# Phonological Feature Representations in Auditory Cortex

Colin Phillips, Thomas Pellathy, University of Delaware
Alec Marantz: MIT

**Abstract**

Although phonemes are the smallest linguistic units that speakers are usually aware of, a good deal of linguistic evidence indicates that sub-phonemic *features* are the smallest building blocks of language. We present evidence from biomagnetic studies that indicate that representations of discrete phonological feature categories are available to left-hemisphere auditory cortex. Sequences of voiced (/bæ, dæ, gæ/) and voiceless (/pæ, tæ, kæ/) consonants were contrasted in a modified auditory mismatch paradigm. Importantly, although sounds contrasted in a many-to-one ratio at the level of phonological features, the use of 12 acoustically diverse tokens of each category ensured that there was no many-to-one ratio at the acoustic level. Therefore, the fact that an auditory cortex mismatch response was elicited confirms that feature representations are available to this part of the brain. Strikingly, however, the ability of auditory cortex to form a feature-level category from groups of diverse sounds appears to be strongly left-lateralized.

## 1. Introduction

Phonemes such as /k/, /æ/ and /t/ are commonly regarded as the building blocks of words, and are probably the smallest linguistic units which speakers may become consciously aware of. However, there is a good deal of evidence that phonemes are not linguistic primitives, but are instead composed of sets of smaller units known as phonological *features*. These features specify place of articulation, manner of articulation, voicing, and a number of other properties, and are probably the atomic units of language.

Historical linguists observed in the 19th century that certain sets of phonemes form 'natural classes', and that these natural classes of sounds tend to undergo the same historical changes. It was not until much later that it was proposed that these natural classes arise because phonemes are composed of smaller feature-sized units (Jakobson, 1941/1968, 1962). This shift in thinking was important, because it implied a commitment to features as the mental representations that underlie equivalence classes. Features have proven to be abundantly useful tools for describing a variety of phonological phenomena – synchronic, diachronic, developmental and neurolinguistic – but the best evidence for feature representations involves demonstrations that groups of phonemes are treated as *identical* at some level of representation.

The current study presents evidence that left hemisphere auditory cortex supports representations in which the members of a phonological natural class are treated as identical. Furthermore, no evidence for such representations is found in the right hemisphere, indicating that these feature-level phonological representations are strongly left-lateralized.

## 2. A Biomagnetic Measure of Identity

In order to probe for neural representations which treat all members of a phonological natural class as identical, we use a modified version of the auditory mismatch paradigm. When a sequence of identical sounds ('standards') is interrupted by infrequent contrasting sounds ('deviants'), the deviants elicit a characteristic mismatch response. The mismatch response has a latency of 150-250 ms and is known to originate in supratemporal auditory cortex (Alho, 1995; Alho *et al*., 1998; Hari *et al*., 1984; Sams *et al*.,1991; Scherg, Vajsar, & Picton, 1989), close to the generator of the auditory N100/M100. This paradigm has been extensively used as a measure of auditory discrimination (Näätänen, Gaillard, & Mäntysalo, 1978; for reviews see Näätänen, 1992; Näätänen & Winkler, 1999), and in a number of studies it has provided evidence of language-specific phonetic discrimination abilities (Aaltonen *et al*., 1997; Dehaene-Lambertz, 1997; Näätänen *et al*., 1997; Winkler *et al*., 1999). We have adapted this paradigm to create a measure of categorization, i.e. a measure of when acoustically different sounds are treated as identical by higher levels of representation (Phillips *et al*., in press; see also Aulanko *et al*., 1993; Dehaene-Lambertz *et al*., 2000).
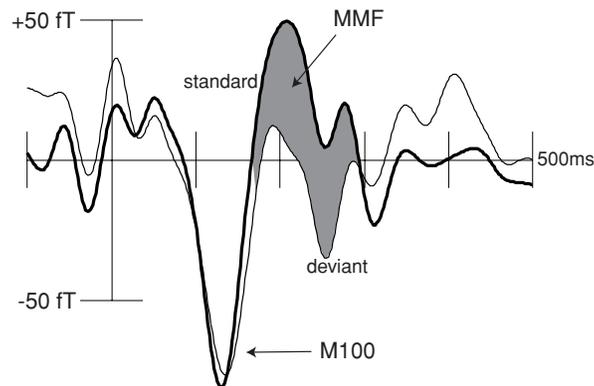


**Figure 1:** Sample averaged MEG response to speech sounds, measured at a single recording site in a mismatch paradigm. Both frequent ('standard') sounds and infrequent ('deviant') sounds elicit a similar auditory M100 response. Responses to standards and deviants typically begin to diverge in the 150-200ms latency range. This divergence is known as the *Mismatch Negativity* (EEG) or the *Magnetic Mismatch Field* (MMF).

