HAL
open science

New evidence for prelexical phonological processing in word recognition

Emmanuel Dupoux, Christophe C Pallier, Kazuhiko Kakehi, Jacques Mehler

To cite this version:

Emmanuel Dupoux, Christophe C Pallier, Kazuhiko Kakehi, Jacques Mehler. New evidence for prelexical phonological processing in word recognition. Language and Cognitive Processes, 2001, 16 (5), pp.491-505. 10.1080/01690960143000191 . hal-02341211

HAL Id: hal-02341211

https://hal.science/hal-02341211

Submitted on 4 Nov 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

New Evidence for Prelexical Phonological Processing in Word Recognition

Emmanuel Dupoux[*]

Christophe Pallier[*]

Kazuhiko Kakehi[**]

Jacques Mehler[*, +]

Abstract

When presented with stimuli that contain illegal consonant clusters, Japanese listeners tend to hear an illusory vowel that makes their perception conform to the phonotactics of their language. In a previous paper, we suggested that this effect arises from language-specific prelexical processes (Dupoux, Kakehi, Hirose, Pallier & Mehler, 1999). The present paper assesses the alternative hypothesis that this illusion is due to a "top-down" lexical effect. We manipulate the lexical neighborhood of non-words that contain illegal consonant clusters and show that perception of the illusory vowel is not due to lexical influences. This demonstrates that phonotactic knowledge influences speech processing at an early stage.

*. Laboratoire de Sciences Cognitives et Psycholinguistique, CNRS-EHESS

**. Graduate School of Human Informatics, Nagoya University, Japan

+. International School for Advanced Studies, Trieste, Italy.

We thank Nicolas Bernard and Takao Fushimi for their help in preparing the stimuli and running the experiment. We also thank James McQueen, Sharon Peperkamp and two anonymous reviewers for very useful comments on a previous version of this manuscript. Part of this work was presented at Eurospeech'99 (Dupoux, Fushimi, Kakehi and Mehler, 1999). Correspondence concerning this article should be addressed to Emmanuel Dupoux, LSCP, EHESS-CNRS, 54 bd Raspail, Paris, 75006. E-mail: dupoux@lscp.ehess.fr.

Most models of spoken word recognition postulate that the acoustic signal is transformed into a prelexical representation, typically a string of phonemes, and that this representation is used to access the lexicon. Such models have to spell out how the acoustic signal is transformed into a prelexical representation and whether this representation is enhanced by one's lexical knowledge. Many studies have established that the mapping between the signal and the prelexical representation is not simple. In their famous "Perception of the Speech Code" paper, Liberman, Cooper, Shankweiler and Studdert-Kennedy (1967) stressed the complexity of the relationships between the acoustic signal and the phonetic message: neighbor phones interact so much that a single acoustic stretch is often ambiguous and requires a larger context to be interpretable (see, for example, Miller and Liberman, 1979; Mann and Repp, 1981; Whalen, 1989). Their proposed solution was that listeners use their knowledge of how speech sounds are produced to decode the speech signal (for example, to compensate for coarticulation). A second source of information is lexical knowledge. Indeed, numerous studies have demonstrated lexical influences on phoneme identification (e.g. Ganong, 1980; Samuel, 1981a, 1987; Frauenfelder, Segui, Dijkstra, 1990). The phenomenon of phonemic restoration attests that lexical knowledge can yield the perception of a phoneme that is not present in the signal (even if acoustics play an important role in the phenomenon; cf. Samuel 1981b). A third source of information that can be used by the speech perception apparatus is phonotactic knowledge. There exists some empirical evidence that listeners tend to assimilate illegal sequences of phonemes to legal ones (Massaro & Cohen, 1983; Hallé, Segui, Frauenfelder & Meunier, 1998). Thus, French listeners tend to hear the sequence /dl/, which is illegal in French, as /gl/, which is legal (Hallé et al. 1998).

