neuro*report*

EVENT-RELATED potentials, recorded through a 128-electrode net, were used to study phonemic processing in the human auditory system. Within a stream of identical syllables, acoustic deviants were introduced that either crossed a phonetic boundary or remained within the same category. Two phonetic boundaries were explored, one present and the other absent in the subjects' native language. A large mismatch negativity was induced by native phonetic deviants, but not by non-native or within-category deviants. We suggest that a language-specific phonemic code has a separate neural representation in sensory memory and can serve as the basis for auditory mismatch detection. The subjects' inability to discriminate non-native phonetic contrasts does not seem related to a late attentional filter, but rather to a genuine loss of auditory discrimination abilities.

Key words: Event-related potentials; Language; Mismatch negativity; Phoneme

Electrophysiological correlates of categorical phoneme perception in adults

Ghislaine Dehaene-Lambertz

Laboratoire de Sciences Cognitives et Psycholinguistique, 54 boulevard Raspail, 75270 Paris cedex 06, and Service de Neuropédiatrie, Centre Hospitalier Universitaire Bicêtre, France

Introduction

Whereas perception of a majority of sounds is continuous, phonetic perception is categorical: listeners are better able to discriminate equivalent acoustic differences when the difference signals a phonetic boundary that when it signals acoustic variants of the same phonetic category. Furthermore, phonetic categorical perception in adults depends on the listeners' maternal language. Whereas infants are able to discriminate almost all phonetic boundaries used in human languages, phonetic perception is reshaped during the first year of life, resulting in a loss of the capacity to discriminate some of the phonetic boundaries not used in the maternal language (e.g. the Hindi contrast retroflex /d/ vs dental /d/ in English native speakers[1] or the /r/ vs /l/ contrast in Japanese adults[2]). Whether these two main characteristics of phoneme perception result from specific linguistic processes or could be explained by general acoustic processes is still debated. In the present experiment, we used event-related potentials to study the cerebral responses to acoustic and phonetic changes and to detect when categorical information useful in the subject's maternal language is taken into account during phoneme processing.

Several experiments have shown specific responses to a novel syllable introduced after a succession of repeated syllables. In particular, a mismatch negativity (MMN) is generated by a mismatch between the physical features of the novel stimulus and the neural

traces of the repeated sound in sensory memory.[3] It has been suggested that this component depends on the acoustic but not on the phonetic differences between standard and deviant stimuli.[4–6] Aaltonen *et al*[7] found a MMN at the introduction of a deviant vowel in a succession of repeated vowels when the standard was at the phonetic boundary, but not when the standard was at the centre of the category. Furthermore, in some brain-damaged patients, a MMN may be evoked by an acoustic change, such as a change in frequency, but not by a phonetic change.[8,9] These observations suggest that activation of specific linguistic processing may contribute to this early electrical component.

In order to increase the sensitivity of the present experiment to phonetic effects, we used an experimental design where the evoked response to the same syllable was studied in different contexts. In order to create the context, syllables were presented in blocks of four. The last syllable of the blocks was kept constant across conditions and was either physically identical to the three first syllables (control trials), or was a variant of the same phonetic category as the preceding syllables (within-category trials) or belonged to a different phonetic category (across-category trials). We studied these three conditions at two phonetic boundaries, one present and the other absent in the maternal language. Spatial resolution of event-related potentials (ERP)s was optimized by using a 128-channel geodesic electrode net.

