

Constraints of vowels and consonants on lexical selection: Cross-linguistic comparisons

ANNE CUTLER

Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

NURIA SEBASTIAN-GALLES

University of Barcelona, Barcelona, Spain

OLGA SOLER-VILAGELIU

Universitat Autònoma de Barcelona, Barcelona, Spain

and

BRIT VAN OOIJEN

Laboratoire de Sciences Cognitives et Psycholinguistique, CNRS, Paris, France

Languages differ in the constitution of their phonemic repertoire and in the relative distinctiveness of phonemes within the repertoire. In the present study, we asked whether such differences constrain spoken-word recognition, via two word reconstruction experiments, in which listeners turned non-words into real words by changing single sounds. The experiments were carried out in Dutch (which has a relatively balanced vowel-consonant ratio and many similar vowels) and in Spanish (which has many more consonants than vowels and high distinctiveness among the vowels). Both Dutch and Spanish listeners responded significantly faster and more accurately when required to change vowels as opposed to consonants; when allowed to change any phoneme, they more often altered vowels than consonants. Vowel information thus appears to constrain lexical selection less tightly (allow more potential candidates) than does consonant information, independent of language-specific phoneme repertoire and of relative distinctiveness of vowels.

Is a *kebra* more like a cobra or a zebra? Each of the real animal names differs from the nonword *kebra* by a single phoneme—*cobra* by the first vowel, *zebra* by the first consonant. But are these two single-phoneme differences equivalent for the listener, or is the nonword more likely to activate one of the real words than the other?

In spoken-word recognition, multiple lexical candidates are activated by the speech input and then compete with one another for recognition. Words compete with other words with which they overlap (McQueen, Norris, & Cutler, 1994; Tabossi, Burani, & Scott, 1995), with words in which they are embedded (Gow & Gordon,

1995; Shillcock, 1990), and with other words that begin the same way (Marslen-Wilson, 1990; Zwitserlood, 1989). Although activation may be inhibited by mismatch in just one phoneme (Marslen-Wilson & Zwitserlood, 1989), there remain strong effects of phonetic similarity, indicating that a heard word produces some activation for words made up of similar-sounding phonemes (e.g., Connine, Blasko, & Wang, 1994; Slowiaczek & Hamburger, 1992) or for words with which it shares phonemes (e.g., Radeau, Morais, & Segui, 1995; Slowiaczek, McQueen, Soltano, & Lynch, in press).

In fact, the speech signal that activates candidate words is not a sequence of separate units but is continuous. Although language researchers describe the information it contains in quantized terms—that is, as a sequence of phonemes—the information cuing an individual phoneme is spread out in time and overlaps with cues to other sounds. Listeners are also able to process the signal continuously and, thus, to extract relevant information about upcoming sounds, as is shown by the sensitivity which they demonstrate to mismatching information in cross-spliced input (Marslen-Wilson & Warren, 1994; McQueen, Norris, & Cutler, 1999; Whalen, 1991). Thus, models of spoken-word recognition are unrealistic if they assume, even for the sake of convenient computational implementation, that listeners hear a string of separate phonemes.

Partial reports of the study were presented to the joint conference of the Acoustical Society of America and the European Acoustics Association, Berlin, in March 1999. This research was supported by grants from the Human Frontier Scientific Program, from Acciones Integradas Hispano-Alemanas, and from the Ministry of Education and Science, Spain. We thank Carolina Cortes, Elena Salillas, Xavier Mayoral, and Salvador Soto in Barcelona and Mattijn Morren, Petra van Alphen, Bram de Kruijff, and Niels Janssen in Nijmegen for technical assistance, Arie van der Lugt for assistance with the statistical analyses, and Debra Titone and Danny Moates for comments on an earlier version of the paper. The full materials for both experiments are available on line (<http://www.mpi.nl/world/persons/profession/anne.html>). Correspondence concerning this article should be addressed to A. Cutler, Max Planck Institute for Psycholinguistics, P. O. Box 310, 6500 AH Nijmegen, The Netherlands (e-mail: anne.cutler@mpi.nl).

This is not to deny the value of phonemic representations. One obvious way in which phonemes are real in lexical activation is that (by definition) they are the minimal unit by which one word differs from another. Thus, the nearest neighbor of any word is any other word that differs from it by one phoneme, and this is clearly important for word activation. But it is not necessarily the case that every phoneme makes an equally salient contribution in the speech signal or constrains word recognition as effectively as every other phoneme.

In fact, it has been claimed that two broad classes of phonemes—namely, consonants and vowels—differ in how rapidly or effectively they constrain lexical recognition. Vowels and consonants differ in some potentially relevant ways; in general, vowels are longer than most consonants, and stressed vowels, at least, are more salient to listeners in that they better resist noise masking (Nooteboom & Doodeman, 1984) and are less often misreported in slips of the ear (Bond & Games, 1980). Although it might seem that vowels would therefore be more useful in lexical activation than are consonants, experimental evidence in fact suggests the reverse. On the basis of results from the word reconstruction task, van Ooijen (1996) claimed that English listeners treat vowels as inherently mutable and as less reliable in constraining lexical access.

The word reconstruction task is a method for assessing which words are activated by input that matches no word perfectly. In this task, listeners hear spoken nonwords that have been constructed to mismatch real words on a single segment. Van Ooijen's (1996) study, in English, used nonwords, all of which could be changed into two real words via the substitution of either a single vowel or a single consonant. For instance, *kebra* in this task could become the real word *cobra* if a change were made in the first vowel or the real word *zebra* if a change were made in the first consonant. The listener's task in word reconstruction is to reconstruct real words—that is, to turn each nonword into a real word by changing just one phoneme of the nonword. Listeners are required to respond, as quickly as possible, with the first real word they find.

The two-alternative task can be implemented in two forms. Listeners can be given a free choice as to which phoneme they change; in this case, the proportion of responses can be compared across phoneme types, and the relative speed of each type of response can also be compared. Alternatively, listeners can be constrained to alter only vowels or only consonants; in this case, the average speed of each type of response can again be compared, and one can also measure the number of errors in each condition.