Among the three above cited sources of information, the influence of phonotactics is the less well established. Both the Massaro & Cohen and the Hallé et al. studies used stimuli in only one language. Therefore, it cannot be excluded that the observed effects were due to universal effects of compensation for coarticulation: it could be that /dl/ is universally harder to perceive than /gl/. A more convincing demonstration of phonotactic effects must involve a cross-linguistic manipulation. A second difficulty is the potential confound between phonotactic and lexical informations. It can be argued that nonwords containing illegal sequences of phonemes typically have fewer lexical neighbors than legal nonwords. As a matter of fact, McClelland and Elman (1986) interpreted the phonotactic effects of Massaro and Cohen as the result of top-down influences from the lexicon (a "lexical conspiracy" effect). They reported on unpublished data where an apparent phonotactic effect (the preference for /dw/ over /bw/ in nonwords with an ambiguous initial phoneme) was reversed in the presence of a strong lexical candidate: /?wacelet/, yielded the perception of /bwacelet/ (from *bracelet*) instead of /dwacelet/ (despite the illegality of the /bw/ cluster). The authors argued that the typical preference of /dw/ over /bw/ is due to the predominance of /dw/ words in the lexicon. They simulated these data, as well as Massaro & Cohen (1983) "phonotactic" result, with a "lexical conspiracy" effect in the TRACE model (but see Pitt & McQueen, 1998 for arguments against this interpretation).

In this paper, we wish to revisit the relative role of phonotactics and lexical knowledge by building on an effect which has been well documented cross-linguistically: the perception of illusory vowels in Japanese. Dupoux, Kakehi, Hirose, Pallier & Mehler (1999) have demonstrated that Japanese listeners, but not French listeners, perceive an /u/ vowel between consonants forming

illegal clusters in Japanese (e.g. between /b/ and /z/)[1]. These data show that the perceptual system of Japanese listeners insert an illusory vowel between adjacent consonants in order to conform to the expected pattern in this language. We called this phenomenon "vowel epenthesis". It suggests that the role of phonotactics is so important as to produce the illusion of a segment which is not actually present in the signal.

Though Dupoux et al. (1999) attributed this effect to phonotactic knowledge, it is not a priori excluded that the illusion results from top-down lexical influences. One may imagine that many Japanese words contain sequences /C_1uC_2/ in which C_1C_2 represent consonant clusters present in the nonword stimuli used in the experiment. It could then be argued that the activations of such lexical items conspire to produce the perception of /u/. Thus, the potential existence of real Japanese words, phonetic neighbors of the non-word stimuli, may have induced participants to report a vowel that is not present in the signal. Some may find excessive the proposal that lexical effects can be as strong as to blindly insert a vowel that is not present in the signal. However, as we noted above, there are well documented demonstrations that the lexicon can fill in missing speech sounds, at least when the underlying signal is degraded or ambiguous (Warren, 1984; Samuel, 1981a). In the present case, the signal is clear, but it contains sequences of phonemes that are illegal for Japanese speakers. The influence of lexical knowledge in such a situation is an open question.

This paper aims at determining the source of the vowel epenthesis effect in Japanese. We try to determine whether the illusory vowel is inserted during the first stage of speech processing, under the influence of phonotactic constraints, or whether it comes from the participants' lexical knowledge. To this aim, we created non-words containing illegal consonant clusters in Japanese. These items produce only one lexical neighbor when a vowel is inserted between the consonants. Specifically, for some items, the lexicon biases for the insertion of the vowel /u/ (like in *sokdo -> sokudo*, speed). In other items, the lexicon call for the insertion of a vowel other than /u/ to generate a word, (like *mikdo -> mikado*, emperor). How do Japanese listeners perceive these illegal non-words? If perceptual processes insert the vowel /u/ within the /kd/ cluster irrespective of the lexical status of the outcome, then they should report hearing an /u/ inside both /sokdo/ and /mikdo/. If, in contrast, their perception is influenced by the nearest real Japanese word, we expect them to report an /u/ and an /a/ respectively.

Experiment 1

Stimuli with a CVCCV pattern were selected so that there is only one possible vowel that can be inserted in the consonant cluster to produce a real Japanese word. The stimuli were then split into two sets, according to the vowel that yields a Japanese word. In the u-Set, the vowel that produced a word was /u/, like in *sokdo->sokudo.* In the non-u-Set, the vowel was /a/, /e/, /i/ or /o/, like in *mikdo->mikado*. Participants were tested on two tasks, transcription and lexical decision. In the transcription task, participants were presented with stimuli and asked to transcribe them into the Roman alphabet. In the lexical decision task, participants decided whether the items were words in Japanese or not. If epenthesis is lexically driven, one expects a strong effect of lexical neighborhood on the perception of the items containing a consonant cluster. In particular, items in the u-Set (e.g. *sokdo)* should be perceived with an epenthetic /u/, and items in the non-u-Set (e.g.