Materials and Methods

To study categorical perception in English and Hindi speaking subjects, Werker and Lalonde[1] synthesized 16 syllables along the voiced place of articulation continuum with a vowel /a/, by varying the starting frequency of the second and third formants. Along this continuum, English listeners perceive samples 1–6 as /ba/ and the others as alveolar /da/ whereas Hindi listeners perceive samples 1–6 as /ba/, samples 7–10 dental /da/ and samples 11–16 as retroflex /Da/. Here, 12 right-handed French adults were tested (mean age 26 years)[1] after giving written informed consent. Four additional subjects were tested and rejected, one because of equipment failure, one because of an high frequency of errors in the control situation (32.9% of 'different' response) and two because they did not perceive any differences between the stimuli. None of them spoke Hindi. Behavioural evidence, presented below, indicated that like English subjects they perceived a phonemic boundary between /ba/ and /da/, but not between /da/ and /Da/

Six syllables were selected for the present experiment (samples 1, 5, 9 and 8, 12, 16: see Ref. 1 for a more precise description of the stimuli). Each syllable was 275 ms in duration. The syllables were presented at an intensity of 69 dB SPL, in block of four syllables with an stimulus-onset interval of 600 ms and an intertrial interval of 4 s. In each trial, the first three syllables were always identical. The identity of the last syllable defined three trial types. In control trials (CO), the last syllable was physically identical to the preceding ones. In within-category (WC) trials it was a variant within the same phonetic category. Finally in across-category (AC) trials, it belonged to a different phonetic category. The physical distance from the preceding syllable was equivalent in WC and AC trials. CO, WC and AC trials were presented at both native and non-native phonetic boundaries, defining six categories of trials: control native (/ba/5 /ba/5 /ba/5 **/ba/5**), WC native (/ba/1 /ba/1 /ba/1 **/ba/5**) and AC native (/da/9 /da/9 /da/9 **/ba/5**); control non-native /Da/12 /Da/12 /Da/12 **/Da/12**), WC non-native (/da/16 /da/16 /da/16 **/Da/12**) and AC non-native (/da/8 /da/8 /da/8 **/Da/12**). Subjects were instructed to detect whether the last syllable changed by making a bimanual yes–no response. The side of the 'yes' response was changed in the middle of the experiment and the order counterbalanced across subjects.

ERP were collected using a 128-channel geodesic electrode net referenced to the vertex, and digitized at 125 Hz over a 3072 ms epoch, including a 150 ms baseline. Trials contaminated by artifacts were automatically rejected. For each subject, the remaining trials were averaged across 'same' and 'different' responses, baseline corrected, transformed to an average reference and digitally filtered (band pass 0.7–20 Hz). Two-dimensional reconstructions of scalp voltage at each time step were computed using spherical spline interpolation.[10] Here, we report only the evoked response to the last syllable.

Results

Behavioural results: French-speaking adults detected a syllable change in 81% of the native AC trials. In all other conditions, the detection rates were low: 4.8% and 12.8% respectively in native CO and WC trials; 5.7%, 13.7% and 18.8% in non-native CO, WC and AC trials. An analysis of variance was computed on detection rates with condition (CO, WC and AC) and phonetic contrast (native and non-native) as within-subject factors. All effects and interactions were significant ($p < 0.001$). There was categorical perception for the native contrast, as shown by the significant difference between native AC and WC trials ($F(1,11) = 132.4$, $p < 0.001$), but none for the Hindi contrast ($F(1,11) = 2.4$, $p = 0.150$). Subjects also showed a small but significant sensitivity to acoustical changes as indicated by significant differences between CO and WC for both native and non-native contrasts ($p < 0.01$).
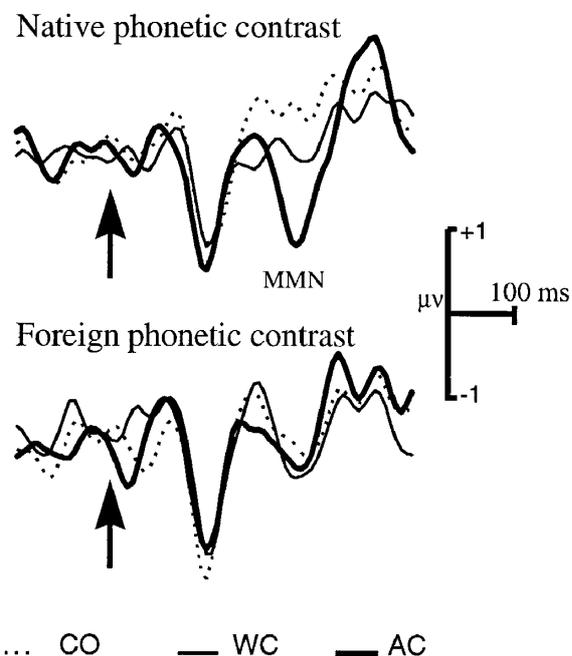

