Short Communication

Categorical effects in fricative perception are reflected in cortical source information

CrossMark

Sol Lago [a,1], Mathias Scharinger [a,b,*,1], Yakov Kronrod [a], William J. Idsardi [a]

[a] Department of Linguistics, University of Maryland, USA
[b] BioCog, Institute for Psychology, University of Leipzig, Leipzig, Germany

ARTICLE INFO

ABSTRACT

Previous research in speech perception has shown that category information affects the discrimination of consonants to a greater extent than vowels. However, there has been little electrophysiological work on the perception of fricative sounds, which are informative for this contrast as they share properties with both consonants and vowels. In the current study we address the relative contribution of phonological and acoustic information to the perception of sibilant fricatives using event-related fields (ERFs) and dipole modeling with magnetoencephalography (MEG). We show that the field strength of neural responses peaking approximately 200 ms after sound onset co-varies with acoustic factors, while the cortical localization of earlier M100 responses suggests a stronger influence of phonological categories. We propose that neural equivalents of categorical perception for fricative sounds are best seen using localization measures, and that spectral cues are spatially coded in human cortex.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

The perception of speech involves the mapping of acoustic information onto discrete and language-specific phonological categories. Previous research has shown that these phonological categories strongly influence how speech is perceived: category information allows for better discrimination of acoustically equidistant sounds if they cross a category boundary than if they do not ('categorical perception', Harnad, 1987). Although within-category discrimination is still possible, it is reduced for plosive consonants such as [p–b], [t–d] and [k–g] and enhanced for vowels (Eimas, 1963; Fry, Abramson, Eimas, & Liberman, 1962; Marklund, Schwarz, & Lacerda, 2014; Pisoni, 1973). The reasons for this contrast remain unsettled but may relate to acoustic or phonological differences between vowels and plosives. Plosives are generally discriminated based on temporal cues such as voice onset time and formant transitions, whereas the discrimination of vowels relies more on their steady-state spectral properties (Drullman, Festen, & Plomp, 1994; Kewley-Port, 1983), which are spatially coded in human cortex (Obleser, Elbert, Lahiri, & Eulitz, 2003; Obleser, Lahiri, & Eulitz, 2004). Alternatively, the contrast might

be driven by phonological or language-specific differences between consonants and vowels; for instance, vowels tend to occupy a prominent position in the syllable hierarchy, they can be stressed and they carry prosodic information (Spencer, 1996).

Fricatives (such as in English 'see' or 'she') can inform the reasons for this contrast since they share properties with both plosive consonants and vowels. Similarly to plosives, fricatives are consonants. However, like vowels, fricatives can be discriminated based on spectral information, as determined by the spectral peaks in the frication noise (Jongman, Wayland, & Wong, 2000). Previous research found inconsistent categorical effects for fricatives, which were discriminated categorically in some studies but not in others (Fujisaki & Kawashima, 1968; Healy & Repp, 1982; Mann & Repp, 1980; Repp, 1981). However, these studies mainly focused on behavioral tasks (for an exception, see Lipski, 2006). Here we examine whether fricative perception shows enhanced within-category discrimination, similarly to vowels, or whether it is strongly categorical, similarly to plosive consonants. We address how acoustic and phonological factors affect the discrimination of a fricative continuum from 'she' to 'see'. We use event-related fields (ERFs) in a Mismatch Negativity (MMN) design (Näätänen, Paavilainen, Rinne, & Alho, 2007; Schröger, 2005) and we model the source of the M100 response, a negative deflection that indicates sound-specific processing (Näätänen & Picton, 1987). Both the MMN amplitude and M100 dipole locations have been previously found to be sensitive to category

---

membership in addition to acoustic differences (Kazanina, Phillips, & Idsardi, 2006; Obleser et al., 2004; Phillips et al., 2000). We show that both behavioral responses and the localization of the M100 component provide reliable indices of categorical processing.

2. Results

In the behavioral tests, participants performed identification and discrimination tasks on a 10-step continuum from [ʃi] to [si]. The MEG experiment comprised four blocks determined by stimuli that belonged to either the [ʃ] category (ʃ1, ʃ2) or to the [s] category (s1 and s2; Fig. 1A). These stimuli served as standards in each of the four blocks. Each block contained two types of deviants: a deviant that belonged to the same phonological category as the standard (*deviant within*) and a deviant that belonged to a different phonological category (*deviant across*), as determined by the identification test. Crucially, when the standards were ʃ1 and s1, deviants across differed from standards in category membership and they were acoustically more distant from standards than deviants within. When the standards were ʃ2 and s2, deviants across differed from standards only in category membership and they were the same acoustic distance apart from standards as deviants within (Fig. 2A).

2.1. Behavioral

In the identification task, the logistic regression revealed that the category boundary across participants corresponded approximately to step 7 in the fricative continuum (50% point = 6.89). In the discrimination task, participants showed evidence of

categorical perception: acoustically equidistant pairs were discriminated more accurately when they crossed the category boundary than when they did not (Fig. 1B). Linear-mixed effects modeling revealed that participants' discrimination was influenced by both acoustic and phonological information: d' scores increased with increasing acoustic distance between the pair members (main effect of acoustic distance: $\hat{\beta} = 0.57$, $SE = 0.09$, $p < 0.01$) and they also increased when the two sounds belonged to a different phonological category than when they belonged to the same category (main effect of phonological category: $\hat{\beta} = 2.17$, $SE = 0.32$, $p < 0.01$). Importantly, there was a negative interaction between these two factors, which showed that acoustic distance had a smaller effect for across-category pairs than for within-category pairs ($\hat{\beta} = -0.47$, $SE = 0.09$, $p < 0.01$). Thus, the effect of acoustic distance was significantly modulated by the category membership of the sounds.