Van Ooijen (1996) used both forms of the task. The listeners in her experiment produced clearly asymmetric response patterns. When they were given a free choice, listeners tended more rapidly to produce responses that changed a vowel than responses that changed a consonant. When they were constrained to alter only vowels or only consonants, listeners were significantly more accu-

rate in performing vowel changes, and the most common type of error that they made was altering a vowel when they were supposed to alter a consonant. Van Ooijen concluded that English listeners treated vowels as, in principle, more mutable than consonants.

Some potential explanations were directly excluded by additional analyses of the results. For instance, featural analyses ruled out the possibility that the difference between the replacing phoneme in the input nonword and its original in the real word was significantly greater for either vowel or consonant substitutions, and no correlation was found between response pattern and featural similarity of original and replacing phoneme. There was also no correlation between asymmetry in vowel versus consonant responses and the opportunity for replacement afforded by the number of consonant versus vowel slots in the individual nonwords. Note, however, that the English vocabulary, as a whole, affords much greater opportunity to turn one word into another by replacing a consonant (e.g., for *cat*: *pat*, *mat*, *cash*, *can*, etc.) than by replacing a vowel (e.g., for *cat*: *kit*, *caught*, *coot*, etc.). Thus, the pattern of results also cannot be explained by a tendency across the vocabulary for vowel replacement to be more likely to yield an alternative word.

Several alternative explanations, however, remain. First, the acoustic closeness among the members of the English vowel repertoire could have led to a cumulative experience by listeners of mistaking one vowel for another more often than mistaking one consonant for another, with the consequence that listeners have learned to give less weight to vowel information in word recognition than to consonant information. Second, asymmetry between the number of vowels and consonants in the phonetic repertoire could be a crucial factor. Most languages, English included, have more consonants than vowels. In fact, the phonemic repertoire of British English is relatively large, and in particular, the number of vowels is high, so that the ratio of vowels (17) to consonants (24) is rather more balanced than in most European languages. Nevertheless, it could have been the case that a simple search for a substitution could be more rapidly completed among the vowel candidates than among the consonant candidates. Third, it could be that vowels are simply more intrinsically variable than consonants, so that listeners are more inclined to alter vowels because they consider vowel information to be less exact. This, the preferred explanation of van Ooijen (1996), could be a universal effect, or it could be specific to British English, perhaps because, in that language, vowels are the main repository of dialectally imposed variation.

In the present study, we attempted to disentangle these alternative suggestions. This is not always a straightforward task. For instance, it would seem that to test the explanation in terms of vowel/consonant repertoire asymmetry, one need only repeat the experiment in a language with the reverse asymmetry—that is, more vowels than consonants. But such languages are very rare, and those that exist tend not to be accessible to psycholinguistic

laboratories. However, the question can be approached from the opposite direction, since many languages, including many easily accessible languages, have vowel/consonant repertoires that are dramatically more asymmetric than that of English. Castilian Spanish, for instance, has five vowels and 20 consonants (Maddieson, 1984; Stockwell & Bowen, 1965). If English listeners' greater willingness to alter vowels than consonants is due to the greater number of consonants than vowels in the language, Spanish listeners should show an even more marked preference in the same direction.

Thus, in the present study, we tested Castilian Spanish listeners in the word reconstruction task. We compared their pattern of results with those of Dutch listeners in the same task. Dutch is a language with even less asymmetry in vowel/consonant repertoire than English: 16 vowels and 19 consonants (Booij, 1995). Listeners in these two populations have been shown in phoneme detection tasks to be sensitive to the constitution of their language's phonetic repertoire and to the effects this has on contextual variability: Dutch speakers show evidence of expecting roughly equal amounts of contextual variability across consonant-vowel (CV) syllables, whereas Spanish speakers expect greater variability in the vowel than in the consonantal portions of the same CV sequences (Costa, Cutler, & Sebastian-Galles, 1998). We expect that if phoneme repertoire asymmetry plays a role in the word reconstruction task, Spanish listeners will show significantly greater preference than Dutch listeners for vowel over consonant alteration.

The same cross-linguistic comparison allows us to test the explanation in terms of acoustic closeness of vowels. The acoustic closeness within the Dutch vowel repertoire resembles that in English: The 16 Dutch vowels make for a crowded vowel space and frequently overlapping realizations. But the Spanish situation is quite different. The 5 Spanish vowels are acoustically distinct, and Spanish listeners are thus unlikely to have more experience of mistaking one vowel for another than of mistaking one consonant for another. Consequently, if the original finding reflected greater acoustic closeness among the vowels than among the consonants of English, we should expect the same effect in Dutch, but no such effect, indeed possibly a reversed effect, in Spanish.

On the other hand, if the original result was due purely to inherent variability of vowel information as such, it might appear generally across languages. Certainly, there is cross-linguistic evidence that listeners exercise caution in phoneme detection responses to vowels both in English and in Spanish (Cutler, van Ooijen, Norris, & Sanchez-Casas, 1996) and that vowels are harder phoneme detection targets than are consonants in both English and Spanish (van Ooijen, Cutler, Sanchez-Casas, 6 Norris, 2000). If, then, listeners expect vowel variability to be greater than consonantal variability in all languages, van Ooijen's (1996) finding should replicate in both Dutch and Spanish word reconstruction.

Experiment 1 applied the word reconstruction task to Dutch. Dutch listeners were presented with Dutch nonwords, such as *kebra*, and we measured the relative speed and likelihood with which they produced real Dutch words, such as *cobra* or *zebra*.

EXPERIMENT 1

Method

Subjects. Forty-six Nijmegen University students, all native speakers of Dutch, took part in the experiment, in return for a small payment. The data for 1 subject were lost.

Materials. Sixty disyllabic nonwords were constructed, each of which could be changed into a real Dutch word by substitution of a single consonant and into another real Dutch word by substitution of a single vowel. An example is *kebra*, or *hieet*, which could become *dieet* (diet) or *hiaat* (hiatus). In half the items, the consonant position for substitution preceded the vowel position (as in *kebra* or *hieet*), and in the remaining half, the vowel position preceded the consonant position (as in *komeel*, which can become *komeet* [comet] or *kameel* [camel]).

