Electrophysiological Correlates of Phonological Processing: A Cross-linguistic Study

G. Dehaene-Lambertz

CNRS UMR 8554 and EHESS, Paris, France and Centre Hospitalier Universitaire Bicetre, France

E. Dupoux and A. Gout

CNRS UMR 8554 and EHESS, Paris, France

Abstract

■ It is well known that speech perception is deeply affected by the phoneme categories of the native language. Recent studies have found that phonotactics, i.e., constraints on the cooccurrence of phonemes within words, also have a considerable impact on speech perception routines. For example, Japanese does not allow (nonnasal) coda consonants. When presented with stimuli that violate this constraint, as in /ebzo/, Japanese adults report that they hear a /u/ between consonants, i.e., /ebuzo/. We examine this phenomenon using event-related potentials (ERPs) on French and Japanese participants in order to study how and when the phonotactic properties of the native language affect speech perception routines. Trials using four similar precursor stimuli were presented followed by a test stimulus that was either identical or different depending on the presence or absence of an epenthetic vowel /u/ between two consonants (e.g., "ebuzo ebuzo ebuzo—ebzo"). Behavioral results confirm that Japanese, unlike French participants, are not able to discriminate between identical and deviant trials. In ERPs, three mismatch responses were recorded in French participants. These responses were either absent or significantly weaker for Japanese. In particular, a component similar in latency and topography to the mismatch negativity (MMN) was recorded for French, but not for Japanese participants. Our results suggest that the impact of phonotactics takes place early in speech processing and support models of speech perception, which postulate that the input signal is directly parsed into the native language phonological format. We speculate that such a fast computation of a phonological representation should facilitate lexical access, especially in degraded conditions. ■

INTRODUCTION

Humans use complex sounds to communicate and convey meaning. The way in which the mapping between the signal and the concepts is realized, however, is heavily language dependent. For instance, some languages use only six consonants to construct words, others more than 80. Some use three vowels, others more than 20. The particular phoneme inventory of the native language has a strong influence on speech discrimination capacities in adults. For instance, adult monolinguals in Japanese have a lot of trouble in distinguishing English /r/ from /l/ sounds, because both are perceived as a single /R/. However, language sound systems differ in ways other than the repertoire of phonemes. They also differ in how particular phonemes can combine in a sequence, i.e., its phonotactic properties. In Japanese, only a rather strict alternation of vowels and consonants is allowed, whereas English allows for clusters of several consonants (e.g., strengths). A lot of research has been devoted to the effect of the phoneme inventory on speech perception, but the role played by the higher-order properties of linguistic signals is only starting to be explored. The focus of this paper is to study, using event-related potentials (ERPs) methodology and a cross-linguistic design, whether phonotactic properties have effects on speech perception routines that are as profound as those triggered by differences in phoneme inventories.

The acquisition of the language phoneme inventory is very quick. At 6 months, infants have established prototypes for the vowels used in their language (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992) and start to loose sensitivity to nonnative vowels (Polka & Werker, 1994). At 12 months, they loose the capacity to discriminate nonnative consonantal contrasts, or at least those that can be assimilated to native categories (Best, McRoberts, & Sithole, 1988; Werker & Tees, 1984). After that, the capacity to perceive foreign phonetic contrasts seems remarkably stable and poor, although discrimination remains possible depending on whether the foreign phonemes are closer or further away from the prototypes in the subjects' native language. This is also true if the foreign phonemes are not easily assimilated to any phoneme in the native language. Pallier, Bosch, and Sebastian (1997) have shown that even very fluent bilinguals who have acquired a second language after age 5, and have used it extensively thereafter, have difficulties with vowel contrasts in the second language.

Electrophysiological studies suggest that the phoneme inventory of a particular language affects very early speech processing: A phonetic representation dependent on the subject's native language is coded in the echoic memory. Näätänen et al. (1997) presented Estonian and Finnish participants with series of the vowel /e/ where items were randomly replaced by a "deviant" vowel (/ö/, /õ/, or /o/). These vowels exist in Estonian and Finnish, except /õ/, which does not exist in Finnish. In Estonian participants, a mismatch negativity (MMN) response was elicited by all the deviant vowels (100 to 240 msec after vowel onset). Its amplitude increased with the acoustic distance between the deviant and the standard vowel. In contrast, in Finnish participants, a significant drop in MMN amplitude was found for the nonnative /õ/ vowel. Such a drop for the nonnative vowel was also documented in the magnetic equivalent of the MMN. Dehaene-Lambertz (1997) used a similar paradigm to study consonant contrasts. She presented French participants with streams of CV syllables where acoustic deviants were introduced that either crossed a phonetic boundary or remained within the same category. Two phonetic boundaries, one present and the other one absent in the subjects' native language, were explored. A large MMN was induced by deviants that crossed the native boundary, but not by nonnative or within-category deviants, 280 msec after syllable onset. In these experiments, the phonological mismatch was not preceded by an acoustical mismatch suggesting that phonetic categorization is computed very early in speech processing (from 100 to 280 msec after stimulus onset). Moreover, this categorization appears to be highly dependent on the subjects' native language. These results are consistent with the view that speech perception involves an early processing stage of phonemic categorization. Such a stage removes all of the irrelevant phonetic details from speech, but retains the linguistically relevant contrasts in the subject's native language. Consequently, two sounds that fall in the same native category are extremely difficult to distinguish (Best et al., 1988).

However, as we said above, properties other and more abstract than the inventory of phonemes differ in language phonologies. Notably, the way in which phonemes can cooccur in words is different. For instance, languages like Japanese have a very limited set of syllable shapes: V, VN, CV, CVN. Japanese syllables cannot have complex onsets (except for consonant-glide onsets) and cannot have codas (except for nasal consonants). In contrast, other languages like French or English allow for much more complex and varied syllable types: CCV, VCC, CCVCC, etc. These contrasting phonotactic properties affect the way in which foreign words are incorporated into a language. For instance (see Table 1), Japanese borrows words from languages with complex syllabic structures by inserting an "epenthetic" vowel [u] or [o] inside consonant clusters or after final consonants in such a way that the outcome conforms to the phonotactics of Japanese.

Table 1. Original Words From a Non-Japanese Language and the Japanese Corresponding Adaptations (From Itô & Mester, 1995)

Original Word	*Japanese Adaptation*
"Fight"	faito
"Festival"	fesutibaru
"Sphinx"	sufiNkusu
"Zeitgeist"	tsaitogaisuto

Does this process of vowel epenthesis occur in the input system, yielding the perception of illusory segments? Or does it reflect much later processing levels, like the influence of orthographic notation or problems in speech production? Several studies have claimed that phonotactic properties directly affect perceptual processes (Dupoux, Kakehi, Hirose, Pallier, & Mehler, 1999; Halle, Segui, Frauenfelder, & Meunier, 1998; Pitt, 1998; Massaro & Cohen, 1983). However, this issue remains controversial. In the following sections, we review these studies and propose to test the various, current, and competing theoretical interpretations using electrophysiological methods.

Phonotactic Effects in Perception

Massaro and Cohen (1983) presented English subjects with syllables starting with an obstruent consonant followed by a liquid. Liquids were synthesized along a continuum between [r] and [l]. The clusters were either legal in English ([sl], [tr], [pl], [pr]) or nonlegal ([sr], [tl], [vl], [vr]). The results show that listeners tend to perceive ambiguous items as legal clusters rather than as illegal ones. Note, though, that these studies were only run on a population of English-speaking subjects, leaving open the possibility that part of the observed effect is due to universal acoustic/phonetic factors. Pitt (1998) recently replicated Massaro and Cohen's results and found that when the illegal clusters are presented within words, similar shifts in consonant identifications are still observed. However, given that these illegal clusters are legal within words (crossroad, maudlin, Atlantic), no such shift should have been found. Therefore, it is possible that part of the original effect was due to some factor other than phonotactics.