2.2. MEG

2.2.1. MMN amplitudes

Participants showed reliable MMN effects in posterior sites in the right and left hemispheres (*MMN right*: $\hat{\beta} = -16.18$, $t = -5.71$, $p < 0.01$; *MMN left*: $\hat{\beta} = 6.65$, $t = 2.65$, $p < 0.01$). In addition, mismatch responses were stronger on the right than on the left, as shown by a significant interaction between the MMN mean absolute amplitude and the laterality factor ($\hat{\beta} = -8.32$, $t = -2.52$, $p < 0.05$). Therefore, only right posterior sites were used in further analyses, although plots for the left hemisphere are included in the Supplementary Materials to provide a full description of the data (Fig. S2).
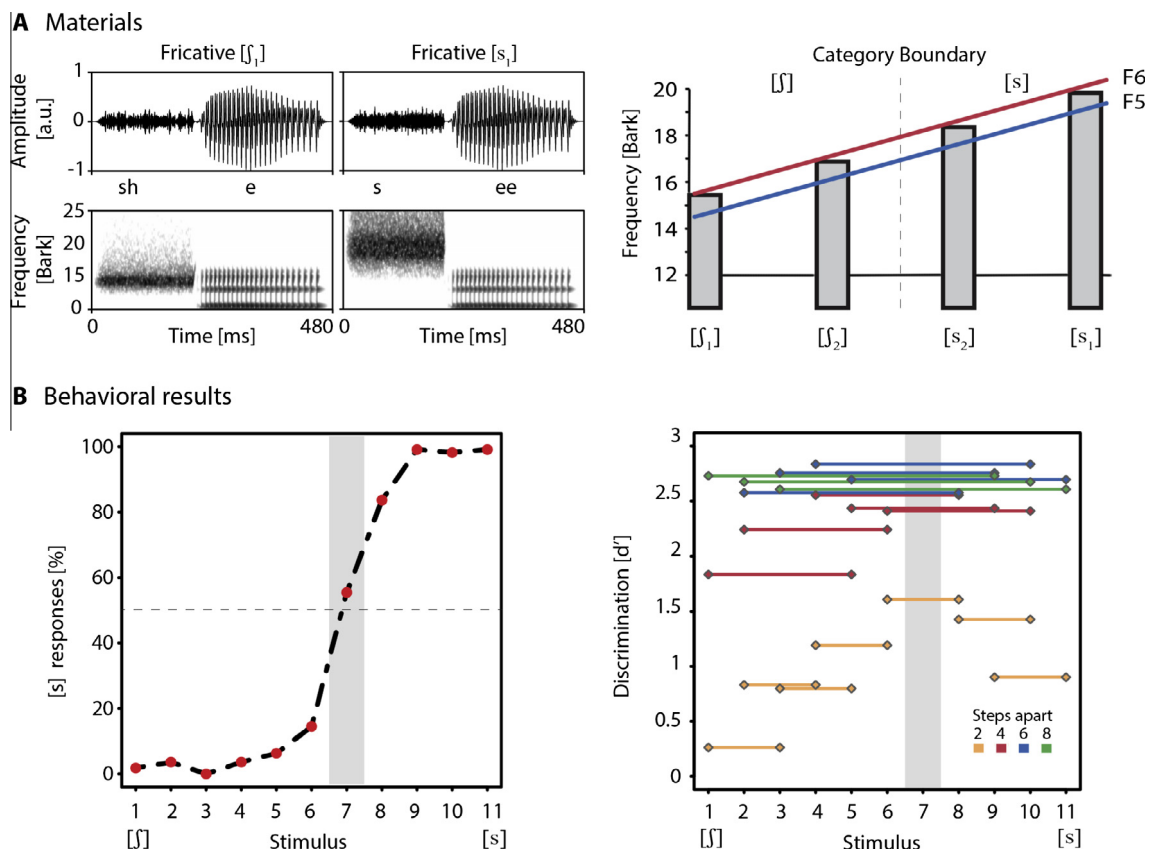

Fig. 1. (A) *left*: Wave form and spectrogram of representative stimulus used in the behavioral and MEG experiments. (A) *right*: Illustration of the MEG experiment design. (B) *left*: Results from the identification task. (B) *right*: Perceptual sensitivity expressed in d' in the discrimination task.