One third of the items contained four phonemes, two vowels and two consonants, so that in these items, the potential opportunity for vowel and consonant substitution was matched (as has been described, most Dutch words contain more consonants than vowels). Also for this subset, half of the items contained earlier consonant substitution positions than vowel substitution positions, and half the reverse; examples are *tula*, which can become *tuba* (tuba) or *tule* (tulle), and *utiek*, which can become *uniek* (unique) or *ethiek* (ethics).

It was not possible to match each pair of words for frequency, but overall frequency (in the CELEX 42-million-word Dutch lexical database; Baayen, Piepenbrock, & van Rijn, 1993) was roughly matched across the sets of words resulting from vowel versus consonant substitutions (mean frequency for vowel substitution, 217; for consonant substitution, 241). For approximately half the items, a vowel substitution resulted in a higher frequency word, whereas for the remaining items, a consonant substitution produced the higher frequency word.

A few items (22%) allowed more than one substitution option, as in *wanding*, which could become *landing* (landing), *randing* (edging), *wending* (turn), or *winding* (coil). The full set of items, with glosses and all word options, is available at the first author's Web address.

Sixty-seven further nonwords were constructed for use as practice and filler items. A few of these resembled the experimental items (e.g. *mader*, which could become *moeder* [mother] or *vader* [father]). The remainder were trisyllabic and were chosen to allow only a simple vowel or consonant substitution (e.g. *telefien* from *telefoon* [telephone] *ofamigie* from *familie* [family]).

The nonwords were divided into three sets, each of which contained 20 of the 60 experimental items. Position of the vowel versus consonant mismatch, relative frequency of the two words, stress pattern, and length of the nonword in phonemes were balanced as closely as possible across the sets. Each set was combined with three separate sets of 14 filler nonwords for use in the three substitution conditions: vowel substitution, consonant substitution, and free choice (vowel or consonant substitution). For each of these conditions, a short practice set was also constructed. All the nonwords were recorded by a female native speaker of Dutch. To ensure that the nonwords were realized so as to resemble the words on which they were based, the speaker spoke the two words in each case prior to the nonword (e.g., *cobra*, *zebra*, *kebra*).

Procedure. The subjects were tested individually or in pairs in individual sound-attenuated booths. The instructions and stimuli were presented over headphones. All the subjects received all ex-

perimental items in the same order, but order of task condition was varied so that 15 subjects performed vowel substitution first, consonant substitution second, and free-choice substitution last, with two other groups of 15 subjects each performing these tasks in different, counterbalanced order. The first set of instructions that subjects received made no mention of the fact that the nature of the task would change during the experiment; examples illustrated the task for that condition only. In the instructions, the subjects were asked to listen carefully to each nonword and to press a response key as soon as they had thought of a real-word substitution and then to speak this word aloud. This method was used instead of voice-triggered timing because the latter is heavily influenced by the nature of word-initial phonemes, which cannot be matched in this task. The subjects had a 10-sec response window per nonword; this interval was established by van Ooijen (1996) to be comfortable for subjects. Spoken responses were recorded. A signal aligned with the onset of each nonword, but inaudible to the listeners, started the clock of a computer running NESU experimental control software; timing was terminated by the subjects' keypress responses. Response times (RTs) were recorded and stored by the computer. After the first and second set of items had been heard, the subjects were informed of the next task and were given new illustrative examples of that task.

Results

Overall Comparisons. The listeners' verbal responses were coded as correct or error, with error subdivided into three groups: responses not complying with the instructions (a word response that could not be derived from the stimulus by substitution of a single phoneme), false alarms (a keypress followed by no verbal response), and, for the vowel and consonant change condition, intrusions (a word formed by substitution of a vowel instead of a consonant or vice versa). Table 1 presents the mean correct RTs and mean overall error rates (including failures to respond, as well as the above types of wrong response) for the three conditions.

Separate analyses of variance (ANOVAs), with subjects and items as random factors, were carried out on the RTs and on the error rates. These analyses showed a significant effect of task condition on both measures [$F_1(2,84) = 28.55, p < .001$, and $F_2(2,118) = 25.53, p < .001$, for RTs; $F_1(2,84) = 104, p < .001$, and $F_2(2,118) = 36.06, p < .001$, for errors]. Subsequent *t* tests showed that all the conditions were significantly different from one another on each measure, across both subjects and items, at the .01 level at least. There was no significant difference between subject groups (i.e., order of presentation of the three task con-

ditions, which was included as a factor in the analyses across subjects) on either measure. The groups factor did not interact with task condition in the RT analysis but did interact significantly in the error analysis: [$F_1(4,84) = 3.42, p < .02$]. All of the groups made the most errors in the consonant substitution condition and the fewest errors in the free-substitution condition, but there was also an effect such that for every task condition, the most errors were made by the group that had that task as its first condition. On both RT and errors, the ordering of conditions (best performance in free substitution, worst in consonant substitution) also held true across groups for the conditions that the listeners heard first.

Similar analyses were carried out with only the subset of 20 items that had equal numbers of vowels and consonants. Again, there was a significant effect of task condition on all measures [$F_1(2,84) = 9.39, p < .001$, and $F_2(2,38) = 5.78, p < .01$, for RTs; $F_1(2,84) = 16.55, p < .001$, $F_2(2,38) = 5.75, p < .01$, for error rates]. For this subset, *t* tests showed that RTs were significantly slower for the consonant substitution condition than for either the vowel substitution condition [$t_1(44) = 2.78, p < .01$; $t_2(19) = 1.93, p < .07$] or the free-choice condition [$t_1(44) = 3.75, p < .001$; $t_2(19) = 3.47, p < .005$], but the latter two did not differ. For this subset, *t* tests on the error rates showed that there were significantly more errors in the consonant substitution condition than in the free-substitution condition [$t_1(44) = 4.1, p < .001$; $t_2(19) = 3.42, p < .005$] and also in the vowel substitution condition than in the free-substitution condition [$t_1(44) = 2.52, p < .02$; $t_2(19) = 2.02, p < .06$], but the difference between the vowel and consonant substitution conditions was significant only across subjects [$t_1(44) = 2.68, p < .01$; $t_2(19) = 1.55, p > .1$]. There was again no effect of order of presentation.