[1] In japanese, the syllabic types are restricted to C(y)V, V, and C(y)VN. As a result, the only legal consonant clusters are of the nasal plus consonants type.

mikdo) should be perceived with a vowel other than /u/ since that is the only way to obtain a Japanese word. In other words, items in both sets should behave similarly in the lexical decision experiment, since all items can produce Japanese words once the appropriate vowel is inserted. In contrast, if vowel epenthesis arises before lexical access, both *sokdo* and *mikdo* items should produce the perception of an illusory /u/, and consequently, only *sokdo* will be processed as a Japanese word. Indeed, while *sokdo* becomes *sokudo* (a word in Japanese) after /u/ epenthesis, *mikdo* becomes *mikudo* (a non-word in Japanese). In order to assess the performance on the test items, we included control words like *sokudo* and *mikado* and non-words like *sokado* and *mikudo* in the lists.

Method

Materials: Two sets of 19 triplets were constructed. The first set was called the u-Set and contains triplets with the following structure: (illegal non-word, word, legal non-word). The first item was a (C)VCCV disyllable containing an illegal consonant cluster in Japanese (e.g., *sokdo*). The other two items had identical segments as the first except for the insertion of a vowel between the middle consonants. The second item of the triplet was a Japanese word obtained by the insertion of an /u/ (e.g., *sokudo*). The third item was a non-word obtained by the insertion of one of the following three vowels: /a/, /e/, /i/ or /o/ (e.g., *sokado*). The second set (the non-u-Set) was similar in all respects to the u-Set, except that the vowels that yielded a word and a non-word, respectively, were swapped around. That is, the illegal non-word *mikdo* yielded the word *mikado* through the insertion of an /a/ and yielded the non-word *mikudo* through the insertion of an /u/. In all cases, there was only one possible way to make a word in Japanese through the insertion of a vowel in the consonant cluster of the first element of a triplet. The items are listed in the Appendix. An additional list of 78 filler items was also constructed, consisting of half words, half legal non-words.

The stimuli were recorded by a female native speaker of Japanese, trained in phonetics, who had been instructed to produce the cluster stimuli without any intervening vowel. All stimuli were checked and when a few glottal pulses appeared between the consonants, these pulses were deleted by digital editing.

Procedure: In the phonetic transcription task, all the items from the 38 triplets were presented through headphones in a different pseudo-random order for each participant. Participants were instructed to type the transcription of the items in the Roman alphabet on a computer keyboard. In the lexical decision task, the same items, with the filler items, were presented through headphones using the EXPE software on a PC computer (Pallier, Dupoux & Jeannin, 1997). Lists were presented in a different pseudo-random order for each participant. Participants were instructed to classify the stimuli into real Japanese words versus non-existent words as fast as they could. If no response was given within 4 seconds after the presentation of an item, the next stimulus was presented. Participants were given a practice session of 10 trials with feedback, followed by the main experiment during which no feedback was given.

Participants: Fourteen native speakers of Japanese were tested on the lexical decision experiment. Seven of them also performed the phonetic transcription task (after completing the lexical decision task). They were all Japanese volunteers recruited in Paris.

Results and Discussion

In this experiment, four triplets in the u-Set and three in the non-u-Set had to be removed because of more than 50% transcription errors or more than 40% lexical decision errors on the words or non-words. Most of these errors fell on /h/-initial words, whose first segment was misperceived as /f/ or /r/, or deleted. This left 15 triplets in the u-Set and 16 triplets in the non-u-Set. No participant was rejected.

First, the phonetic transcription results for the consonant cluster items were analyzed separately for items in the u-Set and items in the non-u-Set. Seventy-nine percent of the transcriptions of the consonant cluster items in the u-Set contained an epenthetic /u/ inserted between the consonants. Only 1% of the responses contained a different vowel (/i/ instead of /u/ in one item). The remaining 20% of the transcriptions reproduced the consonant cluster. In the non-u-Set, 75% of the transcriptions contained an epenthetic /u/, 1% are cases of consonant deletion (/juSi/ instead of /jurSi/), and 24% reproduced the consonant cluster. The difference in /u/ transcription between the two items sets was not significant ($X^2(1)=0.10$, $p>0.10$).

