PHONOLOGICAL REPRESENTATIONS AND REPETITION PRIMING

Christophe Pallier[1], *Nuria Sebastián-Gallés*[2] *and Angels Colomé*[2]

[1]LSCP, EHESS-CNRS, 54 bd Raspail, Paris, France.
[2] Univ. de Barcelona, Dept. psicologia bàsica., Barcelona, Spain
e-mail:pallier@lscp.ehess.fr
http://www.ehess.fr/centres/lscp/
Proceedings of Eurospeech'99, Budapest, Hungary, Sept. 5-9, 1999 (p. 1907-1910)

ABSTRACT

An ubiquitous phenomenon in psychology is the 'repetition effect': a repeated stimulus is processed better on the second occurrence than on the first. Yet, what counts as a repetition? When a spoken word is repeated, is it the acoustic shape or the linguistic type that matters? In the present study, we contrasted the contribution of acoustic and phonological features by using participants with different linguistic backgrounds: they came from two populations sharing a common vocabulary (Catalan) yet possessing different phonemic systems [1]. They performed a lexical decision task with lists containing words that were repeated verbatim, as well as words that were repeated with one phonetic feature changed. The feature changes were phonemic, i.e. linguistically relevant, for one population, but not for the other. The results revealed that the repetition effect was modulated by linguistic, not acoustic, similarity: it depended on the subjects' phonemic system.

1. INTRODUCTION

Are word forms memorized as abstract phonological representations or, rather, as detailed acoustic-phonetic representations (for example, as a set of acoustic exemplars associated to each word [2])? An empirical argument cited in favor of the 'acoustic' hypothesis is the sensitivity of repetition priming to changes in non-linguistic features: for example, when a word presented in noise is repeated twice with the same voice, it is easier to identify than when the voice has changed [2, 3]. The present study assessed the effect of linguistic knowledge on repetition priming: our rationale is that if words are stored in memory in an acoustic format, the repetition effect should *not* be modulated by the language-specific phonological knowledge of the participants. On the other hand, if words are stored using language-specific phonological representations, then the repetition effect should be affected by the subjects' linguistic background.

We tested people from two populations who shared a common vocabulary yet possessed different phonemic systems. All were Spanish-Catalan bilinguals, living in Barcelona (northern Spain); the first population consisted of people raised in Catalan-speaking families, and the second population consisted of people raised in Spanish-speaking families. The latter learned Catalan as a second language quite early (around 4-6 years of age), and became fluent speakers of this language that they are using everyday (they received the same bilingual education as the Catalan natives). In a previous study [1], we established that these two populations behaved differently in phoneme categorization and discrimination experiments involving the Catalan vowels /ɛ/ and /e/ (this contrast is used phonemically in Catalan but not in Spanish). While the Catalan-dominant subjects yielded categorization and discrimination curves that revealed that they perceived the contrast, Spanish-dominant subjects were not sensitive to this contrast.[1] This research showed that even an early and intensive experience with a phonemic contrast is not sufficient to master it.[2]

One aim of this paper is to assess whether the difficulty that Spanish-dominant speakers have with the Catalan /ɛ/ vs. /e/ contrast, in phoneme categorization and discrimination tasks, is also reflected in their word identification abilities. In other words, are Spanish-dominant listeners able or not to distinguish between Catalan words differing only in Catalan-specific contrasts?

A traditional view on speech perception and word recognition provides a ready answer to this question: it assumes that the acoustic speech stream is converted into phonemes (or features, or syllables..., in any case a language-specific phonological representation) before being compared with the lexical representations. If one believes

[1]Henceforth, we will call "Catalan-dominant", the persons who were exposed to Catalan from birth and "Spanish dominant" those who learned Catalan as a second language.

[2]Recall that the 'Spanish-dominant' are very good speakers of Catalan and have been exposed intensively to Catalan since 4-6 years of age. They received a bilingual Catalan-Spanish education. According to the official teaching regulations, Catalan vowels and consonants must be studied during each year of obligatory education (6-16 years). Starting from 8 years old, pupils study the relationship between sound and grapheme corresponding to every vocalic and consonantic sound (open/closed, voiced/voiceless...). Until children are 11 years old, stress is put on recognizing and classifying words according to the auditory discrimination of the degree of openess of a given vowel. As pupils get older, the role of stress is emphasized and the diacritic accent mark is introduced. Sounds continue to be studied through new techniques as application of rules, visual memorizing of the most common words or identification, classification and replacement of vowels.

that the results obtained in the phoneme identification tasks reflect this prelexical representation, one expects, then, that Spanish-dominant people should not be able to distinguish between Catalan minimal pairs of words such as, say, *té* (te, 'has') and *te* (tɛ, 'tea').