In order to test for category representations, sequences of sounds are presented in which there is a many-to-one ratio among standard and deviant sounds at a higher, categorial level, but in which there is no many-to-one ratio at a lower, acoustic level of representation. This dissociation between levels of representation can be achieved by using a number of acoustically different tokens of each phonological category. If the sequence of sounds elicits an auditory cortex mismatch response, then it can be concluded that the category representations in question are available to auditory cortex, since the mismatch response is contingent upon a many-to-one ratio between standards and deviants. Phillips *et al*. (in press) use this approach to demonstrate auditory cortex representations of the phoneme-level categories /d/ and /t/. The syllables /dæ/ and /tæ/ were presented in a many-to-one ratio; however, many acoustically different tokens of /dæ/ and /tæ/ were used, spanning a range of voice onset time (VOT) values. Since the same acoustic

parameter which distinguished the two phonological categories also varied within each category, there was a many-to-one ratio of standards to deviants only at the level of phonological categories. In this way, the mismatch paradigm, which is normally used as a measure of fine-grained discrimination, can be used to test whether auditory cortex is able to group different sounds together on the basis of higher level category representations. Aulanko *et al.* (1993) used a similar approach in studying discrimination of /dæ/ and /gæ/, by using syllables with a variety of different pitches. However, since the within category variation affected only low-frequency spectral information (F0) and the between category variation affected only high-frequency spectral information (F2, F3 transitions), the mismatch response elicited in this study could be explained in purely acoustic terms, without reference to phonological categories.

One might object that the different members of individual phoneme-level categories are hard to discriminate, and therefore that the demonstration of categorization in Phillips *et al.* (in press) is no more than a demonstration of poor discrimination. This concern makes it particularly interesting to test for category representations at the level of phonological features, since feature-level categories group together sets of phonemes which are very easy to discriminate. The use of feature categories therefore makes it possible to distinguish the contribution of categorization and discrimination to the mismatch response. In the current studies, we focused on the feature [voice], contrasting the voiced stops /b, d, g/ and the voiceless stops /p, t, k/. Each of the six phonemes is easily discriminated from all of the others. Furthermore, it is well-known that voice onset time (VOT) is the primary cue to voicing, and that there is a sharp perceptual boundary between VOT values which are perceived as voiced and VOT values which are perceived as voiceless (Liberman *et al.*, 1961). Importantly, however, the perceptual boundary occurs at a different VOT value for different places of articulation. For instance, labial stops (/b,p/) show shorter boundaries, and velar stops (/g,k/) show longer boundaries (Kuhl & Miller, 1978; Lisker & Abramson, 1970; Miller, 1977; Summerfield, 1982).

## 3. Methods

### 3.1. Subjects & Materials

12 healthy normal adults (5 females) participated in all stages of the experiment. All subjects in the MEG experiment were strongly right-handed, as measured by the Edinburgh Handedness Inventory (Oldfield, 1971). All subjects gave informed consent and were paid $10/hour for their participation.

Stimuli consisted of synthesized CV syllables, generated using the cascade vocal tract of the Klatt speech synthesizer for Macintosh (SenSyn, Sensimetrics Inc., Cambridge, MA). Stimuli were drawn from three voice onset time continua: a labial ([bæ-pæ]) continuum, an alveolar ([dæ-tæ]) continuum and a velar ([gæ-kæ]) continuum. Each continuum consisted of 101 synthesized syllables, covering VOT values from 0-100 ms in 1 ms increments. The members of each VOT continuum were identical in all synthesis parameters except for the timing of voicing onset. Place of Articulation was marked by contrasting F2 and F3 onset values, and by different amplitudes of frication in the initial 5ms noise burst and of aspiration for the remainder of the syllable. All other synthesis parameters were held constant, including F1 transition parameters, which are known to significantly affect voicing judgments (Stevens & Klatt, 1974; Lisker, 1975; Summerfield & Haggard, 1977). Full synthesis parameters are available on request.

## 3.2. Design

All subjects in the MEG experiment first participated in an identification task, to determine perceptual boundaries for each subject in all three place of articulation categories. 13 stimuli from each of the 3 VOT continua were classified as voiced or voiceless. This experiment confirmed that perceptual boundaries are shortest for labial stops (22ms), slightly longer for alveolar stops (24ms) and longest for velar stops (33ms), as shown in previous studies (e.g. Lisker & Abramson, 1970; Summerfield, 1982), even when stimuli with different places of articulation are presented in the same sequence. Note that the boundary-shift was smaller here than in some other studies, due to the fact that F1 transitions were held constant across place of articulation.