FIG. 1. Event-related potentials evoked by the last syllable over a 600 ms epoch starting 150 ms before syllable onset (arrow). On this right superofrontal electrode, the mismatch negativity response was particularly large for the native across-category change (AC) relative to the other change conditions (WC : within-category change, CO : control).

The average RT was 745 ms. An ANOVA calculated with the same factors as above revealed significant main effects and interactions ($p < 0.04$). Subjects were faster for native (718 ms) and non-native (736 ms) CO trials and for native AC (711 ms) trials, than for native and non native WC trials (755 and 770 ms) and for non-native AC trials (780 ms). In particular, there was a significant difference between native and non-native AC trials ($F(1,11) = 8.31$, $p = 0.015$), again indicating that subjects found the native AC trials easier than the non-native AC trials.

Electrophysiological results: An initial inspection of interpolated voltage maps as a function of time in each of the six conditions suggested that there was a much greater response to syllable change in the native AC condition than in other types of trials (Figs 1,2). This evoked response began 184 ms after the onset of the syllable and continued through to the behavioural response. The other remarkable event was that around 370 ms, all change trials were different from CO trials (Fig. 2).

For statistical analysis, four time windows were selected by inspecting the time course of a two-dimensional reconstruction of voltage of the difference wave between the native AC and CO conditions, where maximal effects were expected: 120–176 ms (N1), 184–224 ms (P2), 248–320 ms (MMN) and 328–400 ms (P3). For each temporal window and each phonetic contrast, average voltage was submitted to a repeated-measures analysis of variance (ANOVA) with condition (CO, WC and AC trials), hemisphere (left and right) and location (5 electrode groups: inferior frontal, superior frontal, parietal, temporal and occipital) as within-subject factors. Whenever a significant effect or interaction involving condition was found, it was analysed using two planned contrasts: an acoustical contrast, CO *vs* WC, was used to study responses to acoustic changes, and a phonetic contrast, WC *vs* AC, was used to probe specific responses to phonetic changes over and above those elicited by acoustic changes only.

No significant difference between conditions was found on the N1 time window. For the native phonetic contrast, the first significant location × condition interaction was found on P2 time window ($F(8,88) = 2.32$, $p = 0.026$). This was related to a significant effect of the phonetic contrast (location × (WC *vs* AC): $F(4,44) = 3.86$, $p = 0.009$) with no effect of the acoustical contrast (location × (CO *vs* WC): ($F(4,44) < 1$). As shown in figure 2, the responses to the phonetic contrast were present at the inferior frontal location ($F(1,11) = 5.79$, $p = 0.035$) and the parietal location ($F(1,11) = 5.33$, $p = 0.041$). On the following temporal window (MMN), the location × condition interaction was still significant ($F(8,88) =$ 2.78, $p = 0.009$), with a persistent effect of phonetic contrast (location × (WC *vs* AC): $F(4,44) = 2.90$, $p = 0.032$) and no effect of acoustic contrast (location × (CO *vs* WC): $F(4,44) = 1.57$, n.s.). The differences resulting from the phonetic contrast were now present on the superior frontal ($F(1,11) = 5.31$, $p = 0.042$) and temporal location ($F(1,11) = 7.88$, $p = 0.017$). Although the location × (CO *vs* WC) interaction was not significant, there was a significant effect of acoustic contrast at these two locations (superior frontal: $F(1,11) = 6.15$, $p = 0.031$) and temporal: $F(1,11) = 5.43$, $p = 0.040$). Finally on the MMN temporal window, there was a significant location × condition × hemisphere interaction ($F(8,88) =$ 2.09, $p = 0.045$) related to a significant effect of phonetic contrast (location × (WC *vs* AC) × hemisphere: $F(4,44) = 2.64$, $p = 0.046$) with no effect of acoustic contrast ($F(4,44) < 1$). *Post hoc* analyses of the phonetic contrast at each site found no hemisphere effect. On P3 time window, the significant location × condition interaction ($F(8,88) = 4.75$, $p <$ 0.001) was once more related to a significant effect of the phonetic contrast (location × (WC *vs* AC): $F(4,44) = 7.01$, $p < 0.001$), with no effect of the acoustic contrast ($F(4,44) = 2,36$, $p < 0.068$). The effect of phonetic contrast was present at superior frontal ($F(1,11) = 8.43$, $p = 0.014$), temporal ($F(1,11) = 11.59$, $p = 0.006$) and parietal locations ($F(1,11) = 11.35$, $p = 0.006$). On the superior frontal ($F(1,11) = 6.58$, $p = 0.026$) and temporal locations ($F(1,11) = 6.71$, $p = 0.025$) as well as on the occipital location ($F(1,11) = 7.35$, $p = 0.020$), there were significant differences due to the acoustical contrast.

