A Destressing "Deafness" in French?

E. Dupoux
EHESS–CNRS, Paris, France

C. Pallier
Rutgers University Center for Cognitive Science, NJ

N. Sebastian
Dpt Psicologica Basica, Universitat de Barcelona

J. Mehler
EHESS–CNRS, Paris, France

— Reprinted by the authors —
published in *Journal of Memory and Language*, **36**, 406-421, 1997.

Spanish but not French uses accent to contrast between words (e.g., *tópo* vs *topó*). Two populations of subjects were tested on the same materials to determine whether this difference has an impact on the perceptual capacities of listeners. In Experiment 1, using an ABX paradigm, we found that French subjects had significantly more difficulties than Spanish subjects to perform an ABX classification task based on accent. In Experiment 2, we found that Spanish subjects were unable to ignore irrelevant differences in accent in a phoneme-based ABX task, whereas French subjects had no difficulty at all. In Experiment 3, we replicated the basic French finding, and found that Spanish subjects benefited from redundant accent information even when phonemic information alone was sufficient to perform the task. In our final experiment, we showed that French subjects can be made to respond to the acoustic correlates of accent; therefore their difficulty in Experiment 1 seem to be located at the level of short-term memory. The implications of these findings for language-specific processing and acquisition are discussed.

The native speaker's difficulties with some nonnative segmental contrasts is well documented, (e.g., Polka & Werker, 1994). Much less is known about the way in which rhythmical and structural properties of the native language affect encoding of foreign words. This study reports four experiments to assess the claim that nonsegmental properties of the native language greatly affect the way in which unfamiliar items are processed and represented.

Learning a language requires, inter alia, acquiring knowledge of the language's sound pattern. Psycholinguists have mostly assumed that learning the sound pattern of a native language essentially involves learning the segments of the language. This assumption is evident in books and articles on the subject, e.g., Carroll (1960) for an early formulation of this view and Miller and Jusczyk (1990) for a later one. There is no question that segments are an essential ingredient of the sound pattern of language though many other properties of speech are also important. Foreign accents reflect segmental difficulties but also supra-segmental ones. For instance, informal observations suggest that French speakers who acquire English after puberty produce segments that are not prototypical of English, and prosodic output is usually deficient as well. It is not uncommon to hear speakers of French who fail to reduce vowels or who make the wrong vowel prominent. Many mistakes are, in fact, due to the inability of native speakers of French to place stress in the correct place. How is one to explain these consistent mistakes? Why is it that even proficient French speakers of English produce anomalous prosodic structure?

In this paper we try to show that such informal observations as the ones alluded to above reflect the essential processing routines and representation structures elaborated specifically for one's native language. Cutler and Mehler (1993) and also Christophe and Dupoux (in press) argue that prosody is essential during first language acquisition. When the child first learns a language, he or she extracts its rhythmical-periodical properties. Those properties facilitate language acquisition and trigger adjustments on perception and production routines that allow efficient language-specific processing. Moreover, in adults, these adjusted perception and production routines are no longer very flexible and disrupt in certain ways the processing of foreign languages. This helps us to explain how natural languages are processed differently by native and by proficient foreign speakers of those languages.

These considerations lead us to conjecture that speakers of different languages become sensitized, not only to different phonetic contrasts, but also to different prosodic properties. If this hypothesis is correct, determining which rhythmical and prosodic structures create problems for foreign speakers should help us to understand how the native language is acquired and how foreign languages are subsequently learned.

This research was supported by grants from the Human Frontier Science Program, the Human Capital program, the Franco-Spanish Exchange Program for 1993-1995, the Fyssen Foundation and DRET. We thank Jennifer Glos, Xavier Hardy and Olivier Crouzet for help in stimulus construction and data gathering. We thank Stanislas Dehaene and Brit van Ooijen for very useful comments on earlier versions of this paper. Mailing address: Emmanuel Dupoux, 54 Bd Raspail, F-75270 Paris Cedex 06, France. http://www.ehess.fr/centres/lscp/persons/dupoux

Consider the accent differences between Spanish and French. In Spanish, the default position for accent[1]in multisyllabic words is on the penultimate syllable. However, many words have accent in other positions. Indeed, there are minimal pairs that have exactly the same segments but differ in meaning because of the accent location. For example, *bébe* ("drink" present tense) vs *bebé* ("baby"), or *tópo* ("mole") vs *topó* ("met") drawn from Spanish, exemplify this property (see Navarro Tomas, 1946). In French, however, accent does not carry lexical information; there is no such contrast as *bébe* versus *bebé*. In this language, accent is either described as being not specified lexically, or as falling on the last full vowel of content words (Dell & Vergnaud, 1984).

Thus, speakers of Spanish have to process and represent accent to identify the lexical item(s) intended by the speaker. Speakers of French, however, do not need to process accent, at least not in the same way. As far as lexical identification is concerned, only segmental content matters, and accent information could be completely left unspecified in the French lexicon.

The following experiments investigate this hypothesis by testing the ability of French and Spanish speakers either to detect stress differences (Exp. 1 and Exp. 3) or to ignore them (Exp. 2). The final experiment addresses the issue of the level at which French/Spanish differences in performance may arise.

Experiment 1

Our first study tested native speakers of French and of Spanish with an ABX discrimination task involving an accent contrast. Subjects were presented with three items that varied only in accent location and had to press a button to indicate whether the last item corresponded to the first or to the second item. In this experiment, we tested two accent contrasts on trisyllabic CVCVCV items: *bópelo* vs *bopélo* (1^{st} vs 2^{nd} syllable with accent), and *bopélo* vs *bopeló* (2^{nd} vs 3^{rd} syllable with accent).

Note that these accent patterns are all possible in Spanish, although the penultimalte pattern (*bopélo*) may be said to be the unmarked case. Since our two experimental contrasts involve items with accent in the penultimate syllable, there is no reason to expect that, for Spanish subjects, 1^{st} vs 2^{nd} syllable accent would be easier or harder than 2^{nd} vs 3^{rd}. In contrast, French is a language with accent on the last full vowel of content words. One might expect accent-final items to be perceived as natural, whereas items with nonfinal accents should be perceived as foreign-sounding. In this case, 2^{nd} vs 3^{rd} syllable accent involves a difference in legality, and should be easier for the French than 1^{st} vs 2^{nd} (which are both illegal). This is not, of course, the only possibility. French subjects might assimilate all accent patterns to one, or simply, may not represent accent at all. In this case, both contrasts would be equally difficult.