Thus, as was predicted, the listeners found the free-substitution condition easiest: They made fewer errors and responded more rapidly. When the task was constrained to substitution of a particular phoneme type, consonants were significantly harder to substitute than vowels.

Vowel versus consonant substitution conditions. The number of intrusions was tallied for each condition. Fifty-three vowel intrusions occurred in the consonant condition (7 of them involving words with equal numbers of vowels and consonants), and 24 consonant intrusions occurred in the vowel condition (2 of which involved words with equal numbers of vowels and consonants). Wilcoxon matched-pairs signed-ranks tests were carried out across subjects and items. Twenty-two subjects made no intrusions, 2 made equal numbers of vowel and consonant intrusions, 14 made more vowel than consonant intrusions, and 7 made more consonant than vowel intrusions. The difference between the latter two sets was significant ($z = 2.22, p < .02$). Eighteen items received no intrusions, 1 item received an equal number of vowel and consonant intrusions, 30 received more vowel than consonant intrusions, and 11 more consonant than vowel

Table 1
Mean Correct Response Time (RTs, in Seconds),
and Mean Error Rates (%) for the Three
Task Conditions in Experiment 1

Condition	RT	Error Rate
All items		
Free-choice	2.12	15.8
Vowel change	2.38	23.3
Consonant change	3.02	43.1
Subset of items with equal numbers of vowels and consonants		
Free-choice	2.65	23.5
Vowel change	2.81	32.4
Consonant change	3.24	43.4

intrusions. The difference between the latter two sets was again significant ($z = 2.6, p < .005$). In other words, vowels intruded upon consonant substitution significantly more than consonants upon vowels.

In these conditions, we also compared performance separately for items for which the vowel change word had higher versus lower frequency than the consonant change word. For both sets of words, RT was significantly faster and error rate significantly lower in the vowel substitution condition; however, the difference between the two conditions was larger when the vowel change word had higher frequency (458 msec, 22.5% error) than when it had lower frequency (393 msec, 17.9% error).

Free-choice condition. The proportions of correct responses in the free-choice condition involving vowel versus consonant substitutions were computed, and the RTs associated with these responses compared. Table 2 presents the respective proportions and mean RTs for the whole materials set and for the subset of 20 items matched for number of vowels and consonants.

Although the mean RT to vowels for the whole set of materials was faster than that to consonants, this difference was not statistically significant. The difference for the matched subset, however, was significant [$F_1(1,42) = 6.39, p < .02$; this analysis was only carried out across subjects, since many items received only one sort of response]. Wilcoxon tests were carried out on the proportions of vowel versus consonant substitutions in this condition. Across the whole set of materials, 31 subjects made more vowel than consonant responses, and 13 subjects more consonant than vowel responses (with one tie), whereas 37 items received more vowel than consonant responses and 20 items received more consonant than vowel responses (with three ties). Both differences were significant ($z = 3.14, p < .001$, across subjects; $z = 1.89, p < .03$, across items). The same analysis across items for the matched subset only (16 items, more vowel responses; 3 items, more consonant responses; one tie) was also significant ($z = 2.01, p < .025$).

In the free-choice condition, that is, vowel responses were much more likely than consonant responses.

Experiment 1 has thus clearly shown that Dutch listeners can more easily substitute vowels than consonants in order to reconstruct a real word from a nonword. The RTs in the consonant substitution condition were the longest, and the error rates were the highest. The participants

were much more likely to make an erroneous vowel substitution in the consonant substitution condition than vice versa, and when given a free choice as to which type of phoneme to substitute, they were significantly more likely to choose a vowel.

EXPERIMENT 2

In Experiment 2, we conducted a parallel experiment in Spanish, which, as was pointed out above, differs from Dutch in that the phoneme repertoire of Spanish is highly asymmetric, with four times as many consonants as vowels, and differs further in that the five vowels of Spanish are acoustically distinct. If the vowel preference depends on search set size—namely, that it is simply easier for listeners to search for a phoneme substitution candidate in a smaller set (of vowels) than in a larger set (of consonants)—the effect should be highly magnified in Spanish. (Note that this explanation is also rendered somewhat doubtful by the Dutch result, given that the consonant set of Dutch is hardly larger than the vowel set.) If, on the other hand, the vowel preference depends on intrinsic contextual variability of vowels, it should be found universally, and hence, it should appear to an equivalent extent in Spanish and in Dutch (or in English). And finally, if the vowel preference depends on acoustic closeness of vowel candidates, producing a cumulative listener experience of vowel confusions, then in Spanish, with its acoustically distinct vowels, the effect should presumably not appear at all.

Method

Subjects. Sixty-four students of psychology at the University of Barcelona, all native speakers of Spanish, participated as volunteers in the experiment. For parity with Experiment 1, the data for 45 subjects were analyzed, 15 for each order of presentation condition. The selection criterion was overall error rate: The excluded subjects made more errors than the selected subjects.

Materials. Sixty nonwords were constructed as stimuli. Each nonword could be turned into a real Spanish word by changing either one of its vowels or one of its consonants (e.g., *pecto*, which can become *pacto*, [pact] or *recto* [straight]). Again, all the stimuli were disyllabic. In half the materials, the word resulting from a vowel change had a higher frequency of occurrence than the word resulting from a consonant change; in the other half, the reverse was the case. The mean frequency for vowel substitution was 136 per million; for consonant substitution, it was 199 (using LEXESP; Sebastian-Galles, Marti, Cuetos, & Carreiras, 1996). Vowel and consonant substitution position was controlled as in Experiment 1. In 22 items, the number of vowels equaled the number of consonants. The materials are again available at the first author's Web address.

An additional 70 nonwords were constructed to serve as practice and filler items, in the same manner as that for Experiment 1. The material was again divided into three sets. The stimuli were recorded by a female native speaker of Spanish, in the same way as that for Experiment 1.

