BASE-LANGUAGE EFFECTS ON WORD IDENTIFICATION IN BILINGUAL SPEECH:
EVIDENCE FROM CATEGORICAL PERCEPTION EXPERIMENTS*

JUDITH BÜRKI-COHEN
Northeastern University
FRANÇOIS GROSJEAN
Université de Neuchâtel
and
JOANNE L. MILLER
Northeastern University

The categorical perception paradigm was used to investigate whether French-English bilinguals categorize a code-switched word as French or English on the basis of its acoustic-phonetic information alone or whether they are influenced by the base-language context in which the word occurs, that is, by the language in which the majority of words are spoken. Subjects identified stimuli from computer-edited series that ranged from an English to a French word as either the English or the French endpoint. The stimuli were preceded by either an English or a French context sentence. In accord with previous studies (Grosjean, 1988), it was found that the base language had a contrastive effect on the perception of a code-switched word when the endpoints of the between-language series were phonetically marked as English and French, respectively. When the endpoints of the series were phonetically unmarked and thus compatible with either language, however, no effect of the base language was found; in particular, we failed to find the assimilative effect that has been observed with other paradigms (Grosjean, 1988; Soares and Grosjean, 1984; Macnamara and Kushnir, 1971). The current results provide confirming evidence that the perception of a code-switched word is influenced by the base-language context in which it occurs and, moreover, that the nature of the effect depends on the acoustic-phonetic characteristics of the code-switched word. In addition, the finding that a contrastive effect occurs across all paradigms used to date, but that an assimilative effect occurs in only some paradigms, suggests that these two context effects may arise at different stages of processing.

Key words: Bilingualism, code-switching, categorical perception paradigm, French

INTRODUCTION

Speech perception involves extracting from the environment acoustic information that can be mapped onto linguistic information represented in our minds. As complex as this task is for monolingual listeners, it may be even more complex for bilingual listeners. This is because bilinguals talking among themselves, although they usually agree on one "base language" carrying the bulk of their conversation, have a tendency to switch occasionally to their other shared language – or "code" – for a word, a phrase, or even an entire sentence (see Grosjean, 1982, for a discussion of linguistic and cultural constraints on code-switching).

The perception of code-switched words by bilinguals has been investigated using a variety of experimental paradigms. Soares and Grosjean (1984) used the phoneme-triggered lexical decision task (Blank, 1980) to study lexical access by bilinguals in both a monolingual situation (where only one language is used) and a bilingual situation (where two languages are used, that is, where code-switching occurs). They found that bilinguals in a monolingual situation responded to word targets in both their languages as fast as monolinguals. However, their response times to code-switched words in a bilingual situation were significantly slower. Moreover, in both situations, the bilingual subjects took longer to respond to nonwords than did the monolingual subjects. These results suggest that during lexical access bilinguals use a general strategy of first searching the base-language lexicon and, only if no entry is found, then searching the lexicon of their alternate language. This assimilative effect of the base language may also explain the results of a study by Macnamara and Kushnir (1971), who found that bilingual subjects understand monolingual passages faster than bilingual passages.

More recently, Grosjean (1988) used the gating paradigm (Grosjean, 1980) to investigate the underlying process of lexical access in a bilingual situation. He presented. bilingual subjects with increasingly longer segments of English (and French filler) words embedded in a French sentence. In particular, he used words that were phonotactically marked as English and had no similarly sounding words in French (such as "slash", whose initial consonant cluster is very rare in French), and phonotactically unmarked words that had French counterparts (such as "pick", with the French counterpart "pique"). All words were monosyllabic and had the same uniqueness point, that is, the point in the left-to-right sequence of phonemes at which the word distinguishes itself from every other word (Marslen-Wilson, 1984). Up to the third gate (80 msec into the word), subjects guessed almost exclusively French words. Over subsequent gates, however, the number of English candidates increased. Moreover, this increase was much steeper for the words that were phonotactically marked as English than for the words with French counterparts. By the fifth gate (160 msec into the word), significantly more English candidates were proposed for the phonotactically marked than for the unmarked words. These results provide additional evidence for an assimilative effect of the base language. However, they also indicate that there may be limits to the assimilation of acoustic input to the base language. Under certain conditions, the base language may have the opposite effect, serving as a contrasting background against which conflicting acoustic information is especially conspicuous. In this case, the initial search of the base-language lexicon

would be immediately aborted in favor of accessing the alternate-language lexicon. Indeed, the words that were phonotactically marked as English were recognized at earlier gates not only compared to the phonotactically unmarked words that had French counterparts, but also compared to a set of phonotactically unmarked words that had no French counterparts.