French and Spanish differ in more than placement of accent. First, French has been described as a syllable-timed language. Spanish is also a syllable-timed language, but it also involves an alternating pattern of prominence peaks that reflects a higher order prosodic unit: the foot. Nonetheless, previous psycholinguistic investigations found that the two languages processed syllabic structures in very similar ways (Bradley, Sánchez-Casas, & García-Albea, 1993; Mehler, Dommergues, Frauenfelder, & Segui, 1981; Pallier, Sebastian-Gallés, Felguera, Christophe, & Mehler, 1993; Sebastian-Galles, Dupoux, Segui, & Mehler, 1992). Second, French has 14 vowels, and Spanish only five. Similarily, Spanish and French differ in the consonantal inventory and in allowed consonant clusters. In order to cope with these differences, our materials only used a subset of segments and syllable structure legal in both languages.

The following experiments use the exact same stimuli for French and Spanish hearers. However, since recording either French or Spanish speakers might have too obviously advantaged one population over the other because of the fine difference in the phonetic realization of phonemes in the two languages, we decided to record the stimuli using speakers of a third language: Dutch. Dutch is a stress-timed language, which allows accent to appear in difference places in words. The phoneme subset that we selected for French and Spanish contains phonemes that are all pronouncable in Dutch. In an informal pretest, Dutch realization of these phonemes were judged to be slightly foreign-sounding but highly acceptable both to French and Spanish listeners.

Method

Materials. Twelve CVCVCV triplets of the form *(bópelo, bopélo, bopeló)* were constructed (see Appendix A). Each member of a triplet had the same phonemic content, and the main accent falling on the first, second or third syllable. None of the vowels were reduced, and they belonged to the set [e],[i],[o],[a],[u]. All items were nonwords both in French and Spanish.

The 12 triplets were recorded by three speakers of Dutch (two females, one male) and checked by a phonetician. They were instructed to mark word accent clearly on the first, second or third syllable, while keeping the other vowels unreduced. Each item was digitized at 16KHz at 16bits (on an OROS AU22 A/D Board), digitally edited, and stored on a computer disk.

From the recorded experimental triplets, 96 trials were constructed. Each trial consisted of three stimuli: A, B and X, the first two said by the two female speakers and the third by the male speaker. The A and B items always had the same segmental content, but differed in accent. Two contrasts were tested: 1^{st} vs 2^{nd} and 2^{nd} vs 3^{rd}. For a given contrast, A and B received the two accents in the two possible orders: (e.g., *bópelo bopélo*, or *bopélo bópelo*). This resulted in 4 different A-B combinations for each experimental triplet. The X item always had the same segmental content as A or B, but on half of the trials, X had the same accent as A and on the other half,

[1] In this paper we use the term "accent" to refer to phonetic/perceptual prominence in a string of syllables. We stay neutral as to whether such prominence is correlated or not with the notion of metrical stress as used in phonology.

it had the same accent as B. The overall design was 2 x 2 x 2: Accent-Contrast x Accent-Order x X-identity.

The 96 experimental trials were split into two blocks (with each condition and stimuli represented equally often in each block). An extra practice block of 10 trials was constructed that contained the same conditions but with different items from the experimental blocks.

Procedure. Each experimental trial presented the three stimuli (A, B and X), separated by an interval of 500 ms. Subjects were instructed to listen to words in a foreign language. They were told that the first two stimuli were different, and that the third one was identical to the first or to the second. They were required to press a button on their left or on their right to indicate whether X was identical to A or to B respectively. Subjects were given a 4000 ms deadline to respond. The next trial started 1000 ms after each response, or after the deadline.

In the ten practice trials, subjects received feedback as to whether their response was correct or not. Feedback consisted in the word “Correct”, “Incorrect”, or the string “The response is A/B” when subjects failed to respond before the deadline. It was displayed for 1000 ms and then erased from the screen. For incorrect responses, the same trial was presented again, until the response was correct. In the two experimental blocks of 48 trials, no feedback was presented. The blocks were separately randomized for each subject. A short pause separated the two experimental blocks. Responses were recorded and reaction time was measured from the onset of the X stimuli by the EXPE software package (Pallier, Dupoux & Jeannin, 1997).

Subjects. Sixteen French and 16 Spanish subjects participated in the experiment. The French subjects were students recruited in Paris. There were 10 men and 6 women, 14 right handers and 2 left handers in the group. Their median age was 25. The Spanish subjects were students at the University of Barcelona who received course credit for the experiment. There were 4 men and 12 women, 13 right handers and 3 left handers. The median age was 20. None of the subjects understood the other language nor did they understand Dutch.

Results

One item (*vasuma*) generated many errors in this experiment and in the next, and was therefore removed from subsequent analyses.

Mean reaction times and error rates broken down by Language (French or Spanish), Accent Contrast (1^{st} vs 2^{nd} or 2^{nd} vs 3^{rd}) and Response Type (A-Response or B-Response) are displayed in Table 1. These two dependent variables were subjected respectively to two ANOVAs, one with subjects and one with items as random variables. Language was a between-subject factor, Response Type and Accent Contrast were within-subject factors, and all three factors were within-item.

In the errors analyses, French subjects made significantly more errors than Spanish subjects (19% vs 4%, $F1(1,30)=17$, $p<.001$; $F2(1,10)=158$, $p<.001$). No effect of Accent Contrast was found (Fs<1), and this factor did not interact with Language. A-Responses yielded more errors than B-Responses (13% vs 10%, $F1(1,30)=4.0$, $p<.05$; $F2(1,10)=20$, $p<.001$), but this effect was mostly due to French subjects, and mostly for the accent 2^{nd} versus 3^{rd} contrast. The three-way interaction between Response Type, Language, and Accent-Type was only significant in the subjects analysis ($F1(1,30)=4.9$, $p<.03$; $F2(1,10)=4.3$, $p<.1$), and the two-way interactions between Response Type and Language, and Response Type and Accent Type were only significant in the items analysis ($p<.05$). No other interaction was significant.

The reaction time analyses revealed that French subjects were slower than Spanish subjects, although the effect was only significant in the items analysis (94ms, $F1(1,30)=1.8$, $p>.1$; $F2(1,10)=77$, $p<.001$). There were no effects of Accent Type. Finally, A-Responses were slower than B-Responses (37ms, $F1(1,30)=7.5$, $p<.01$; $F2(1,10)=9.6$, $p<.01$) and there was no significant interaction with the other factors.

Discussion

French and Spanish subjects were asked to perform an ABX task based on two accent contrasts. Spanish Subjects made very few mistakes but French subjects had difficulty performing the task. This difference can be seen mostly in the error rate, although there was also a trend in reaction times.

