The representation of native and non-native lexical items in early bilinguals

Núria Sebastián-Gallés and Laura Bosch
Departament de Psicologia Bàsica
Universitat de Barcelona (Spain)
E-mail: sebastia@psico.psi.ub.es, lbosch@psi.ub.es

ABSTRACT

In a series of two experiments the representation of L2 words and non-words was analyzed. Catalan-Spanish bilinguals, differing in their L1, were asked to perform a lexical decision task on Catalan words and non-words. Non-words were based on real words, but with one vowel replaced. Critically, this vowel change could involve a Catalan contrast which Spanish natives find difficult to perceive. The results confirmed and extended our previous results in that in spite of early and intensive exposure, Spanish-dominant bilinguals fail to perceive some Catalan-contrasts and that this failure has consequences at the lexical level.

1. INTRODUCTION

Research in second language speech perception and bilingualism has focused a lot of effort on analyzing the perceptual difficulties of bilinguals in mastering the different phonological aspects of the second language (see for a review [1, 2]. In general, this research hints that although highly proficient bilinguals manage to master many aspects of their second language, this mastery is very rarely perfect. It could be the case that bilinguals have (relatively) impoverished mechanisms of speech perception and that their lexical representations are not so rich in relevant information, when compared with the processing and representation of their first language.

The aim of the present research is to further investigate this issue. We will do it by studying how words are recognized by Spanish-Catalan bilinguals, in particular we will further explore how these bilinguals store in the mental lexicon words that contain vowel contrasts that are difficult to perceive.

Our previous research has shown that Sp-dominant bilinguals, in spite of high degree of competence, have difficulties in perceiving some Catalan contrasts (absent in Spanish) [3-6]. In particular, [6] compared the performance of highly competent Spanish-Catalan bilinguals in an adaptation of the gating task. Two populations were studied, bilinguals who learnt only Catalan during the first years of their life, and bilinguals who learnt only Spanish in this period *but* who did not significantly differ from the Catalan-dominant population in their performance at the last gate. The results showed that Spanish-dominant bilinguals needed more information (more gates) to identify the correct answer.

Recently we have shown that these difficulties also extend to the lexicon [4]. In this study, Spanish-Catalan bilinguals, were asked to perform a lexical decision task. The repetition priming effects to pairs of words differing minimally in contrasts existing in both languages of bilinguals or in just one of their languages were compared. Pallier et al. took advantage of the fact that lexical decision times are faster on a repeated word trial than on its first presentation. The authors observed that Spanish-dominant participants (maternal language Spanish and unlikely to perceive certain Catalan-specific contrasts) considered the occurrence of the second member of a Catalan-specific minimal pair equivalent to an exact repetition of the first item; that is, the reaction times (RTs) to words when preceded by themselves in the list (RTs to the second occurrence of /netə/ in a list) were the equivalent to the RTs when they were preceded by their minimal pair (in this case /netə/ preceded by /nɛtə/ in the same list). However, Catalan-dominant bilinguals whose maternal language was Catalan and who can perceive these contrasts, did not show the repetition effect for these pairs (none of the bilinguals showed repetition priming for minimal pairs of words containing phonemes existing in both languages: /pɔt-bɔt/). In short, Spanish-dominant bilinguals, but not Catalan-dominant bilinguals treated Catalan-specific minimal pairs as homophones. .The conclusion of this paper is that minimal pairs are stored as homophones. However, the number of items was quite small (8 minimal pairs)

In the present paper we try to extend and replicate the results of that study using a different paradigm and materials. The Pallier et al. study had the advantage that the homophony effects were studied indirectly, but methodologically it restricted the study to a small number of stimuli (minimal pairs). Here we will use a different methodology that will allow us to use a larger set of stimuli and thus, also explore other issues.

If Spanish-dominant bilinguals store Catalan words containing either /e/ or /ɛ/ as a single category, they should show great difficulties in discriminating between real words and non-words made by exchanging these vowels. For example, the Catalan word for "window" is "finestra" pronounced /finestrə/. If these bilinguals do not perceive the contrast /e-ɛ/, and their lexical representations do not distinguish between the two, they should have difficulties in deciding whether /finestrə/ or /finɛstrə/ is the real word. However, they should not show these difficulties when the vowel change involves vowels used in Spanish.