A MEG Design

I [ʃ₁] [ʃ₁] [ʃ₁] [ʃ₁] [ʃ₂] [ʃ₁] [ʃ₁] [ʃ₁] [s₁]   **II** [s₁] [s₁] [s₁] [s₂] [s₁] [s₁] [s₁] [s₁] [ʃ₁]

III [ʃ₂] [ʃ₂] [ʃ₂] [ʃ₁] [ʃ₂] [ʃ₂] [ʃ₂] [ʃ₂] [s₂]   **IV** [s₂] [s₂] [s₂] [s₂] [s₁] [s₂] [s₂] [s₂] [ʃ₂]

standard   deviant within   deviant across   standard   deviant within   deviant across
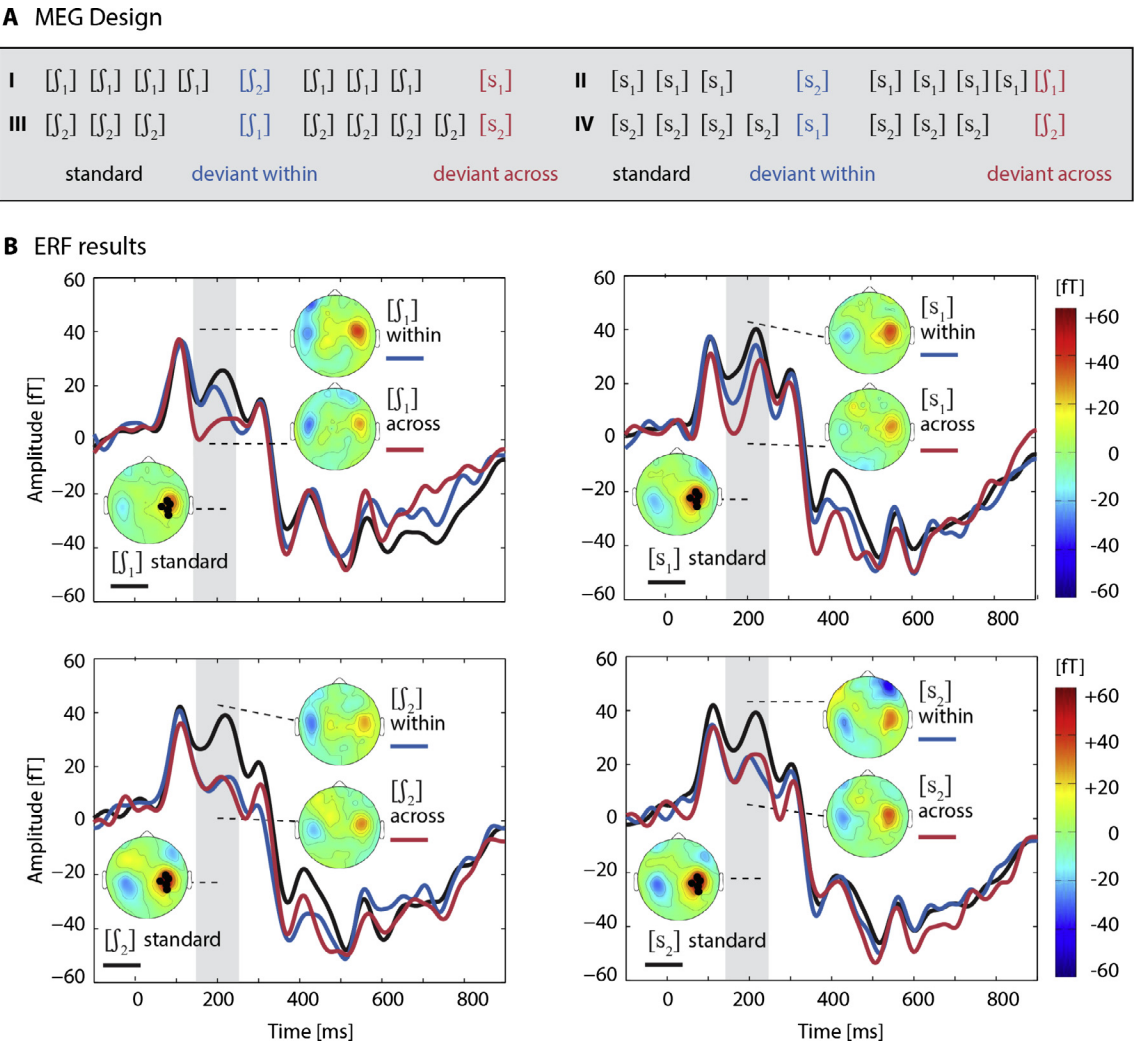
B ERF results

Fig. 2. (A) Optimal passive oddball paradigm used in the MEG study. Each standard was interspersed with two types of deviants, belonging to the same or a different phonological category as the standard. For ʃ₁ and s₁ stimuli, deviants across and within differed in category membership and acoustic distance; for ʃ₂ and s₂ stimuli, deviants across and within only differed in category membership. (B) MMN patterns at posterior right channels. The MMN window (150–250 ms) is highlighted in gray, and topographic plots are calculated as average responses in the MMN time window. The 5 channels selected for analysis are shown in black.

Statistical analyses in the 150–250 ms time window revealed that MMN responses were modulated by acoustic distance but not by category membership alone. For pooled responses to ʃ₁ and s₁ sounds, where deviants across belonged to a different phonological category and were acoustically more distant from standards than deviants within, deviants across elicited marginally stronger MMN responses than deviants within ($\hat{\beta} = -9.17$, $t = -1.76$, $p = 0.08$). In contrast, deviants across and within did not differ for ʃ₂ and s₂ sounds, where both deviants were

acoustically equidistant from standards, and deviants across differed from deviants within only in their category membership ($\hat{\beta} = 0.72$, $t = 0.16$, $p = 0.87$). MMN responses across right posterior sites are shown in Fig. 2B.