Based on the perceptual boundaries established in the identification task, 4 different tokens of each of the 6 phonological categories were selected for each subject for the MEG experiment (i.e. 12 voiced, 12 voiceless, cf. Figure 2A). In the *phonological feature condition*, standard stimuli were chosen randomly from the 12 members of one voicing category on 87.5% of trials, and deviant stimuli were chosen randomly from the 12 members of the other voicing category on 12.5% of trials. Standards and deviant stimuli were presented in pseudorandom order, with the additional constraint that deviant stimuli were always followed by at least one standard. Interstimulus intervals varied randomly from 550-950ms. Due to the substantial acoustic variation among the 12 members of each voicing category, there was no many-to-one ratio at the level of acoustics (Figure 3A). Furthermore, due to the variation in perceptual boundaries across place-of-articulation, some stimuli from the voiced category had longer VOT values than some of the stimuli from the voiceless category. However, when the stimulus sequence is represented in terms of the categories [+voice] and [-voice], there is a clear many-to-one ratio, of the kind required to elicit a mismatch response (Figure 3B). Thus, the presence or absence of an mismatch response in this design can be used as a test of whether the phonological feature representation in Figure 3B is available to auditory cortex. A sequence containing 700 standards and 100 deviants was presented, lasting around 15 minutes.

In order to guard against artifactual differences between the responses to standards and deviants due to acoustic differences between voiced and voiceless sounds, the standard and deviant categories were reversed in the second half of the phonological feature condition, which consisted of a further 700 standards and 100 deviants. This made it possible to compare responses to the same stimuli as standards and deviants. The order of standards and deviants was counterbalanced, such that half of the subjects heard voiced standards first, and half of the subjects heard voiceless standards first.
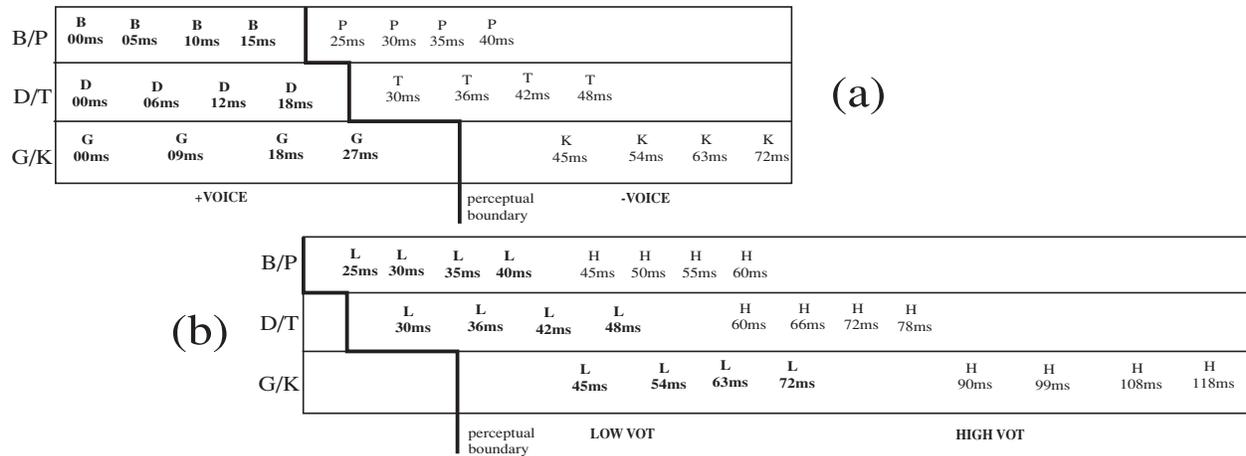
**Figure 2:** Typical Set of Stimuli for the Phonological Feature Condition (a), and the Acoustic Condition (b). Groups of stimuli are distinguished in the figure by the use of bold and plain text.
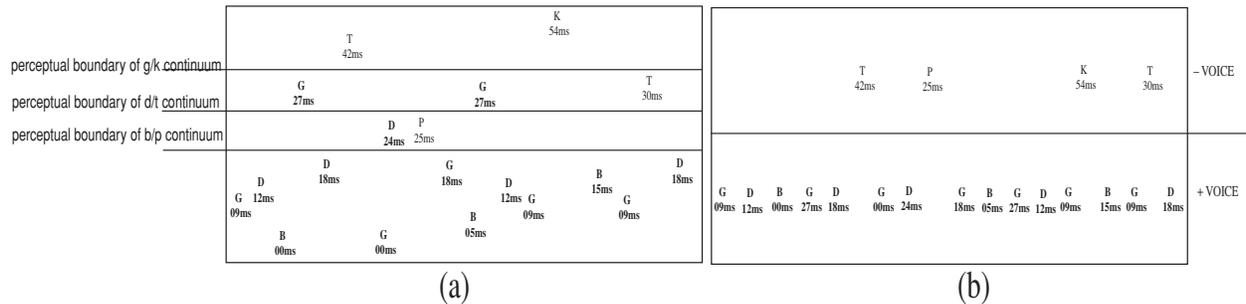