ISBN 1-876346-48-5

2. EXPERIMENT 1

2.1. METHOD

2.1.1. Participants

Eighty Spanish-Catalan bilinguals participated in this experiment. All of them had been born in Catalonia (most of them in Barcelona or its metropolitan area). Half of them were raised as Spanish monolinguals until the age of four at the latest (when schooling starts). During the first years of their lives, contact with Catalan was casual. The other half of the bilinguals were a mirror image , with Catalan as their first language. All participants had received a bilingual education and they declared themselves to be very fluent in the two languages, in both listening and reading. At the moment of testing, they were all undergraduate Psychology students at the University of Barcelona, where most of the lectures (60%) are given in Catalan. They participated in the experiment in exchange of course credits

2.1.2. Materials

Thirty-four Catalan words containing the vowel /e/ and thirty-four Catalan words containing the vowel /ɛ/ were selected. Because of recording problems, two words from the e set and one from the ɛ set were discarded. Words varied in length (from one to four syllables). The two sets were matched were word frequency (average tokens per million for e-words: 722.27, sd. 1281.77; for ɛ-words, average: 662.19, sd. 1270.83, t-test<1[7]). The corresponding non-words were created by replacing vowel /e/ with /ɛ/.Thus, the word "galleda" (meaning "bucket"), pronounced /gəʎ**ɛ̲**ðə/, generated the non-word /gəʎ**e̲**ðə/ and the word "ulleres" (meaning "glasses"), pronounced /uʎ**ɛ̲**rəs/, generated the non-word /uʎ**e̲**rəs/. Because Catalan features vowel reduction, /e/ and /ɛ/ can only occur in stressed positions, so these changes were restricted to stressed syllables.

Forty control words were selected. These words also varied between one and four syllables long and were matched for frequency with the experimental items (average= 378.69, sd. 649.81, t-test p<.14). Thirty-five non-words were created from similar words to the controls by replacing their stressed vowel. These changes never involved vowels /e/ or /ɛ/. For example, the word "llençol" (meaning "sheet" –bed), pronounced /ʎənsə̲l/ had its stressed vowel changed to make /ʎənsa̲l/. Two lists were generated, so that one member of each pair appeared in each list

2.1.3. Procedure

Participants were tested in individual sound-attenuated booths. They were asked to press a "yes" button immediately on hearing a word stimulus and a "no" button on hearing a non-word. Instructions specifically warned them that the stimuli would be very similar to words and that they were made by changing a single vowel; they were particularly told that in many cases the change would involve the replacement of a sound /e/ by a sound /ɛ/ and vice-versa. Feed-back was provided during the training phase, in which some words and non-words containing the /e-ɛ/ exchange were included. Each trial first comprised an asterisk (fixation point) displayed for 300 ms. in the center of the computer screen. Immediately after the asterisk, the stimulus was played binaurally. The onset of the auditory stimulus triggered the timing mechanism. The asterisk reappeared to start the next trial 750ms after the subject had responded. Subjects had 1500 ms. to respond, otherwise the asterisk appeared indicating the onset of the following trial. The order of presentation was fully randomized for each subject. Participants were encouraged to respond as quickly as possible and to maintain their response fingers over the response buttons so as not to lose time. Half of the participants were tested with one of the lists and the other half with the other.

2.2. RESULTS

Error rates for the experimental stimuli were very high, in particular for participants with Spanish as their first language (see figure 1). In fact, these participants showed a bias to consider most experimental non-words as real words. Because of this high error rate and bias it was decided to only analyze the reaction times for the control stimuli (for which both populations showed very low error rates) and to carry out the accuracy analyses using the A' statistics. Furthermore, because tha A' data was not normally distributed, non-parametric analyses were performed (Wilcoxon matched-pairs signed-ranks tests).

2.2.1.Reaction time analyses

Reaction times above 3000 ms. and below 200 ms. were discarded, this represented less than 4% of all data. Two ANOVA, one by subjects and one by items were carried out on the reaction times of the control conditions. The analyses showed that words were responded to faster than non-words; F1(1,78)=164.73, p<.0001 and F2=(1,73)= 40.43, p<.0001 (words= 1004 ms., non-words= 1160 ms.). Differences between Catalan and Spanish dominant bilinguals did not reach significance in the subject analysis (F(1,78)= 2.458, p>.12), but they were significant in the item analysis (F2(1,73)= 70.203, p<.0001 (Catalan = 1055 ms. Spanish = 1108 ms.).