Interestingly, there was no noticeable difference between the accent 1^{st} vs 2^{nd} contrast and the 2^{nd} vs 3^{rd} contrast, either for the French or for the Spanish subjects. For the Spanish, both comparisons involve accent-second items (the unmarked case) and two other accent patterns that are possible in this language. It is, therefore, not very surprising that the two contrasts were equally easy for the Spanish population. However, one might expect the French listeners to find the accent-final items more prototypical than the other two accent patterns. Yet there was no difference between 1^{st} vs 2^{nd} and 2^{nd} vs 3^{rd} contrasts, suggesting that the difficulty that French subjects have with accent is general.

Experiment 1 demonstrated that Spanish subjects are overall more efficient than French subjects in discriminating accent contrasts. However, it might be that the Dutch materials contained vowels or consonants that were more prototypical for the Spanish than for the French subjects. In order to control for these potential effects, it would be useful to reverse the previous pattern of results, that is, to find a situation that should be easy for the French to perform and difficult for the Spanish.

Experiment 2

In this experiment, we asked subjects to *ignore* accent, and to respond on the basis of segmental information only. Accent was varied orthogonally with phonemic structure so that relying on accent information would lead to chance performance. The reasoning is that French subjects who have difficulty hearing accent in the first place, would have no difficulty in ignoring it. In contrast, for the Spanish speakers, accent plays an important role in word identification. Hence, re-

Table 1
Mean reaction time, standard error, and error rate to ABX judgments based on two accent contrasts (Experiment 1): accent 1^{st} versus 2^{nd}, and accent 2^{nd} vs 3^{rd}.

Language	Response Type	Mean	(se)	Err	Mean	(se)	Err	Mean	(se)	Err
		accent 1^{st} vs 2^{nd} *(bópelo bopélo)*			accent 2^{nd} vs 3^{rd} *(bopélo bopeló)*			Mean		
FRENCH										
	X=A	1230	(57)	21%	1261	(54)	23%	1245	(39)	22%
	X=B	1226	(61)	19%	1228	(49)	14%	1227	(39)	16%
	Mean	1228	(41)	20%	1244	(36)	18%	1236	(27)	19%
SPANISH										
	X=A	1163	(52)	5.1%	1178	(49)	3.9%	1171	(35)	4.5%
	X=B	1120	(41)	3.1%	1107	(39)	3.9%	1114	(28)	3.5%
	Mean	1142	(33)	4.1%	1143	(31)	3.9%	1142	(23)	4.0%

quiring Spanish subjects to ignore accent should be difficult. This experiment used the same materials as Experiment 1, and the same subjects population.

Method

Materials. Twelve CVCVCV quadruplets of the form *(bópelo, bopélo, sópelo, sopélo)* were constructed (see Appendix B). They resulted from the crossing of two contrasts: an accent contrast (accent in the first or accent in the second syllable) and a phoneme contrast (a single phoneme change in one of the syllables). The segmental content of the first two members of the quadruplets was the same as in Experiment 1. The phoneme change occurred equally often in the first, second or third syllable. All items were nonwords both in French and in Spanish.

The 12 quadruplets were recorded by the same three speakers of Dutch as in Experiment 1 and were digitized in the same way. From the recorded experimental quadruplets, 192 experimental trials were constructed. Each experimental trial consisted of three stimuli: A, B and X, the first two spoken by the two females and the third by the male. The A and B items always differed in one phoneme *and* in placement of accent. In half of the trials, A had accent first, and in the other half it had accent second. In half of the trials, A was a stimulus from the Experiment 1 set, and in the other half it was from the 'phoneme-change' set. This resulted in 4 different A-B combinations for each experimental quadruplet (e.g. *bópelo-sopélo*, *bopélo-sópelo*, *sópelo-bopélo*, and *sopélo-bópelo*). X always had the same segmental content as A or B, and had the same accent as A or B. This resulted in 4 different cases that we will consider as the crossing of two factors: whether X had the segments of A or B (Response Type), and whether X had an accent congruent with its segments (Congruency). As far as the last factor is concerned, there are two cases: in one condition, X was identical to A or B both in terms of accent *and* phoneme. This was called the Accent Congruent Condition. In the other, X had the segments of one stimulus, and the accent of the other. This was the Accent Incongruent Condition. The overall design was thus 4 x 2 x 2: A-B combinations x Response Type x Congruency.

The 192 experimental trials were split into four blocks of 48 trials each (A, B, C and D). Each condition and stimulus was represented equally often in each block. An extra practice block of 10 trials was constructed that contained the same conditions but with different items from the experimental blocks.

Procedure. Stimulus presentation and feedback were as in Experiment 1. Subjects were instructed to listen to words in a foreign language in which accent contrasts were irrelevant. They were told that the first two stimuli were words that differed by one speech sound, and that the third one was identical to the first or to the second in terms of their speech sounds. They were required to press a button on their left or on their right to indicate whether X was identical to A or to B respectively.

After the instructions, subjects performed the ten practice trials, receiving feedback as to whether their response was correct or not. As in Experiment 1, subjects were not allowed to pass on to the experimental session unless they had correctly responded to all the items in the practice session. After the practice trials had been completed, subjects received two experimental blocks of 48 trials (half of the subjects received Blocks A and B, and half received Blocks C and D), with no feedback. The blocks were separately randomized for each subject. A short pause separated the two blocks. Response deadline, intertrial interval and reaction time measurement were as in Experiment 1.

Subjects. Sixteen French and sixteen Spanish subjects participated in the experiment. They were drawn from the same population as in Experiment 1. The French subjects included 11 men and 5 women, 13 right handers and 3 left handers. Their median age was 22. The Spanish subjects were 2 men and 14 women, 15 right handers and 1 left hander. The median age was 21. None of the subjects understood the other language.

Table 2
Mean reaction time, standard error, and error rate to ABX judgments based on a phonemic difference as a function of whether accent was congruent or incongruent with phonemic information (Experiment 2).

Language	Response Type	Mean	(se)	Err	Mean	(se)	Err	Mean	(se)	Err
		Congruent (*bópelo sopélo bópelo*)			Incongruent (*bópelo sopélo bopélo*)			Mean		
FRENCH										
	X=A	1072	(69)	6.5%	1062	(67)	8.5%	1067	(47)	7.5%
	X=B	1016	(58)	5.1%	1056	(57)	6.2%	1036	(40)	5.7%
	Mean	1044	(44)	5.8%	1059	(43)	7.4%	1051	(31)	6.6%
SPANISH										
	X=A	1258	(82)	13.9%	1252	(64)	17.9%	1255	(51)	15.9%
	X=B	1173	(57)	3.7%	1233	(61)	11.1%	1203	(42)	7.4%
	Mean	1216	(50)	8.8%	1243	(44)	14.5%	1229	(33)	11.6%

Results

As in Experiment 1, the item *vasuma* generated many errors and was removed from subsequent analyses.