2.2.2. M100 dipoles

M100 mean dipole locations of the four standard types differed along the lateral–medial axis (Fig. 3). We focused on this axis since it has been previously found to reflect a tonotopic distribution of
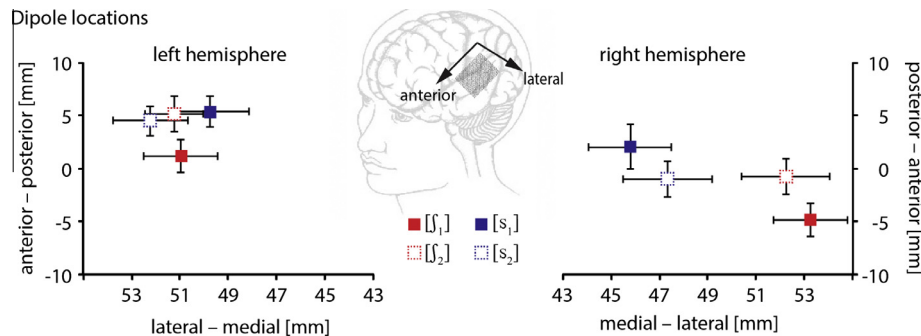

Fig. 3. Mean estimated ECD locations for the M100 component (80–130 ms) in left and right hemispheres.

M100 dipole sources, with higher frequencies corresponding to more medial coordinates and lower frequencies corresponding to more lateral coordinates (Pantev et al., 1995). While no clear patterns were observed in the left hemisphere, in the right hemisphere the location of the standard sounds showed a pattern consistent with categorical perception: although the acoustic distance between ʃ$_1$–ʃ$_2$, ʃ$_2$–s$_2$, and s$_2$–s$_1$ was identical, the locations of the dipoles in response to these sounds showed smaller distances between members of the same category than between members of different categories.

This observation was supported by pairwise comparisons between the dipoles of these sounds. There was not a statistical difference between the locations of ʃ$_1$ and ʃ$_2$ ($\hat{\beta} = 1.26$, $t = 0.60$, $p = 0.56$) or s$_2$ and s$_1$ ($\hat{\beta} = 1.53$, $t = 0.93$, $p = 0.37$). In contrast, the difference between ʃ$_2$ and s$_2$ approached significance ($\hat{\beta} = 2.82$, $t = 1.89$, $p = 0.08$). This suggests that although the acoustic distance between these sounds was identical to the distance between ʃ$_1$–ʃ$_2$ and s$_2$–s$_1$, their perception, as reflected in their location along the medial–lateral axis, differed more strongly when they belonged to different phonological categories. However, the interaction between acoustic information (expressed as the averaged F5 and F6 formant values of each stimulus) and phonological category (expressed as membership to the /s/ or /ʃ/ categories) did not reach significance when entered in the full model ($\hat{\beta} = 0.23$, $t = 0.14$, $p = 0.89$). Thus, the results from the pairwise comparisons should only be taken as suggestive and they are discussed further in the General Discussion.

3. Discussion

This study examined the relative contribution of acoustic and phonological information to the perception of fricative sounds. Participants' behavior showed clear evidence of categorical perception: in the right hemisphere, acoustically equidistant pairs were discriminated more accurately when they crossed a category boundary than when they did not. In addition, the localization of M100 responses showed sensitivity to phonological information: M100 sources to acoustically equidistant exemplars from the [ʃi–si] continuum showed that within-category sounds elicited dipoles whose coordinates clustered together along the medial–lateral axis. Importantly, when the same acoustic distance between sounds crossed a category boundary, there was a trend for the resulting dipoles to differ in location along the medial–lateral axis.

These results suggest that category information can result in a 'warping' of dipole locations: equidistant physical stimuli were mapped onto phonological representations resulting in smaller within than across-category differences along the medial–lateral axis. This indicates that phonological information can influence acoustic processing, a conclusion that is in line with previous findings that tonotopic organization in auditory cortex is modulated by top-down information (e.g. Oh, Kwon, Yang, & Jeong, 2013). A broader implication of these findings is that specific phonemes may be represented by neuronal populations that are separated enough to resolve to different dipoles. This hypothesis is consistent, for instance, with findings on visual perception and object recognition, where optical imaging and single-cell recordings have suggested that inferotemporal cortex is composed of a mosaic of neuronal patches that respond to basic shapes, whose combination can then encode arbitrary objects (e.g. McCandliss, Cohen, & Dehaene, 2003; Tanaka, 2003; Tsunoda, Yamane, Nishizaki, & Tanifuji, 2001).

These results, however, should be taken with caution, since they were not supported by an interaction between acoustic information and phonological category in the full statistical analysis of

dipole locations. The lack of an interaction is likely to have resulted from the limited amount of data in combination with the high collinearity between these two factors. Collinearity problems are pervasive in designs that include a continuous predictor and a categorical predictor that is a transformation on the same scale (Friedman & Wall, 2005; Wurm & Fisicaro, 2014). Further research and a different experimental design will be needed to further dissociate the contribution of phonological and acoustic factors to dipole locations.