More recently, Halle et al. (1998) studied the perception of word initial clusters in French. The clusters *[dl] and *[tl] are illegal word onsets contrary to [gl] and [kl]. These illegal clusters tend to be perceived as [gl] and [kl] in an open-response paradigm as well as in a forced choice paradigm. A gating study, however, demonstrates that during the initial portions of the stimuli, subjects perceive the phonemes as dental stops, but as information for the liquid becomes available in larger gates, perception switches and subjects identify the initial phoneme as velar. Such misperception is also found in a speeded phoneme detection study on nonwords beginning with the cluster [dl]: /d/ is missed in 69% of the cases, whereas /g/ is detected in 80%. As in Massaro and Cohen's study, one of the shortcomings of this work is that it is not cross-linguistic and, hence, it is difficult to evaluate whether the observed effects are truly language dependent or are simply universal acoustic/phonetic effects.

Dupoux et al. (1999) have studied the phenomenon of vowel epenthesis in Japanese using a cross-linguistic design. They used a continuum of stimuli ranging from ebzo (no vowels between the consonants) to ebuzo (a full vowel between the consonants). The Japanese, but not the French participants, reported the presence of a vowel [u] between consonants, even in stimuli containing no acoustic correlate of vowels. French participants, in contrast, had problems discriminating items with different vowel lengths (ebuzo versus ebuuzo), a distinctive contrast in Japanese, but not in French. These authors also used a speeded ABX discrimination paradigm and found that Japanese participants had trouble discriminating between the endpoint VCCV and VCuCV stimuli. The confusion between ebuzo and ebzo in Japanese was found even with participants who were quite proficient in French. These results suggest that phonotactics play a role important enough as to provide an illusory perception of segments.

Finally, in terms of language acquisition, Jusczyk, Friederici, Wessels, Svenkerud, and Jusczyk (1993) and Jusczyk and Luce (1994) found that between 6 and 9 months, infants develop a preference for phoneme sequences that are typical to their native language. They prefer to listen to lists of native words than to foreign words even when the two languages are as close as Dutch and English. The main differences between these two languages come from the phoneme inventory and phonotactics while word-prosody is rather similar in both languages. Although this does not constitute direct proof that phonotactics play an early role in speech perception, at least, it shows that phonotactics characteristics of the native language may be in place during the first year of life, and as early as the phoneme inventory.

In brief, many experiments suggest that phonotactics play a role both in perception and in acquisition. However, it is not clear at which processing level phonotactics interfere with speech perception. The next section examines the various options at hand.

Phonotactics in Models of Speech Perception

In the following, we distinguish three classes of models that make contrasting claims with about the effects of high-order phonological properties such as phonotactics.

Segmental Models

Segmental models claim that speech sounds are first categorized in terms of discrete segments (phonemes or bundle of features) (Marslen-Wilson & Warren, 1994; McClelland & Elman, 1986; Eimas & Corbit, 1973). A representation in terms of a sequence of segments is derived and then used to retrieve word forms in the lexicon. In such models, whenever a listener is presented with a sequence of phonemes that belong to the phonetic inventory of the language in question, a stable representation should be obtained, regardless of whether or not that particular sequence is used in the language. The only way these models could account for phonotactic effects would be to claim that the lexicon uses feedback to provide information, or that the effects occur in another part of the processing system. McClelland and Elman (1986) claim that the phonotactic effects found by Massaro and Cohen (1983) can be modeled by appealing to top-down word to phoneme influences during perception (but see a reply in Massaro & Cohen, 1991). Marslen-Wilson and Warren (1994) claims that nonwords are perceived through the activation of the lexicon; in such a view, phonotactic effects emerge as a rather late process of finding nonwords through analogy. In brief, in this first class of models, phonotactics can only have a rather late effect, which reflects either the joint activation of many words in the lexicon, or rely on a postaccess mechanism.

Hierarchical Models

As above, these models propose that a segmental representation is derived first. In addition, however, this representation is used to construct a hierarchical phonological representation that contains higher-order units such as morae, syllables, and feet (Pallier, 1994; Church, 1987; Frazier, 1987). Such a phonological representation is obtained from the segments using language-specific rules. In Pallier, Sebastian-Gallés, Felguera, Christophe, and Mehler (1993), evidence was found that listeners build a structured representation containing segments and syllables on-line (see also Pallier, 1994). In such models, one could then propose that incorrect or illegal phonological forms are automatically regularized by the parsing device. Hence, an illegal nonword like ''ebzo'' for Japanese speakers would be corrected by the parser as ''ebuzo'' via the insertion of a vowel. In such a model, because of the time it takes for the parser to operate, some delay in the detection (or correction) of an illegal form should be predicted. Hence, one should expect an early segmental representation level, and a somewhat later ''regularized'' phonological representation level.

Coarse Coding Models

Coarse coding models postulate that the input signal is directly parsed into large processing units. For example, Mehler, Dupoux, and Segui (1990) have proposed SARAH, a model based on an array of syllable detectors. In this model, speech sounds are categorized into syllable-sized units. The repertoire of syllables includes all the syllables used in the language. Similar proposals have been made for triphones (Wicklegren, 1969), diphones (Klatt, 1979), and semisyllables (Dupoux, 1993; Fujimira, 1976). In such a view, an account of phonotactically based assimilation goes as follows: Faced with a foreign language, the perceptual system tries to parse the signal using, say, the available native syllabic categories. In Japanese, there are no syllable categories containing consonant clusters or coda consonants. A stimulus like /ebzo/ therefore activates categories for ''e'' and ''zo''. It also activates, to a lesser extent, all syllables that start with /b/: ''bu,'' ''ba,'' ''be,'' ''bi,'' and ''bo''. The ''bu'' interpretation is favored, maybe because in Japanese, the [u] vowel is frequently shortened or devoiced and shows considerable allophonic variation (see Keating & Hoffman, 1984; Beckman, 1982). Hence, the prototype for ''bu'' is fairly compliant and should emerge as the best match. Phonotactic, in this final class of models, probably has a very early effect, and is not distinguishable from that of phoneme inventories.

As we saw above, these three classes of models make rather clear-cut predictions with respect to the time-course of processing for phonotactic information. We propose to test these hypotheses by using high-density ERPs, which are an ideal tool for exploring the time-course of cerebral processing.

Hypotheses and Design

The goal of this experiment was to find out when phonotactic properties of one's native language influence speech perception. In order to study this, we used vowel epenthesis as described above, in a cross-linguistic design. We used a mismatch detection task in which a series of four similar precursor stimuli were presented followed by a fifth stimulus either similar to the previous stimuli (control condition) or different (deviant condition) (see Table 2). The stimuli were minimal pairs of nonwords, where the only difference was the presence or absence of the epenthetic vowel /u/ between two consonants (igumo versus igmo). The behavioral predictions based on Dupoux et al. (1999) were that French subjects would detect the deviant stimulus, whereas Japanese subjects would not. In order to prevent acoustic information from influencing the detection of the deviant stimulus, acoustic variability was introduced in the precursors: These stimuli were randomly drawn from a large set of stimuli recorded by six different female Japanese speakers. A male voice was used for the test stimulus. In order to control for the temporal characteristics of the test item, the duration of the individual phonemes was matched across trials by using the resynthesized version of a Japanese male speaker. Finally, we introduced distractor trials in which the test item was deviant for both the French and the Japanese subjects (igimo). These trials were introduced so that the Japanese subjects would be given clear cases of a ''different'' response during the experiment.

ERPs were recorded during the behavioral task. To detect when Japanese phonotactics modify the electrophysiological responses in the Japanese, as compared to the French participants, we compared the ERPs in the control versus deviant conditions for both groups.