**Figure 3:** Acoustic (a) versus phonological (b) distribution of stimuli in the half of the Phonological Feature Condition in which +voice sounds are standards.

A second, *acoustic condition* was run, in order to address a possible alternative explanation of the results of the feature condition. A mismatch response elicited in the feature condition could, in principle, be explained in terms of the acoustic properties of the stimuli rather than in terms of their phonological feature categories. A mismatch response could be due to *ad hoc* formation of groups of 'shorter VOT' and 'longer VOT' syllables, with no need to refer to phonological categories. Therefore, the design of the acoustic condition was identical to the feature condition in all respects, except that VOT values were increased across-the-board, such that there was no longer a many-to-one ratio at the phonological level, because all stimuli now fell in the voiceless range of the continuum. The 12 members of the 'low VOT' category in the acoustic condition were identical to the voiceless category in the phonological feature condition. The relation of the members of the 'high VOT' category to the low VOT category was identical to the relation of the voiceless category to the voiced category in the feature condition (Figure 2B). VOT values were longer for alveolar and velar stimuli, due to the longer perceptual boundaries for these categories. Since category boundaries and natural VOT values are presumably perceptually equivalent across different places of articulation, VOT-space is effectively 'stretched' for labial and velar sounds.

5

3.3. MEG Recordings

MEG recordings were made using a 64-channel whole-head array of first order axial gradiometers (Kanazawa Institute of Technology, Kanazawa, Japan) at the KIT/MIT MEG laboratory. Subjects lay on their back inside a magnetically shielded room. Most recordings were made between 2am and 5am in order to avoid artifacts due to a nearby subway line. Recordings were made with a sampling rate of 500 Hz and a bandwidth of 200 Hz. Subjects listened passively to the stimulus sequence, which lasted around 30 minutes for each condition, and which was presented binaurally at a comfortable sound level. Although many studies using the mismatch paradigm have focused on responses to non-attended contrasts, a number of recent studies have demonstrated mismatch responses elicited by attended sound contrasts (e.g. Alho *et al.*, 1998; Dehaene-Lambertz, 1997; Woldorff, Hillyard, Gallen, Hampson, & Bloom, 1998). Recordings were selectively averaged off line, using a 100 ms pre-stimulus baseline, and filtered using a 0.5 Hz to 40 Hz bandpass filter.

Subjects were excluded from further analysis if they failed to show a clearly defined auditory M100 response in a pretest consisting of a sequence of 100 1 kHz tones. 4 subjects were excluded for this reason (in 3 out of 4 cases this was because the brain responses were masked by daytime environmental noise). Data from one additional subject had to be excluded from the acoustic condition, due to equipment failure.

An important principle of source localization when using axial MEG sensors is that neural activity generates a dipolar pattern of magnetic fields at the scalp, with positive and negative extremes of this field observed on either side of the underlying source. For example, the auditory M100 and the MMF are known to originate in auditory cortex in the supratemporal plane, but they are most clearly observed at more anterior and posterior MEG channels. Therefore, in order to focus on evidence for an auditory cortex MMF, all further analyses of the MEG recordings were based on 4 groups of 5 MEG channels (i.e. 20 channels total): an anterior and a posterior group in the left hemisphere, and their right hemisphere counterparts. The selected channels were those channels at which the auditory M100 was strongest in the grand average waveforms.

## 4. Results

4.1 Phonological Feature Condition

Mean field strengths at each channel subset were entered into a repeated measures ANOVA, with stimulus-type (standard/deviant), phonological feature category (voiced/voiceless), channel group (anterior/posterior) and hemisphere (left/right) as factors. Statistical analyses of the mismatch response (MMF) are based on mean field strengths at a 170-230ms latency interval. We also report results from an earlier 60-100ms interval and a later 300-400ms interval where they are relevant.

