable stimuli *so* and *fo* in Experiment 3. This finding further suggests that the MMN results of Experiments 1 and 2 reflect differences in neurophysiological responses to particular phoneme combinations rather than acoustic-phonetic processes at the level of single phonemes. Given the amount of nonlinearities in the auditory system and the acoustic complexity of natural speech stimuli, however, a partial contribution of 'directionality' to the observed effects cannot be excluded completely.

Visual inspection of the MMN effects in each of the three experiments also renders unlikely an interpretation of our findings in terms of individual phoneme frequencies (Figs. 3–6). Whereas the frequencies of the phonemes /s/ and /k/ are relatively similar, /f/ has a much lower frequency (see Table 1). Thus, if the strength of MMN effects would depend on differences in the distributional frequencies of /s/, /k/, and /f/ alone, we would expect a relatively similar MMN for *notsel*, *so*, and *notkel*, but a much smaller MMN for *notfel* and *fo*, which is clearly not the case. In sum, the present MMN results suggest enhanced auditory cortical responses to phoneme clusters that frequently occur in the speech input. These stronger neurophysiological responses may relate to the extensively reported behavioral finding

that stimuli with a high phonotactic probability are easier to acquire, recognize and memorize (see Auer and Luce, 2003).

Why would a non-word with a high phonotactic probability lead to an enhanced MMN response? One possible explanation follows from Hebbian associative learning principles, that is, frequently co-occurring events lead to the formation of neural memory representations (Cruikshank and Weinberger, 1996; Polk and Farah, 1998; Pulvermüller et al., 2001). Previous electrophysiological studies revealed neural changes in animal primary auditory cortex dependent on the distributional frequencies of simple acoustic features (Nelken, 2004; Zhang et al., 2001). The present MMN findings suggest that the frequent exposure to certain phoneme sequences during development, i.e. those with a high phonotactic probability, may lead to enhanced auditory cortical responses and, possibly, to the formation of auditory cortical memory traces. Alternatively, our results may reflect a combination of experience-dependent phonological and basic acoustic influences related to universal principles of phonotactics. It is important to note that these two factors are not independent. Phoneme combinations that are perceptually more distinctive and/or easier to articulate tend to occur more frequently across languages (Hume and Johnson, 2001) and may thus have a higher phonotactic probability.

In order to disentangle the contribution of language-specific and universal processes, it would be interesting to compare the processing of a single set of stimuli in a cross-linguistic study with speakers of languages with different distributional properties for the relevant sounds. However, a cross-linguistic comparison of phonotactic *probabilities* (as opposed to phonotactic *violations*) would be very difficult, because, for example, the same phoneme typically has a different phonetic implementation in different languages (Cho, 2004; Cho and Ladefoged, 1999; Cho and McQueen, 2005). Moreover, in case comparable phoneme sequences do occur in different languages, their phonotactic probabilities tend to be similar. Accordingly, behavioral studies on phonotactic probability typically do not try to distinguish language-specific from universal phonotactics (e.g. Pitt and McQueen, 1998).

Only a few visual studies previously investigated neural correlates of phonotactic probability (Pylkkänen et al., 2002; Stockall et al., 2004). These studies used magneto-encephalography (MEG) to measure brain activity during performance of visual lexical decision tasks and found modulations of event-related responses between 200 and 400 ms after stimulus onset. In both studies, an MEG response around 350 ms showed an earlier latency for words and non-words with a high as compared to low phonotactic probability, indicating facilitated neural processing. Yet, it is not clear whether similar results would occur during spoken language processing because our MMN data show phonotactic probability effects between 160 and 200 ms.

Our ERP results suggest auditory cortical correlates of phonotactic probability, but do not provide a precise anatomical localization. Previous studies have shown that the location of the neural MMN sources depends on the nature of the auditory stimuli involved, with prominent generators located in the superior temporal cortex and in particular the planum temporale (Giard et al., 1990; Liebenthal et al., 2003; Molholm et al., 2005). A recent functional magnetic resonance imaging (fMRI) study indicated that the neural processing of language-specific phonotactic constraints involves the left superior temporal and the left anterior supramarginal gyri (Jacquemot et al., 2003). Although the present MMN findings indirectly suggest that similar auditory cortical areas may provide a neural substrate for the processing of phonotactic probabilities, further studies using different imaging techniques like fMRI will be needed to provide a more precise characterization of the underlying network of brain areas.