Procedure. The subjects were tested individually in sound-attenuated booths. The experiment was programmed in EXPE control software (Pallier, Dupoux, & Jeannin, 1997) and run on a HP-486 microcomputer. The subjects sat in front of the computer and listened to the stimuli through headphones. The instructions were presented on the computer screen, both at the beginning of the ex-

Table 2
Proportions of Correct Responses in the Free-Choice Condition Involving Vowel Versus Consonant Changes, With Mean Response Times (RTs, in Seconds), in Experiment 1

Condition	Proportion	RT
All items		
Vowel responses	57.4	2.07
Consonant responses	42.6	2.20
Subset of items with equal numbers of vowels and consonants		
Vowel responses	63.2	2.45
Consonant responses	36.8	2.90

periment and at the beginning of each experimental condition. Within condition, stimuli were randomized individually for each subject. The subjects typed their word responses via the computer keyboard. Otherwise, the procedure was the same as that in Experiment 1.

Results

Overall comparisons. The listeners' written responses were coded in the same manner as that for Experiment 1. At this point, it was noticed that some of the materials inadvertently allowed the same response as other items or a similar response (e.g., consonant substitution for *ecto* was *esto* [this (masc.)] and for *esla* was *esta* [this (fern.)]). For each such item, the one that any individual subject had heard second was individually determined and discarded. Because of the individual randomization, this procedure resulted in the loss of no items overall, but in the loss of five responses per subject (either one or two per condition). Table 3 presents the mean correct RTs and mean overall error rates for the three conditions.

Separate ANOVAs, with subjects and items as random factors, were again carried out on the RTs and on the error rates. These analyses again showed a significant effect of task condition on both measures [$F_1(2,88) = 20.48$, $p < .001$, and $F_2(2,14) = 12.68$, $p < .001$, for RTs; $F_1(2,88) = 60.03$, $p < .001$, and $F_2(2,118) = 24.12$, $p < .001$, for errors]. Subsequent t tests showed that RTs were significantly faster and error rates significantly lower for the free-choice condition than for the consonant substitution condition [$t_1(44) = 6.61$, $p < .001$, and $t_2(56) = 4.73$, $p < .001$, for RTs; $r_1(88) = 11.12$, $p < .001$, and $t_2(59) = 7.14$, $p < .001$, for errors], and RTs were significantly faster for the free-choice condition than for the vowel substitution condition [$t_1(44) = 4.86$, $p < .001$; $t_2(57) = 2.75$, $p < .01$]. The difference between error rates in the free-choice condition and those in the vowel substitution condition was just significant across subjects but insignificant across items. RTs in the vowel and consonant conditions differed significantly from one another across items, but not across subjects; error rates in the vowel substitution condition were significantly lower [$t_1(44) = 8.38$, $p < .001$; $t_2(59) = 4.21$, $p < .001$]. Again, the ordering of conditions was the same as that in Table 3 for the conditions that the listeners heard first on the error measure, although on RT, the order of vowel and consonant substitution was reversed.

Similar analyses were carried out with only the subset of 22 items that had equal numbers of vowels and consonants. There was again a significant effect of task condition on both RTs and errors [$F_1(2,88) = 9.82$, $p < .001$, and $F_2(2,38) = 5.49$, $p < .005$, for RTs; $F_1(2,88) = 18.67$, $p < .001$, $F_2(2,42) = 6.76$, $p < .005$, for errors]. For this subset, t tests showed that RTs were significantly slower for the consonant substitution condition than for the free-substitution condition [$t_1(44) = 4.18$, $p < .001$; $t_2(19) = 3.36$, $p < .005$], whereas RTs in the vowel substitution condition differed from those in the other two conditions only across subjects, not across items. For this subset, t tests on the error rates showed that there were significantly more errors in the consonant substitution condition than in the free-substitution condition [$t_1(44) = 9.6$, $p < .001$; $t_2(21) = 3.77$, $p < .001$], whereas the differences in error rates between the vowel substitution condition and the other two conditions was again significant across subjects but did not significantly differ across items.

Thus again, as was predicted, the listeners found the free-substitution condition easiest: They made fewer errors and responded more rapidly. As in Experiment 1, furthermore, there were signs of an advantage for vowel over consonant substitutions.

Vowel versus consonant substitution conditions.

The number of intrusions was tallied for each condition. One hundred and fourteen vowel intrusions occurred in the consonant condition (52 of them involving words with equal numbers of vowels and consonants), and 50 consonant intrusions occurred in the vowel condition (22 of which involved words with equal numbers of vowels and consonants). Wilcoxon matched-pairs signed-ranks tests were carried out across subjects and items. Twelve subjects made no intrusions, 2 made equal numbers of vowel and consonant intrusions, 21 made more vowel than consonant intrusions, and 10 made more consonant than vowel intrusions. The difference between the latter two sets was significant ($z = 2.84$, $p < .005$). Eleven items elicited no intrusions, 4 elicited an equal number of vowel and consonant intrusions, 33 received more vowel than consonant intrusions, and 12 received more consonant than vowel intrusions. The difference between the latter two sets was again significant ($z = 3.06$, $p < .005$). The same analysis was carried out for the subset of 22 items with equal numbers of vowels and consonants. Seventeen subjects made no intrusions, 2 made equal numbers of vowel and consonant intrusions, 18 made more vowel than consonant intrusions, and 8 made more consonant than vowel intrusions. The difference between the latter two sets was significant ($z = 2.55$, $p < .02$). Four items elicited no intrusions, 1 elicited an equal number of vowel and consonant intrusions, 15 received more vowel than consonant intrusions, and 2 received more consonant than vowel intrusions. The difference between the latter two sets was again significant ($z = 2.36$, $p < .03$). Just as in Experiment 1, therefore, vowels were here significantly more likely to intrude upon consonant substitution than the reverse.

Table 3
Mean Correct Response Times (RTs, in Seconds)
and Mean Error Rates (%), for the Three
Task Conditions in Experiment 2

Condition	RT	Error Rate
All items		
Free-choice	1.81	23.2
Vowel change	2.24	29.3
Consonant change	2.55	52.0
Subset of items with equal numbers of vowels and consonants		
Free-choice	1.69	24.0
Vowel change	1.98	36.5
Consonant change	2.61	56.3

Again, we compared performance in these conditions separately for items for which the vowel change word had higher versus lower frequency than the consonant change word. For both sets of words, RT was again significantly faster and error rate significantly lower in the vowel substitution condition, although again the difference between the two conditions was larger when the vowel change word had higher frequency (805 msec, 27.2% error) than when it had lower frequency (189 msec, 17.2% error).