In this type of experimental paradigm, where a deviant item is presented after a succession of similar items, a mismatch process between the features of the novel stimulus and the neural traces of the preceding stimuli in sensory memory leads to an early discrimination effect (MMN) (Näätänen, 1990). If a separate phonological coding, independent of acoustical format, is carried out as suggested by Dehaene-Lambertz (1997) and

Table 2. Experimental Conditions and Predictions for the Behavioral Responses in Japanese and French Participants

Condition	*Precursor Items*				*Test Item*	*Predictions: Japanese*	*Predictions: French*
Control	igumo igmo	igumo igmo	igumo igmo	igumo igmo	igumo igmo	Same	Same
Deviant	igumo igmo	igumo igmo	igumo igmo	igumo igmo	igmo igumo	Same	Different
Distractor	igumo igmo	igumo igmo	igumo igmo	igumo igmo	igimo igimo	Different	Different

Näätänen et al. (1997), French subjects should display a MMN response, time-locked to the point of deviance between the context and the test stimuli (third phoneme). It was also expected that French subjects would show a deviance effect at the time-window of the late positive component (LPC). Indeed, this component is sensitive to the conscious detection of a deviant item and to the decision process involved in making a response. It is directly correlated to behavioral results, which must show that French subjects have no problem in discriminating /igmo/ from /igumo/.

After isolating discrimination responses in French subjects, we have studied how this is affected by the subjects' native language. Here, the predictions are quite straightforward. If the phonotactic effect takes place late, early electrical components should be identical for French and Japanese subjects. Indeed, we should find an early MMN for the "ebuzo" versus "ebzo" contrast in Japanese and French subjects, even if later processes mask this mismatch effect and prevent Japanese subjects from consciously detecting the change. If, on the other hand, the phonotactic effect takes place very early on, a very reduced or no MMN response should be found for Japanese subjects.

In brief, in order to find an electrical component sensitive to a specific language, we first isolated the time-windows for which the control versus deviant comparison was significant in French subjects. We then tested whether or not the effect for Japanese subjects was significant in this time-window. Furthermore, we computed the condition (control versus deviant) × language (French versus Japanese) interaction. If the electrical component tested is indeed language-specific, language × condition interaction should be significant.

RESULTS

Behavioral Responses

The analysis of the percentage of "different" responses showed a main effect for language ($F(1,22) = 382.8, p < .0001$), a main effect for condition ($F(1,22) = 2{,}026.9, p < .0001$), and a significant language × condition interaction ($F(1,22) = 1{,}814.1, p < .0001$) (Table 3). A post hoc analysis revealed that this interaction is mostly due to the deviant condition. Indeed, French subjects perceived a change in 95.15 percent of the trials, whereas Japanese subjects perceived it in only 8.88 percent of the trials ($F(1,22) = 1{,}045.8, p < .0001$). In the control condition, performances were nearly identical for French and Japanese, but the Japanese subjects made more errors than the French ($F(1,22) = 5.13, p = .034$).

French subjects responded slower than the Japanese (1174 versus 1023 msec; $F(1,22) = 4.31, p = .05$). In both population, subjects were slower in the deviant condition (1174 msec) than in the control condition (1023 msec; $F(1,22) = 13.5, p = .001$). Interaction between language and condition was significant ($F(1,22) = 4.5, p = .045$).

Electrophysiological Results

At the vertex, the classical auditory components N1 P2 were recorded, followed by two negativities N2 and N3 due to the multisyllabic structure of the test item. Then, a slow positivity, the LPC due to the active response, was recorded. Because the control and deviant items began to diverge around 237 msec following the item onset,[1] i.e., at the introduction of the third phoneme, we indicated the latencies for the electrical components relative to this time. The inspection of the time-course of two-dimensional reconstructions of t test values in the comparison of deviant with control trials isolated three time-windows when significant differences were present in French subjects: 139 to 283 msec (376 to 520 msec post item onset) including N2 and N3 for the first response, 291 to 419 msec (528 to 656 msec post item onset) for the second response, 523 to 651 msec (760 to 888 msec post item onset) for the third response (peak of the LPC). In Japanese subjects, deviant minus control differences were weaker in size and duration than in French subjects. When differences were observed, they occurred during the time-windows isolated in French subjects.

First Response: 139 to 283 msec Postdeviance Onset

The first condition effect in French subjects consisted of a sharper negativity in deviant than in control trials at 147 msec (N2) and 251 msec (N3) above frontal electrodes. The polarity of this effect was reversed above temporal electrodes. As illustrated in Figure 1, the cartography of the difference between deviant and control (t test) was very close to the topography of a MMN as described in the literature, i.e., negativity above the fronto-central region with a reverse of polarity at the mastoids. In Japanese subjects, no difference between conditions was evident.

To study this first response, we have chosen a central pair and the mastoid electrodes, where the deviant minus control difference was respectively negative and positive (Figure 1). For both pairs of electrodes, the condition effect was significant in French (central pair: $F(1,11) = 7.55, p = .019$ and mastoid pair: $F(1,11) = 5.23, p = .043$), but not in Japanese subjects (central pair: $F(1,11) = 1.23, p = .291$ and mastoid pair: $F(1,11) < 1$). The language × condition interaction was significant for the central electrodes ($F(1,22) = 8.17, p = .009$), but not for the mastoid electrodes $F(1,22) = 1.46, p = .240$). There was no main effect for hemisphere nor any significant interaction of hemisphere with any of the other factors. However, as illustrated by Figure 1, the condition effect was stronger in French subjects for the left mastoid ($F(1,11) = 5.63, p = .037$) than the right ($F(1,11) = 1.75, p = .417$).

Table 3. "Different" Responses and Reaction Times in Japanese and French Participants

Experimental Condition	*Control*	*Deviant*	*Distractor*
Japanese participants			
"Different" responses (%)	6.1	8.9	99.5
Reaction times (msec)	1,030	1,016	996
French participants			
"Different" responses (%)	1.7	95.1	99.2
Reaction times (msec)	1,134	1,215	1,128

Second Response: 291 to 419 msec Postdeviance Onset

A second response was present around 300 msec in French subjects and lasted about 130 msec: While there was a beginning of posterior positivity for the control condition, it was delayed in the deviant condition. The subtraction deviant minus control was very asymmetric in French subjects with a marked positivity above the right-frontal region and a medial-posterior negativity, more important on the right side (Figure 2). Although the voltage cartographies in Japanese subjects seemed similar to those in French subjects, the t test maps showed much weaker effects in size and duration.

An inferior-frontal pair (at the maximum of the positivity) and an occipital pair (at the maximum of the negativity) were analyzed for this second time-window. In French subjects, there was no main effect of condition for the infero-frontal electrodes, but a significant hemisphere × condition interaction ($F(1,11) = 10.28$, $p = .008$) due to a significant condition effect present only for the right-infero-frontal electrode ($F(1,11) = 19.17$, $p = .001$ for the right electrode, $F(1,11) < 1$ for the left). In Japanese subjects, there was no effect of condition for this pair ($F(1,11) = 1.62$, $p = .125$), nor any significant hemisphere × condition interaction ($F(1,11) < 1$). The language × condition × hemisphere interaction was