Reaction times and error rates were subjected to subjects and items ANOVAs (see means in Table 2). Language (French or Spanish) was a between-subject factor, Response Type (A-Response or B-Response), and Congruency (Congruent vs Incongruent) were within-subject factors. All three factors were between items.

The error analyses revealed an effect of Language, with Spanish subjects making more errors than French subjects (11.6% vs 6.6%, $F1(1,30)=4.0$, $p<.05$; $F2(1,10)=17$, $p<.002$). The Incongruent Condition generated more errors than the Congruent Condition (10.9% vs 7.3%, $F1(1,30)=7.4$, $p<.01$; $F2(1,10)=8.7$, $p<.01$), but this was mostly due to the Spanish subjects. However, the interaction between Language and Congruency was only significant in the items analysis ($F1(1,30)=2.4$, $p>.1$; $F2(1,10)=6.8$ $p<.03$). A-responses generated more errors than B-responses (11.7% vs 6.5%, $F1(1,30)=11.6$, $p<.002$; $F2(1,10)=11.5$, $p<.007$), although this was mostly due to Spanish subjects. The interaction between Response Type and Language was significant in both analyses ($F1(1,30)=4.8$, $p<.04$; $F2(1,10)=10.0$, $p<.01$). No other interaction reached significance.

The reaction time analyses showed that Spanish subjects had significantly longer reaction times than French subjects (178ms, $F1(1,30)=4.1$, $p<.05$; $F2(1,10)=59$, $p<.001$). Congruent responses did not differ from Incongruent responses ($p>.1$). A-responses yielded slower reaction times than B-responses (41ms, $F1(1,30)=4.2$, $p<.05$; $F2(1,10)=6.2$, $p<.03$). There was an interaction between Congruency and Response type which was only significant in the items analysis ($F1(1,30)=3.8$, $.05<p<.1$; $F2(1,10)=5.7$, $p<.04$), and no other interaction was significant.

Discussion

We found that French subjects were both faster and more accurate than Spanish subjects, reversing the pattern observed in Experiment 1. This suggests that the French can concentrate on the segments and ignore irrelevant variations in accent. In contrast, Spanish subjects were slower and made significantly more errors, suggesting that they cannot ignore irrelevant accent variations. Although the Spanish listeners made more errors in the Incongruent than in the Congruent condition, the effect was not reflected in reaction times, suggesting that ignoring accent was difficult for the subjects and globally increased their reaction times.

Figure 1 shows the overall results of Experiments 1 and 2. An overall ANOVA reveals a significant interaction between tasks and Language for both Reaction Times ($p<.02$ in both items and subjects analyses) and errors ($p<.001$ in both items and subjects analyses), confirming that subjects of the two languages process phoneme and accent information in a very different way.

Experiment 3

In the previous two experiments, we compared French and Spanish subjects on ABX tasks that focused either on the accent or on the phoneme. We found that French listeners had difficulties responding on the basis of accent whereas the Spanish listeners had trouble ignoring accent. However, these differences all rest on comparisons between groups of subjects drawn from similar but different populations (students in Paris vs students in Barcelona). Another shortcoming of the previous experiments is that the differences found might be due to the way in which subjects understood the task, rather than to the way in which they perceived the stimuli. This is especially true in Experiment 2 where subjects were asked to ignore accent. French and Spanish subjects may differ in their understanding of the notion of accent.

In order to evaluate these possibilities, we carried out a further experiment comparing directly for each individual subject responses based on accent and responses based on phonemes, while keeping the other dimension constant. We predicted that unlike Spanish Subjects, French subjects would have greater difficulty in discriminating accent as com-

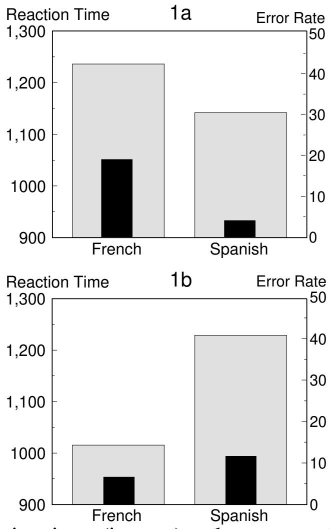

Figure 1. Reaction times (in grey) and error rates (in black) to ABX judgments in French and Spanish Subjects. Fig1a: ABX on accent only; phonemes fixed (Experiment 1). Fig1b: ABX on phonemes only; accent varied orthogonally (Experiment 2).

pared to phonemes. We also included a *redundant* condition in which both accent and phoneme information were associated with the same reponse. Because Spanish subjects represent both phonemic and accent information, they should be faster in the redundant condition than in either the phoneme condition or the accent condition. French subjects, in contrast, should respond in the redundant condition as if accent was unchanged. Because of the higher number of conditions in this experiment, and because we found no differences between 1^{st} vs 2^{nd} accent locations in the previous experiments, we decided to limit ourselves to just these two positions.

Method

Materials. Twelve new CVCVCV quadruplets of the form *(fídape, fidápe, lídape, lidápe)* were constructed (see Appendix C). They resulted from the crossing of two contrasts: an accent contrast (accent in the first or accent in the second syllable) and a phoneme contrast (a single phoneme change in one of the syllables). The phoneme change occurred equally often in the first, second, or third syllable. The vowel set was [a], [e], [i], [o], [u] and the consonant set [p], [t], [k], [b], [d], [g], [f], [s], [l], [m], [n] which are all common segments in Spanish, French (and Dutch). All items were nonwords, both in French and in Spanish. The 12 quadruplets were recorded by two speakers of Dutch (one male, one female), and were digitized as in Experiments 1 and 2.

From the recorded experimental quadruplets, 288 experimental trials were constructed. Each experimental trial consisted of three stimuli: A, B and X. A and B were spoken by the same female speaker and the X stimulus by the male speaker. A and B either differed by one phoneme only (*fídape, lídape*: Phoneme Condition) by accent only (*fídape, fidápe*: Accent Condition) or by both accent and phoneme (*fídape, lidápe* Redundant Condition). The trials were counterbalanced such that all members of a quadruplet (*fídape, fidápe, lídape, and lidápe*) appeared in both position A and B, for each of the above defined matching conditions. This resulted in 12 different A-B combinations for each experimental quadruplet. X was identical (both in accent and in phoneme) to either A or B. There were two response types (A-responses trials, and B-responses trials). The overall design was thus 3 x 4 x 2: Matching Condition x Quadruplet Counterbalancing x Response Type.

The 288 experimental trials were split into four blocks of 72 trials each (A, B, C and D). Each condition and stimulus were represented equally often in each block. A practice block of 10 trials was constructed that contained the same conditions but with different items as the experimental blocks.