Second, the lexical decision data was analyzed in terms of percent 'word' responses. The data are displayed in Table 1. As one can see, words were labeled as words in 94% of the cases, and non-words in only 7% of the cases. The results of the cluster items depended on set type. 71% of the cluster items in the u-Set are labeled as words, a score significantly different from chance ($p<.007$). Planned contrasts showed that this score differed significantly from the score obtained with control non-words ($F1(1,13)=222$, $p<.001$; $F2(1,14)=120$, $p<.001$) and control words ($F1(1,13)=23$, $p<0.001$; $F2(1,14)=8$, $p<0.02$). In contrast, the cluster items in the non-u-Set were classified as words in only 8.0% of the trials, a score significantly below chance ($p<.001$). This score was significantly different from the score with the control words ($F1(1,13)=868$, $p<.001$; $F2(1,15)=1329$, $p<.001$), but did not differ from the score with the control non-words (both $Fs<1$).

Table 1: percent "word" response to cluster and control items in Experiment 1.

	Non-words	Words	Clusters
u-Set			
Percent “ word ” response	2.8%	91.0%	71.4%
Example	*sokado*	*sokudo*	*sokdo*
non-u-Set			
Percent “ word ” response	10.3%	96.9%	8.0%
Example	*mikudo*	*mikado*	*mikdo*

Third, we analyzed the RTs for responses falling into the `predicted' category. For items in the u-Set the predicted category was “Word” for both the Japanese words and for illegal cluster items. For the non-u-Set items the predicted category was “Non-Word” for Japanese non-words and for illegal cluster items. The reaction times are shown in Table 2. An ANOVA revealed that the control words yielded significantly faster latencies than control non-words ($F1(1,13)=28$, $p<.001$; $F2(1,30)=49$, $p<.001$). Cluster items in the u-Set tended to yield faster RTs than control non-words ($F1(1,13)=3.3$, $.05<p<.1$; $F2(1,14)=4.5$, $p=.053$), but did not differ significantly from control words (both $Fs<1$) In contrast, cluster items in the non-u-Set yielded slower RTs than the control words ($F1(1,13)=23$, $p<.001$; $F2(1,15)=62$, $p<.001$), but did not differ significantly from the control non-words ($F1(1,13)=2.9$, $p>.1$; $F2(1,15)=3.3$, $.05<p<.1$). Thus, the latencies to respond to cluster items depended on the items in the lexical category to which they are assimilated.

Table 2: Reaction times (ms), standard error, and percent error to cluster and control items in Experiment 1.

	Non-words			Words			Clusters		
	RT	SE	Err	RT	SE	Err	RT	SE	Err
u-Set	1223	33	2.8%	1046	33	9.0%	1093	77	28.6%
non-u-Set	1315	68	10.2%	933	31	3.1%	1234	61	8.0%

The overall pattern of results for this experiment was straightforward. The transcription task yielded similar rates of /u/ insertion irrespective of the presence or absence of a neighboring lexical item with that particular vowel. We found no insertion of a vowel other than /u/ (except for one participant who reported an /i/ in "*namda*" producing the Japanese word "*namida*"; this amounts to 1% of the responses). Thus, globally, the presence of biasing lexical neighbors did not influence transcriptions. The same conclusion arose from the analysis of the lexical decision data. Participants consistently classified cluster items as if they had inserted an /u/ prior to lexical access.

The reported results could be attributed to the fact that the stimuli were recorded by a native Japanese speaker. Even though the speaker was trained in phonetics, at times she could not avoid inserting short vowels between the cluster consonants. Her speech was digitally edited to remove pitch pulses corresponding to an inter-consonantal vowel. However, it could be that traces of coarticulation remained in the adjacent consonant allowing Japanese participants to reconstruct the underlying vowel. To control for this possibility, in the next experiment, we used a speaker whose maternal language allows consonant clusters.

Experiment 2

In order to remove the possibility that the effects obtained in the previous experiment are due to potential traces of /u/ produced by the Japanese speaker, the stimuli were recorded anew, but this time they were spoken by a native speaker of French. He was instructed to imitate words and non-words produced by a Japanese speaker, and produced the cluster stimuli without any intervening vocalic element. If the effects obtained in the previous experiment were due to coarticulation, much less vowel epenthesis should obtain in the new experiment. If, in contrast, the results were a true phonologically-driven phoneme restoration, the same results should obtain.