5. Conclusion

We used a reversed oddball paradigm to investigate neurophysiological correlates of phonotactic probability. We found significantly enhanced MMN responses to non-words with a high vs. low phonotactic probability suggesting auditory cortical tuning to statistical regularities of phoneme combinations in the language input. Furthermore, this finding was not replicated when we presented a single phoneme contrast in the neutral context of phonotactically comparable stimuli. These findings indicate the relevance of the language environment in shaping the neural system underlying speech perception.

References

Auer Jr ET, Luce PA. In: McLennan CT, Luce PA, Mauner G, Charles-Luce J, editors. Probabilistic phonotactics in spoken word recognition. University at Buffalo. Working papers on language and perception; 2003. p. 164–202.

Baayen RH, Piepenbrock R, Gulikers L. The CELEX lexical database (Release 2) [CD-ROM]. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania [Distributor]; 1995.

Bailey TM, Hahn U. Determinants of wordlikeness: phonotactics or lexical neighborhoods? J Mem Lang 2001;44:568–91.

Boersma P, Weenink D. Praat 4.0: a system for doing phonetics with the computer [Computer software]. Amsterdam, The Netherlands: Universiteit van Amsterdam; 2002.

Cheour M, Ceponiene R, Lehtokoski A, Luuk A, Allik J, Alho K, Naatenen R. Development of language-specific phoneme representations in the infant brain. Nat Neurosci 1998;1:351–3.

Cho T. Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English. J Phonetics 2004;32:141–76.

Cho T, Ladefoged P. Variations and universals in VOT: evidence from 18 languages. J Phonetics 1999;27:207–29.

Cho T, McQueen JM. Prosodic influences on consonant production in Dutch: effects of prosodic boundaries, phrasal accent and lexical stress. J Phonetics 2005;33:121–57.

Coady JA, Aslin RN. Young children's sensitivity to probabilistic phonotactics in the developing lexicon. J Exp Child Psychol 2004;89: 183–213.

Cruikshank SJ, Weinberger NM. Evidence for the Hebbian hypothesis in experience-dependent physiological plasticity of neocortex: a critical review. Brain Res Rev 1996;22:191–228.

Dehaene-Lambertz G. Electrophysiological correlates of categorical phoneme perception in adults. Neuroreport 1997;8:919–24.

Dehaene-Lambertz G, Baillet S. A phonological representation in the infant brain. Neuroreport 1998;9:1885–8.

Dehaene-Lambertz G, Dupoux E, Gout A. Electrophysiological correlates of phonological processing: a cross-linguistic study. J Cogn Neurosci 2000;12:635–47.

Eulitz C, Lahiri A. Neurobiological evidence for abstract phonological representations in the mental lexicon during speech recognition. J Cogn Neurosci 2004;16:577–83.

Gage N, Poeppel D, Roberts TP, Hickok G. Auditory evoked M100 reflects onset acoustics of speech sounds. Brain Res 1998;814:236–9.

Gathercole SE, Frankish CR, Pickering SJ, Peaker S. Phonotactic influences on short-term memory. J Exp Psychol Learn Mem Cogn 1999;25: 84–95.

Giard MH, Perrin F, Pernier J, Bouchet P. Brain generators implicated in the processing of auditory stimulus deviance: a topographic event-related potential study. Psychophysiology 1990;27:627–40.

Hume E, Johnson K. A model of the interplay of speech perception and phonology. In: Hume E, Johnson K, editors. The role of speech perception in phonology. New York: Academic Press; 2001. p. 3–26.

Jacobsen T, Horvath J, Schröger E, Lattner S, Widmann A, Winkler I. Pre-attentive auditory processing of lexicality. Brain Lang 2004a;88:54–67.

Jacobsen T, Schröger E, Alter K. Pre-attentive perception of vowel phonemes from variable speech stimuli. Psychophysiology 2004b;41: 654–9.

Jacquemot C, Pallier C, LeBihan D, Dehaene S, Dupoux E. Phonological grammar shapes the auditory cortex: a functional magnetic resonance imaging study. J Neurosci 2003;23:9541–6.

Jäncke L, Wustenberg T, Scheich H, Heinze HJ. Phonetic perception and the temporal cortex. Neuroimage 2002;15:733–46.

Jusczyk PW. How infants begin to extract words from speech. Trends Cogn Sci 1999;3:323–8.

Liebenthal E, Ellingson ML, Spanaki MV, Prieto TE, Ropella KM, Binder JR. Simultaneous ERP and fMRI of the auditory cortex in a passive oddball paradigm. Neuroimage 2003;19:1395–404.