In the present study, we used the categorical perception paradigm to gain additional evidence that both base language and phonetic structure affect the perception of code-switched words. The categorical perception paradigm, which has been widely used to investigate monolingual speech perception (see Repp, 1984, for an overview), was developed by Liberman and his colleagues at Haskins Laboratories (Liberman, Harris, Hoffman, and Griffith, 1957). Modifying formant transitions by means of their newly constructed pattern playback system, Liberman *et al.* synthesized an acoustic continuum ranging from [be] to [de] to [ge]. When subjects were asked to identify the stimuli of the series, they divided them into three sharply defined categories, [be], [de], and [ge]. When asked to discriminate between stimuli, they discriminated best those stimuli labeled differently in the identification task, and worst those stimuli labeled as belonging to the same category.

The categorical perception paradigm has proven particularly fruitful for the study of context effects in speech perception. It has been established that the location of the category boundary, defined as the point along the series where a stimulus word is identified with equal probability as belonging to either category, is not solely based on the acoustic information contained in the stimulus, but can be influenced by information contained in the preceding sentence, such as speaking rate, syntactic structure, and semantic plausibility (see Repp and Liberman, 1987, for a discussion of these findings).

In the present study, we used the categorical perception paradigm as a tool to investigate context effects in the perception of speech by bilinguals, in particular, to assess whether bilinguals identify a stimulus as a member of the lexicon of one or the other language on the basis of the acoustic information contained in the stimulus alone, or whether they are influenced by the language in which the preceding words are spoken, that is, by the base language.[1] For this purpose, we constructed stimulus series ranging from an English to a French word such that the closer a stimulus of the series was to the French endpoint of the series, the more of the French word it contained, and the closer a stimulus was to the English endpoint, the more of the English word it contained. Our first question was whether the perception of such a between-language series would be influenced by the base language in which the words were presented. While we expected that the English and French endpoints, and the near-endpoint stimuli, would be identified as the English or French word regardless of context, we hypothesized that the base language would have an effect on the identification of the stimuli in the middle of the series.

---

[1] In previous applications of the categorical perception paradigm to the investigation of speech perception in bilinguals, researchers were mainly interested in comparing the identification results of bilinguals with those of monolinguals (see, among others, Elman, Diehl, and Buchwald, 1977).

Our second question was whether the nature of the base-language effect, if it exists, would depend on the acoustic-phonetic make-up of the code-switched word, in particular, on the compatibility of the code-switched word with the phonetic structure of the base language. We thus tested two between-language series, a "language-neutral" and a "language-selective" series. The endpoints of the "language-neutral series", English "day" and its French counterpart "dé", were chosen so that they would be similar to each other with no prominent phonetic cues marking them as either English or French. They were thus considered to be relatively compatible with either language. For this series, we expected to replicate the assimilative effect of the base language found by Grosjean with English words that sounded similar to French words (see above), so that the middle stimuli would be perceived as French in French context and as English in English context. On the other hand, the endpoints of the "language-selective series", English "ray" and French "ré", were chosen so that their phonetic make-up (viz., their initial consonant) would betray them immediately as English or French. They were thus compatible with only one of the two languages. Based on Grosjean's gating results with phonotactically marked English words in French context (see above), we expected that in French context, the middle stimuli of this series would be perceived as English, whereas in English context, they would be perceived as French. In other words, the base language would serve as a contrasting background against which the middle stimuli would be evaluated.

In summary, then, we predicted both an effect of the base language on the perception of code-switched words and an interaction of this base-language effect with the acoustic-phonetic make-up of the code-switched items. The base language should have an assimilative effect on the perception of language-neutral stimuli and a contrastive effect on the perception of language-selective stimuli.

Before proceeding to test our two experimental questions, it was important to establish that the typical categorical perception results obtained with within-language series would generalize to our two English-French between-language series. Our first experiment, therefore, consisted of a forced-choice identification task performed on the "day–dé" and the "ray–ré" series in isolation, without base-language context. Based on the standard identification results for within-language series, we expected subjects to divide both series into two discrete categories, one corresponding to the English endpoint, the other to the French endpoint. To assess further the similarities between the perception of between- and within-language series, we also asked subjects to discriminate pairs of stimuli from within each series. Based on the standard results for within-language series, we expected subjects to discriminate best those stimuli that they categorized differently.