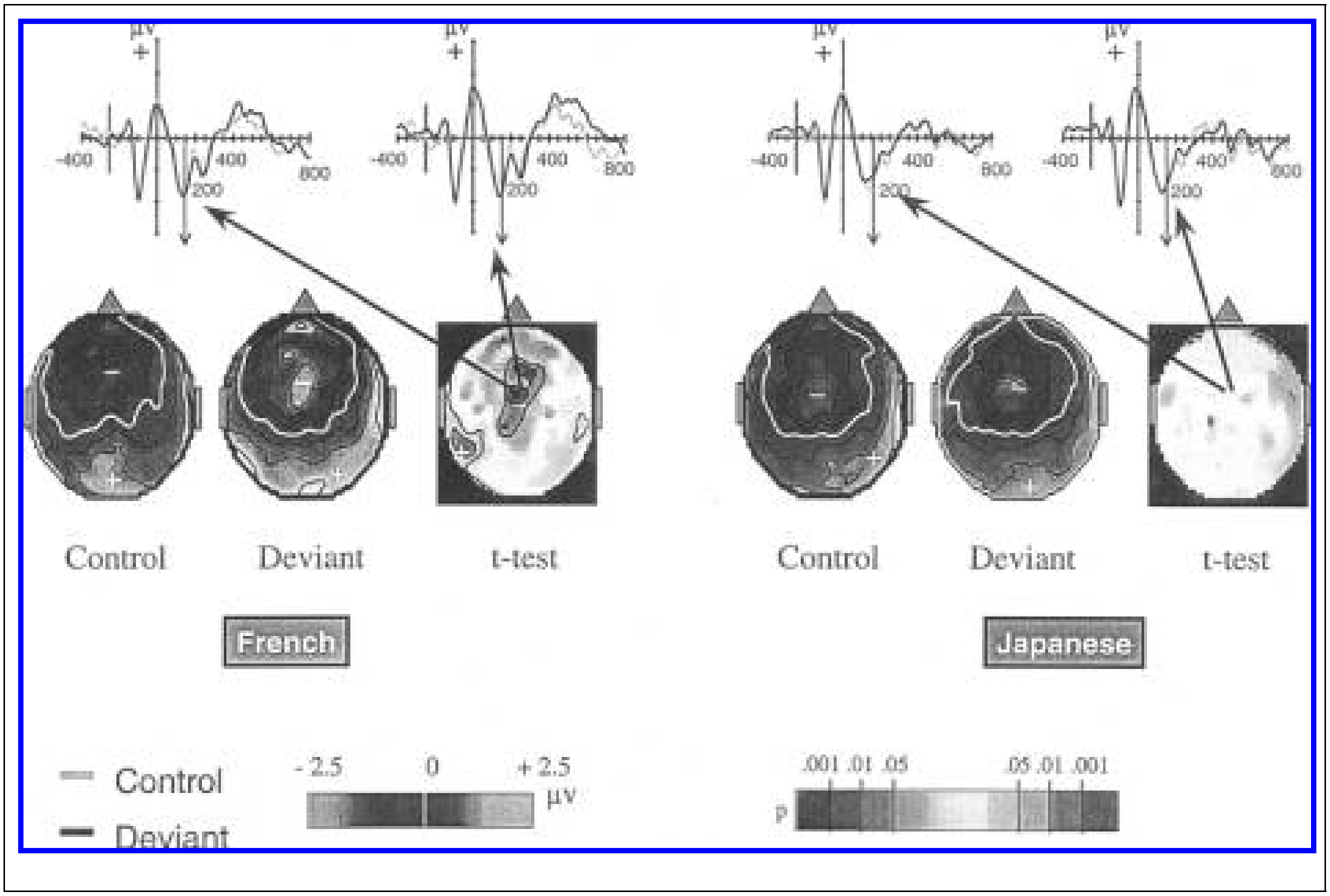

Figure 1. ERPs to the last item of the trials in French (left) and Japanese (right) participants. Top: ERPs from two symmetrical central electrodes. The first bar indicates the item onset and the second bar the onset of the deviance between control and deviant items. Bottom: Maps of evoked responses to control and deviant conditions at 164 msec following deviance onset (arrow on the waveforms) and maps of statistical significance (t test) of deviant versus control item at the same time. In French participants, a more important negativity for deviant than for control items is recorded at the fronto-central site with a polarity reversal over the temporal regions, whereas no difference between conditions is present in Japanese.

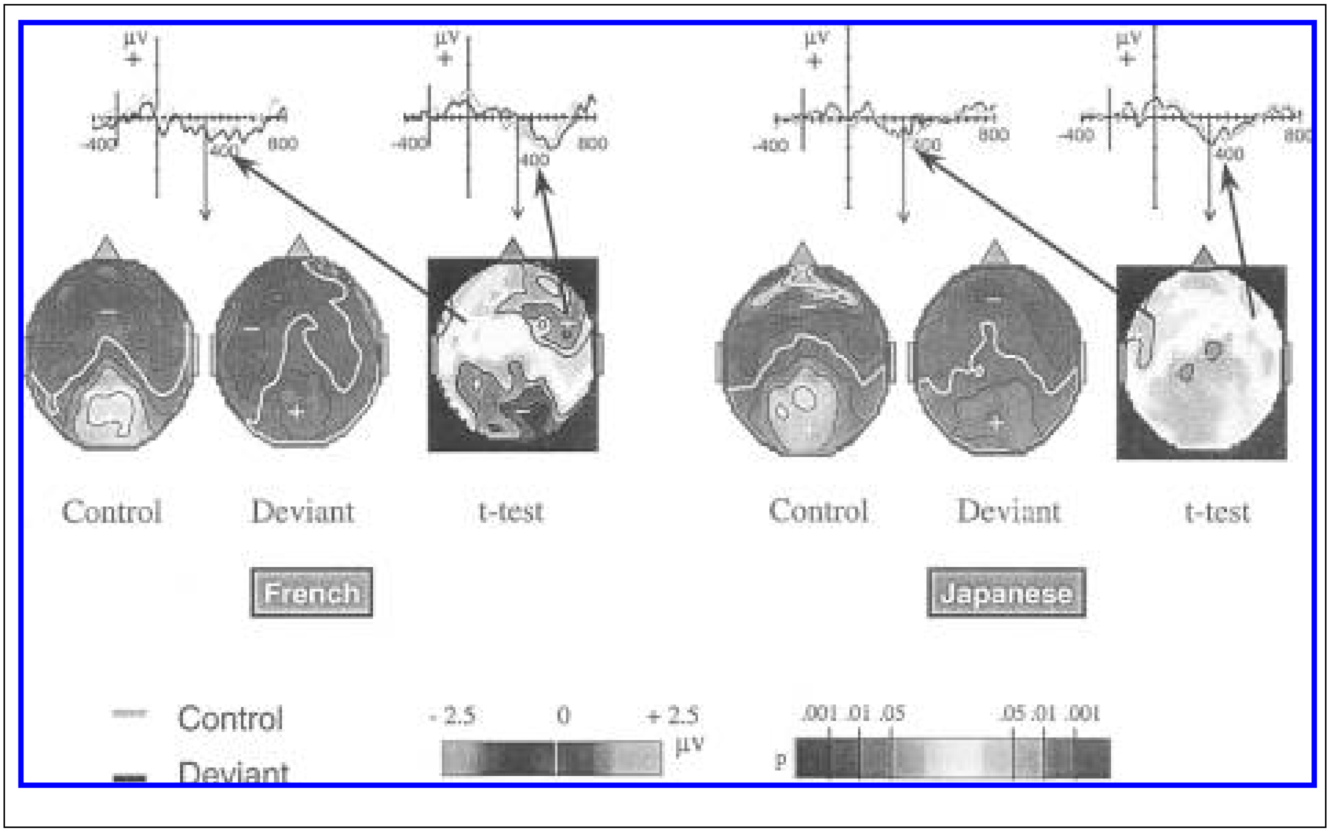

Figure 2. ERPs to the last item of the trials in French (left) and Japanese (right) participants. Top: ERPs from two symmetrical infero-frontal electrodes. The first bar indicates the item onset and the second bar the onset of the deviance between control and deviant items. Bottom: Maps of evoked responses to control and deviant conditions at 315 msec following deviance onset (arrow on the top waveforms) and maps of statistical significance (*t* test) of deviant versus control item at the same time. In both groups, while there is a beginning of posterior positivity for the control condition, it is delayed in the deviant condition. However, the condition effect is much more important for French than for Japanese subjects yielding a significant condition by language interaction.

significant ($F(1,22) = 5.43, p = .029$) due to a significant condition × language interaction for the right electrode alone ($F(1,22) = 8.40$, $p = .008$ and $F(1,22) < 1$ respectively over the right and left electrodes). Over the occipital pair, there was a condition effect in both the Japanese and French subjects ($F(1,11) = 14.37, p = .003$ in French and $F(1,11) = 4.9, p = .049$ in Japanese). The language × condition interaction was again significant ($F(1,22) = 5.62$, $p = .027$). For this occipital pair, there was no significant interaction of hemisphere with any other factor.

Third Response: 523 to 651 msec Postdeviance Onset

Finally, the third response in French subjects was related to a larger and longer LPC for the deviant than for the control condition. As with the second response, it was a relatively slow response that lasted around 220 msec. In Japanese subjects, the LPC was weaker than in French subjects with little difference between deviant and control conditions (Figure 3).