Free-choice condition. Table 4 presents the respective proportions of correct responses in this condition involving vowel versus consonant substitutions and the mean RTs associated with these responses, for the whole set of materials and for the subset of 22 items matched for number of vowels and consonants.

The mean RT to vowels and consonants was not significantly different, either in the whole set of materials or in the matched subset. Wilcoxon tests were again carried out on the proportions of vowel versus consonant substitutions. Across all items, 32 subjects made more vowel than consonant responses, and 10 subjects made more consonant than vowel responses (with three ties), whereas 39 items received more vowel than consonant responses and 18 items more consonant than vowel responses (with two ties). The advantage for vowel over consonant substitutions was significant across both subjects ($z = 4.05, p < .001$) and items ($z = 2.5, p < .02$). The same analysis for the matched subset only also was significant across both subjects (33 subjects more vowel responses, 4 subjects more consonant responses, eight ties; $z = 4.64, p < .001$) and items (14 items more vowel responses; 5 items more consonant responses; two ties; $z = 2.09, p < .05$). Thus again, just as in Experiment 1, the free-choice condition produced a significantly greater likelihood of vowel responses than of consonant responses.

Cross-Linguistic Comparisons

Finally, joint analyses of Experiments 1 and 2 were conducted to explore the significance of the differences across experiments.

Overall comparisons. Combined ANOVAs were conducted for the RTs and for the errors. The Spanish subjects had faster RTs than the Dutch subjects, but this difference was significant only across subjects [$F_1(1,88) = 15.33, p < .001$]; across items, the significance of the difference was marginal. It is reasonable to assume that this

difference in speed of response resulted from a difference in the speed-accuracy tradeoff adopted by the two subject groups, since the proportion of errors was significantly higher for the Spanish group than for the Dutch group [$F_1(1,88) = 19.99, p < .001$; $F_2(1,118) = 7.02, p < .01$]. Importantly, however, on neither measure was there any trace of an interaction between language group and task condition (three of the four F s < 1). The interaction was likewise insignificant on both measures in a combined analysis across only each experiment's subset of items with equal numbers of vowels and consonants.

Vowel versus consonant substitution conditions. A chi-squared test showed that the pattern of intrusions in these two constrained conditions did not approach being significantly different in the two language groups, either for all items or for the matched subsets.

Free-choice condition. A cross-group analysis of the number of vowel versus consonant substitutions in this condition revealed no significant intergroup differences. An analysis of the RTs for vowel and consonant responses, respectively, revealed a significant interaction between language group and vowel versus consonant RT in the whole set of materials [$F_1(1,88) = 16.42, p < .001$], but this interaction was not significant across items and did not reach significance across either subjects or items in an analysis over only the matched subsets.

Thus, the two language groups effectively produced the same patterns of response in all aspects of the word reconstruction task.

GENERAL DISCUSSION

Our two experiments have shown a clear parallelism in the way Dutch listeners and Spanish listeners perform the word reconstruction task. Both these populations, like the English listeners in the earlier study of van Ooijen (1996), found it demonstrably easier to find a real word from a nonword input by altering a vowel than by altering a consonant. Three significant asymmetries emerged: The listeners' responses were faster and more accurate when they were constrained to substitute vowels than when they were constrained to substitute consonants; they made more vowel intrusions upon consonant substitution than vice versa; and when allowed a free choice, they made significantly more vowel than consonant alterations. The same pattern of results was observed in the materials as a whole and in a subset of the materials in which, unusually for European languages, the opportunity for vowel versus consonant change was exactly matched.

We cannot, of course, make a definitive claim on the basis of evidence from just three languages that the vowel preference is universal; the results are certainly consistent with a universal preference, however, and we are inclined to adopt it as a provisional assumption. It is, nevertheless, clear that we can reject several alternative hypotheses raised in the introduction.

First, we can reject an explanation based on the acoustic similarity of the members of a crowded vowel reper-

Table 4

Proportion of Correct Responses in the Free-Choice Condition Involving Vowel Versus Consonant Changes, With Mean Response Times (RTs, in Seconds), in Experiment 2

Condition	Proportion	RT
All items		
Vowel responses	60.97	1.85
Consonant responses	39.03	1.71
Subset of items with equal numbers of vowels and consonants		
Vowel responses	67.93	1.79
Consonant responses	32.07	1.55

toire. Because English has many similar vowels, English listeners might have built up experience of having to alter initially faulty hypotheses about vowels. This should not be true of Spanish, since the five vowels of Spanish are acoustically distinct from one another. Our results clearly showed that the vowel preference held as strongly in Spanish as in Dutch (and in English), and in consequence, we consider the acoustic similarity hypothesis disproven.

Second, we can also reject the hypothesis that ascribed the English vowel preference to asymmetry in the numbers of vowels versus consonants in the language. The prediction for this case was a significantly stronger vowel preference in Spanish (with its 4:1 consonant-to-vowel ratio), but an attenuated preference or none at all in Dutch (with its relatively balanced ratio). We did not find such an asymmetry: The vowel preference patterned similarly in both languages.

Note that language-specific opportunities for contextual variability do affect listening. Costa et al.'s (1998) Spanish-Dutch comparisons examined the relative size of variability effects in phoneme detection, whereby predictable contexts (e.g., detection of /p/ in *so lo mo po*) produce faster detection responses than variable contexts (e.g., /p/ in *su le mi po*). Costa et al. found that the effects of vowel variability on consonant detection and of consonant variability on vowel detection were virtually equivalent in Dutch, whereas in Spanish, the effects were asymmetrical: Consonant variability had more effect on vowel detection than vowel variability had on consonant detection. This result suggests that listeners are, in fact, highly sensitive to the potential for contextual variability offered respectively by a balanced versus an asymmetric consonant/vowel ratio; such sensitivity does not seem to determine word reconstruction performance, however.