Method

The method was identical to the one employed in Experiment 1.

Materials: The same sets of u- and non-u-triplets as in the previous experiments were used. The stimuli were recorded by a native speaker of French who imitated a native speaker of Japanese as follows: The Japanese speaker first read a given triplet, and the French speaker repeated, trying to imitate the segmental and suprasegmental features of Japanese except in the cluster condition, where the consonant cluster was pronounced as such. Each triplet was recorded three times, and the best tokens were selected by a French and Japanese listener. The stimuli were then digitally recorded (16kHz).

Procedure: The procedures for the phonetic transcription task and lexical decision task were the same as in the previous experiment.

Participants: Fifteen native speakers of Japanese, recruited in Paris, participated in the lexical decision experiment. Seven of them also performed the phonetic transcription task.

Results and Discussion

Overall, 5 triplets in the non-u-Set and 5 triplets in the u-Set were removed because the words or non words yielded more than 40% errors or are incorrectly transcribed by more than half of the participants. This leaved 14 triplets in the u-Set and 14 triplets in the non-u-Set (the items are listed in the Appendix). Two participants who made more than 50% errors on the non-words (both for distractors and test items) were removed from subsequent analysis.

First, the phonetic transcription results for the consonant cluster items were analyzed separately for items in the u-Set and items in the non-u-Set. In the u-Set, sixty-five percent of the transcriptions of the consonant cluster items contained an epenthetic /u/ inserted between the consonants. Only 1% of the responses contained a different vowel (/o/ instead of /u/ in one item), and in 3% the second consonant was deleted. The remaining 30% of the transcriptions reproduced the consonant cluster. In the non-u-Set, fifty-nine percent of the transcriptions contained an epenthetic /u/, 9% a different vowel (the vowel /i/), and 31% reproduced the consonant cluster response. The /i/ response arose in two items: *rekSi* and *rikSi*. It is interesting to note that the /i/ responses turns these two items into real Japanese words (*rekiSi* and *rikiSi*, respectively). Note also that in these items, 21% of the responses were still /u/ insertions. The difference in /u/ insertion rates between the u-Set and the non-u-Set did not reach significance ($X^2(1)=0.52$, $p>0.10$).

Second, the lexical decision data were analyzed in terms of percent 'word' response. The data are displayed in Table 3. As one can see, words were labeled as words (93% of the cases), and non-words were not (8% of the cases). The results of the cluster items depended on the type of set. An average 70% of the cluster items in the u-Set were labeled as words, a score significantly different from 50% ($p<.007$). Planned contrasts showed that this score was significantly different from the score with the control non-words ($F1(1,12)=103$, $p<.001$; $F2(1,13)=108$, $p<.001$) and with the control words ($F1(1,12)=12$, $p<0.003$; $F2(1,13)=13$, $p<0.001$). In contrast, the cluster items in the non-u-Set were classified as words in only 18.7% on average (significantly below 50%, $p<.001$). This score was significantly different from the score with the control words ($F1(1,12)=341$, $p<.001$; $F2(1,13)=109$, $p<.001$), and from the control non-words although the difference was significant only for the participants analysis ($F1(1,12)=11.7$, $p<.005$; $F2(1,13)=1.7$, $p>.1$).

Table 3: percent "word" response to cluster and control items in Experiment 2.

	Non-words	Words	Clusters
u-Set			
Percent “ word ” response	6.0%	91.8%	70.4%
Example	*sokado*	*sokudo*	*sokdo*
non-u-Set			
Percent “ word ” response	10.4%	96.1%	18.7%
Example	*mikudo*	*mikado*	*mikdo*

Third, we analyzed the RTs for the responses in the dominant category, that is, responses as 'word' for cluster items in the u-Set and responses as 'non-word' for items in the non-u-Set. These responses are shown in Table 4, together with the RTs for the control words and non-words. An Anova revealed that the control words yielded significantly faster latencies than control non-words ($F1(1,12)=16$, $p<.002$; $F2(1,13)=91$, $p<.001$). Cluster items in the u-Set did not differ significantly from control words (both $Fs<1$), but yielded significantly faster RTs than the control non-words in the participants ' analysis ($F1(1,12)=4.7$, $p<.05$; $F2(1,13)=2.7$, $p>.1$). In contrast, cluster items in the non-u-Set yielded slower RTs than the control words ($F1(1,12)=28.0$, $p<.001$; $F2(1,13)=56.5$, $p<.001$), but did not differ significantly from the control non-words ($F1(1,12)=1.9$, $p>.1$; $F2(1,13)=3.0$, $p>.1$). In other words, the latencies to respond to cluster items were again similar to the latencies for the items in the lexical category to which they are assimilated.