However, as we said above, a number of researchers have expressed doubts about the use of abstract phonological representations in word recognition [2, 3, 4]. They have proposed, instead, that word-forms are stored in the brain as one or several detailed acoustic traces, and that word recognition involves a "direct" comparison between these memorized acoustic patterns and the one elicited by the current acoustic signal. As no intermediate phonological representation is involved, it is said that lexical access is "direct".[3] The direct access hypothesis was first proposed, to our knowledge, by D. Klatt [5, 4] who listed a series of arguments against the need for pre-lexical phonological representation. In a nutshell, he claimed that the cost of the inevitable pre-lexical classification errors would not be worth the reduction in complexity in the lexical search algorithm. If detailed acoustic/phonetic informations indeed "percolate" to the lexicon, then Spanish-dominant listeners may have different representations for Catalan words such as *té* and *te*.

Recent experimental data has been presented to support the direct access hypothesis. Specifically, these experiments have shown that non-contrastive, and indeed non-linguistic, information is kept in memory when words are processed. For example, it is well-established that subjects presented with lists that contains repeated words improve their performance at the second occurrence of a word, in about any task. Several studies, in the auditory modality, found that this so-called "repetition effect" is modulated by the acoustic similarity between the original and the target words. Thus, for example, a change in speaker-voice decreased the amplitude of repetition effect in recognition [6] and in identification tasks [2, 7]. This demonstrates unequivocally that when a spoken stimulus is processed, some of its non-linguistic characteristics are kept in memory, which can help further processing of similar items. This type of evidence, among others, led Pisoni [3] to state, "that indexical and linguistic attributes of speech are not neatly partitioned into two independent channels of information by the nervous system." According to him, these results support an episodic, exemplar-based theory of word recognition (see also [2, 9]. According to "episodic" theory, the brain records detailed traces of every event which impinges upon the senses and objects are represented in memory by groups of such perceptual traces [10]. In the case of words, each would be associated to many acoustic tokens, and word recognition would consist in finding the nearest match in this vast memory.[4]

If word recognition is indeed based on detailed acoustic representations, then it may be the case that Spanish-dominant subjects, despite their difficulties with the /ɛ/-/e/ contrast in phoneme identification tasks, are able to identify correctly Catalan words that form minimal pairs along this contrasts, e.g. *té* vs. *te*.

There are other reasons to question the idea that the performance in word recognition should be immediately predictable on the basis of the results obtained in phoneme identification and classification tasks. It is doubtful that these tasks tap purely prelexical processes; they may actually be controlled by metalinguistic processes, as studies on phonological awareness suggest [14]. Also, it is known that the speech processing system can be sensitive to cues which are not available to consciousness: [15] has demonstrated that Japanese listeners, despite their notorious difficulties to distinguish between american /r/ and /l/ are sensitive to this contrast in adaptation experiments (see also [16]). It is not a priori impossible that some of this acoustic/phonetic information can "percolate" to the lexicon. Finally, in the neuropsychological literature, there are reports of patients who perform badly in phoneme identification and discrimination tasks yet have no trouble recognizing words; and the reverse pattern is also attested. Therefore, it is conceivable that the Spanish-dominant subjects, despite their lack of sensitivity to the /ɛ/-/e/ distinction in phoneme identification and discrimination tasks, might still be sensitive to it when identifying words. The following experiment was designed to assess this possibility.

2. EXPERIMENT

The subjects had to perform an auditory lexical decision task on lists that contained minimal pairs of Catalan words. Three contrasts existing in Catalan but not in Spanish were used: /ɛ/-/e/, /o/-/ɔ/, and /s/-/z/. Thus, for example, the stimulus sɛba was followed, latter down the list, by the stimulus seba; these two words differ only in a contrast that is hard for the Spanish-dominant group. The predictions were the following: If the comparison underlying the repetition effect is based on language-specific phonological representations, Spanish-dominant subjects, but not Catalan-dominant subjects, should treat sɛba as a repetition of seba. If, however, the comparison uses a representation that encodes fine acoustic details, then the Spanish-dominant and the Catalan-dominant participants should behave in the same way.

2.1 Method

2.1.1 Material

All the stimuli used in the experiment were Catalan words or pseudo-words. We considered three phonemic con-

[3]The metric of comparison is, of course, critical. It must emphasize linguistically relevant characteristics of the signal. Word identification supposedly occurs using a nearest neighbor rule in the metric space of all-word forms.