Procedure. The scheduling of practice and experimental trials was as in Experiments 1 and 2. Subjects were instructed to listen to words in a foreign language. They were told that the first two stimuli were different words, and that the third was identical to the first or to the second. They were required to press a button on their left or on their right to indicate whether X was identical to A or to B, respectively.

In the ten practice trials, subjects received feedback as to whether their response was correct. For incorrect responses, the same trial was presented immediately again until the response was correct. After practice, subjects received two experimental blocks of 72 trials (half of the subjects received Block A and B, and half received Block C and D), with no feedback. The blocks were separately randomized for each subject. A short pause separated the two blocks. Responses were recorded and reaction times measured from the onset of the X stimuli by the EXPE software package.

Subjects. Twenty French and twenty Spanish subjects participated in the experiment. The French subjects were students from the Ecole Polytechnique, Palaiseau. There were 15 men and 1 woman, 11 right handers and 5 left handers. Their median age was 22. The Spanish subjects were psychology students at the University of Barcelona who received course credit for their participation. There were 3 men and 13 women, 15 right handers and 1 left hander, with a median age of 20. About half of the French subjects had learned Spanish as a third language (the second one being English) at the age of 14. None of the Spanish subjects had learned French. None of the subjects knew Dutch.

Results

Reaction times and error rates were subjected to subjects and items ANOVAs (see means in Table 3). Language (French or Spanish) was a between-subject factor, Response Type (A-Response or B-Response) and Matching Condition (Phoneme, Accent and Redundant) were within-subject factors. All three factors were within-item.

The two populations did not differ in their mean error rate (all ps$>$.1). However, there was a main effect of Matching Condition ($F1(2,76)=21$, $p<.001$; $F2(2,22)=17$, $p<.001$)

Table 3
Mean reaction time, standard error, and error rate to ABX judgments in a redundant, phoneme only and accent only condition (Experiment 3).

Lang	Response Type	Mean	(se)	Err	Mean	(se)	Err	Mean	(se)	Err
		Redundant *(fídape lidápe fídape)*			Phoneme *(fídape lídape fídape)*			Accent *(fídape fidápe fídape)*		
FRENCH										
	X=A	1005	(41)	2.7%	1019	(44)	2.1%	1234	(54)	13.7%
	X=B	1004	(47)	2.9%	1023	(39)	3.3%	1099	(51)	7.9%
	mean	1005	(31)	2.8%	1021	(29)	2.7%	1167	(38)	10.8%
SPANISH										
	X=A	958	(36)	7.9%	965	(31)	8.1%	1017	(36)	13.9%
	X=B	874	(30)	3.5%	965	(30)	3.1%	935	(32)	2.7%
	mean	916	(24)	5.7%	965	(21)	5.6%	976	(24)	8.3%

as well as an interaction with Language ($F1(2,76)=5.3$, $p<.007$; $F2(2,22)=7.0$, $p<.004$). The Accent Condition yielded more errors than the Phoneme Condition, especially for the French subjects (10.8% vs 2.7%, both $p<.001$) and to a lesser extent for the Spanish subjects (8.3% vs 5.6%, $p<.05$). The Phoneme and Redundant Conditions did not differ in terms of error rates. A-responses yielded more errors than B-responses, (13.8% vs 5.3%, $F1(1,38)=11.6$, $p<.002$; $F2(1,11)=16.2$, $p<.002$) and Response Type interacted both with Language ($F1(1,38)=4.9$, $p<.03$; $F2(1,11)=20$, $p<.001$) and with Matching Condition ($F1(2,76)=10.5$, $p<.001$; $F2(2,22)=4.9$, $p<.01$). No other interaction was significant.

Overall, Spanish subjects had significantly faster reaction times than French subjects (191ms, $F1(1,38)=4.8$, $p<.03$; $F2(1,11)=169$, $p<.001$). The Matching Condition factor was significant ($F1(2,76)=27$, $p<.001$; $F2(2,22)=15$, $p<.001$), as was an interaction with Language ($F1(2,76)=10$, $p<.001$; $F2(2,22)=25$, $p<.001$). To understand the interaction, we ran a series of post-hoc F-tests. The Accent Condition was significantly slower than the Phoneme Condition for French listeners (146ms, $p<.001$ both by items and by subjects), but not for Spanish listeners (both $ps>.1$). In contrast, the Redundant condition was significantly faster than the Phoneme Condition for Spanish (49ms, $p<.05$ by subjects, and $p<.02$ by item), but not for French listeners ($ps>.1$). A-responses were significantly slower than B-responses (108ms, $F1(1,38)=25$, $p<.001$; $F2(1,11)=34$, $p<.001$). The Response Type factor interacted with the Matching Condition ($F1(2,76)=12$, $p<.001$; $F2(2,22)=6.0$, $p<.01$), and there was a three-way interaction between Response Type, Matching Condition and Language ($F1(2,76)=4.6$, $p<.01$; $F2(2,22)=3.5$, $p<.05$). These interactions can be accounted for by the fact that the difference between accent and phoneme is much larger in A-Responses than in B-Responses (and mostly due to the French listeners). Also, the difference between the Redundant and Phoneme Conditions is larger in B-responses than in A-Responses (and mostly due to the Spanish listeners).

Discussion

In this experiment, French subjects were slow in the Accent Condition and relatively fast in the two other conditions. The error data corroborate the pattern observed in the response latencies. Spanish subjects were fastest in the Redundant Condition and equally slow in the Accent and in the Phoneme Conditions.

This pattern of data supports our hypothesis. French subjects have difficulty in perceiving accent and do not use accent in the Redundant Condition. Consequently, their performances are identical in the Phoneme and in the Redundant Condition. Spanish subjects, in contrast, perceive both accentual information and phonemic information with equal ease. These two conditions therefore show similar reaction times and errors. When accentual and phonemic information are redundant, Spanish subjects are faster.[2]

Notice that in the present experiment, the accent condition produces errors in French listeners (10%), but relatively less so than in Experiment 1 (20%). One possible reason for this is that only one accent contrast was used in the present experiment, whereas Experiment 1 used two contrasts: 1^{st} vs 2^{nd} and 2^{nd} vs 3^{rd}. Having only to deal with one accent contrast may allow French subjects to focus more successfully way on the acoustic information necessary to perform the task. The other possibility is that we only used two voices in the present experiment, thereby allowing subjects to compare A and B stimuli in terms of rather low level characteristics (since they were spoken in the same voice). In Experiment 1, in contrast, all three stimuli were spoken in a different voice, thereby inducing subjects to use a more abstract level of representation. This predicts that if we used only one speaker, French sub-

[2] Note that the redundancy gain is only apparent in the reaction times, not in the errors. This could be due to the fact that there is some trial-to-trial variation as to whether the phonemic information is available earlier or later than the accent information. If in congruent trials, subjects use the first available information, they will then be faster on average than when they can only use one type of information. However accuracy will not be affected.

jects may use low level representations and have even less trouble with accent contrasts. We examine this possibility in the next experiment.