For the non-native Hindi contrast, no significant effect was found before the P3 time window. At that moment, a significant location × condition interaction was found ($F(8,88) = 2.16$, $p = 0.038$). This was related to an effect of the acoustic contrast (location × (CO *vs* WC): $F(4,44) = 3.02$, $p = 0.027$) with no effect of the phonetic contrast ($F(4,44) > 1$). This effect was present only at the occipital location ($F(1,11) = 5.96$, $p = 0.033$).

Discussion

Both behavioural and electrophysiological responses indicated that, for equal acoustic distance, subjects showed a particular sensitivity to an acoustic change that signaled a phonemic boundary in their native language. Such a change was detected with high accuracy and generated an early specific evoked response by about 200 ms. Other types of acoustic changes that did not coincide with a phonemic change were detected only rarely, although they did induce slightly elevated hit rates and response times relative to a no-change control. These performance differences
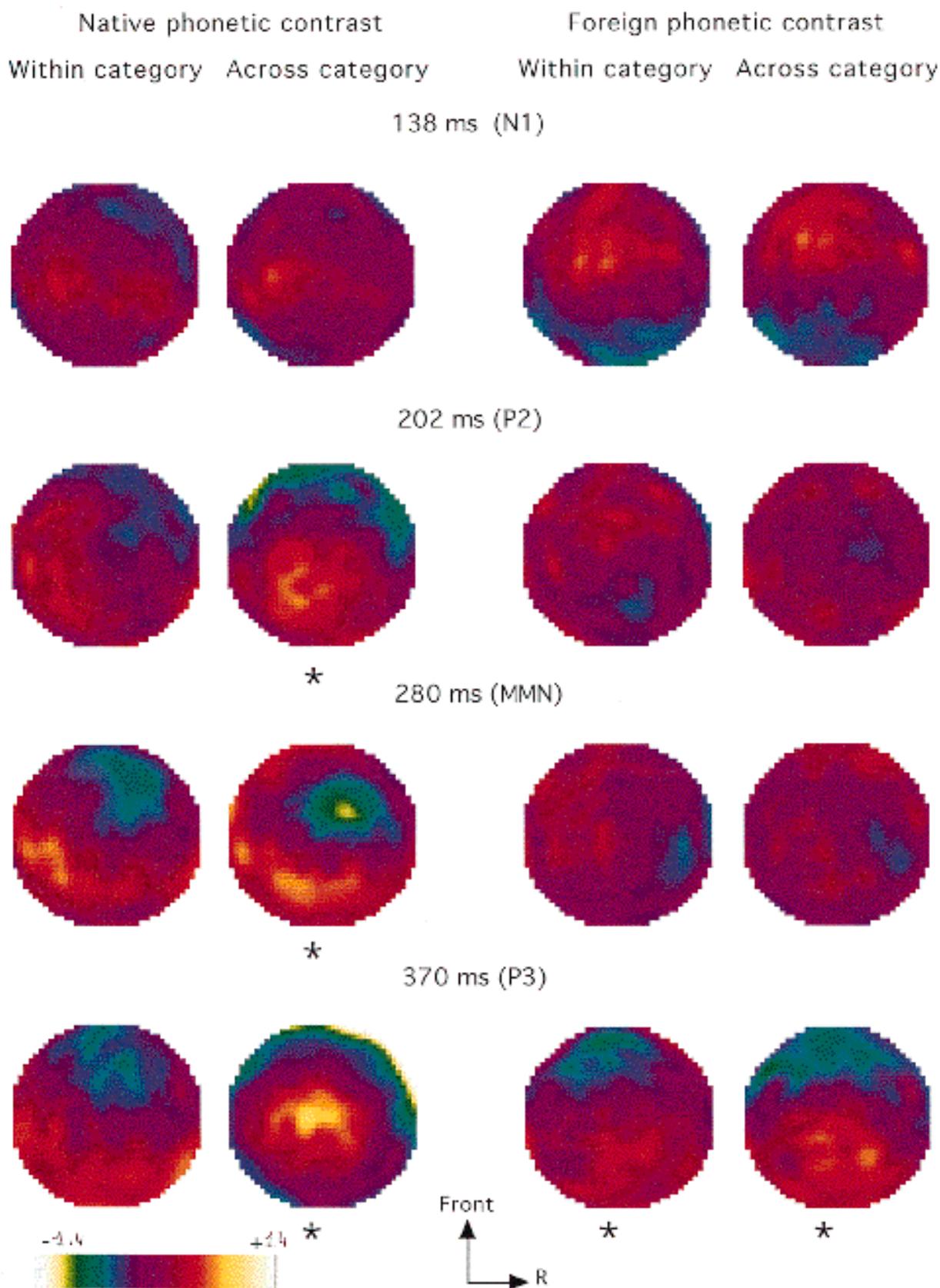
FIG. 2.   Topography of the grand-averaged substraction of ERPs to deviant syllables minus ERPs to ontrol trials at the maximum of each electrical component. Voltages are colour-coded using the scale at bottom left (mV). *$p < 0.05$ for the interaction of location and change.

were reflected in a significant difference in evoked responses between all change trials and the control trials in the P3 time window.