In contrast with the localization results, we found effects of acoustic distance but no evidence of categorical perception in participants' MMN responses. These results differ from previous studies on plosive consonants. MMN responses to plosive consonants tend to be strongly modulated by category information, in contrast with MMN responses to vowels, which usually show sensitivity to within-category differences (e.g., *plosives*: Kazanina et al., 2006; Phillips et al., 2000; *vowels*: Hacquard, Walter, & Marantz, 2007; Marklund et al., 2014; Näätänen et al., 1997). The lack of a modulatory effect of category information for fricative sounds suggests that the contrast between consonants and vowels might be due to the extent to which vowels and plosive consonants rely on spectral and temporal information. Previous MMN studies on plosive consonants have mostly focused on voicing contrasts, which are primarily based on differences in temporal structure, that is, on voice-onset time differences that describe the lag between the release of consonant closure and onset of quasi-periodic vocal fold vibration.

In contrast, the main acoustic contrast between the phonological categories of [ʃ] and [s] in our study was conveyed in the steady state F5 and F6 formant locations. Note that participants could not rely on secondary cues such as formant transitions (cf. Jongman et al., 2000) since fricatives were consistently followed by the same vowel. We hypothesize that the MMN showed stronger sensitivity to acoustic than phonological differences because the main cue for fricative discrimination was spectral, rather than temporal. A greater sensitivity to acoustic information for spectral processing is consistent with studies on spectrally based speech sound differences, which have usually found similar MMNs for within- and across-category contrasts (Kasai et al., 2001; Marklund et al., 2014; Sharma & Dorman, 1998; Sharma, Kraus, McGee, Carrell, & Nicol, 1993).

A strong reliance on spectral cues may also explain the right-lateralization of MMN responses to fricative sounds in our study. Although previous studies have sometimes found a right-hemispheric dominance for the MMN (e.g., Paavilainen, Alho, Reinikainen, Sams, & Näätänen, 1991; Rinne, Alho, Ilmoniemi, Virtanen, & Näätänen, 2000) responses to contrasts between plosive consonants are typically left-lateralized (e.g. Kazanina et al., 2006). However, the processing of spectral information tends to rely more on the right hemisphere (e.g. Zatorre & Belin, 2001), which may explain the right-lateralization of the MMN responses that we observed in our study. Alternatively, the pre-attentive nature of the MMN design in combination with the frication noise of the stimuli may have induced participants to process the sounds in a non-speech mode in the MEG task, yielding stronger responses in the right hemisphere, which is usually more driven by acoustic processing. In contrast, the behavioral task may have induced a more phonologically driven processing mode since participants were explicitly asked to categorize the sounds. A third possibility is that the absence of a left-lateralized MMN component might have been due to the particular nature of the phonological contrast between [s–ʃ]. These sounds are both coronal, which may be a non-primary distinction and might be underspecified with regard to place of articulation (e.g. Lahiri & Reetz, 2010).

Overall, our results suggest that categorical perception is not an 'all-or-nothing' effect, but rather that it depends on the level of

processing and on the nature of the acoustic contrast under study. To this end, the behavioral findings suggest that attentional measures of speech processing reflect the influence of phonological information, while earlier, pre-attentive MMN responses index a stronger influence of acoustic (spectral) information. Crucially, however, our M100 dipole findings illustrate that phonological information is available at very early processing stages, and that category membership might result in warped spatial representations in cortex.

4. Methods

4.1. Participants

Behavioral data was obtained from eleven students at the University of Maryland (9 females, age range = 19–31 years, mean age = 21.8, 10 right-handed) and MEG data from sixteen volunteers (9 females, age range = 19–22 years, mean age = 20.8, all right-handed). All participants were native English speakers with normal or corrected-to-normal vision, provided informed consent and received course credit or financial compensation for their participation. They had no self-reported hearing impairments or neurological disorders. All procedures followed the guidelines of the local ethics committee from the University of Maryland and were in accordance with the Declaration of Helsinki.

4.2. Materials

Stimuli were synthesized according to a previously established protocol (McQueen, Jesse, & Norris, 2009) using the Klatt synthesizer as implemented in McMurray, Tanenhaus, and Aslin (2002), together with the phonetic sound software PRAAT (Boersma & Weenink, 2011). The synthesized stimuli consisted of a 210 ms fricative portion, followed by the vowel [i] with a duration of 270 ms resulting in the words [ʃi] ('she') and [si] ('see', Fig. 1A). The fricative portion of the sounds differed along a 10-step continuum from a clear instance of [ʃ] to a clear instance of [s]. The fricative continuum was constructed by manipulating the location of the 5th and 6th formants (spectral peaks, F5 and F6) in Bark space, which was chosen to account for the non-linearity of the frequency resolution of the human auditory system (Zwicker, 1961). For F5, formant values ranged from 14.5 to 19.5 Bark and for F6, from 15.5 to 20.5 Bark (Table S1). These values approximate the mean spectral locations for [ʃ] and [s] in natural productions (Jongman et al., 2000).

For the MEG experiment, we selected four sound positions that evenly spanned the category boundary determined by the behavioral tests. These four positions were labeled ʃ$_1$, ʃ$_2$, s$_2$, and s$_1$, with ʃ$_1$ and s$_1$ referring to clear category members and ʃ$_2$ and s$_2$ referring to members located closer to the category boundary. For each of these positions we selected three different tokens 1/8 Bark apart from each other, in order to add acoustic variability and to encourage MEG participants to map the experimental stimuli onto phonological categories as opposed to specific sounds (cf. Phillips et al., 2000). Onset and offset ramps (5 ms) for all stimuli ensured that acoustic artifacts were minimized. Average sound intensities were normalized to 70 dB (based on the mean power in Pa^2/s) in order to guarantee reliable auditory delivery at approximately 60 dB SPL in the MEG scanner.