	Words	Non-words
Catalan dominant	965 (118)	1144 (155)
Spanish dominant	1043 (196)	1174 (171)

Table 1. Average reaction times for each control condition and bilingual population (sd in parenthesis)

The interaction of these two factors was significant in both analyses (F(1,78)=1.131, p<.041 and F2(1,73)=13.59, p<.0001). Post-hoc analyses of means showed that Catalan-dominant participants responded faster to words than Spanish-dominant ones (t1(78)= 2.149, p<.04, t2(39)= 8.748, p<.0001). The difference between the two populations for non-words was only significant in the item analyses (t1(78)= 0.823, n.s., t2(34)= 3.248, p<.003). Table 1 shows the average and standard deviations for each population and condition.

2.2.2.Error analyses

Figure 1 shows the average A' statistics for each bilingual population. Wilcoxon matched-pairs signed-ranks tests were computed across subjects and items for each type of stimuli: control, ɛ-type and e-type. The analyses showed significant differences between both populations for the control stimuli, but only in the item analysis (by subjects: z= -1.594, p<.111; by items: z= -4.545, p<.0001).

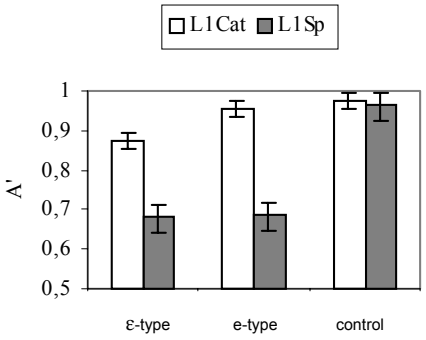

Figure 1. Average A' statistics for each bilingual population and stimuli type.

The two populations differed in both ε-type and e-type responses in the subjects and items analyses. For ε-type: by subjects: z= -6.930, p<.0001; by items: z= -6.711, p<.0001. For e-type: by subjects: z= -7.300, p<.0001; by items: z= -6.955, p<.0001.

When individual data were analyzed, the results showed quite different distributions for Spanish-dominant and Catalan-dominant bilinguals in the experimental stimuli. In particular for the e-closed stimuli, while only one Catalan bilingual had a score of less than .86, only four Spanish-dominant bilinguals scored above this. That is, only ten percent of Spanish-dominant bilinguals fell within the Catalan range (if the Catalan dominant bilingual with a score of .80 is considered as a category limit, then seven Spanish Catalan bilinguals fell within the Catalan range, that is 17.5%). The distribution of e-open responses did not yield so clear a separation of the two populations. While no Catalan dominant participant had an A' score of less than .75, roughly one third (32.5%) of the Spanish dominant bilinguals scored above this lower cut-off point.

2.3. DISCUSSION

The results of this experiment are quite clear-cut. They show that in spite of a high command of their second language, Spanish-dominant bilinguals are far from Catalan dominant bilinguals, when their accuracy in recognizing stimuli including vowels /e/ and /ɛ/ is analyzed. However, the fact there were significant differences between the two populations (although only by items) in the reaction times and in the accuracy data (A') for *control* items as well, complicates matters. However, the situation in which the two populations were tested cannot be considered fully equivalent. Spanish-dominant bilinguals were asked to distinguish between Catalan sounds that are particularly difficult to them, thus it is not surprising that they were slower when making their decisions regarding the control items as well. Furthermore, because of their false recognition of experimental non-words as words, the percentage of "no" responses was much lower than for Catalan-dominant bilinguals. This may have influenced their criteria to give a positive or a negative response, a bias that might have influenced their responses to the control stimuli.

If the differences observed between the two populations were due to the above reasons, testing these participants with just the control stimuli should eliminate them.

3. EXPERIMENT 2

3.1. METHOD

Forty new participants from the same population as in the previous experiment were tested. Half of them were only exposed to Spanish during the first years of their live and half of them to Catalan.

Materials and experimental procedure were the same as in the previous experiment, except that only the control stimuli were used here, thus, only one list was used (all participants were tested with the same stimuli).