EXPERIMENT 1: IDENTIFICATION AND DISCRIMINATION OF BETWEEN-LANGUAGE SERIES IN ISOLATION

Materials

Two English-French series were constructed, one with language-neutral endpoints
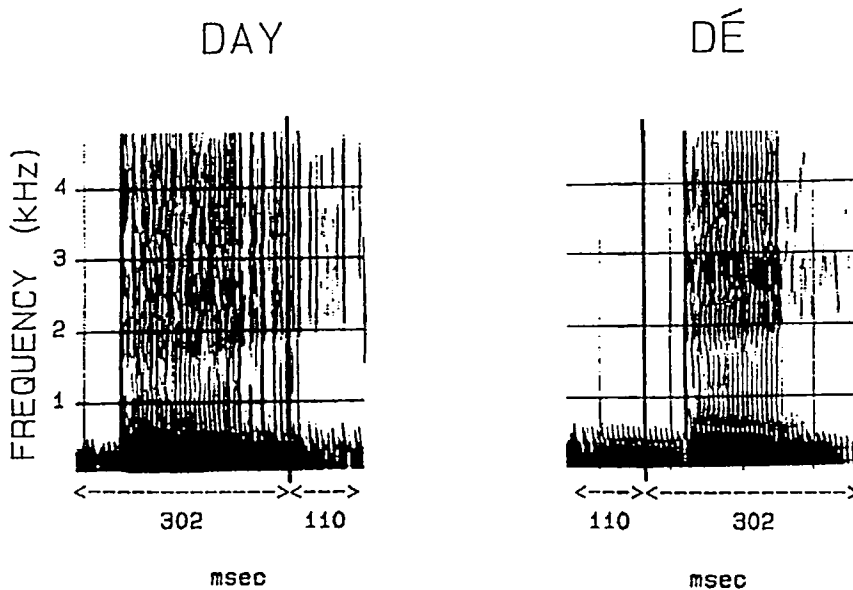
DAY DÉ

Fig. 1.   Spectrograms of the first and last stimulus of the series with language-neutral endpoints, ranging from the naturally produced English "day" (left) to the naturally produced French "dé" (right). The vertical lines mark the points at which the English vowel and the French prevoicing were cut back before constructing the hybrid stimuli.

phonetically compatible with both languages and the other with language-selective endpoints compatible with only one language.

Language-neutral series. English "day" and French "dé" (thimble, dice) were chosen as endpoints that are phonetically compatible with either language. A number of tokens of each word were spoken by a bilingual speaker of British English and Parisian French (FG), who was judged by several colleagues of the authors to have no foreign accent in either language. These tokens were recorded on audiotape and subsequently digitized (sampling rate of 20 kHz, low-pass filtering at 9.8 kHz) and measured by means of a computer waveform editing program implemented on a DEC PDP 11/44 minicomputer. Two well-articulated endpoints of nearly identical duration were chosen, and the cursors at the end of the words were set so that both endpoints measured 412 msec. The spectrograms of these two endpoints are shown in Figure 1. As can be seen from a comparison of the formant trajectories of the two vowels, our speaker pronounced "day" with a fairly high vowel and only minimal diphthongization. The quality of the two vowels is therefore quite similar, even though their duration differs (348 msec vs. 241 msec) (see Ladefoged, 1975, for a discussion of the variability of the diphthong [eɪ] in different forms

of English). As for the consonants, apart from the difference in place of articulation (alveolar vs. dental), they also appear to be distinguished mainly by a durational property, namely, the length of their prevoicing (64 msec vs. 171 msec). Differences in place of articulation (alveolar vs. dental) occur in either language occasionally as a consequence of coarticulation (Ladefoged, 1975), just as do variations in prevoicing and vowel length and quality. Thus the two stimuli do not contain any salient features that would exclude them from being considered as a member of the lexicon of the other language. While still distinguishable, they are about as similar as any words taken from the two languages can be.

Apart from the two endpoints, the 16-member series consisted of 14 hybrid stimuli constructed by replacing acoustic segments of the English endpoint "day" with acoustic segments of the French endpoint "dé". This was done by segmenting the digitized end-points into parts, and then concatenating these parts. All cuts were made at zero-crossings to avoid discontinuities in the waveform. To produce the 14 hybrid stimuli, first the English vowel was cut back by 110 msec to more closely approximate the length of the French vowel, and the French prevoicing was cut back by approximately 110 msec to more closely match the length of the English prevoicing. The cutoff points are indicated on the spectrograms in Figure 1. Then, increasingly longer final segments of English "day" were replaced with increasingly longer final segments of French "dé", so that the proportion of English gradually decreased, while the proportion of French gradually increased. Segments were incremented by approximately 20 msec, so that the first hybrid stimulus consisted of 282 msec or 94% English followed by 19 msec or 6% French, the second hybrid stimulus consisted of 265 msec or 87% English followed by 40 msec or 13% French, and so on up to the 14th and last hybrid stimulus, which contained 18 msec or 6% English followed by 280 msec or 94% French. Each hybrid stimulus had a total duration of 300 msec (within a few msec to allow for the fact that all cuts were made at zero-crossings). The 16-member series thus ranged from the original English "day" to the original French "dé", with 14 shorter hybrid stimuli in between. The stimuli within the series sounded remarkably natural, with no salient perceptual discontinuities within stimuli.