4.3. Procedure

4.3.1. Behavioral experiments

Identification and discrimination tasks were carried out in the MEG scanner in order to ensure similar testing conditions across experiments. Participants were lying in supine position and wore non-magnetic Etymotics in-ear headphones (ETYMOTIC Research Inc., Elk Grove Village, IL). Instructions were delivered visually and projected on a screen mounted 20 cm above participants' heads. A two-button response device registered participants' responses.

In the identification task, participants heard sounds from the [ʃi–si] continuum and were instructed to classify the fricative portion as either [ʃ] or [s] by means of button presses. Response-button assignment was counter-balanced across participants. Each trial began with a 500 ms blank screen. Then a fixation cross appeared for 300 ms and was immediately followed by a sound from the continuum. Each sound was repeated ten times across the experiment in a randomized order. If a button press was not registered 4 s after trial onset, the next trial began.

In the discrimination task, participants gave same/different judgments to sound pairs taken from the continuum. We varied the acoustic distance between the sounds in each pair, which could be 2 (e.g. stimuli 1–3), 4 (e.g. 1–5), 6 (e.g. 2–8), 8 (e.g. 1–9) or 10 steps apart (1–11). Each pair was repeated 3 times across the experiment. To counterbalance, we introduced a similar number of identical pairs, resulting in a 46:54 same-to-different ratio. Each trial started with a fixation cross that was displayed for 300 ms and was immediately followed by the experimental sounds, separated by a 300 ms inter-stimulus interval (ISI). The next trial began when participants provided a judgment, or if no response was recorded after 2 s. Trials were separated by 500 ms during which the screen remained blank.

4.3.2. MEG experiment

Stimuli were presented in a modified many-to-few optimal passive oddball paradigm (Näätänen, Pakarinen, Rinne, & Takegata, 2004; Partanen, Vainio, Kujala, & Huotilainen, 2011). Frequently-presented standard stimuli were interspersed with two types of infrequent deviant stimuli approximating a 7:1:1 ratio. Per block, the total number of standards approximated 708 and the number of deviants approximated 208 (104 per each deviant type), summing up to a maximum of 916 stimuli presentations. Stimuli were pseudo-randomized such that no more than 5 and no less than 2 standards occurred in a sequence. Trial duration randomly varied between 900 and 1250 ms. Altogether, each block lasted approximately 16 min. Block order was counterbalanced across participants.

Participants passively listened to the sounds at a comfortable listening level (∼60 dB). In order to maintain wakefulness, they watched a self-selected movie without sound. Continuous MEG was registered from 157 channels (KIT, Japan), with three fiducial and two pre-auricular markers attached to participants' heads. Prior to MEG acquisitions, markers were digitized with a Polhemus Fastrak Digitizer system, together with an approximation of their head shapes, registering an average of 1000 head-surface points outside the scanner. Fiducials and pre-auricular markers were co-registered inside the MEG before and after each experimental session in order to account for head movements and to enable M100-dipole modeling. MEG acquisition used a sampling frequency of 1000 Hz, an online 200-Hz low-pass together with a DC high-pass filter and a 50 Hz notch filter for the removal of ambient electric oscillations.

4.4. Data analysis

4.4.1. Behavioral

Statistical analyses were carried out with R, an open source programming language and environment for statistical computing (R Development Core Team, 2014). We used the *lme4* package

(Bates, Maechler, Bolker, & Walker, 2014) and computed *p*-values using the Satterthwaite's approximation for denominator degrees of freedom with the *lmerTest* package (Kuznetsova, Bruun Brockhoff, & Bojesen Christensen, 2014). All statistical models reported here included by-participant random intercepts in order to account for inter-subject variability.

In the identification task, identification scores were calculated as the mean of participants' responses for each step in the continuum. The category boundary across participants was obtained by fitting a mixed-effects logistic regression to participants' responses (Jaeger, 2008). The boundary was then estimated as the point where identification performance was at chance (=50% point). In the discrimination task, we determined perceptual sensitivity by calculating *d'* scores for each sound pair (Macmillan & Creelman, 2005). The log-linear correction method described in Hautus (1995) was used to avoid the appearance of non-finite values in the case of extreme false alarm or hit rates.

4.4.2. MEG

MEG data were first cleaned from vibration and sensor-noise using the routines by de Cheveigné and Simon (2007, 2008) implemented in MATLAB (ver. 7.9, The Mathwork Inc.). Data were subsequently pre-processed using Fieldtrip (Oostenveld, Fries, Maris, & Schoffelen, 2011). The epochs of interest comprised a pre-stimulus baseline of 100 ms and a following 900 ms window. SQUID jump, muscle, and eye-related artifacts were determined using Fieldtrip routines. Trials with artifacts were rejected if any of the artifact detection algorithms exceeded threshold of 20 pT. This led to the exclusion of not more than 15% of standard or deviant trials. One participant was excluded from further analyses due to an extreme percentage (>50%) of rejected trials.