According to Näätänen,[3] auditory novelty detection is based on a comparison of the deviant stimulus with the trace of previous stimuli kept in sensory memory. This process is signaled by the MMN. Its maximum amplitude is over the frontal regions, with a polarity inversion at mastoid electrodes.[4,11] The voltage distribution is asymmetrical, with predominant right superior frontal negativity and left temporal positivity.[12] Our data match this description in detail. As illustrated in figure 1, a large MMN was observed for the across-category change at the native boundary, but it was much smaller for the within-category change near the native boundary and it was missing altogether for both types of changes near the Hindi boundary. Because the amount of acoustic difference was identical in all these conditions, acoustic parameters alone cannot explain these results. We therefore hypothesize that phonemic characteristics are coded and stored in sensory memory and contribute to the mismatch process.

Giard *et al*[12] showed that changes in different auditory parameters (e.g. duration, frequency or intensity) generate distinct MMNs that appear to originate in part from distinct neural populations within the supratemporal plane of the auditory cortex. The present experiment suggests that yet another neural population in auditory cortex codes for phonemic attributes. Indeed, at the native boundary, the interaction of hemisphere, location and condition was significant only for the phonemic change, but not for the within-category change, suggesting that brain regions involved in linguistic processing may be different and more asymmetrical than those involved in processing a non-linguistic change.