For this last response, we have chosen the central and the inferior-frontal pairs already analyzed during the preceding time-windows. The voltage amplitude was higher in French than in Japanese subjects yielding a main language effect for the infero-frontal electrodes ($F(1,22) = 18.63, p < .001$), and a trend for the central pair ($F(1,22) = 3.05, p = .095$). In French, there was a condition effect for the central ($F(1,11) = 12.27$, $p = .005$) and the infero-frontal pairs ($F(1,11) = 11.95, p = .005$). In Japanese, there was no condition effect for both pairs. The language × condition interaction was significant for both pairs (central pair: $F(1,22) = 4.44, p = .047$ and infero-frontal pair: $F(1,22) = 6.75, p = .016$). There was also in French subject a significant condition × hemisphere interaction for the central pair ($F(1,11) = 5.83, p = .034$) due to a predominant response over the right side ($F(1,11) = 15.52$, $p = .002$ over the right electrode and $F(1,11) = 2.24$, $p = .163$ over the left). Above this same right side, there was a condition effect in Japanese subjects ($F(1,11) = 6.67$, $p = .025$) with no significant hemisphere × condition interaction. However, the language × condition interaction was again significant ($F(1,22) = 5.91, p = .024$) for this right central electrode.

DISCUSSION

In this experiment, a deviant stimulus was introduced after the presentation of four similar stimuli. Because

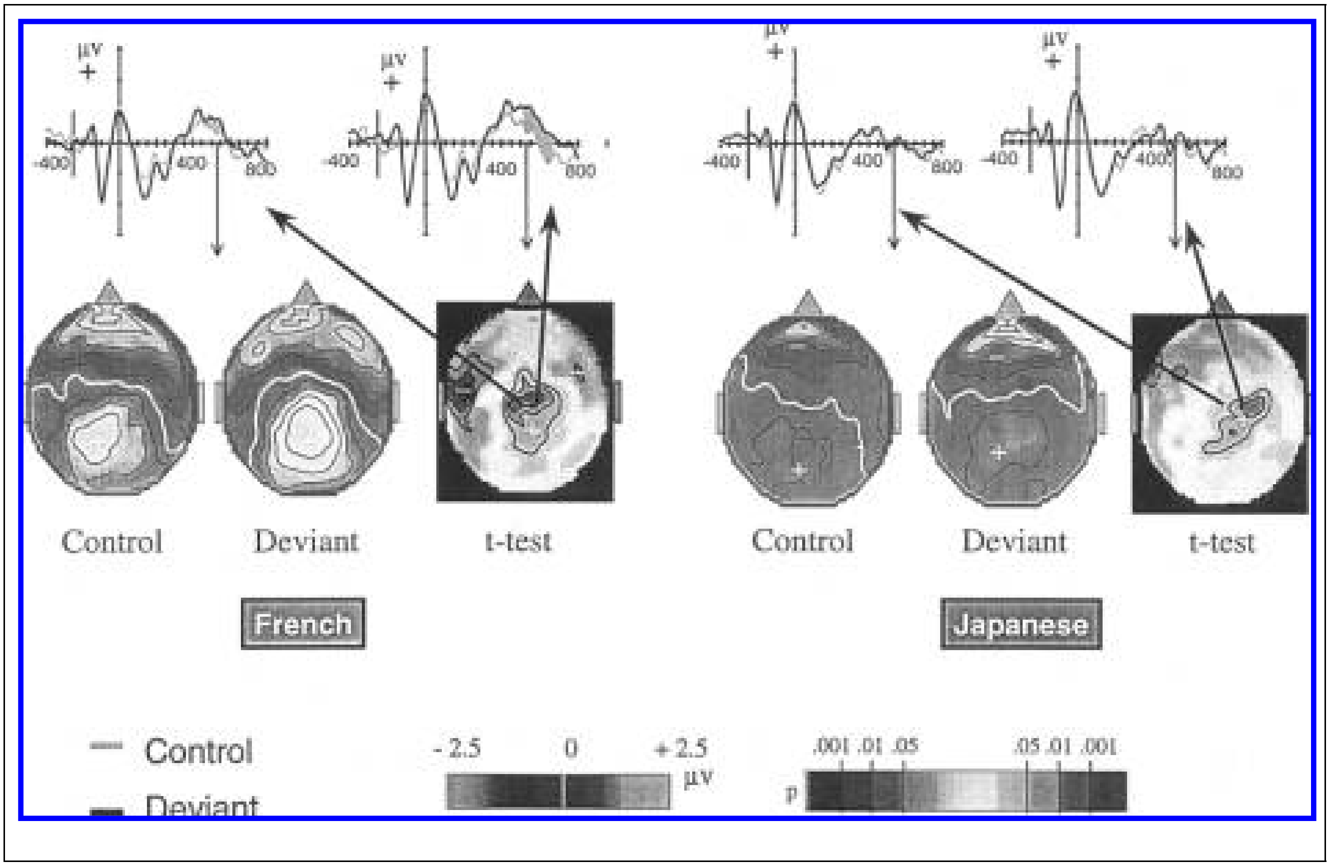

Figure 3. ERPs to the last item of the trials in French (left) and Japanese (right) participants. Top: ERPs from the two same symmetrical central electrodes presented in Figure 1. The first bar indicates the item onset and the second bar the onset of the deviance between control and deviant items. Bottom: Maps of evoked responses to control and deviant conditions at 531 msec following deviance onset (arrow on the top waveforms) and maps of statistical significance (*t* test) of deviant versus control item at the same time. A larger and longer LPC was elicited for the deviant than for the control condition, especially in French participants.

our goal was to study linguistic representations independent of acoustical representations, we used precursor items spoken by four different voices. The similarity between items could thus be computed only at a linguistic level. This experimental situation is very close to a natural situation where listeners have to normalize speech across different speakers. The behavioral results demonstrated that Japanese and French subjects do not react similarly to the same stimuli, and that performances are strongly influenced by the subjects' native language. There was a huge language by condition interaction. Japanese subjects almost never heard a difference between items like /igmo/ and items like /igumo/. These results clearly indicate that the phonotactic rules of the native language deeply modify speech perception and confirm that vocalic epenthesis in Japanese is not only a production phenomena, but also a by-product of perception processes. This is confirmed by the electrophysiological results that we discuss below.

For French subjects, the deviant versus control comparisons produced three responses related to the introduction of a novel linguistic item. The statistical analyses showed that these responses were either not present in Japanese or were of shorter duration and weaker. Moreover, the language × condition interaction was significant for all three responses. We will first consider the meaning of these three responses in French subjects in the context of the existing literature. We will then discuss the cross-linguistic evidence (electrophysiological results) and their significance for speech perception models.

In the literature, there are reports of a mismatch negativity or MMN that is elicited whenever there is a mismatch between the features of a perceived stimulus and the representation in the sensory memory left by the stimuli immediately preceding it. This component is specific to the auditory modality and involves generators that are predominantly located in the planum temporale (Giard, Perrin, Pernier, & Bouchet, 1990). Recent experiments have shown that the MMN is not only elicited when an acoustical mismatch is detected, but also when the differences between the precursors and the test item concern more abstract properties, like phonetic categories (Dehaene-Lambertz, 1997; Näätänen et al., 1997), or expectations build by subjects during the experimental run (Cowan, Winkler, Teder, & Näätänen, 1993). Because of its latency (139 to 283 msec postdeviance onset) and its topography (centro-frontal negativity synchrone of temporal positivity, which tends to be asymmetric in favor of the left mastoid), the first significant response observed in French subjects seems very close to a MMN, as it is described in the literature. Note, however, the differences between the procedure used in our experiment and that of classical MMN experiments. First, our participants were required to make an active comparison between the test item and the precursors whereas MMN is usually recorded in a passive listening paradigm. Second, our stimuli were more complex than the stimuli usually used in experiments in which MMNs are elicited: The context was highly variable, and the stimuli were multi-syllabic. It is, in fact, quite surprising to find such similarities between our first response and MMN, and it would be interesting to determine whether this response is subserved by the same generators as those classically involved in MMN: Attention to the stimuli in our experiment may have enhanced automatic processes, and selectively favored one type of processing rather than another. In any event, our results demonstrate that across different speakers, different speech rates, and, hence, across important acoustic differences, French subjects were able to establish a more abstract representation that is used to compare test and precursor items. Such comparison was done at an early latency, indeed, a latency comparable to those found in comparison of simple acoustical dimensions.