In the first three experiments, we consistently found differences according to Response Type. Overall, A-Responses were more difficult than B-responses. This may reflect the fact that judging immediate identity (B-Responses) may be performed on a shallow memory store, whereas judging identity when there is some intervening material (A-Responses) requires holding several stimuli in memory, and keeping track of the order of each stimulus. Overall, A-Response situations may just be more confusing and open to different strategies than B-Response situations.

Apart from this general effect, response type interacts with the matching conditions in interesting ways. In particular the difficulty that French subjects have with accent (relative to some baseline) is always larger in A-responses than in B-responses (Exp 1 and 3). One possible explanation for this is that, as the task gets harder, existing differences in difficulty are amplified. Even in B-responses, the French subjects have difficulty with accent. One other possible explanation is that, for the French subjects, accent information is mostly represented in some acoustic store, which decays rather quickly and is poorly represented or not represented at all in a more abstract and longer term storage. In B-response situations, French subjects can judge accent identity based on this low level store. In A-response situations, the strategy would be less probable.

This raises the issue of the level at which the French deficit should be located. Is it the case that French subjects are impaired in hearing the acoustic correlates of accent, or do they not code accent in some short term memory buffer? The next experiment was designed to examine this question.

Experiment 4

The results obtained so far suggest that French subjects have difficulty with accent. However, we do not know at which level this difficulty arises. Is it the case that French subjects have lost perceptual sensitivity to accent contrast or rather that they have trouble with representing and storing in working memory accent patterns that are otherwise accurately perceived? In order to explore this issue, we placed French subjects in a simplified situation in which they only had to discriminate between two stimuli that were spoken by the same speaker. Subjects heard two stimuli (separated by a tone of varying duration) and had to judge whether the stimuli were the same tokens or varied in their pronunciation. There were two Different conditions, one in which the stimuli differed in accent (*fídape, fidápe*), and one in which they differed in one phoneme (*fídape, lídape*). In the Same condition, subjects heard two different recordings of the same stimuli.

In order to assess the potential effect of memory load, a pure 2KHz tone of either 200 ms or 2200 ms was inserted between each member of a stimulus pair. The idea is that for long intervening tones, echoic/acoustic memory traces of the first stimulus in a pair should be less accessible than for short tones, forcing subjects to rely on a more central representation. If French subjects are impaired in the early processing of the acoustic correlates of accent, they should have problems with the Accent Condition (compared to the Phoneme Condition), irrespective of tone duration. However, if the problem is linked to later processing/encoding stages, one would expect to see no problem with accent in this simplified task, or only in the long tone condition.

Method

Materials. The 12 experimental quadruplets of Experiment 3 were used and 288 new experimental trials were constructed. Each experimental trial consisted of two stimuli, A and X, which were spoken by the same Dutch female as in Experiment 3. A and X were either identical in accent and phoneme (although they were different tokens, *fídape, fídape*), differed by one phoneme only (*fídape, lídape*: Phoneme Condition) or by accent only (*fídape, fidápe*: Accent Condition). The trials were counterbalanced so that all members of a quadruplet (*fídape, fidápe, lídape, and lidápe*) appeared in position A, for each of the above defined matching conditions. This resulted in 12 different A-X combinations for each experimental quadruplet. Each A-X combination could appear in two ISI conditions, separated by a 200 ms 2kHz intervening pure tone, or a 2200 ms 2kHz pure tone. The overall design was thus 3 x 4 x 2: Matching Condition x Quadruplet Counterbalancing x ISI.

The 288 experimental trials were split into four blocks of 72 trials each (A, B, C, and D). Each condition and stimulus were represented equally often in each block, except that blocks A and C had only short ISI conditions, and blocks B and D had only long ISI conditions. A practice block of 10 trials was constructed that contained the same conditions as the experimental blocks but with different items.

Procedure. Each experimental trial consisted of the following sequence of three events: stimulus A, tone, stimulus B, each separated by an interval of 300ms. Subjects were instructed to listen to words in a foreign language separated by a tone. They were told that the two words were either identical or pronounced in a different way. They were required to press a button on their right or on their left to indicate whether X was identical to or different from A, respectively. The deadline and intertrial intervals were the same as in the previous experiments.

In the ten practice trials, subjects received feedback as to whether their response was correct or not. For incorrect responses, the same trial was presented immediately again, until the response was correct. After practice, subjects received two experimental blocks of 72 trials with no feedback presented (half received Blocks A and B, and half received Blocks C and D). The blocks were separately randomized for each individual subject. Subjects were distributed into two groups according to whether they received the short ISI or the long ISI first (orders A-B, and C-D, for short ISI first, and B-A and D-C for long ISI first). A short pause separated the two blocks. Responses were recorded and reaction times measured from the onset of the X stimuli by the EXPE software package.

Subjects. Twenty French students recruited in Paris participated in the experiment. There were 10 women, and 10 men in the group. Their median age was 30. None of the French subjects had learned Spanish as a foreign language or knew Dutch.

Results

Table 4 shows reaction times and error rates on Same and Different responses across the different conditions. We analysed the Different responses, with ISI and Matching Condition as within subject factors and Group as a between subject factor. All three factors were within-item.

In the error analysis, there was no main effect but there was an interaction between condition and ISI in the items analysis ($F1(1,19)=2.5$, $p>.1$; $F2(1,11)=10.3$, $p<.008$). This interaction was due to the fact that, in the long ISI condition, the Accent condition tended to generate more errors than the Phoneme condition (4.3% vs 1.7%, $F1(1,19)=1.4$, $p>.1$; $F2(1,11)=17.4$, $p<.002$), whereas no such trend was found in the short ISI condition (2.1% vs 2.7%, ns).

Reaction times in the Accent Condition were not significantly different from those in the Phoneme Condition ($F1(1,19)=1$, $p>.1$; $F2(1,11)<1$). Long ISIs resulted in longer reaction times than short ISIs but this difference was only significant in the items analysis (41ms, $F1(1,19)=2.6$, $p>.1$; $F2(1,11)=25.2$, $p<.001$). The two factors did not interact ($p>.1$).

Discussion

In this experiment, we had subjects perform an AX discrimination task on either accent or phoneme contrasts. The results show that French subjects *can* discriminate stimuli that differ only in accent, and do so with few errors (3.2%) compared to what we found in Experiments 1 and 3 (19% and 11%). The performance was similar for accent or phoneme contrasts and remained accurate when the two stimuli were separated by a 2200 ms tone, although there was a trend towards more errors for accent contrasts in the long tone condition. This suggests that French subjects can detect the acoustic correlates of accent and maintain a working memory of these correlates for more than 2 seconds. In this simplified paradigm, French subjects have very accurate access to the acoustic correlates of accent contrasts.