For the identification task, 15 randomized blocks of the 16 stimuli (240 stimuli in all) were recorded on audiotape, with an interstimulus interval (ISI) of 2 seconds and an interblock interval (IBI) of 5 seconds. The discriminability of the stimuli was assessed with a standard ABX task, using only the 14 hybrid stimuli; that is, the endpoints (110 msec longer than the hybrids) were not used in the discrimination task. The twelve two-step pairs (stimulus 2 and 4, 3 and 5, and so on up to stimulus 13 and 15) were tested in four triads for each pair, ABA, ABB, BAA, and BAB, resulting in 48 triads. Each triad was tested four times, resulting in a total of 192 triads. The triads were randomized and recorded in four blocks of 48 triads each, with an ISI of 750 msec, an ITI (intertriad interval) of 3 seconds, and an IBI of 5 seconds. Each pair was thus tested 16 times.

Language-selective series. English "ray" and French "ré" (musical note) were chosen as endpoints that are unambiguously marked by their phonetics as to which language they belong. A number of tokens of the two words were recorded from the same bilingual speaker who had produced the endpoints of the language-neutral series. Again, the tokens

RAY   RÉ

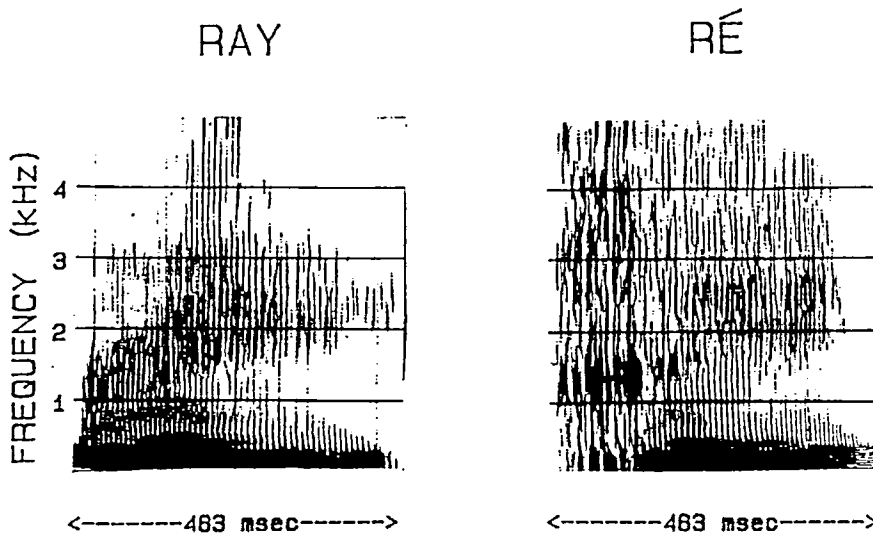<-------483 msec------>   <-------483 msec------>

Fig. 2. Spectrograms of the first and last stimulus of the series with language-selective
endpoints, ranging from the naturally produced English "ray" (left) to the.
naturally produced French "ré" (right).

were digitized and measured. Two well-articulated English and French endpoints of
almost identical duration were found and the cursors at the end of the words set so that
both endpoints measured 463 msec. The spectrograms of the two endpoints are shown in
Figure 2.

These spectrograms illustrate the distinct acoustic properties of the English semivowel
[ɹ] (an alveolar approximant) and the French [ʀ] (a uvular fricative) (Ladefoged, 1975).
Neither of these sounds occurs in the other language. The vowel is again not very
different across the two languages. As was done for the language-neutral series, the eleven
hybrid stimuli of the 13-member language-selective series were constructed by replacing
final segments of the English "ray" with final segments of the French "ré". The last
232 msec of the English vowel were replaced with the last 234 msec of the French vowel
to produce the first hybrid stimulus, which consisted of 50% English followed by 50%
French. For the rest of the series, which involved the substitution of the consonantal
portion, the proportion of English was decreased, and the proportion of French was
increased, in steps of approximately 20 msec. The second hybrid stimulus thus contained
207 msec or 45% English followed by 258 or 55% French, the third hybrid stimulus 188
msec or 41% English followed by 274 msec or 59% French, and so on up to the eleventh
and last hybrid stimulus, which contained 22 msec or 5% English followed by 441 msec

or 95% French. The entire series, ranging from the unaltered English "ray" to the unaltered French "ré", thus contained 13 stimuli, each measuring 463 msec (plus or minus a few msec to accommodate for zero-crossing cuts). Again, the stimuli sounded very natural, with the hybrid stimuli appearing to be somewhat unclearly pronounced, rather than artificially distorted.