Table 4: Reaction times (ms), standard error, and percent error to cluster and control items in Experiment 2.

	Non-words			Words			Clusters		
	RT	SE	Err	RT	SE	Err	RT	SE	Err
u-Set	1231	78	6.0%	1055	40	8.2%	1084	37	29.6%
non-u-Set	1241	64	10.4%	949	34	3.8%	1323	86	18.7%

In the above analysis of the transcription task, we found that the stimuli recorded by the French speaker produced 10% more cluster responses than in the previous experiment (a significant difference, $p<.001$). This suggests that the /u/ epenthesis effect can be enhanced when coarticulation cues remain in the adjacent consonants. Yet, the overall /u/ response rate was still very high (62%), a rate quite similar to that reported in Dupoux et al. (1999). In addition, two items in the non-u-Set yielded an identification of /i/ instead of /u/ (namely, *rekSi* and *rikSi*). A post-hoc analysis showed that the lexical decision responses for these two items were 77% and 54% "word" responses, respectively, whereas all the other items in the non-u-Set yielded a majority of non-word responses. Could it be that these two items reflect a lexical influence on the epenthesis effect? If so, the insertion of /i/ is due to the presence of the Japanese words *rekiSi* and *rikiSi*. However, none of the other items in the non-u-Set behaved similarly. Despite the existence of lexical items with non-u vowels between the medial consonant, all the other items elicited the perception of an /u/ and a non-word response. An alternative interpretation for *rekSi* and *rikSi* is that there are cases in Japanese of /i/ epenthesis. Shinohara (1997) discussed some of these cases, and while she argued that /i/ epenthesis may not be a productive phenomenon, it is worthwhile noting that the majority of the existing cases of loan words with /i/ insertion arise in the context of voiceless stop-fricative clusters (i.e. *textile* -> /tek<u>i</u>sutairo/), hence the same context as in *rikSi*. Interestingly, there were 5 items in the non-u-Set where the insertion of /i/ should have yielded a word. None of these items had a voiceless stop-fricative cluster, and none of them gave rise to the perception of an /i/. This suggests that the insertion of /i/ in *rekSi* and *rikSi* may also have a prelexical origin. Further research should uncover whether these cases are bona fide prelexical epenthesis effects or lexical influences.

General Discussion

We created Japanese non-words that contained illegal consonant clusters, and manipulated their lexical neighborood. The stimuli were either digitally edited utterances produced by a native Japanese talker (Experiment 1), or natural utterances produced by a native speaker of a language

that allows consonant clusters (Experiment 2). In both cases, and irrespective of the lexical status of the perceptual outcome, Japanese listeners reported hearing a vowel /u/ between the consonants. For example, they transcribe *sokdo* as `sokudo' (a real Japanese word), and *mikdo* as `mikudo' (a nonword), taking no account of the existence of a lexical neighbor with a different vowel (*mikado*). In a speeded lexical decision task, participants classified the stimuli like *sokdo* as real words, and the stimuli like *mikdo* as non-words. Moreover, lexical decision times with the stimuli containing the illegal consonant cluster were as fast as with the stimuli with an original vowel /u/.

These results allow us to reject a purely lexical account of vowel epenthesis, and support the interpretation of epenthesis as a prelexical process. Interestingly enough, our data do not reveal any influence of lexical information at all. Lexical influences should have produced more epenthesis in items from the u-Set than in items from the non-u-Set, but this was not the case : the rates of /u/ epenthesis were not greater in items with /u/ neighbors (*sokdo*) than in items with non-/u/ neighbors (*mikdo*).