Comparison of responses to standards and deviants in the phonological feature condition indicates that a mismatch response was elicited in the left hemisphere (Figure 4). The presence of the mismatch response is indicated by the divergence of responses to standards and deviants starting around 100ms after stimulus onset. This early divergence of responses to standards and

deviants has been observed in other mismatch studies in which subjects attended to the stimuli (e.g. Woldorff *et al*., 1998). The fact that the divergence has a dipolar distribution across anterior and posterior channel groups indicates that the generator of the mismatch response is close to the generator of the auditory M100, in supratemporal auditory cortex. Statistically, the dipolar mismatch response is seen in the interaction of stimulus type and channel group in the left hemisphere ($F(1,7) = 16.54$, $p < 0.0001$). This interaction is also significant when analyses are run on each voicing category individually: voiced, $F(1,7) = 8.02$, $p < 0.01$; voiceless, $F(1,7) = 9.32$, $p < 0.01$. No other main effects or interactions involving stimulus-type were significant at this time interval. The stimulus-type x channel group interaction is already significant at an earlier 60-120ms latency interval ($F(1,7) = 5.96$, $p < 0.05$). As shown in Figure 4, this is due to the dipolar divergence of responses to standards and deviants beginning at around 100ms. The interaction remains significant at a later 300-400ms latency interval ($F(1,7) = 15.93$, $p < 0.0001$).

Results in the right hemisphere are strikingly different. The stimulus-type x channel group interaction does not approach significance at the MMF time interval ($F(1,7) = 0.03$, $p = 0.85$) or at the later 300-400ms time interval. There is a significant 3-way interaction of stimulus-type, category and channel group ($F(1,7) = 7.65$, $p < 0.01$). This effect was due to the fact that responses (to both standards and deviants) were more positive at anterior channels in runs in which standards were voiceless, and more negative in runs in which standards were voiced. The effect was not related to any dipolar mismatch response.
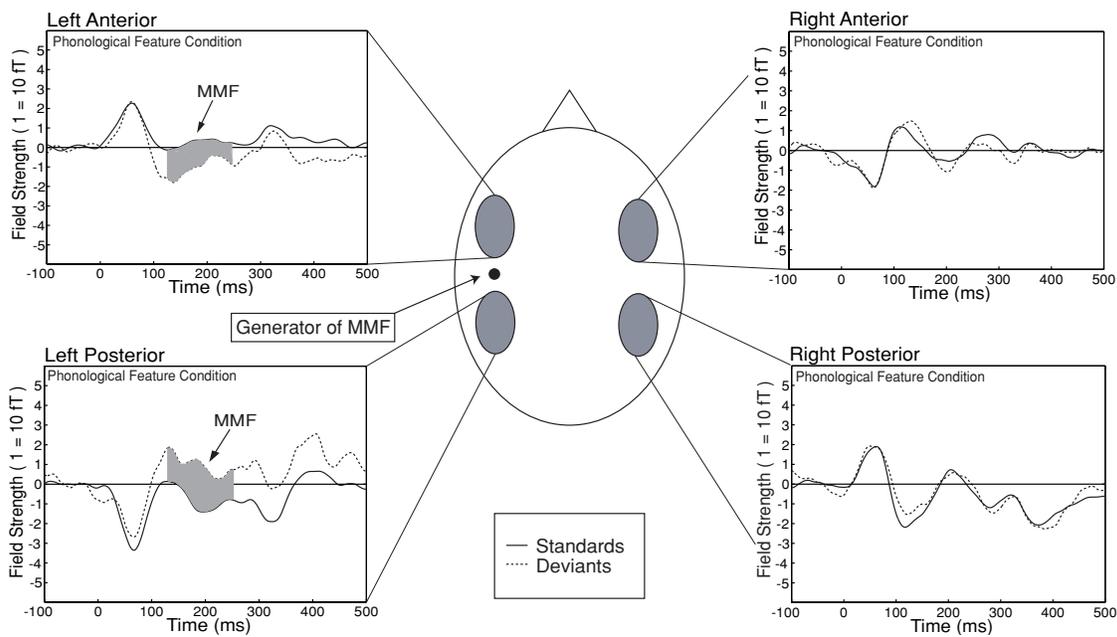


**Figure 4**: Grand average waveforms derived from groups of 5 posterior and anterior temporoparietal channels in the phonological feature condition, combining responses to voiced and voiceless sounds. The polarity reversal in the left hemisphere mismatch field from anterior to posterior recording sites indicates a dipolar mismatch field generated by a source in supratemporal auditory cortex. Corresponding right hemisphere channels show no mismatch response**.**

7

The fact that a MMF was elicited in left-hemisphere auditory cortex indicates that a phonological feature representation of the kind shown in Figure 3B is rapidly available to auditory cortex. We discuss the hemispheric contrast further in Section 5 below.