The identification test tape consisted of 15 randomized blocks of the 13 stimuli, for a total of 195 stimuli. There was an ISI of 2 seconds and an IBI of 5 seconds. The discriminability of the stimuli was assessed in a standard ABX task. Stimulus 1 was not compared to any other stimulus because of the jump from 0% to 50% French between the first and the second stimulus. The remaining ten two-step pairs (stimulus 2 and 4, 3 and 5, and so on, up to 11 and 13) were tested in four triads per pair, ABA, ABB, BAA, and BAB, resulting in 40 triads. Each triad was tested four times, resulting in a total of 160 triads. The triads were randomized and recorded in four blocks of 40 triads each, with an ISI of 750 msec, an ITI of 3 seconds, and an IBI of 5 seconds. Each pair was thus tested 16 times.

Subjects

Eight French-English bilingual adults (six females and two males) whose first language was French served as subjects. They all were very fluent in English and French and had used both languages on a daily basis for at least seven years, that is, since they had moved to the United States (Boston area). With the exception of one subject, none of them had been exposed to English as a second language before secondary school. There were individual differences in the degree of foreign accent in their spoken English. The main criteria for qualifying as a subject were daily use of both languages and membership in a bilingual community that frequently code-switches.

Procedure

The eight subjects were run individually on both the language-neutral and the language-selective series, with the order of the two series counterbalanced across subjects. Subjects were tested in one session in their home environment. Oral and written instructions were given in French. Subjects listened to the stimuli over binaural headphones at a comfortable listening level.

In the identification task, the subjects were asked to circle the word corresponding to the perceived item on a response sheet (Dé vs. Day; Ré vs. Ray). They were instructed to answer on all trials, even if the words seemed to be somewhere between "day" and "dé", or "ray" and "ré". In the subsequent ABX discrimination task, they were asked to indicate, by writing down a 1 or a 2, whether they perceived the third stimulus (X) in the triad as the same as the first (A) or the second (B) stimulus. They were instructed to answer even if the stimuli all sounded the same to them. None of the subjects asked questions about the origin of the stimuli.

Results and discussion

The group results for the eight subjects in the identification and discrimination tasks

are displayed in Figure 3 ("day–dé") and Figure 4 ("ray–ré"). Percent identification as the English endpoint ("day" or "ray") of the series and percent correct discrimination are plotted against the stimulus number along the series. The data points in the discrimi-nation curve refer to discrimination between the two neighboring stimuli. For example, a value of 50% at stimulus number 3 means that stimuli 2 and 4 were discriminated correctly half the time, which, for this task, is chance.

As can be seen, performance on the identification task was extremely orderly for both series. Furthermore, for both series, subjects divided the stimuli into two rather sharply defined categories, and did not perceive the hybrid stimuli as gradually going from English to French. Nor was there any indication that they perceived any discontinuities in the construction of the stimuli: For example, despite the fact that the endpoints of the "day–dé" series were 110 msec longer than each of the hybrid stimuli, the stimuli next to the endpoints were identified virtually as often as "day" or "dé" as the endpoints themselves (virtually 100% of the time). Similarly, the jump from 0% to 50% French between the first and the second stimulus of the "ray–ré" series had no observable consequences as they were both identified as "ray" 100% of the time.

The mean category boundary[2] lies at the stimulus value of 9.37 for the "day–dé" series, where the stimulus would be approximately 55% French, and at 7.50 for the "ray–ré" series, where the stimulus would be approximately 75% French. Visual inspection of the spectrograms of the "day–dé" stimuli does not reveal any obvious reason for the specific location of the category boundary. Inspection of the spectrograms for the stimuli in the "ray–ré" series, however, reveals that the hybrid stimuli were labeled as French "ré" as soon as the characteristic uvular frication of the French [ʁ] appears, i.e., at stimulus 8.