A second response was evident in French subjects between 291 and 419 msec postdeviance onset. Its topography is quite peculiar with positivity over the right-frontal region and negativity over the medial-posterior regions. Frontal asymmetry is very clear as shown by Figure 2 and by the highly significant hemisphere × condition interaction. Such an effect has not been reported previously in experiments where MMN was recorded. However, because of the relative complexity of our comparison task, it is possible that in order to respond, subjects not only used the phonetic representation stored in the sensory memory, but also relied on other processing in order to do a second check before responding. The difference in topography between the first and second response (see Figures 1 and 2) eliminates a second loop in the same process. In the following, we speculate that this second effect could be related to higher level processes, that mediate, e.g., metaphonological awareness. Several experiments have indeed described an electrical response related to phonological processing, called phonological mismatch or PMM (Connolly & Phillips, 1994). This response has been elicited for pairs of words or nonwords in metaphonological tasks like rhyme judgement (Perez-Abalo, Rodriguez, Bobes, Gutterriez, & Valdes-Sosa, 1994) or when subject's expectation about the word at the end of a sentence was violated (Connolly & Phillips, 1994). A phonological mismatch response has also been recorded after visual presentation of words or pictures (Perez-Abalo et al., 1994). Although it is difficult to judge from the single electrode presented in that paper whether it is the same or a close neural network that is involved under the visual and auditory conditions, the functional similarities of the responses in the two modalities suggest that the PMM may reflect a phonological representation independent of the modality of the stimulus presentation. The PMM latency is longer than that of the MMN (270–300 msec in Connolly & Phillips (1994), 250–450 msec in Praamstra, Meyer, & Levelt, 1994) on alliteration of pairs of words and compatible with our second response (290 to 400 msec). The topographies are more difficult to compare across experiments because of the small number of electrodes used in the recording system of these experiments, and the different choices of a reference. However, none of these experiments has reported the frontal asymmetry that we have. More data are needed to better define the process that elicits a PMM and whether our second effect is similar to what is described in the literature as a PMM.

Finally, we found a late positive complex (LPC), which is known to be due to the conscious detection of a less frequent stimulus and is modulated by the response decision. Because the behavioral results showed a major effect of deviance detection in French subjects, a significant difference between deviant and control was expected at this level. Indeed, there was a larger and longer positivity over the central regions for the deviant condition as compared to the control.

In conclusion, we have identified three electrical responses related to the detection of a phonemic deviance in complex items in French subjects. We will now see how native language modifies these responses and discuss the Japanese results as related to the French data.

ERPs demonstrated that native language interacts with phonological representation and goes as deep as sensory memory: For all three responses identified in French subjects, a significant language × condition interaction was found. First, no early MMN effect was found in Japanese subjects, even with instruction focussing subjects' attention towards the auditory stimuli and a deviance detection task. This suggests that phonotactics play a very early role that probably goes back to the coding of phonetic properties. Second, at the time-window of the second response, Japanese and French voltage cartographies look similar and a weak, but significant condition effect is present over the occipital electrodes in Japanese. Finally, there was also a weak condition effect at the third time-window in this group.

It might seem paradoxical that the first response is sensitive to the subjects' native language while the following responses appear to be dependent on universal coding. Two possibilities could explain this paradox. First, since ERPs rely on the precise timing of a response related to input, a variable response might not appear in the average. The mismatch response is probably not all or nothing processing. It is therefore possible that from time to time, deviant items are effectively coded as deviant in the sensory memory of Japanese subjects, but this computation is neither as specific nor as automatic as that of French subjects and would disappear through the averaging process. However, the second response may amplify differences between deviant and control by summing up the previous deviance computations across a longer time-window and thus demonstrate a condition effect. The other possibility is that the second response processes information coming from different subsystems, e.g., a prototypicality system, which may emit an error signal when illegal clusters are presented. There is also the possibility of a phonetic system that keeps track of the phonemes presented, but this process may well be encapsulated and not easily accessible for mismatch processes. Although each subsystem on its own would emit a signal too weak to be noticed, together they would be strong enough to create a deviance effect at the second time-window. In any case, the difference in sensitivity of the first two responses to native language confirms the fact that the phonological representations tagged by these responses are different.

What are the theoretical implications of these results for speech perception models? The important result here is that at the time of the first mismatch response in French subjects, Japanese subjects show no evidence of a deviance effect. We could thus conclude that a fast and automatic coding of the speech input exists that relies mainly on the formats authorized by the native language. These results are compatible with Coarse Coding models like Sarah proposed by Mehler et al. (1990), but are difficult to explain using Segmental or Hierarchical models. However, the condition effect observed in Japanese during the second response suggests that some information about the input could be recovered from other processing systems.

From a behavioral point of view, it has already been demonstrated that the Japanese are able to access this information. For example, Dupoux et al. (1999) found in Experiments 1 and 2 that, although Japanese listeners tend to report that they hear a vowel /u/ in stimuli like /igmo/, they do so at a significantly lower rate than when the /u/ is really present (65–70% versus 95%). Similarly, in Experiment 3 they found that Japanese listeners make 32% errors in an ABX task involving stimuli like /igmo/ and /igumo/. Although such an error rate was significantly higher than for control French listeners (6%), it was still better than chance, suggesting that Japanese subjects have residual abilities to distinguish a real from an illusory vowel.What our current study suggests is that this residual capacity is slower, and that it relies on a different network than the one used with native contrasts.

In conclusion, language phonotactics deeply affect speech perception. Such fast computation of a phonological representation might be useful in a noisy environment or in the event of mispronunciation in order to reconstruct the correct item and to facilitate lexical access. Because infants are sensitive to the phonotactic constraints of their native language around 9 months, we need to study how theses rules are implemented during language learning. Kuhl et al. (1992) have suggested that exposure to a particular language, and thus to particular phonemes, increases responses to the language prototypes phonemes while decreasing responses to the adjacent phonemes that are not present in the environment. If we extend this concept from phoneme category to phoneme combinations, the learning brain may in fact compute larger units, that include several phonemes. This would accelerate automatic responses to frequent combinations of phonemes, but would prevent the stabilization of the representation of phoneme combinations that are never encountered.

METHOD

Subjects

Twelve Japanese and 12 French subjects were recruited in Paris and tested individually, after giving written informed consent. They were all right-handed according to self-report and the Edinburgh inventory. None of them had a history of neurological or psychiatric disease, or a hearing deficit. At the end of the experiment, they filled out a detailed biographical questionnaire about their experience with foreign languages.

Japanese Subjects

Three men and nine women (age: 20 to 36, median 28.4) were paid for participating in the experiment. They had all begun to study English after the age of 12, mostly by reading (according to the questionnaire, 80% of the teaching was in written mode and 20% was spoken). They had all begun to study French after the age of 18 (except one subject at 16), mostly by reading (according to the questionnaire, 60% of the teaching was written and 40% was spoken).

French Subjects

Three men and nine women (age: 18 to 39, median 25) were volunteers and were not paid for their participation. None spoke Japanese. They had started studying English after the age of 12. Some also knew German, Spanish or Italian.

Stimuli

The stimuli for this experiment consisted of 18 items or six triplets of the form $V_1C_1C_2V_2$ (igmo), $V_1C_1UC_2V_2$ (igumo), and $V_1C_1IC_2V_2$ (igimo). Each triplet was uniquely defined by the particular combination of the $V_1C_1C_2V_2$, which we call a radical. There were six radicals: igmo, igna, ikno, ikma, okna, and ogma.[2] All 18 resulting stimuli were nonwords in both Japanese and French.