Epochs (−100 to 900 ms) were then filtered with a finite impulse response (FIR) low-pass filter of 20 Hz, and a high-pass filter of 0.5 Hz. Baseline correction was done by subtracting the mean amplitude of the pre-stimulus interval (−100 to 0 ms) from the epoch. Trials were averaged separately for the four standard and eight deviant types. The MMN time-window (150–250 ms) was determined from visual inspection of the deviant grand average waveform. Measured from deviance onset, this window corresponded to the typical time window of the MMN (Näätänen et al., 2007; Winkler, 2007). For subsequent analyses, we selected 5 channels from each of the four quadrants, left anterior, left posterior, right anterior, right posterior, at which absolute amplitudes were maximal for the 8 deviant conditions averaged together.

The main statistical analysis examined whether the mean amplitude of responses to across- and within-category deviants differed in the 150–250 ms time window. For this purpose, fricatives types $\int_1$–s_1 were averaged together, as for both sounds across- and within-category deviants differed in category membership and acoustic distance to standards. Similarly, responses to fricatives types $\int_2$–s_2 were averaged together as for both sounds across and within-category deviants differed only in category membership.

4.4.3. M100 dipole fitting

Equivalent current dipoles (ECDs) were calculated for the M100 component of the standard stimuli. ECDs were not calculated for the MMN component. In contrast with the N1, which is commonly modeled by one generator per hemisphere (e.g., Näätänen & Alho, 1995a; Näätänen & Picton, 1987), the MMN is standardly assumed to have at least two different sources per hemisphere, an assumption that would be violated by a single-dipole analysis (Alho, 1995; Näätänen & Alho, 1995b). In addition, we only calculated dipole fits to standard sounds because their higher number guaranteed more reliable estimates due to a lower signal-to-noise ratio. The M100 was computed in a time window between 80 and 130 ms

post-stimulus onset on the basis of 32 channels in each hemisphere (16 anterior and 16 posterior). The number of channels was increased in order to improve the fit of the ECD solutions. Fitting was done as described in previous MEG studies (Obleser et al., 2004; Scharinger, Poe, & Idsardi, 2011) using the program MEGLAB (Kit, Japan). For each standard type in each hemisphere, this resulted in mean coordinates of the corresponding dipole, with *x*-values describing the medial–lateral axis, *y*-values the anterior–posterior axis, and *z*-values the superior-inferior axis. We conducted pairwise comparisons between mean coordinates in the medial–lateral axis using mixed-effects linear models with by-subject random intercepts.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bandl.2015.02.003.

References

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

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. *R package version 1.1–6*. <http://CRAN.R-project.org/package=lme4>.

Boersma, P., & Weenink, D. (2011). *PRAAT: Doing phonetics by Computer (ver. 5.2.24)*. Amsterdam: Institute for Phonetic Sciences.

de Cheveigné, A. d., & Simon, J. Z. (2007). Denoising based on time-shift PCA. *Journal of Neuroscience Methods, 165*(2), 297–305.

de Cheveigné, A. d., & Simon, J. Z. (2008). Sensor noise suppression. *Journal of Neuroscience Methods, 168*(1), 195–202.

Drullman, R., Festen, J. M., & Plomp, R. (1994). Effect of temporal envelope smearing on speech reception. *The Journal of the Acoustical Society of America, 95*(2), 1053–1064.

Eimas, P. D. (1963). The relation between identification and discrimination along speech and non-speech continua. *Language and Speech, 6*, 206–217.

Friedman, L., & Wall, M. (2005). Graphical views of suppression and multicollinearity in multiple linear regression. *The American Statistician, 59*, 127–136.

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

Fujisaki, H., & Kawashima, T. (1968). Identification and discrimination of synthetic speech sounds. In *Paper presented at the 6th international congress on acoustics*, Tokyo.

Hacquard, V., Walter, M. A., & Marantz, A. (2007). The effects of inventory on vowel perception in French and Spanish: An MEG study. *Brain and Language, 100*(3), 295–300.

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

Hautus, M. J. (1995). Corrections for extreme proportions and their biasing effects on estimated values ofd? *Behavior Research Methods, Instruments, & Computers, 27*(1), 46–51.

Healy, A. F., & Repp, B. H. (1982). Context independence and phonetic mediation in categorical perception. *Journal of Experimental Psychology. Human Perception and Performance, 8*(1), 68–80.

Jaeger, F. T. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory & Language, 59*, 434–446.

Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *The Journal of the Acoustical Society of America, 108*, 1252–1263.

Kasai, K., Yamada, H., Kamio, S., Nakagome, K., Iwanami, A., Fukuda, M., et al. (2001). Brain lateralization for mismatch response to across- and within-category change of vowels. *Neuroreport, 12*(11), 2467–2471.

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

Kewley-Port, D. (1983). Time-varying features as correlates of place of articulation in stop consonants. *Journal of the Acoustical Society of America, 73*, 322–335.

Kuznetsova, A., Bruun Brockhoff, P., & Bojesen Christensen, R. (2014). lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). *R package version 2.0–6*. <http://CRAN.R-project.org/package= lmerTest>.

Lahiri, A., & Reetz, H. (2010). Distinctive features: Phonological underspecification in representation and processing. *Journal of Phonetics, 38*, 44–59.

Lipski, S. C. (2006). *Neural correlates of fricative contrasts across languages*. Stuttgart: University of Stuttgart. <http://elib.uni-stuttgart.de/opus/volltexte/2007/2892/>.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide*. Mahwah, NJ: Erlbaum.