Our observations contrast with some earlier ERP experiments in which the MMN was found to depend more on acoustic changes than on phonemic differences.[4–6] Several methodological features, however, may account for this discrepancy. Most importantly, all previous ERP studies of phonemic processing have used only a small number of electrodes, while our present results are based on a dense array of 128 electrodes. This improved scalp coverage made it possible to detect differences between conditions that might have been overlooked otherwise. If we had considered only the Fz electrode, for instance, as reported in Ref. 5, we would have found no difference between phonemic and acoustic changes at the native boundary ($F(1,11) = 1.73$, n.s.).

Another potential difference with previous studies is that ERPs were recorded in an active change detection task with a variety of distinct syllables, thus preventing subjects from focusing on small acoustic changes. Behavioural experiments have shown that attention and task conditions can induce acoustic or linguistic processing of the same stimuli.[13] Here, attention orientation toward linguistic characteristics might have amplified the evoked responses to the phonemic change, thus allowing it to become significant.

Perhaps the best indication that subjects were engaged in a language-specific analysis of the stimuli comes from the complete absence of a MMN to changes near the non-native boundary. The non-native across-category change signals a phonemic boundary for Hindi subjects and for babies during the first year of life.[1] However, our French subjects appeared totally oblivious to it, to such an extent that there was no suggestion of an MMN in this condition (Fig. 2). Thus, our results, in agreement with previous observations,[7] suggest that no MMN is elicited when the stimuli fall well within a single phonemic category, as was the case for all /da/ and /Da/ stimuli for French subjects.

Two alternative accounts have been proposed for the loss of discrimination of foreign phonetic contrasts in adults. Either there is a radical perceptual reorganization, leading to a shrinkage of the perceptual distance between sounds within a native phonemic category[14] or alternatively, early phonetic perception remains unchanged and universal, but a late attentional filter prevents subjects from attending to non-native contrasts.[15] Our data provide no evidence for the latter hypothesis. If French speakers possessed a universal phonetic processing stage still capable of perceiving the Hindi contrast, one would have expected ERPs to reflect its differential activation during non-native within-category and across-category trials. The present results rather suggest that the inability to discriminate foreign phonetic contrasts may be due to a genuine reduction in sensory discrimination abilities, even if adults may be later trained, to some extent, to re-acquire a non-native contrast.[16]

Conclusion

In the present experimental conditions, a large mismatch negativity to syllables was observed only when a phonetic change occured and was relevant in the subject's native language. We suggest that a language-specific phonemic code has a separate neural representation in sensory memory and can serve as the basis for auditory mismatch detection. The subjects' inability to discriminate non-native phonetic contrasts does not seem related to a late attentional filter, but rather to a genuine loss of auditory discrimination abilities.

References

1. Werker JF and Lalonde CE. *Dev Psychol* **24**, 672–683 (1988).
2. Miyawaki K, Strange W, Verbrugge R *et al. Percept Psychophys* **331-340** (1975).
3. Näätänen R. *BBS* **13**, 201–288 (1990).
4. Sams M, Aulanko R, Aaltonen *et al. J Cogn Neurosci* **2**, 344–357 (1990).
5. Sharma A, Kraus N, McGee T *et al. Electroencephalogr Clin Neurophysiol* **88**, 64–71 (1993).
6. Maiste AC, Wiens AS, Hunt MJ *et al. Ear Hearing* **16**, 68–90 (1995).
7. Aaltonen O, Paavilainen P, Sams M *et al. Scand J Log Phon* **17**, 107–112 (1992).
8. Aaltonen O, Tuomainen J, Laine M *et al. Brain Lang* **44**, 139–152 (1993).
9. Kraus N, McGee T, Carrell TD *et al. Ear Hearing* **16**, 19–37 (1995).
10. Perrin F, Pernier J, Bertrand D *et al. Electroencephalogr Clin Neurophysiol* **72**, 184–187 (1989).
11. Tiitinen H, May P, Reinikainen K *et al. Nature* **372**, 90–92 (1994).
12. Giard MH *et al. J Cog Neurosci* **7**, 133–143 (1995).
13. Pisoni DB. *Percept Psychophys* **13**, 253–260 (1973).
14. Kulh PK, Williams KA, Lacerda F *et al. Science* **255**, 606–608 (1992).
15. Werker JF and Tees RC. *JASA* **75**, 1866–1878 (1984).
16. Werker JF and Logan JS. *Percept Psychophys* **37**, 135–144 (1985).