The stimuli were presented in blocks of five items. The first four items constituted the precursors and the last the test item. The material therefore consisted in two sets of stimuli: one for the precursor items and one for the test items. In both sets, stimuli were selected from a larger group by three French phoneticians and six naive Japanese subjects. We only retained those stimuli that were intelligible and that sounded reasonably natural in both languages.

The precursor set consisted in 72 different, naturally produced items (six radicals by two forms (igmo or igumo) by six speakers). They were selected from a set of 648 stimuli produced by six female Japanese speakers. For each speaker, nine utterances for each radical and both forms /igumo/ and /igmo/ were recorded in a sound attenuated room. The stimuli were digitized at 16 kHZ/16 bits on an OROS AU 22 board and processed on a waveform editor. Six items for each of the six radicals and for the $V_1C_1UC_2V_2$ (igumo) and the $V_1C_1C_2V_2$ (igmo) forms were selected from this large set, depending on the distribution of Japanese speakers in both groups. In the $V_1C_1UC_2V_2$ (igumo) stimuli set, we selected the items where the /u/ vowel was consistent with the French prototype. Indeed, in these stimuli, the production of the Japanese /u/ vowel varied between the /u/ and the /y/ French prototype. In the $V_1C_1C_2V_2$ stimuli, although Japanese speakers knew foreign languages, they could not be prevented from inserting a very short /u/ vowel into the consonant cluster. Stimuli were edited with a waveform editor, and the vocalic part was progressively removed until a French judge found that the /u/ vowel could no longer be perceived.

The test items consisted in 18 synthetic items—one item for each of the six radicals in each of the three types: $V_1C_1UC_2V_2$ (igumo), $V_1C_1C_2V_2$ (igmo), and $V_1C_1IC_2V_2$ (igimo). They were synthesized with a MBROLA speech synthesizer (Dutoit, Pagel, Pierret, Bataille, & Vreken, 1996), using the natural productions of a male Japanese speaker as a model. The algorithm for the speech synthesis used a male voice and a French diphone database. The test stimuli were edited with a speech editor to make sure there was no "schwa" vowel inserted in the consonant cluster; any evidence of a schwa was removed to obtain an unambiguous consonant cluster. The phonemes of the resynthesized stimuli had the same duration and the same pitch as the original natural one. Across the six radicals, the durations of the first three phonemes were measured. The first phoneme, V1, started 9 to 11 msec after the onset of the file. The second phoneme, C1, started 130 to 136 msec after the onset of the file. The third phoneme (U, C2, or I, depending on the form of the stimulus) started 236 to 240 msec after the onset of the file. Mean stimulus durations were: 746 msec for $V_1C_1UC_2V_2$ stimuli, 653 msec for $V_1C_1C_2V_2$, and 647 msec for $V_1C_1IC_2V_2$.

Procedure

Trials consisted in blocks of five items separated by 600 msec of silence: Four precursor items were followed by one test item. The precursors could either be in the igmo or in the igumo form, the test item defining the condition. Six types of trials were randomly presented. In the control condition, the test item was identical to the precursors (igmo... → igmo or igumo... → igumo). In the deviant condition, the test item was different from the precursors (igmo... → igumo or igumo...igmo). In the distractor condition, the test item was always of the igimo type (igumo... → igimo or igmo... → igimo). For each trial, a radical was randomly selected and then the four precursors were randomly chosen from the six possibilities in their radical group. The 36 different trials (six radicals × two forms × three conditions) were repeated 10 times each, in different random order. Reaction times were measured from the onset of the test item with a maximum response delay allowed of 3 sec. The next trial was presented 1,200 msec after the behavioral response. The entire set of 360 trials was divided into 20 trial blocks separated by a short pause to allow subjects to relax. The total duration of the experiment was 1 hr.

Stimuli and trial presentation, randomization, and response measurement were effected using the EXPE software package on a PC compatible with a Proaudio Spectrum 16 D/A Board (Pallier & Dupoux, 1997). Stimuli were presented binaurally via loud speakers, on a 65 dB HP.

Subjects were tested individually in a quiet room. Following electrode application, they were seated in a comfortable chair and were instructed to move as little as possible and fix a point in front of them during each block of recordings. The subjects were informed that they would hear lists of five stimuli, each of which would be made up of Japanese nonwords. The first four items would be identical and uttered by female Japanese speakers while the fifth would be spoken by a male voice. They were instructed to indicate if the last item was different by making a bimanual same–different response. The side associated with the "same" response was changed in the middle of the experiment, and the order counterbalanced across subjects.

Before starting the ERP recording, there was a training session of 12 randomly chosen trials, consisting of six controls and six distractors. No deviant trial was presented. It was considered that both Japanese and French subjects would respond "same" in the control condition and "different" in the distractor condition. Subjects received visual feedback about whether they had made a correct response or not. Training session results were excluded from the data analysis.

Recording System

ERPs were collected using a 128-channel geodesic electrode net (Tucker, 1993) referenced to the vertex. This device consists 128 Ag/AgCl electrodes encased in sponges moistened with a salty solution. The net was applied in anatomical reference to the vertex and the cantho-meatal line. Vertical eye movements and blinks were monitored via two frontal and two infra-orbital electrodes and two canthal electrodes were used to check for horizontal eye movements.

Scalp voltages were recorded during the entire experiment, amplified, filtered between 0.1 and 39.2 Hz, and digitized at 125 Hz. Then, the EEG was segmented into epochs starting at 1500 msec before the onset of each test stimulus and ending 1500 msec after it. These epochs were automatically edited to reject trials contaminated by significant eye movements (deviation higher than 70 μV on the horizontal and vertical para-ocular electrodes), or body movements (local deviation higher than 70 μV and global deviation higher than 100 μV). The artifact-free trials were then averaged for each subject across the thee experimental conditions (control, deviant, distractor). Averages were corrected through a 150-msec baseline (1500 to 1350 msec before stimulus onset), transformed into reference-independent values using the average reference method, and then digitally filtered between 0.5 and 20 Hz. Two-dimensional reconstructions of scalp voltage at each time-step were computed using a spherical spline interpolation (Perrin, Pernier, Bertrand, & Echallier, 1989).

Data Analysis

Behavioral Data

The "ikuma" test stimulus was removed when results were analyzed. Because of its acoustical quality, it was not perceived in the same way as the other test items in the vowel condition: When "ikuma" was the test stimulus, there were 60% errors in the control condition in Japanese subjects. In order to keep a fully factorial design, all "ikm" radical trials in both populations were removed. Trials in which subjects did not respond before the deadline were excluded from the analysis (less than 1% of the trials).

Two separate analyses of variance (ANOVA) were carried out for the percentage of "different" responses and for reaction times with language (French or Japanese) as the between-subject factor, and condition ("deviant" or "control") as the within-subject factor.

Electrophysiological Results

The goal of this experiment was to detect when Japanese phonotactics modify the electrophysiological responses of Japanese as opposed to French subjects. We thus adopted the following strategy to analyze the ERPs. First, the inspection of the time-course of two-dimensional reconstructions of t test values in the comparison of deviant with control trials was been used to isolate the time-windows for which significant differences were present in French subjects. Second, we checked if significant effects were also present in Japanese subjects using similar t test value maps. Third, for each selected time-window, two pairs of symmetrical electrodes were chosen, one pair at the maximum of the positivity and the second at the maximum of the negativity of the difference between deviant and control. For each time-window, average voltage was computed for the pair of electrodes selected and for the two conditions (deviant versus control) and submitted to an analysis of variance (ANOVA) with condition (deviant and control), hemisphere (left and right) as within-subject factors and language (Japanese or French) as between-subjects factor. Only the analyses done on the trials in which the behavioral responses were the dominant ones for each condition and each population are reported here.[3]

Acknowledgments

This study was supported by the Ministere Français de la Santé et de la Recherche PHRC 1995 No. AOM95011, the Groupement d'Intéret Scientifique Sciences de la Cognition No. PO 9004, the Fondation Evian, and the Mc Donnell Foundation.