Mann, V. A., & Repp, B. H. (1980). Influence of vocalic context on perception of the [zh]-[s] distinction. *Perception & Psychophysics, 28*(3), 213–228.

Marklund, E., Schwarz, I.-C., & Lacerda, F. (2014). Mismatch negativity at Fz in response to within-category changes of the vowel /i/. *Neuroreport, 25*(10), 756–759.

McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: Expertise from reading in the fusiform gyrus. *Trends in Cognitive Sciences, 7*, 293–299.

McMurray, B., Tanenhaus, M. K., & Aslin, R. N. (2002). Gradient effects of within-category phonetic variation on lexical access. *Cognition, 86*(2), B33–B42.

McQueen, J. M., Jesse, A., & Norris, D. (2009). No lexical–prelexical feedback during speech perception or: Is it time to stop playing those Christmas tapes? *Journal of Memory and Language, 61*(1), 1–18.

Näätänen, R., & Alho, K. (1995a). Generators of electrical and magnetic mismatch response in humans. *Brain Topography, 7*, 315–320.

Näätänen, R., & Alho, K. (1995b). Mismatch negativity – A unique measure of sensory processing in audition. *The International Journal of Neuroscience, 80*, 317–337.

Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Ilvonen, A., et al. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature, 385*, 432–434.

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology, 118*(12), 2544–2590.

Näätänen, R., Pakarinen, S., Rinne, T., & Takegata, R. (2004). The mismatch negativity (MMN): Towards the optimal paradigm. *Clinical Neurophysiology, 115*, 140–144.

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*(4), 375–425.

Obleser, J., Elbert, T., Lahiri, A., & Eulitz, C. (2003). Cortical representation of vowels reflects acoustic dissimilarity determined by formant frequencies. *Brain Research. Cognitive Brain Research, 15*(3), 207–213.

Obleser, J., Lahiri, A., & Eulitz, C. (2004). Magnetic brain response mirrors extraction of phonological features from spoken vowels. *Journal of Cognitive Neuroscience, 16*(1), 31–39.

Oh, J., Kwon, J. H., Yang, P. S., & Jeong, J. (2013). Auditory imagery modulates frequency-specific areas in the human auditory cortex. *Journal of Cognitive Neuroscience, 25*, 175–187.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*. http://dx.doi.org/10.1155/ 2011/156869, 1-9.

Paavilainen, P., Alho, K., Reinikainen, K., Sams, M., & Näätänen, R. (1991). Right hemisphere dominance of different mismatch negativities. *Electroencephalography and Clinical Neurophysiology, 78*(6), 466–479.

Pantev, C., Bertrand, O., Eulitz, C., Verkindt, C., Hampson, S., Schuierer, G., et al. (1995). Specific tonotopic organizations of different areas of the human auditory cortex revealed by simultaneous magnetic and electric recordings. *Electroencephalography and Clinical Neurophysiology, 94*(1), 26–40.

Partanen, E., Vainio, M., Kujala, T., & Huotilainen, M. (2011). Linguistic multifeature MMN paradigm for extensive recording of auditory discrimination profiles. *Psychophysiology, 48*, 1372–1380.

Phillips, C., Pellathy, T., Marantz, A., Yellin, E., Wexler, K., Poeppel, D., et al. (2000). Auditory cortex accesses phonological categories: An MEG mismatch study. *Journal of Cognitive Neuroscience, 12*, 1038–1105.

Pisoni, D. B. (1973). Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception & Psychophysics, 13*(2), 253–260.

R Development Core Team (2014). R: A Language and Environment for Statistical Computing (Version 3.1). Vienna, Austria.

Repp, B. H. (1981). Two strategies in fricative discrimination. *Perception and Psychophysics, 30*, 217–227.

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.

Scharinger, M., Poe, S., & Idsardi, W. J. (2011). A three-dimensional cortical map of vowel space. Evidence from Turkish. *Journal of Cognitive Neuroscience, 23*(12), 3972–3982.

Schröger, E. (2005). The mismatch negativity as a tool to study auditory processing. *Acta Acústica, 91*, 490–501.

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

Sharma, A., Kraus, N., McGee, T., Carrell, T., & Nicol, T. (1993). Acoustic versus phonetic representation of speech as reflected by the mismatch negativity event-related potential. *Electroencephalography and Clinical Neurophysiology, 88*(1), 64–71.

Spencer, A. (1996). *Phonology: Theory and description* (1. publ. ed.). Oxford: Blackwell.

Tanaka, K. (2003). Columns for complex visual object features in inferotemporal cortex: Clustering of cells with similar but slightly different stimulus selectivities. *Cerebral Cortex, 13*, 90–99.

Tsunoda, K., Yamane, Y., Nishizaki, M., & Tanifuji, M. (2001). Complex objects are represented in macaque inferotemporal cortex by the combination of feature columns. *Nature Neuroscience, 4*, 832–838.

Winkler, I. (2007). Interpreting the mismatch negativity. *Journal of Psychophysiology, 21*(3–4), 147–163.

Wurm, L. H., & Fisicaro, S. (2014). What residualizing predictors in regression analyses does (and what it does not do). *Journal of Memory and Language, 72*, 37–48.

Zatorre, R. J., & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex, 11*, 946–953.

Zwicker, E. (1961). Subdivision of the audible frequency range into critical bands. *The Journal of the Acoustical Society of America, 33*(2), 248.