[4]The main difference with Klatt's original proposal is that not only one but several acoustic traces are stored for each word. One appeal of episodic theories is that they have been able to account, better than others, for the performance in several tasks [11, 12, 13]. For example, token frequency and training effects are elegantly explained in such a framework.

trasts that exist in Catalan but not in Spanish: [ɛ] vs. [e], [ɔ] vs. [o] and [s] vs. [z]. For each of these contrast categories, we selected sixteen words yielding eight minimal pairs. In addition, we selected eight additional minimal pairs of word using various contrasts that exist in Spanish as well as in Catalan (forming a "varied contrasts" category). Then, we created sixty-four Catalan pseudo-words yielding thirty-two minimal pairs following the same pattern as for the words. The words and pseudo-words were mostly bisyllabic though mono- and tri-syllabic items were also included. Finally, we selected 152 words and pseudo-words to be used as filler items. We recorded a Catalan speaker reading those stimuli at a pace of one word every two seconds.

We created four counterbalanced lists of 280 stimuli in the following way: In each list, one member of each minimal pair (e.g. *casa* from (*casa*-*caza*, 'marry'-'house')) appeared and was followed, 8 to 20 items further down in the list, either by the other token in the minimal pair (e.g. *caza*), or by itself (e.g. *casa*). The members of a given pair appeared in the same positions across the four lists; but in two lists it was the same token (the first or the second member of the pair) that was repeated, while in the other two lists, both members of the pair appeared. The first 24 items of the lists were fillers.

2.1.2 Procedure

Participants were assigned to each of the four lists according to their order of arrival and so that the number of "Catalan-dominant" and "Spanish-dominant" subjects were balanced within and across the four lists. They were tested in individual booths, seated in front of a computer that controlled the display of the instructions, played the stimuli off the hard disk and recorded responses [17]. Stimuli were presented through headphones every 2.5 seconds; for each stimulus, the participant had to decide as quickly and accurately as possible if it was a Catalan word or not, indicating his/her response by pressing one of two response buttons. Responses were recorded until 1.5 sec. after the offset of the stimulus; Response time was measured from the onset of the stimulus.

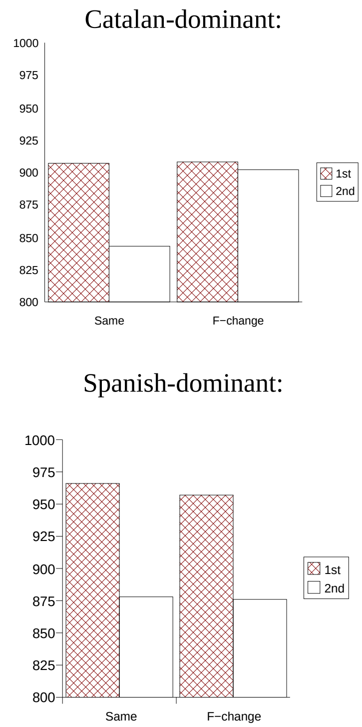
2.1.3 Participants

Sixty-four undergraduate Psychology students from the University of Barcelona participated in the experiment and received extra course credits. Half (32) were born in Catalan speaking families, and half (32) were born in Spanish speaking families. The latter had learned Catalan at most at age 6 year and were, like the former, fluent bilingual speakers of Spanish and Catalan.

2.2 Results

Correct responses accounted for 91.5% of the data set and were retained for analysis. In the first batch of analysis, we excluded the "varied contrasts" pairs and considered only the pairs involving the /e/, /o/ /s/ phonemes which we collapsed into a single category. Average reaction-times were computed for each subject and each item. Figure 1 shows, for each population, the average reaction times on the first and second occurences of words, as a function of the relationship between the first and the second member of the pair ('same token' or 'feature change'). The repetition effect is measured by the difference between the 2^{nd} and the 1^{st} occurrence.

Figure 1: Reaction times on first and second occurrences of words



Subjects' and Items' mean reaction-times were submitted to analyses of variance with the factors Language-dominance (Catalan vs. Spanish), Repetition (1st occurrence or 2nd occurrence of a member in a pair), Pair type (same tokens or feature-change), Lexicality (words vs. non-words), and List of stimuli. The factor Language produced a significant effect in the subject analysis (F1(1,56)=13, p<.001), but not in the item analysis (F2(1,56)=1.1). Lexicality yielded a 51 msec main effect (pseudo-words being slower to respond to than words), and produced several significant interactions with the other factors so that we decided to split the analyses of words and non-words. Actually, there were not any significant effect in the analysis of pseudo-words, thus analyses for pseudo words stops here.