Reprint requests should be sent to Ghislaine Dehaene-Lambertz, Laboratoire de Sciences Cognitives et Psycholinguistique, 54 boulevard Raspail, 75270 Paris cedex 06, France. Phone: (33) 1-49-54-22-62. Fax: (33) 1-45-44-98-35. Email: ghis@lscp.ehess.fr.

The color figures can be found at www.ehess.fr/centres/ilscp/persons/ghis/buzofig.pdf.

Notes

1. The speech synthethiser was programmed to begin the third phoneme at 237 msec after item onset. However, because of coarticulation, differences could already be seen in the sonograms and perceptible 50 msec earlier. However, if differences were audible, the identity of the third phoneme was not perceptible on average before 239 msec. These values give a bracket for deviance onset.
2. These radicals were selected on the basis of a reanalysis of previous results (Dupoux et al., 1999). We chose from the consonants that had given the most robust effects in the past.
3. The same analyses were also computed for all trials independent of the behavioral response. Because there were few trials for which the behavioral response was different from what was expected (around 9% in Japanese and 5% in French subjects), averaged voltages and statistical results were almost identical to the results presented here. In particular, no early significant difference between control and deviant condition could be identified in Japanese subjects and the language × condition interaction was still significant for the central pair ($F(1,22) = 7.6, p = .011$) for the first time-window.

REFERENCES

Beckman, M. E. (1982). Segmental duration and the ''mora'' in Japanese. *Phonetica, 39,* 113–135.

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

Church, K. W. (1987). Phonological parsing and lexical retrieval. *Cognition, 25,* 53–69.

Connolly, J. F., & Phillips, N. A. (1994). Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. *Journal of Cognitive Neuroscience, 6,* 256–266.

Cowan, N., Winkler, I., Teder, W., & Näätänen, R. (1993). Memory prerequisites of mismatch negativity in the auditory event-related potential (ERP). *Journal of Experimental Psychology: Learning, Memory and Cognition, 19,* 909–921.

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

Dupoux, E. (1993). Prelexical processing: The syllabic hypothesis revisited. In G. T. M. Altmann & R. Shillcock (Eds.), *Cognitive models of speech processing: The second sperlonga meeting* (pp. 81–114). Hove East Sussex, UK: Lea & Febiger.

Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., & Mehler, J. (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance, 25,* 1568–1578.

Dutoit, T., Pagel, V., Pierret, N., Bataille, F., & Vreken, O. V. D. (1996). *The MBROLA project: Towards a set of high-quality speech synthesizers free of use for non-commercial purposes.* Paper presented at the ICSLP'96, Philadelphia.

Eimas, P. D., & Corbit, J. D. (1973). Selective adaptation of linguistic feature detectors. *Cognitive Psychology, 4,* 99–109.

Frazier, L. (1987). Structure in auditory word recognition. *Cognition, 25,* 157–187.

Fujimira, O. (1976). Syllables as concatenated demisyllables and affixes. *Journal of the Acoustical Society of America, 59,* S55.

Giard, M. H., Perrin, F., Pernier, J., & Bouchet, P. (1990). Brain generators implicated in the processing of auditory stimulus deviance: A topographic event-related potential study. *Psychophysiology, 27,* 627–640.

Halle, P., Segui, J., Frauenfelder, U., & Meunier, C. (1998). Processing of illegal consonant clusters: A case of perceptual assimilation? *Journal of Experimental Psychology: Human Perception and Performance, 24,* 592–608.

Itô, J., & Mester, A. (1995). Japanese phonology. In J. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 817–838). Oxford, UK: Blackwell.

Jusczyk, P. W., Friederici, A., Wessels, J., Svenkerud, V., & Jusczyk, A. (1993). Infants' sensitivity to the sound pattern of native language words. *Journal of Memory and Language, 32,* 402–420.

Jusczyk, P. W., & Luce, P. A. (1994). Infants' sensitivity to phonotactic patterns in the native-language. *Journal of Memory and Language, 33,* 630–645.

Keating, P., & Hoffman, M. (1984). Vowel variation in Japanese. *Phonetica, 41,* 191–207.

Klatt, D. H. (1979). Speech perception: A model of acoustic–phonetic analysis and lexical access. *Journal of Phonetics, 7,* 279–312.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experiences alter phonetic perception in infants by 6 months of age. *Science, 255,* 606–608.

Marslen-Wilson, W., & Warren, P. (1994). Levels of perceptual representation and process in lexical access: Words phonemes and features. *Psychological Review, 101,* 653–675.

Massaro, D. W., & Cohen, M. M. (1983). Phonological constrainst in speech perception. *Perception and Psychophysics, 34,* 338–348.

Massaro, D. W., & Cohen, M. M. (1991). Integration versus interactive activation: The joint influence of stimulus and context in perception. *Cognitive Psychology, 23,* 558–614.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18,* 1–86.

Mehler, J., Dupoux, E., & Segui, J. (1990). Constraining models of lexical access: The onset of word recognition. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 236–262). Cambridge: MIT Press.

Näätänen, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral and Brain Sciences, 13,* 201–288.

Näätänen, R., Lehtokovski, A., Lennes, M., Cheour, M., Huotilainen, M., Iovonen, A., Vainio, M., Alku, P., Ilmoniemi, R. J., Luuk, A., Allik, J., Sinkkonen, J., & Alho, K. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature, 385,* 432–434.

Pallier, C. (1994). *Role de la syllabe dans la perception de la parole: études attentionnelles.* These de doctorat de l'Ecole des hautes études en sciences sociales [available from the author], Paris.

Pallier, C., Bosch, L., & Sebastian, N. (1997). A limit on behavioral plasticity in spech perception. *Cognition, 64,* B9–B17.

Pallier, C., & Dupoux, E. (1997). EXPE: An expandable programming language for on-line psychological experiments. *Behavior Research Methods, Instruments, and Computers, 29,* 322–327.

Pallier, C., Sebastian-Gallés, N., Felguera, T., Christophe, A., & Mehler, J. (1993). Attentional allocation within the syllabic structure of spoken words. *Journal of Memory and Language, 32,* 373–389.

Perez-Abalo, M. C., Rodriguez, R., Bobes, M. A., Gutteriez, J., & Valdes-Sosa, M. (1994). Brain potentials and the availability of semantic and phonological codes over time. *NeuroReport, 5,* 2173–2177.

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

Pitt, M. (1998). Phonological processes and the perception of phonotactically illegal consonant clusters. *Perception and Psychophysics, 60,* 941–951.

Polka, L., & Werker, J. F. (1994). Developmental changes in perception of non-native vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 421–435.

Praamstra, P., Meyer, A. S., & Levelt, W. J. M. (1994). Neuropsychological manifestations of phonological processing: Latency variation of a negative ERP component timelocked to phonological mismatch. *Journal of Cognitive Neuroscience, 6,* 204–219.

Tucker, D. (1993). Spatial sampling of head electrical fields: The geodesic electrode net. *Electroencephalography and Clinical Neurophysiology, 87,* 154 –163.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganisation during the first year of life. *Infant Behavior and Development, 7,* 49–63.

Wicklegren, W. A. (1969). Context-sensitive coding associative memory and serial order in (speech) behavior. *Psychological Review,76,* 1–15.

This article has been cited by:

1. G. Dehaene-Lambertz , T. Gliga . 2004. Common Neural Basis for Phoneme Processing in Infants and Adults. *Journal of Cognitive Neuroscience* **16**:8, 1375-1387. [Abstract] [PDF] [PDF Plus]
2. Alice Mado Proverbio, Liza Vecchi, Alberto Zani. 2004. From Orthography to Phonetics: ERP Measures of Grapheme-to-Phoneme Conversion Mechanisms in Reading. *Journal of Cognitive Neuroscience* **16**:2, 301-317. [Abstract] [PDF] [PDF Plus]