## 4.2 Acoustic Condition

In contrast to the phonological feature condition, in the acoustic condition no mismatch response was observed in either hemisphere at the MMF latency interval. The stimulus-type x channel group interaction was not significant in either the left hemisphere ($F(1,6) = 0.01$, $p = 0.92$) or the right hemisphere ($F(1,6) = 1.39$, $p = 0.24$), as shown in Figure 5.
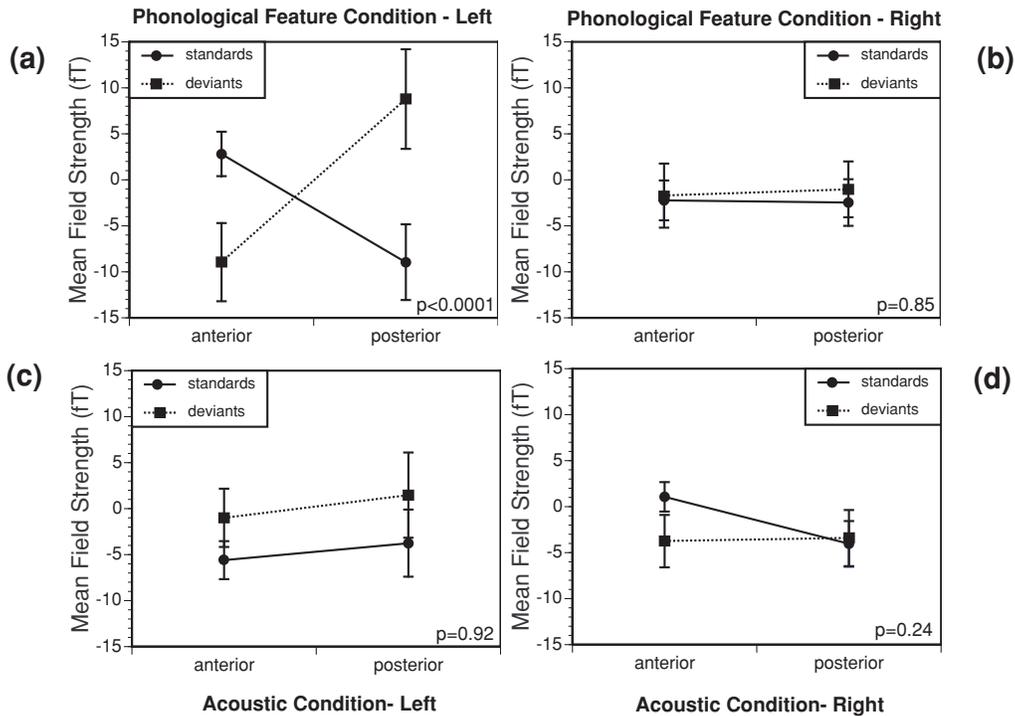


**Figure 5**: Mean field strength in 170-230ms time interval at anterior and posterior temporoparietal channel groups. The interaction of stimulus-type and channel group characteristic of an auditory cortex mismatch response is significant only in the left hemisphere in the phonological feature condition.

In the left hemisphere only the stimulus-type x channel group interaction was significant in the later 300-400ms time interval ($F(1,6) = 4.84$, $p < 0.05$), reflecting a dipolar difference between standards and deviants with an onset around 250ms. This effect was due to an interaction involving the high VOT stimulus group only, and was completely absent for the low VOT stimulus group. Thus, although broadly similar in distribution to the mismatch response observed in the phonological feature condition, this effect was weaker and had an onset around 150ms later than the phonological mismatch response. This late effect of the acoustic contrast is compatible with studies of discrimination which have found responses to non-native speech contrasts that are weaker and much later than the mismatch responses elicited by native language

contrasts (Dehaene-Lambertz, 1997; Dehaene-Lambertz *et al.*, 2000). Taken together, these late responses suggest that whereas phonologically coded information may be accessed very rapidly, acoustically coded information is available more slowly, at least at the SOAs used in our studies (around 1 second).

## 5. Discussion

There were two main findings in these studies, based on the two main contrasts observed in the results. First, the contrast between the phonological feature condition and the acoustic condition provides evidence for feature-level category representations available to auditory cortex, which are available by around 100ms after stimulus onset. Second, the contrast between the left and the right hemisphere in the phonological feature condition suggests an asymmetry in the ability of auditory cortex to represent feature-categories.