In the analyses restricted to words, there was a triple interaction between Language, Repetition and Pair type (F1(1,56)=4.2, p<.05; F2(1,21)=4.2, p=.05). Analysis restricted to each language showed that the Repetition by Pair type interaction was significant for the Catalans (F1(1,28)=8, p<.01; F2(1,21)=6, p<.05; cf. top panel of figure 1), but not for the Spanish (both F¡1; cf. bottom panel of figure 1). Indeed the repetition effect was significant in every case except for the Catalans, in the 'Feature-change' condition.

We also conducted focused analyses for each contrast (ɛ-e, ɔ-o, s-z, m-n), separately for the words and for the non-words. Each analysis was an anova declared with the within-subjects factors Repetition and Type of Pair (there were 16 such anovas). No significant effect arose in the anovas concerning pseudowords. Therefore we turn to words. First, for the Spanish group: the main effect of Repetition was significant in the three Catalan categories, and it did not interact with Type of pair; in the Common category (m-n), it was the opposite pattern: there was a significant interaction, but no main effect. Second, for the Catalan group, there were significant main effects of Repetition in all categories except /s-z/, and there were significant interactions due to Repetition by Type of Pair in all categories except /e-ɛ/.

3. CONCLUSION

The main outcome of this study is that the facilitatory repetition effect is based on a phonological match rather than on an acoustic match. In other words, the repetition effect in auditory lexical decision is based on a language-specific metric of similarity: for example, "sèba" is similar to "séba" for Spanish-dominant subjects but not for Catalan-dominant subjects.

This result demonstrates that words are not memorized *only* under acoustic forms, but rather use language-specific linguistic representations. It also extends the finding of [1] with a paradigm that taps words' representations: Spanish-Catalan bilinguals who learnt Catalan at an early age do not have the same phonological representations as native Catalan speakers.

4. REFERENCES

[1] C. Pallier, L. Bosch, and N. Sebastián-Gallés. A limit on behavioral plasticity in speech perception. *Cognition*, 64(3):B9–B17, 1997.

[2] Stephen D. Goldinger. *Words and Voices: implicit and explicit memory for spoken words*. Ph.d., Indiana University, 1992.

[3] D. B. Pisoni. Some thoughts on "normalization" in speech perception. In K. Johnson and J. W. Mullenix, editors, *Talker Variability in Speech Processing*. Academic Press, San Diego, 1996.

[4] D. H. Klatt. Lexical representations for speech production and perception. In T. Myers, J. Laver, and J. Anderson, editors, *The Cognitive Representation of Speech*, pages 11–31. North-Holland Publishing Company, Amsterdam, 1981.

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

[6] F. I. M. Craik and K. Kirsner. The effect of speaker's voice on word recognition. *Quaterly Journal of Experimental Psychology*, 26:274–284, 1974.

[7] S. D. Goldinger. Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 22(5):1166–1183, 1996.

[8] D. B. Pisoni. Talker normalization in speech perception. In Y. Tohkura, E. Vatikiotis-Bateson, and Y. Sagisaka, editors, *Speech Perception Production and Linguistic Structure*, pages 143–151. IOS Press, Amsterdam Holland, 1992.

[9] S. M. Sheffert and C. A. Fowler. The effects of voice and visible speaker change on memory for spoken words. *Journal of Memory and Language*, 34:665–685, 1995.

[10] L. L. Jacoby and L. R. Brooks. Nonanalytic cognition: memory, perception, and concept learning. In G. Bower, editor, *The Psychology of Learning and Motivation*, pages 1–47. Academic Press, New York, 1984.

[11] D. L. Hintzman. "schema abstraction" in a multiple-trace memory model. *Psychological Review*, 93:411–428, 1986.

[12] R. M. Nosofsky. Tests of an exemplar model for relating perceptual classification and recognition memory. *Journal of Experimental Psychology: Human Perception and Performance*, 17:700–708, 1991.

[13] R. M. Nosofsky. Exemplar-based accounts of relations between classification, recognition, and typicality. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 14:700–708, 1988.

[14] J. Morais, L. Cary, J. Alegria, and P. Bertelson. Does awareness of speech as a sequence of phones arise spontaneously ? *Cognition*, 7:323–331, 1979.

[15] V. A. Mann. Distinguishing universal and language-dependent levels of speech perception: evidence from japanese listeners' perception of english "l" and "r". *Cognition*, 24(3):169–196, 1986.

[16] D. B. Pisoni and J. Tash. Same-different reaction times to consonants. *Journal of the Acoustical Society of America*, 55(A):436, 1974.

[17] C. Pallier, E. Dupoux, and X. Jeannin. EXPE: an expandable programming language for on-line psychological experiments. *Behavior Research Methods, Instruments, & Computers*, 29(3):322–327, 1997.