We can further discount the possibility that the English result could reflect the fact that the strongest cues to English dialectal differences are found in vowels. British English listeners are indeed accustomed to guessing speakers' regional origins from their vowels; but in Dutch and Spanish, dialectal cues pattern somewhat differently than in English. In Dutch, both vowels and consonants differ across regional variants, but listeners are, if anything, more accustomed to guessing regional origins from consonants—the hard *g* of the northern Netherlands versus the soft *g* of the south and Belgium, the devoiced initial fricatives of the western cities, and so on. The same is true of Spanish, where fricatives, in particular, give clues to regional origin; a Castilian speaker's dental fricative contrasts with the alveolar fricative used by a speaker from the south of Spain or from the Canary Islands. Thus, neither Dutch nor Spanish listeners should feel more inclined to alter vowels than consonants simply on the basis of experience in guessing dialects. The parallelism of the present results for Dutch and Spanish with van Ooijen's (1996) results from English suggest that dialectal variation was not crucial to the original English finding.

Only one hypothesis considered in the introduction explicitly predicted the observed result—that is, a significant vowel preference in both Dutch and Spanish—and that is the suggestion that listeners' performance is determined by intrinsic differences between the information provided by vowels and that provided by consonants. Some vowel-consonant differences appear cross-linguistically and may reflect such inherent differences. For instance, in vowel detection, a negative correlation between target duration and response time—the longer the vowel, the faster it is detected—has been observed in English (with its many confusable vowels), but also in Spanish (with its few, distinct vowels); this effect has been attributed (Cutler et al., 1996; van Ooijen et al., 2000) to listener caution in vowel detection, owing to accumulated experience of vowel variability in context. Such a hypothesis would then ascribe the vowel substitution preference in word reconstruction also to listeners' awareness that vowels are intrinsically more susceptible to contextually determined variation than consonants are. The vowel preference should then, indeed, be universal and should be unaffected by whether a language has many vowels or few, a symmetric or an asymmetric consonant/vowel ratio, and vowels that are confusable or distinct. Certainly, it should be observable, as we have shown, in Dutch and Spanish, as well as in English.

We note here that supporting evidence is available from an investigation by Marks, Moates, Bond, and Vazquez (2000), who also studied word reconstruction by Spanish speakers—namely, Spanish-English bilinguals (with Spanish as first language) residing in the United States. They, too, observed a significant vowel preference, in agreement with the results reported here. Additional evidence is available from Japanese; using a similar task, in which listeners reconstructed words by altering a consonant-vowel portion, Cutler and Otake (2000) found that preserved consonant information was more useful to listeners than was preserved vowel information. In this latter study, listeners were not explicitly manipulating phonemes, and yet a phoneme-based difference appeared in the relative utility of the information within the altered portion.

Thus, there is clearly an underlying motivation that is equivalent cross-linguistically for the vowel substitution preference in the word reconstruction task. Listeners find it easier to change a vowel than a consonant to construct a lexical hypothesis from a nonword input, regardless of the relative number or distinctiveness of the phonemes in their native language. Some aspect of listeners' experience with spoken language has led to this apparently robust behavioral preference. There are several ways in which a readiness to alter vowels could be a learned behavior in speech recognition. If phonetic context causes, on average, more variability in vowels than in consonants, listeners might, as a result, build up experience of being unsure about which vowel they have heard and of having to alter initial hypotheses about vowels when these hypotheses do not activate words. That is, listeners may sim-

ply have learned to regard vowel information as less reliable than consonant information. Another reason could be related to the structure of words in vocabularies and the consequences of this for the relative availability of lexical neighbors. As was discussed above, the opportunity to turn one word of any language into another by replacing a consonant is, in general, much greater than the same chance with replacement of a vowel. Listeners could have discovered that making a mistake about a consonant is likely to activate a lexical neighbor by accident, whereas making a mistake about a vowel involves rather less risk of unwanted neighbor activation. On this account, listeners may simply have learned that vowel mutability incurs less cost for the word recognition process than consonant mutability would. On either account, listener experience is consistent with intrinsically mutable vowel hypotheses.

Other evidence supports both of these suggestions. Vowel variability is well attested even in an invariant consonantal context (Hillenbrand, Getty, Clark, & Wheeler, 1995; Peterson & Barney, 1952), and the perceptual results of these studies show that listeners fail to agree on outlying tokens of vowel types. Thus, as Rosner and Pickering (1994) describe it, "the variability that a speaker accepts exceeds the variability of productions identified with high accuracy by a listener" (p. 101). Koopmans-van Beinum (1980) concluded similarly, on the basis of a finding that only one third of vowels excerpted from spontaneous Dutch conversation were correctly identified, that "perception of vowels in free conversation is based for a small part only on the stationary section of the vowel" (p. 151). Even in a language with few vowels, variability is far greater in connected speech than in isolated word production (Keating & Huffman, 1984).

Also, it is easy to establish that the relative contributions to vocabulary structure of vowels and consonants are not balanced. For the CELEX English database, we examined words from 2 to 15 phonemes in length; there were 2.2 times as many neighbors resulting from a consonant replacement (e.g., *pat* as a neighbor for *cat*) as from a vowel replacement (e.g., *kit* as a neighbor for *cat*). The same calculation for Dutch in CELEX produced 1.72 neighbors from consonant replacement for every neighbor from vowel replacement, whereas for a Spanish lexical database of over 75,000 words (Sebastian-Galles et al., 1996), there were 2.07 neighbors from consonant replacement for every neighbor from a vowel replacement. Thus, the asymmetry is comparable in the three languages in which the vowel preference in word reconstruction has been observed.

Both hypotheses are difficult to test directly. It is, as far as we know, impossible to find—at least for purposes of psycholinguistic testing—vocabularies in which replacing a vowel produces a greater likelihood of turning one word into another than does replacing a consonant. Thus, we cannot test in a language in which this hypothesis would predict an absent or reversed preference in

word reconstruction. Without computationally available lexica, which exist for only relatively few languages so far, it is even hard to establish less asymmetry than in the languages examined above. Likewise, it is difficult to test directly the relative weight of vowel variability against consonant variability (also, of course, widespread) in listeners' perceptual processing of speech.

However, our demonstration of a vowel substitution preference in languages as crucially different as Dutch and Spanish has enabled us to dispose of several hypotheses. Certainly, the preference appears in listeners from diverse language backgrounds, and it may well be universal. Regardless of past linguistic experience, listeners think a *kebra* is more like a *cobra* than like a *zebra*.