Mitterer H, Blomert L. Coping with phonological assimilation in speech perception: evidence for early compensation. Percept Psychophys 2003; 65:956–69.

Molholm S, Martinez A, Ritter W, Javitt DC, Foxe JJ. The Neural Circuitry of Pre-attentive Auditory Change-detection: an fMRI Study of Pitch and Duration Mismatch Negativity generators. Cereb Cortex 2005;15: 545–51.

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

Näätänen R, Lehtokoski A, Lennes M, Cheour M, Huotilainen M, Iivonen A, Vainio M, Alku P, Ilmoniemi RJ, Luuk A, Allik J, Sinkkonen J, Alho K. Language-specific phoneme representations revealed by electric and magnetic brain responses. Nature 1997;385: 432–4.

Näätänen R, Tervaniemi M, Sussman E, Paavilainen P, Winkler I. Primitive intelligence in the auditory cortex. Trends Neurosci 2001;24:283–8.

Nelken I. Processing of complex stimuli and natural scenes in the auditory cortex. Curr Opin Neurobiol 2004;14:474–80.

Nuwer MR, Comi G, Emerson R, Fuglsang-Frederiksen A, Guerit JM, Hinrichs H, Ikeda A, Luccas FJ, Rappelsburger P. IFCN standards for digital recording of clinical EEG. International Federation of Clinical Neurophysiology. Electroencephalogr Clin Neurophysiol 1998;106: 259–61.

Obleser J, Lahiri A, Eulitz C. Auditory-evoked magnetic field codes place of articulation in timing and topography around 100 milliseconds post syllable onset. Neuroimage 2003;20:1839–47.

Phillips C, Pellathy T, Marantz A, Yellin E, Wexler K, Poeppel D, McGinnis M, Roberts T. Auditory cortex accesses phonological categories: an MEG mismatch study. J Cogn Neurosci 2000;12: 1038–55.

Picton TW, Alain C, Otten L, Ritter W, Achim A. Mismatch negativity: different water in the same river. Audiol Neurootol 2000;5:111–39.

Pitt MA, McQueen JM. Is compensation for coarticulation mediated by the lexicon? J Mem Lang 1998;39:347–70.

Polk TA, Farah MJ. The neural development and organization of letter recognition: evidence from functional neuroimaging, computational modeling, and behavioral studies. Proc Natl Acad Sci USA 1998;95: 847–52.

Pulvermüller F, Kujala T, Shtyrov Y, Simola J, Tiitinen H, Alku P, Alho K, Martinkauppi S, Ilmoniemi RJ, Naatenen R. Memory traces for words as revealed by the mismatch negativity. Neuroimage 2001;14:607–16.

Pylkkänen L, Stringfellow A, Marantz A. Neuromagnetic evidence for the timing of lexical activation: an MEG component sensitive to phonotactic probability but not to neighborhood density. Brain Lang 2002;81:666–78.

Saffran JR, Aslin RN, Newport EL. Statistical learning by 8-month-old infants. Science 1996;274:1926–8.

Schröger E. Measurement and interpretation of the mismatch negativity. Behav Res Methods, Instr Comp 1998;30:131–45.

Semlitsch HV, Anderer P, Schuster P, Presslich O. A solution for reliable and valid reduction of ocular artifacts, applied to the P300 ERP. Psychophysiology 1986;23:695–703.

Stockall L, Stringfellow A, Marantz A. The precise time course of lexical activation: MEG measurements of the effects of frequency, probability, and density in lexical decision. Brain Lang 2004;90:88–94.

Storkel HL. Learning new words: phonotactic probability in language development. J Speech Lang Hear Res 2001;44:1321–37.

Trask RL, Trask RL. A dictionary of phonetics and phonology. London: Routledge; 1996.

Vitevitch MS, Luce PA. Probabilistic phonotactics and neighborhood activation in spoken word recognition. J Mem Lang 1999;40:374–408.

Winkler I, Lehtokoski A, Alku P, Vainio M, Czigler I, Csépe V, Aaltonen O, Raimo I, Alho K, Lang H, Iivonen A, Naatenen R. Pre-attentive detection of vowel contrasts utilizes both phonetic and auditory memory representations. Cogn Brain Res 1999;7:357–69.

Zhang LI, Bao S, Merzenich MM. Persistent and specific influences of early acoustic environments on primary auditory cortex. Nat Neurosci 2001; 4:1123–30.