We can offer two reasons why the accent distinction may have been easier in the present paradigm than in Experiments 1, 2 and 3. One would be due to the fact that the present AX paradigm only involves two stimuli (the tone being ignored), which decreases the memory load compared to the ABX paradigm. Subjects may only have used some kind of immediate buffer[3] rather than having to store the identity and sequential order of three trisyllabic items (which is almost a supra-span situation in terms of number of syllables). The other reason may be that the present situation involved no change in talker, allowing the task to be performed at a less abstract level than in the previous two experiments. Hence, subjects might now use an acoustic representation to discriminate accent, a strategy less available when there was a speaker change.

Both accounts converge on the same idea. We suggest that acoustic information is processed in basically the same way in French and Spanish subjects. However, in order to retain such information in a short term memory store, this information has to be recoded into a more abstract level. We further suggest that language-specificity comes into play at this level. After acoustic information is processed, it is recoded in a different linguistic format by speakers of different languages. Word accent plays no lexical role in French, and is hence not represented at this level. Why this should be so will be discussed below.

General discussion

In this study, we have documented the poor ability of French speakers to deal with contrasts in accent. French subjects, contrary to Spanish subjects, have difficulty in making discriminations based on accent as indicated by slow reaction times and numerous errors (Experiment 1). This deficiency is coupled with the absence of any beneficiary or detrimental effect of concomitant accent variations when they perform judgments based on phonemes (Experiments 2 and 4). In contrast, Spanish subjects readily discriminate among accent patterns (Experiment 1). Moreover, they can be shown to extract and represent such information automatically, even when they are asked to focus on segmental information. Orthogonal variations of accent slow them down when they have to perform a task based on phonemes only (Experiment 2). Mutatis mutandis, Spanish subjects respond faster when accent and phonemes are redundantly varied compared to cases when only one dimension is available for decision (Experiment 3). For Spanish subjects, accent appears to be a nondetachable aspect of phonological information, whereas for the French, this information is not represented, at least not at the level that is supposedly tapped by this task. Stress is probably used at a different level by French speakers, e.g., for finding word or phonological phrase boundaries. This would be compatible with claims made by Frasier (1987) and Church (1987) that noncontrastive/allophonic information is used to parse speech into phonological constituents. Our final study shows, however, that French subjects are not altogether insensitive to differences in lexical stress, in that, under appropriate circumstances, they can detect the acoustic correlates of accent (Experiment 4).

This research sheds new light on past studies of language-specific processing. Languages are often characterized by the set of phonemes that they use to distinguish words. For example, English uses the contrast between [r] and [l] (*race* versus *lace*), whereas Japanese does not. These differences affect the ability of adults and young infants to distinguish speech sounds (Best, McRoberts, & Sithole, 1988; Goto, 1971; Mann, 1986; Miyawaki et al., 1981; Werker & Tees, 1984a). Similarly, languages are characterized by higher order units (such as syllables, or morae) which allow the

[3] Note however, that any such immediate buffer seems to be not affected at all by an intervening 2200 ms tone.

Table 4
Mean reaction time, standard error, and error rate to acoustic same-different AX judgments, in an identity, phoneme change and accent change condition (Experiment 4).

Condition	Mean	(se)	Err	Mean	(se)	Err
	Short ISI			Long ISI		
			"Same" Responses			
Identity Condition *(fídape fídape)*	943	(29)	5.4%	972	(32)	5.2%
			"Different" Responses			
Phoneme Change *(fídape lídape)*	919	(27)	2.7%	942	(32)	1.7%
Accent Change *(fídape fidápe)*	918	(30)	2.1%	978	(42)	4.3%

specification of restrictions on the co-occurrence of individual segments. Evidence for cross-linguistic variation in the way these structures are used in perception have been found (Bradley et al., 1993; Cutler, Mehler, Norris, & Segui, 1983; Cutler & Norris, 1988; Mehler, Segui, & Frauenfelder, 1981; Otake, Hatano, Cutler, & Mehler, 1993; Pallier et al., 1993; Sebastian-Galles et al., 1992; Zwitserlood, Schriefers, Lahiri, & Donselaar, 1993).

Our study shows that language-specific effects are not restricted to differences in the segmental inventory (or in the inventory of higher order units). Suprasegmental information, such as accent is also treated in different ways by listeners of different languages.[4] We believe it is likely that our results will generalize to other accent distributions. For instance, we predict that listeners who speak languages with fixed initial accent will behave like French listeners. We also predict that our results will extend to other dimensions (tone, pitch accent, etc).

Current models have viewed speech perception as consisting essentially in the discovery of phonemes arranged in their sequential order (McClelland & Elman, 1986; Norris, 1994). Although suprasegmental information was never explicitly excluded from these models, studies dealing with segments are much more common. On the basis of the above reported studies we are inclined to favor a view according to which segmental information should be complemented with a variety of abstract representations of prosodic dimensions. Moreover, the way in which these dimensions are represented is not universal. Depending on the language's phonology, dimensions that have a contrastive value will be specified in full detail, and others that have a fixed or predictable distribution will not be represented at all. Such a view is compatible with studies showing that knowledge about the use of suprasegmental structure in the native language begins to be acquired very early by the young infant, perhaps earlier than knowledge about permissible segmental units (Jusczyk, Cutler, & Redanz, 1993; Mehler et al., 1988; Mehler, Bertoncini, Dupoux, & Pallier, 1994).

Eimas and his colleagues published a number of experiments that show that shortly after birth infants can discriminate contrasts that are used in a natural language (Eimas, Miller, & Juszcyk, 1987). Werker and her colleages, pursued this line of work and demonstrated that before the end of their first year infants begin to neglect some contrasts absent from the language they are mastering (for instance, Werker & Tees, 1984a, Polka & Werker, 1994, but some contrasts remain, see Best, McRoberts & Nomathemba, 1988). Jusczyk, Friederici, Wessels, Svenkerud, and Jusczyk (1993) found that infants show some sensitivity to their language's phonotactic properties around the age of 9 months. Notice, however, that Kuhl and her colleagues (Kuhl, 1991) have shown that the vowels are learned around 6 months, that is, a few months prior to the consonants.