The fact that a left-hemisphere mismatch field was elicited in the phonological feature condition shows that there are category representations available to auditory cortex which treat the different phonemes in voicing categories as identical. Within each voicing category a wide variety of VOT values was used, with even an overlap in the VOT values used for voiced and voiceless stimuli, due to the effect of place-of-articulation on perceptual boundaries. Since VOT varied both within and between categories, there was no constant acoustic property which distinguished standard and deviant stimuli. Identical acoustic variability was present in the acoustic condition, yet there was no evidence of a mismatch response, lending further support to the notion that feature-level category representations are responsible for the mismatch response in the feature condition. These findings are noteworthy both because they provide clear neurolinguistic evidence for feature category representations, and also because they indicate that these are representations which are supported by relatively low-level cortical systems and are accessed very rapidly.

Related arguments for feature-level representations have been made in the infant perception literature by Hillenbrand (1983; but cf. Moroff, 1985) and by Jusczyk *et al.* (1999), based on the grouping of different phonemes belonging to nasal and manner feature categories. However, the use of single tokens of naturally recorded syllables in these studies made it possible for the infants to use constant acoustic properties to form category groupings. In our studies, we excluded this possibility by using substantial within-category variation in VOT, which was exactly the same acoustic property used to cue the between-category difference.

Note that our results also address a concern raised by Winkler *et al.* (1999), who suggest that evidence of grouping of sounds in mismatch studies reflects not phonological categorization, but rather failure of phonetic discrimination. Such concerns are valid in the case of phoneme-level categories, particularly stop consonants, where it is hard to distinguish effects of categorization from effects of poor discrimination. However, the concern disappears in the case of feature-level categories, since the different phonemes in a given feature category are easily discriminated from one another.

The second main result of our studies was the hemispheric contrast seen in the feature condition. Whereas a clear mismatch response was elicited in the left hemisphere, no mismatch response at all was observed in the right hemisphere. Our finding of a dipolar mismatch response only in the left hemisphere stands in striking contrast with a growing number of studies which report bilateral mismatch responses (Ackermann, Lutzenberger & Hertrich, 1999; Alho,

Connolly, Choeur *et al*., 1998; Alho, Winkler, Escera *et al*., 1998; Levänen *et al*., 1996; Näätänen & Alho, 1995; Opitz *et al*., 1999; Paavilainen *et al*. 1991). A number of these studies have reported hemispheric differences in specific parameters of the mismatch response (either timing or amplitude), and more commonly report that the mismatch response is *stronger* in the right hemisphere than the left (Levänen *et al*., 1996; Opitz *et al*., 1999; Paavilainen *et al*., 1991). Importantly, however, whereas all of these other studies examine *discrimination* of acoustic or phonetic contrasts, our studies test for representations which *group together* easily discriminable sounds into phonological feature categories. Therefore, the strong lateralization of the mismatch response in our results may reflect the fact that our studies require listeners to access a higher level of representation (i.e. phonological feature categories) than the studies of discrimination.

## References

Aaltonen, O., Eerola, O., Hellström, A., Uusipaikka, E., & Lang, A. (1997). Perceptual magnet effect in the light of behavioral and psychophysiological data. *Journal of the Acoustical Society of America, 101,* 1090–1105.

Ackermann, H., Lutzenberger, W. & Hertrich, I. (1999). Hemispheric lateralization of the neural encoding of temporal speech features: a whole-head magnetoencephalographic study. *Cognitive Brain Research*, *7*, 511–518.

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

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

Alho, K., Winkler, I., Escera, C., Huotilainen, M., Virtanen, J., Jaaskelainen, I. P., Pekkonen, E., & Ilmoniemi, R. J. (1998). Processing of novel sounds and frequency changes in the human auditory cortex: Magnetoencephalographic recordings. *Psychophysiology*, *35*, 211–224.

Anderson, S. (1985). *Phonology in the twentieth century*. Academic Press, New York.

Aulanko, R., Hari, R., Lounasmaa, O. V., Näätänen, R. & Sams, M. (1993). Phonetic invariance in the human auditory cortex. *Neuroreport, 4,* 1356–1358.

Chomsky, N. & Halle, M. (1968). *The Sound Patterns of English*. New York: Harper & Row.

Dehaene-Lambertz, G. (1997). Electrophysiological correlates of categorical phoneme perception in adults. *Neuroreport, 8,* 919–24.

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

Hari, R., Hämäläinen, M., Ilmoneimi, R., Kaukoranta, E., Reinikainen, K., Salminen, J., Alho, K., Näätänen, R., & Sams, M. (1984). Responses of the primary auditory cortex to pitch changes in a sequence of tone pips: neuromagnetic recordings in man. *Neuroscience Letters, 50,* 127-132.