This interpretation, however, is made more complex because of *high vowel devoicing* in Japanese. In several dialects, high vowels (i.e., /i/ and /u/) are often devoiced between voiceless obstruents, or utterance finally after voiceless obstruents. Several factors seem to influence the likelyhood of devoicing: speech rate, pitch accent placement, consonant type (stop versus fricative), as well as sociological effects (see Kondo, 1997; Varden 1998). Similarily, the precise phonetic implementation of devoicing seems to be non-homogeneous, ranging from something close to a wispered vowel (with formant structure), to fricative noise (without formant structure), going even to pure deletion (Tsuchida, 1997 ; Varden 1998). Abstracting away from these details, Japanese listeners may encounter words like *sokudo* in phonetic forms that are similar to /sokdo/. Could it be that this phenomenon accounts for our results?

Two arguments can be offered to counter this view. First, a post-hoc analysis revealed that the eight items containing voiceless obstruent clusters did not yield more epenthesis of /u/ than the others items (Experiment 1: 78% vs. 76%, respectively, $p>0.10$; Experiment 2: 65% vs. 61%, $p>0.10$). Second, the vowel /i/ devoices as readily as the vowel /u/. Nevertheless, /i/ epenthesis is quite infrequent, both in our experimental results and in the pattern of foreign word borrowings (see for instance Shinohara, 1997). If high vowel devoicing were the sole basis for the vowel epenthesis effect in perception, we should have found equal amounts of /u/ and /i/ epenthesis.

Nevertheless, let us assume, counterfactually, that high vowel devoicing applies only to /u/ and within all the consonant clusters used in our material. One may then suppose that items in the u-set (like *sokudo*) have two possible phonetic forms (e.g. /sokdo/ and /sokudo/) stored in the mental lexicon of Japanese. It is certainly not the case, however, that /mikdo/ is stored as a variant *mikado* (because /a/ does not devoice). Therefore, a lexical feedback hypothesis cannot explain the results of the transcription task where participants transcribed /mikdo/ as *mikudo* as readily as they transcribed /sokdo/ as *sokudo*. In brief, the existence of vowel devoicing in Japanese does not alter our conclusion that it is the phonological context, not the lexical context, that triggers the illusory perception of /u/.

It turns out that the transcription task provides a critical piece of evidence against lexical involvement. This task could be criticized as being rather off-line and metalinguistic ; without denying that, we note that Dupoux et al (1999) found very similar results with off-line transcription and more on-line tasks such as speeded ABX discrimination. Furthermore, Dehaene-Lambertz,

Dupoux & Gout (2000), using evoked potentials in an oddball paradigm, found that Japanese participants showed no mismatch negativity response (MMN) to a change from *ebzo* to *ebuzo*, whereas French participants show an early MMN reponse to the same change (140-280msec after the offset of /b/). This further supports our hypothesis that phonotactic constraints modulate the brain response at a very early processing stage, prior to lexical access.

Overall, our results suggest that models of speech perception must take phonotactic information into consideration. Questions still remain about what could the mechanisms be. We can foresee two alternative accounts.

The first account postulates a rule-based process that inspects the phonetic representation of utterances (Church, 1987; Frazier, 1987) and inserts a vowel whenever there is an illegal sequence of consonants (as in (1)). The rules would look like those typically proposed in phonological accounts of loan word adaptations.

(1) $\varnothing \rightarrow$ u / C_C

A problem with this account is that there is no independant evidence for this rule in the phonology of standard Japanese. It is thus not clear how it would be acquired by children given that they are not presented with pairs of foreign words together with their Japanese adaptations.

The second account postulates a pattern matching process that assimilates foreign sounds to the phonetically closest prototypes of the native language. It is similar to the Perceptual Assimilation Model (Best, 1994, see also Takagi & Mann, 1994), with the important modification that whole syllables rather than individual segments, are used as prototypes. Phonetic pattern matching explains why /u/ is used as an epenthetic vowel, instead of /i/ or /a/. Indeed, /u/ tends to be the shortest vowel (Beckman, 1982) and the one with the most formant variability (Keating and Huffman, 1984). Hence, /ku/ is the closest syllabic prototype that matches the segment /k/ in /mikdo/.

Of course, cross-linguistic research is needed to further work out and test these alternative accounts and also to understand how and when phonotactic constraints are acquired by young infants.

References

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

Beckman, M. (1982). Segment duration and the « mora » in Japanese, *Phonetica*, **39**, 113-135.

Best, C. T. (1994). The emergence of native-language phonological influence in infants: A perceptual assimilation model. In J. Goodman & H. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (p. 167-224). Cambridge, MA: MIT Press.