3.2. RESULTS AND DISCUSSION

Reaction times above 3000 ms. and below 200 ms. were discarded, this represented less than 3% of all data. Two ANOVA, one by subjects and one by items were carried out on the reaction times. The analyses showed that words were responded to faster than non-words; F1(1,38)=55.285, p<.0001 and F2=(1,73)= 38.425, p<.0001. Differences between Catalan and Spanish dominant bilinguals did not reach significance in any of the analyses (both F<1).

The interaction of these two factors was only significant in the item analysis (F(1,38)= 2.831, p>.10 and F2(1,73)= 16.085, p<.0001). Post-hoc analyses of means showed no differences between populations for either words or non-words in the subject analyses (both t1s <1), although the analysis by items showed significant differences both for words (t2(39)= 2,395, p<.03) and non-words (t2(34)= 3.208, p<.003). Table 2 shows the average and standard deviations for each population and condition.

	Words	Non-words
Catalan dominant	934 (190)	1058 (170)
Spanish dominant	906 (114)	1103 (193)

Table 2. Average reaction times for each condition and bilingual population (sd in parenthesis)

The analysis of the accuracy data (A') showed no significant differences between the two groups of bilinguals (z= 1.753, p>.08 by subjects and z=1.237, p>.21, by items); the average A' were .988 for the Catalan-dominant group and .982 for the Spanish-dominant group.

The present data shows that the differences observed in the previous experiment in the control stimuli between Spanish and Catalan-dominant bilinguals were due to the experimental situation, thus both populations can be considered to be equivalent in their knowledge of the Catalan lexicon.

4. CONCLUSIONS

The present research replicates and extends the conclusions of [4]. Spanish-dominant bilinguals have great difficulties in distinguishing between words and non-words differing only by a perceptually subtle Catalan-specific contrast.

However, there was an additional finding in experiment 1 in that Catalan-dominant bilinguals did not perform similarly with both experimental categories. Indeed, these participants had greater difficulties in distinguishing between

ISBN 1-876346-48-5

words and non-words of the /ɛ/ category than between words and non-words of the /e/ category. We believe that this effect may be due to their exposure to the Catalan dialect spoken by Spanish-dominant bilinguals. These bilinguals do not only fail to perceive some of the Catalan-specific contrasts, but also they do not produce them, thus, in their spoken dialect the /e-ɛ/ contrast does not exist. Because of the linguistic situation in Catalonia where individuals often speak both languages depending on the listener, Catalan-dominant bilinguals are frequently exposed to a dialectal variation where this contrast is not present. The asymmetry observed in their behavior indicates that this exposure modifies their lexical representations. It remains an open question whether their perceptual capacities are also modified because of this exposure. Nevertheless, our previous work indicates that Catalan-dominant bilinguals are very good in discrimination tasks, where lexical knowledge was not involved.

REFERENCES

[1] N. Sebastián-Gallés and J. Kroll, "Phonology in Bilingual Language Processing: Acquisition, Perception, and Production", in *Phonetics and Phonology in Language Comprehension and Production: Differences and Similarities*, A. Meyer and N. Schiller, Eds. Berlin: Mouton de Gruyter, 2003.

[2] N. Sebastián-Gallés and L. Bosch "Phonology and Bilingualism", in *Handbook of Bilingualism: Psycholinguistic Approaches*, J.F. Kroll and A.M.B. de Groot, Eds. Oxford: Oxford University Press. In press.

[3] L. Bosch, A. Costa and N. Sebastian-Gallés, "First and second language vowel perception in early bilinguals". *European Journal of Cognitive Psychology*, vol 12, pp. 189-222, 2000.

[4] C. Pallier, A. Colomé and N. Sebastián-Gallés, "The influence of native-language phonology on lexical access: Exemplar-based vs. abstract lexical entries". Psychological Science, 2001. **12**: p. 445-449.

[5] C. Pallier, L. Bosch and N. Sebastián-Galleés, "A limit on behavioral plasticity in vowel acquisition". *Cognition*, vol. 64, pp. B9-B17,1997.

[6] N. Sebastián-Gallés and S. Soto-Faraco, "On-line processing of native and non-native phonemic contrasts in early bilinguals". *Cognition*, vol. 72, pp. 112-123,1999.

[7] N. Sebastián-Gallés *et al.*, *LEXESP: Léxico informatizado del Español*. Barcelona (Spain): Edicions Universitat de Barcelona. 2000,