Converging evidence for the location of the category boundary for each series is provided by the discrimination data. For the "day–dé" series, the best mean discrimi-nation performance was obtained for stimuli 8–10, precisely those stimuli closest to, but on either side of, the mean category boundary (9.37). For the "ray–ré" series, the best mean discrimination performance was obtained for stimulus pair 7–9, which also straddles the mean category boundary (7.50).

In summary, our results indicate that subjects divide between-language series into sharply defined categories and discriminate best those stimuli that they categorize differently. Experiment 1 thus shows that the perception of a computer-edited series between monosyllabic words in two different languages is very similar to the perception of a series between two words in the same language.

---

[2] All boundary values were calculated by fitting a linear regression line to the data in the boundary region of the individual subject's identification function (i.e., excluding the data in the tails of the function) and taking as the boundary the stimulus value that corresponded to 50% identification as the English endpoint ("day" or "ray").
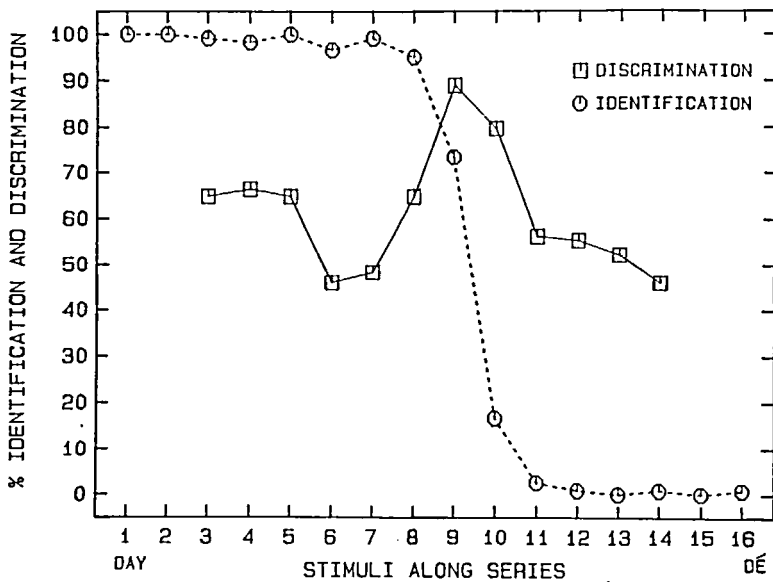
Fig. 3. Group results for the identification as "day" (circles) and correct discrimination (squares) of the language-neutral series, "day–dé". Each square shows the discrimination performance for the two neighboring stimuli.
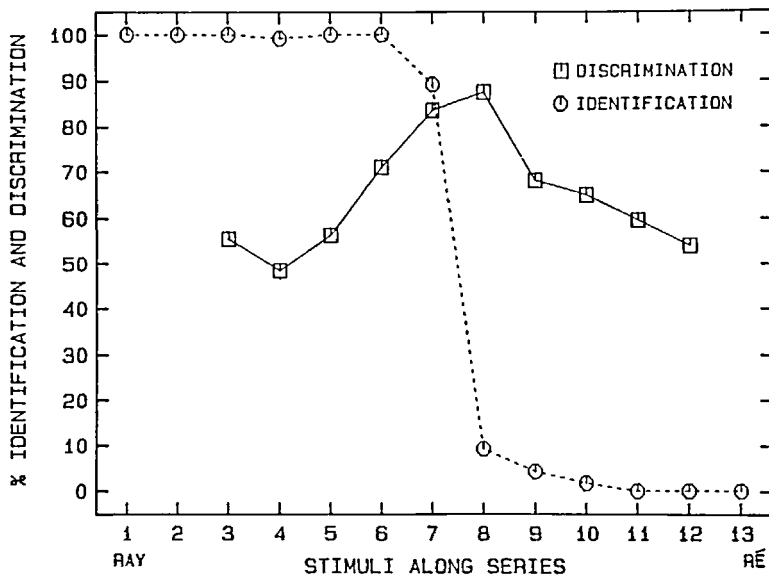

Fig. 4. Group results for the identification as "ray" (circles) and the correct discrimination (squares) of the language-selective series, "ray–ré". Each square shows the discrimination performance for the two neighboring stimuli.

EXPERIMENT 2: IDENTIFICATION OF BETWEEN-LANGUAGE SERIES IN
ENGLISH AND FRENCH CONTEXTS

Experiment 1 demonstrated that a between-language series is perceived very much like
a series within one language. We could thus proceed to test our two experimental
questions, namely, whether categorization of the stimuli is influenced by the base-
language context and, if so, whether the nature of this base-language effect depends
on the acoustic-phonetic make-up of the words used to construct the series.