Dehaene-Lambertz, G., Dupoux, E., and Gout, A. (2000). Electrophysiological correlates of phonological processing: a cross-linguistic study, *Journal of Cognitive Neuroscience*, **12**(4), 635-647.

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**(6), 1568-1578.

Dupoux, E., Fushimi, T., Kakehi, K., & Mehler, J. (1999). Prelexical locus of an illusory vowel effect in Japanese. Proceedings of Eurospeech '99, 7th European Conference on Speech Communication and Technology.

Ganong, W. F. (1980). Phonetic categorisation in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance*, **6,** 110-125.

Frauenfelder, U. H. and Segui, J. and Dijkstra, T. (1990). Lexical effects in phonemic processing: Facilitatory or inhibitory ? *Journal of Experimental Psychology: Human Perception and Performance*, **16,** 77-91.

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

Hallé, 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**(2), 592-608.

Keating, P.A., & Huffman, M.K. (1984). Vowel variation in Japanese, *Phonetica*, **41**, 191-207.

Kondo, M. (1997). *Mechanisms of vowel devoicing in Japanese.* Ph. D. dissertation, University of Edinburgh.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Revie*w, **74**, 431-461.

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

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

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

Miller, J. L., & Liberman, A. M. (1979). Some effects of later-occurring information on the perception of stop consonants and semivowels. *Perception & Psychophysic*s, **25**, 457-465.

Pitt, M.A., & McQueen, J.M. (1998). Is compensation for coarticulation mediated by the lexicon? *Journal of Memory and Language*, **39**, 347-370.

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

Samuel, A.G. (1981a) Phonemic restauration: Insights from a new methodology. *Journal of Experimental Psychology: General*, **110,** 474-494.

Samuel, A. G. (1981b). The role of bottom-up confirmation in the phonemic restoration illusion. *Journal of Experimental Psychology: Human Perception and Performance*, **7,** 1131-1142.

Samuel, A. G. (1987). Lexical uniqueness effects on phonemic restoration. *Journal of Memory and Language*, **26,** 36-56.

Shinohara, S. (1997). *Analyse phonologique de l'adaptation japonaise de mots étrangers.* Unpublished doctoral dissertation, Université de la Sorbonne Nouvelle.

Takagi, N., & Mann, V. (1994). A peceptual basis for the systematic phonological correspondences between Japanese loan words and their English source words. *Journal of Phonetics*, **22**, 343-336.

Tsuchida, A. (1997). *The phonetics and phonology of japanese vowel devoicing.* Ph.D. dissertation. Cornell University.

Varden, J.K. (1998). *On high vowel devoicing in standard modern Japanese: implications for current phonological theory.* Unpublished Ph.D. dissertation. University of Washington.

Warren, R.M. (1984). Perceptual restoration of obliterated sounds. Psychological Bulletin, 96, 371-383.

Whalen, D. H. (1989). Vowel and consonant judgements are not independent when cued by the same information. *Perception & Psychophysic*s, **46**(3), 284-292.

Appendix

Items in the u-Set

bakro-bakuro-bakaro, harka-haruka-haraka, hirma-hiruma-hiroma,
huksi-hukusi-hukisi, hikme-hikume-hikime, kaksa-kakusa-kakasa,
kokdo-kokudo-kokado, kokmu-kokumu-kokamu, kokso-kokuso-kokaso,
kokti-kokuti-kokiti, komgi-komugi-komigi, magro-maguro-magaro,
samsa-samusa-samosa, sokdo-sokudo-sokido, soksi-sokusi-sokasi,
tuksi-tukusi-tukesi, yakba-yakuba-yakaba, yakza-yakuza-yakaza,
yursi-yurusi-yurisi

Items in the non-u-Set

arsi-arasi-arusi, hiktu-hiketu-hikutu, huksa-hukasa-hukusa,
kikme-kikime-kikume, kormo-koromo-korumo, mikdo-mikado-mikudo,
namda-namida-namuda, nikbi-nikibi-nikubi, omsa-omosa-omusa,
reksi-rekisi-rekusi, riksi-rikisi-rikusi, sakba-sakaba-sakuba,
sekri-sekiri-sekuri, sikti-sikiti-sikuti, takra-takara-takura,
taksa-takasa-takusa, tegru-tegaru-teguru, wakme-wakame-wakume,
waksa-wakasa-wakusa