Hillenbrand, J. (1983). Perceptual organization of speech sounds by infants. *Journal of Speech and Hearing Research, 26*, 268-282.

Jakobson, R. (1941/1968). *Child Language, Aphasia, and Phonological Universals* (German version, 1941; English translation, Mouton, The Hague, 1968).

Jakobson, R. (1962). *Selected Writings, vol. 1*. The Hague: Mouton.

Jusczyk, P., Goodman, M. B., & Baumann, A. (1999). Nine-month-olds' attention to sound similarities in syllables. *Journal of Memory and Language, 40*, 62-82.

Kraus, N, McGee, T., Carrell, T., Zecker, S., Nicol, T. & Koch, D. (1996). Auditory neurophysiologic responses and discrimination deficits in children with learning problems. *Science*, *273*, 971-973.

Kuhl, P. K. & Miller, J. D. (1978). Speech perception by the chinchilla: Identification functions for synthetic VOT stimuli. *Journal of the Acoustical Society of America, 63*, 905-917.

Levänen, S., Ahonen, A., Hari, R., McEvoy, L., & Sams, M. (1996). Deviant auditory stimuli activate human left and right auditory cortex differently. *Cerebral Cortex*, *6*, 288–296.

Liberman, A. M., Harris, K. S., Kinney, J. A., & Lane, H. L. (1961). The discrimination of relative-onset time of the components of certain speech and non-speech patterns. *Journal of Experimental Psychology, 61,* 379–388.

Lisker, L. (1975). Is it VOT or a first-formant transition detector? *Journal of the Acoustical Society of America, 57*, 1547-1551.

Lisker, L. & Abramson, A. (1970). The voicing dimension: Some experiments in comparative phonetics. *Proceedings of the 6th International Congress of Phonetic Sciences*, Prague, 1967, pp. 563-567.

Miller, J. L. (1977). Nonindependence of feature processing in initial consonants. *Journal of Speech and Hearing Research, 20*, 519-528.

Moroff, D. (1985). Do infants perceive similarity within feature classes? A critique of Hillenbrand. *Journal of Speech and Hearing Research, 28*, 316.

Näätänen, R. (1992). *Attention and Brain Function*. Hillsdale, NJ: Erlbaum.

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

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

Näätänen, R., Lehtoskoskl, A., Lennes, M., Cheour, M., Huotilainen, M., Ilvonen, A., Vainio, M., Alku, P., Ilmoniemi, R., 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. & Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychological Bulletin, 125*, 826-859.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia, 9*, 97–113.

Opitz, B., Mecklinger, A., von Cramon, D. Y., & Kruggel, F. (1999). Combining electro-physiological and hemodynamic measures of the auditory oddball. *Psychophysiology*, *36*, 142–147.

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

Phillips, C., Pellathy, T., Marantz, A., Yellin, E, Wexler, K, Poeppel, D., McGinnis, M. & Roberts, T. (in press). Auditory cortex accesses phonological categories: an MEG mismatch study. *Journal of Cognitive Neuroscience*.

Sams, M., Kaukoranta, E., Hämäläinen, M., & Näätänen, R. (1991). Cortical activity elicited by changes in auditory stimuli: Different sources for the magnetic N100m and mismatch responses. *Psychophysiology, 28,* 21–29.

Scherg, M., Vajsar, J., & Picton, T. W. (1989). A source analysis of the late human auditory evoked potentials. *Journal of Cognitive Neuroscience, 1,* 336–355.

Stevens, K. N. & Klatt, D. H. (1974). Role of formant transitions in the voiced-voiceless distinction for stops. *Journal of the Acoustical Society of America, 55*, 653-659.

Summerfield, Q. (1982). Differences between spectral dependencies in auditory and phonetic temporal processing – relevance to the perception of voicing in initial stops. *Journal of the Acoustical Society of America, 72*, 51-61.

Summerfield, Q. & Haggard, M. P. (1977). On the dissociation of spectral and temporal cues to the voicing distinction in initial stop consonants. *Journal of the Acoustical Society of America, 62*, 435-448.

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

Woldorff, M., Hillyard, S., Gallen, C., Hampson, S., & Bloom, F. (1998). Magneto-encephalographic recordings demonstrate attentional modulation of mismatch-related neural activity in human auditory cortex. *Psychophysiology, 35,* 283–292.

colin@udel.edu, pellathy@udel.edu, marantz@mit.edu