As we discussed in the introduction, there is evidence that bilingual subjects use a
general strategy of trying to map incoming acoustic information onto representations of
the base language (Soares and Grosjean, 1984; Grosjean, 1988). In other words, the base
language seems to have an assimilative effect on the perception of a code-switched word.
Grosjean (1988) showed, however, that this assimilative effect of the base language
may be limited to code-switched words that are phonetically compatible with the base
language. In the case of words carrying distinct acoustic-phonetic cues revealing their
membership in the alternate-language lexicon, the base language may have a constrastive
effect. On the basis of this evidence, we made the following predictions. First, we
expected the middle stimuli of our series to be perceived differently depending on the
preceding base language. Second, we expected the nature of this base-language effect to
vary with the nature of the series. For the language-neutral series, we expected the base
language to have an assimilative effect on the middle stimuli, so that the middle stimuli
would be perceived as the English endpoint in English context and as the French end-
point in French context. For the language-selective series, however, we expected the base
language to have a contrastive effect on the middle stimuli, so that they would be
perceived as the English endpoint in French context and as the French endpoint in
English context. In other words, we expected the middle stimuli of the language-neutral
series to be perceived as a base-language word but the middle stimuli of the language-
selective series to be perceived as a code-switched word.

Materials

An English and a French context sentence similar in meaning and length were chosen.
In light of the task to be performed by the subjects, English "We have to categorize
(ray/day)" and its French translation "Il faut qu'on catégorise (ré/dé)" seemed
appropriate. It was necessary to construct new series since words spoken in context do
not have exactly the same acoustic properties as the same words spoken in isolation.
The accent-free French-English male bilingual who produced the endpoint stimuli in
Experiment 1 also produced the stimuli for this experiment. He read 10 tokens each of
the English base-language sentence ending with English "day" or "ray", and 10 tokens
each of the French base-language sentence ending with French "dé" or "ré".

Language-neutral series. To construct the language-neutral series in context, we
digitized the 10 English sentences ending with "day" and the 10 French sentences ending
with "dé" and measured the duration of the context up to the closure before "day"/
"dé", the duration of the closure, the voice-onset time (VOT) of "day"/"dé" and the
total duration of "day"/"dé" The English and French sentences closest to the average

on all measures were chosen to construct the series, which was done by digital splicing. English "day" (with a VOT of 17 msec and a duration of 318 msec, cut back slightly from 324 msec to be exactly twice as long as the French endpoint) and French "dé" (with a VOT of 29 msec and a duration of 159 msec) were spliced out to provide the two end-points. Then, increasingly longer initial segments of "day" were replaced with increasingly longer initial segments of "dé" to create the hybrid stimuli of the series. Since "day" was twice as long as "dé", the "day" segments were incremented in steps of 20 msec, whereas the "dé" segments were incremented by 10 msec (within 1 or 2 msec to allow for zero-crossing cuts). For the first hybrid stimulus, consisting of 3% French followed by 97% English, the initial 21 msec of "day" were replaced with the initial 10 msec of "dé". To construct the second hybrid stimulus, consisting of 6% French followed by 94% English. 39 msec of "day" were replaced with 18 msec of "dé" and so on, until all of "day" had been replaced with "dé" from beginning to end, and the duration of the stimuli along the series had decreased from 318 msec for the English endpoint to 159 msec for the French endpoint. The series, including the two endpoints, had 17 stimuli, which were each concatenated to the English (1214 msec) and to the French (1212 msec) context sentences.[3]

Language-selective series. The language-selective series in context was constructed in a similar fashion. The 10 English sentences "We have to categorize ray" and the 10 French sentences "Il faut qu'on catégorise ré" were digitized, and the duration of the context up to "ray"/"ré" and the duration of "ray"/"ré" itself were measured. One sentence in each language with contexts and stimulus words closest to the average on both measures was chosen. "Ray" (298 msec) and "ré" (224 msec) were then spliced out to create the series. The first stimulus of the series was the unaltered "ray". To produce the hybrid stimuli, increasingly longer initial segments of "ray" were replaced with increasingly longer initial segments of "ré". Since "ré" was only 75% of the length of "ray", roughly every 13.5 msec of "ray" were deleted and replaced with 10 msec of "ré". The second stimulus contained 13 msec of "ré" followed by 282 msec of "ray", the third stimulus contained 21 msec of "ré" followed by 272 msec of "ray", and so