Mehler and colleagues have argued that infants learn about the prosody much before they learn about the segments of their own language. Indeed, Mehler et al. (1988) have shown that very young infants react differentially to novel sentences from their native language very shortly after birth. Moreover, infants still display such a differential reaction when the utterances are filtered to ensure that most of the segmental information is removed and only the prosodic information remains available. Moon, Cooper, and Fifer (1993) report similar results. Thus, infants individuate prosodic properties of their native language before they begin to specify its segmental aspects. Incidentally, it can be argued that the early individuation of prosodic properties protects the infant who is confronted from birth by multilingual input. Jusczyk et al. (1992) have shown that infants around 6 months are sensitive to sentence level prosodic boundaries. Jusczyk et al. (1993) have shown that by 6 months, English infants start to be sensitive to the typical stress pattern of their language (that is, strong-weak). We conjecture that around the same age, French infants should start losing sensitivity to accent contrasts.

Although much remains to be done, we believe that prosodic information will be added to our current models of adult speech perception and improve our understanding of

[4] Cutler and Norris (1988) found that the distribution of stress plays a role in word segmentation in English. However, they claim that this effect is not really suprasegmental since it rests on a contrast between reduced and nonreduced vowels, a segmental distinction. In English, it is uncertain whether purely suprasegmental properties affect word recognition (see Cutler & Clifton, 1984).

how speech is acquired and used by human infants.

References

Best, C. T., McRoberts, G. W., & Sithole, M. N. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by english-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 345–360.

Bradley, D. C., Sánchez-Casas, R. M., & García-Albea, J. E. (1993). The status of the syllable in the perception of spanish and english. *Language and Cognitive Processes*, *8*(2), 197–233.

Carroll, J. B. (1960). *Psycholinguistics: a book of readings.* New York: Holt, Rinehart and Winston, Inc.

Christophe, A., & Dupoux, E. (in press). A bootstraping approach to word segmentation: the role of prosodic structure. *The Linguistic Review*.

Church, K. W. (1987). Phonological parsing and lexical retrieval. *Cognition*, *25*, 53–70.

Cutler, A., & Clifton, C. J. (1984). The use of prosodic information in word recognition. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance x.* Hillsdale, NJ: Erlbaum.

Cutler, A., & Mehler, J. (1993). The periodicity bias. *Journal of Phonetics*, *21*, 103–108.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1983). A language specific comprehension strategy. *Nature*, *304*, 159–160.

Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 113–121.

Dell, F., & Vergnaud, J.-R. (1984). Les développements récents en phonologie : quelques idées centrales. In F. Dell, D. Hirst, & J.-R. Vergnaud (Eds.), *Forme sonore du langage* (pp. 1–42). Paris: Hermann.

Eimas, P. D., Miller, J. L., & Juszcyk, P. W. (1987). On infant speech perception and the acquisition of language. In S. Harnad (Ed.), *Categorical perception: the groundwork of cognition* (pp. 161–195). Cambridge, UK: Cambridge University Press.

Goto, H. (1971). Auditory perception by normal japanese adults of the sounds 'r' and 'l'. *Neuropsychologia*, *9*, 317–323.

Jusczyk, P., Cutler, A., & Redanz, N. (1993). Preference for predominant stress patterns of english words. *Child Development*, *64*, 675–687.

Jusczyk, P., Friederici, A., Wessels, J., Svenkerud, V., & Jusczyk, A. (1993). Infant's recognition of foreign versus native language words. *Journal of Memory and Language*, *32*, 402–420.

Jusczyk, P., Hirsh-Pasek, K., Kemler-Nelson, D., Kennedy, L., Woodward, A., & Piwoz, J. (1992). Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology*, *24*, 252-293.

Kuhl, P. K. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics*, *50*, 93–107.

Mann, V. A. (1986). Distinguishing universal and language-dependent levels of speech perception: Evidence from japanese listeners' perception of english [l] and [r]. *Cognition*, *24*, 169–196.

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

Mehler, J., Bertoncini, J., Dupoux, E., & Pallier, C. (1994). The role of suprasegmentals in speech perception and acquisition. *Dokkyo International Review*, *7*, 343–376.

Mehler, J., Dommergues, J. Y., Frauenfelder, U., & Segui, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior*, *20*, 298–305.

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants,. *Cognition*, *29*, 143–178.

Mehler, J., Segui, J., & Frauenfelder, U. (1981). The role of the syllable in language acquisition and perception. In T. Myers, J. Laver, & J. Anderson (Eds.), *The cognitive representation of speech.* Amsterdam: North Holland.

Miller, J. L., & Jusczyk, P. W. (1990). Seeking the neurobiological bases of speech perception. In P. D. Eimas & A. M. Galaburda (Eds.), *Neurobiology of cognition* (pp. 111–137). Cambridge, MA: MIT Press.

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A., Jenkins, J., & Fujimura, O. (1981). An effect of linguistic experience: the discrimination of /r/ and /l/ by native speakers of japanese and english. *Perception & Psychophysics*, *18*, 331–340.

Moon, C., Cooper, R., & Fifer, W. (1993). Two-day-olds prefer their native language. *Infant Behavior and Development*, *16*, 495-500.

Norris, D. (1994). Shortlist: a connectionist model of continuous speech recognition. *Cognition*, *52*, 189–234.

Otake, T., Hatano, G., Cutler, A., & Mehler, J. (1993). Mora or syllable? speech segmentation in japanese. *Journal of Memory and Language*, *32*, 258–278.

Pallier, C., Dupoux, E., & Jeannin, X. (1997). Expe5: 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 syllabic structure of spoken words. *Journal of Memory and Language*, *32*, 373–389.

Polka, L., & Werker, J. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(2), 241–435.

Sebastian-Galles, N., Dupoux, E., Segui, J., & Mehler, J. (1992). Contrasting syllabic effects in catalan and spanish. *Journal of Memory and Language*, *31*, 18-32.

Tomas, T. N. (1946). *Estudios de fonologia espanola.* Madrid: Gredos.

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

Werker, J. F., & Tees, R. C. (1984b). Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America*, *75*, 1866–1878.

Zwitserlood, P., Schriefers, H., Lahiri, A., & Donselaar, W. van. (1993). The role of the syllable in the perception of spoken dutch. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *19*(2), 260–271.

Appendix: Materials

Experiment 1:

`baveta, bopelo, detoma, lumisa, mepado, metilo, picadu, povami, rimato, someta, tamido, vasuma.`

Experiment 2:

```
bopelo-sopelo, povami-lovami,
tamido-pamido, vasuma-fasuma,
detoma-deboma, lumisa-ludisa,
picadu-piradu, rimato-ripato,
mepado-mepato, metilo-metibo,
someta-somefa, baveta-baveka.
```

Experiments 3 and 4:

```
bedapi-nedapi, kidosa-kimosa,
dabitu-dalitu, nelufo-nepufo,
dufoga-dufola, nibako-tibako,
fidape-lidape, nolaku-nosaku,
fubeno-fubeto, poleda-poleka,
kebuli-pebuli, tamido-tamipo.
```