General Summary

In this study the electrical activity of the brain was recorded through 128 electrodes while French subjects listened to a series of four syllables, in which the last syllable sometimes changed. The changes could be acoustic, with no perceived change in phoneme, or they could result in a new phoneme (e.g. /ba/ /ba/ /ba/ /da/) which was or was not used in the subject's native language. The results showed a specific brain electrical response only in the event of a native phonemic change, suggesting that phonemic information is rapidly extracted and represented in sensory memory in the human auditory system.

Native phonetic boundary   Foreign phonetic boundary
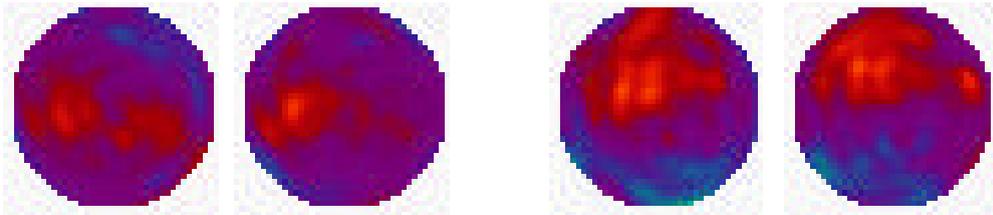
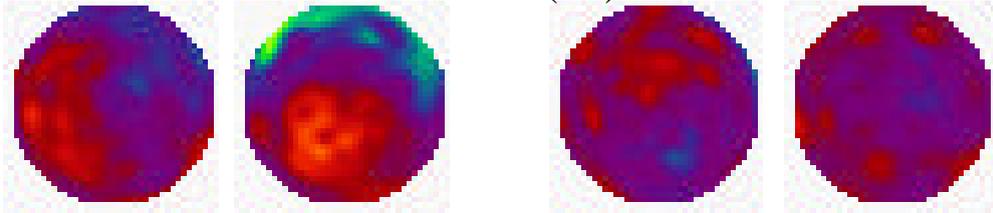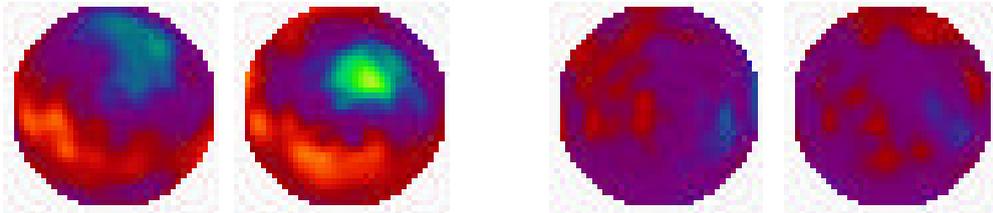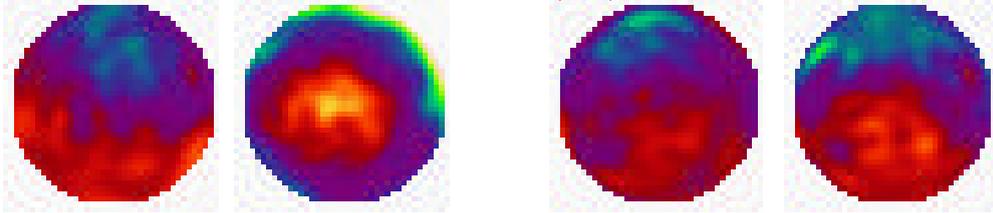Within   Across   Within   Across

138 ms (N1)

202 ms (P2)

*

280 ms (MMN)

*

370 ms (P3)

*   *   *

Front

R

-1.4 $+1.4$

μv