REFERENCES

- BAAYEN, R. H., PIEPENBROCK, R., & VAN RUN, H. (1993). *The CELEX lexical database* [CD-ROM]. Philadelphia: University of Pennsylvania, Linguistic Data Consortium.
- BOND, Z. S., & GARNES, S. (1980). Misperceptions of fluent speech. In R. Cole (Ed.), *Perception and production of fluent speech*. Hillsdale, NJ: Erlbaum.
- Boou, G. (1995). *The phonology of Dutch*. Oxford: Oxford University Press.
- CONNINE, C. M., BLASKO, D. G., & WANG, J. (1994). Vertical similarity in spoken word recognition: Multiple lexical activation, individual differences, and the role of sentence context. *Perception & Psychophysics*, *56*, 624-636.
- COSTA, A., CUTLER, A., & SEBASTIAN-GALLES, N. (1998). Effects of phoneme repertoire on phoneme decision. *Perception & Psychophysics*, *60*, 1022-1031.
- CUTLER, A., & OTAKE, T. (2000). *Rhythmic categories in spoken-word recognition: Evidence from Japanese*. Manuscript submitted for publication.
- CUTLER, A., VAN OOIJEN, B., NORRIS, D., & SANCHEZ-CASAS, R. (1996). Speeded detection of vowels: A cross-linguistic study. *Perception & Psychophysics*, *58*, 807-822.
- Gow, D. W., & GORDON, P. C. (1995). Lexical and prelexical influences on word segmentation: Evidence from priming. *Journal of Experimental Psychology: Human Perception & Performance*, *21*, 344-359.
- HILLENBRAND, J., GETTY, L. A., CLARK, M. J., & WHEELER, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America*, *97*, 3099-3111.
- KEATING, P. A., & HUFFMAN, M. K. (1984). Vowel variation in Japanese. *Journal of Phonetics*, *41*, 191-207.
- KOOPMANS-VAN BEINUM, F. J. (1980). *Vowel contrast reduction: An acoustic and perceptual study of Dutch vowels in various speech conditions*. Unpublished doctoral thesis, University of Amsterdam.
- MADDIESON, I. (1984). *Patterns of sounds*. Cambridge: Cambridge University Press.
- MARKS, E. A., MOATES, D. R., BOND, Z. S., & VAZQUEZ, L. (2000). *Vowel mutability: The case of Spanish*. Manuscript submitted for publication.
- MARSLÉN-WILSON, W. D. (1990). Activation, competition, and frequency in lexical access. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 148-172). Cambridge, MA: MIT Press.
- MARSLÉN-WILSON, W. D., & WARREN, P. (1994). Levels of perceptual representation and process in lexical access: Words, phonemes, and features. *Psychological Review*, *101*, 653-675.
- MARSLÉN-WILSON, W. D., & ZWITSERLOOD, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception & Performance*, *15*, 576-585.
- MCQUEEN, J. M., NORRIS, D. G., & CUTLER, A. (1994). Competition in

- spoken word recognition: Spotting words in other words. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 20, 621-638.
- MCQUEEN, J. M., NORRIS, D. G., & CUTLER, A. (1999). Lexical influence in phonetic decision-making: Evidence from subcategorical mismatches. *Journal of Experimental Psychology: Human Perception & Performance*, 25, 1363-1389.
- NOOTEBOOM, S. G., & DOODEMAN, G. J. N. (1984). Speech quality and the gating paradigm. In M. P. R. van den Broecke & A. Cohen (Eds.), *Proceedings of the 11th International Congress of Phonetic Sciences* (pp. 481-485). Dordrecht: Foris.
- PALLIER, C., DUPOUX, E., & JEANNIN, X. (1997). EXPE: An expandable programming language for on-line psychological experiments. *Behavior Research Methods, Instruments, & Computers*, 29, 322-327.
- PETERSON, G. E., & BARNEY, H. L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America*, 24, 175-184.
- RADEAU, M., MORAIS, J., & SEGUI, J. (1995). Phonological priming between monosyllabic spoken words. *Journal of Experimental Psychology: Human Perception & Performance*, 21, 1297-1311.
- ROSNER, B. S., & PICKERING, J. B. (1994). *Vowel perception and production*. Oxford: Oxford University Press.
- SEBASTIAN-GALLES, N., MARTI, M. A., CUETOS, F., & CARREIRAS, M. (1996). *LEXESP: Base de datos informatizada de la lengua española* [LEXESP: Computerized database of the Spanish language]. Universitat de Barcelona, Departament de Psicologia Basica.
- SHILLCOCK, R. C. (1990). Lexical hypotheses in continuous speech. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguists and computational perspectives* (pp. 24-49). Cambridge, MA: MIT Press.
- SLOWIACZEK, L. M., & HAMBURGER, M. (1992). Prelexical facilitation and lexical interference in auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 18, 1239-1250.
- SLOWIACZEK, L. M., MCQUEEN, J. M., SOLTANO, E. G., & LYNCH, M. (in press). Phonological representations in prelexical speech processing: Evidence from form-based priming. *Journal of Memory & Language*.
- STOCKWELL, R. P., & BOWEN, J. D. (1965). *The sounds of English and Spanish*. Chicago: University of Chicago Press.
- TABOSSI, P., BURANI, C., & SCOTT, D. (1995). Word identification in fluent speech. *Journal of Memory & Language*, 34, 440-467.
- VAN OOUEN, B. (1996). Vowel mutability and lexical selection in English: Evidence from a word reconstruction task. *Memory & Cognition*, 24, 573-583.
- VAN OOUEN, B., CUTLER, A., SANCHEZ-CASAS, R., & NORRIS, D. (2000). *Vowel and consonant detection in English and Spanish*. Manuscript submitted for publication.
- WHALEN, D. H. (1991). Subcategorical phonetic mismatches and lexical access. *Perception & Psychophysics*, 50, 351-360.
- ZWITSERLOOD, P. (1989). The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition*, 32, 25-64.

(Manuscript received May 10, 1999;
revision accepted for publication October 10, 1999.)