---

[3] Note that in Experiment 1 we started substitution from the end of the word whereas, in this experiment, we started substitution from the beginning. The reason has to do with the acoustic-phonetic characteristics of "dé" and "day" when spoken in context, as compared to isolation. In context the two words were spoken with no prevoicing, and the VOT value for "dé" was longer than that for "day". The consequence was that, had we started substitution from the end of the word, we would have created some stimuli that contained a sequence of burst/aspiration and vowel of "day", followed by aspiration and vowel of "dé". This unacceptable sequence could be avoided simply by starting substitution from the beginning of the word, as we did. The reason that substitution from the end of the word was not a problem in Experiment 1 was that both "dé" and "day" were spoken with prevoicing in isolation, so that it was possible to equate the length of the prevoicing (and vowel) before constructing the hybrid stimuli, while still keeping the release burst intact. With equally long segments of prevoicing and vowel, substitution from the end of the word never created an unacceptable sequence of the sort described above.

on. By the 12th and last stimulus (263 msec) of the series, the initial 49% of English "ray" was replaced with 50% of French "ré". The original "ré" (39 msec shorter than the last stimulus of the series) was not included in the series. To obtain the test sentences, each stimulus was concatenated to the English (1126 msec) and the French (1121 msec) context sentences.

Test tapes

One English and one French test tape for both series were recorded. Each of the four tapes contained 10 blocks of 12 ("ray–ré") or 17 ("day–dé") sentences. The sentences were randomized within blocks, with an ISI of 2 seconds and an IBI of 6 seconds.

Subjects

Twelve French-English bilinguals (ten females and two males) from the same population as in Experiment 1 served as subjects. One of the subjects had also participated in Experiment 1.

Procedure

Each subject participated in all four conditions of the experiment, with the stimuli presented in two sessions, one for each language context. Which language context and which series was tested first were counterbalanced across subjects. Subjects were again tested individually in their homes. The English and French sessions were held on different days, but no longer than one week apart. During each session, written and oral instructions were given in the respective base language. The experimenter (the first author) was a multilingual speaker who naturally code-switched in all her interactions with the subjects to create a setting in which code switches do occur. Subjects were told that they would be presented with repeated tokens of the context sentence followed by either the English or the French word, and that they were to circle the perceived item on a response sheet (Dé vs. Day; Ré vs. Ray). They were asked to respond even if they were not sure of what they had heard, that is, even if the item seemed to be somewhere between "day" and "dé", or "ray" and "ré". The test sentences were presented over binaural headphones at a comfortable listening level.

Results and discussion

Figure 5 displays the group results for the language-neutral series, "day–dé", in context. As can be seen, subjects divided the language-neutral series into two well-defined categories, with no significant shift of the category boundary as a function of the base language (t (11) = 0.38, p > 0.10). The mean category boundaries were located at stimulus values of 8.46 and 8.54 in the English and French conditions, respectively. This corresponds to a stimulus consisting of approximately 30% French followed by 70% English. Contrary to our prediction of an assimilative effect of the base language, the stimuli of the language-neutral series were categorized independently of the base language in which they were presented.

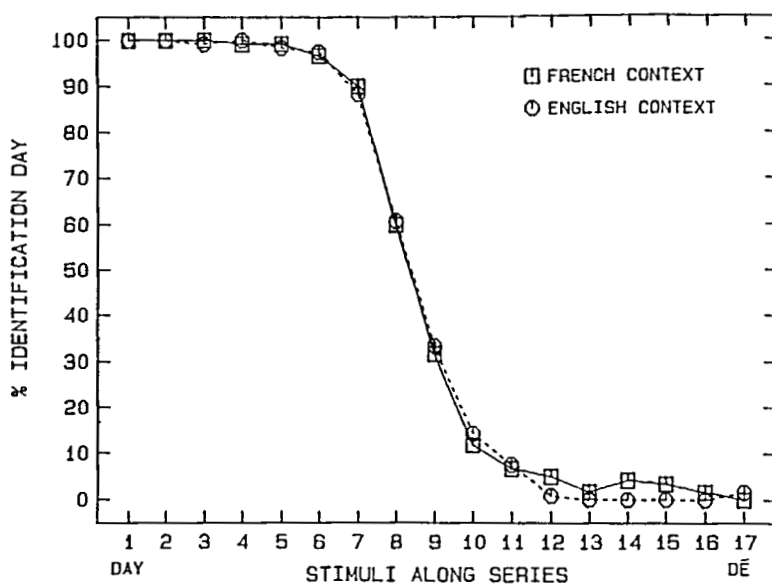Figure 6 displays the group results for the language-selective series, "ray–ré". Again,

Fig. 5. Group results for the identification as "day" of the "day–dé" series in English (circles) and French (squares) context.
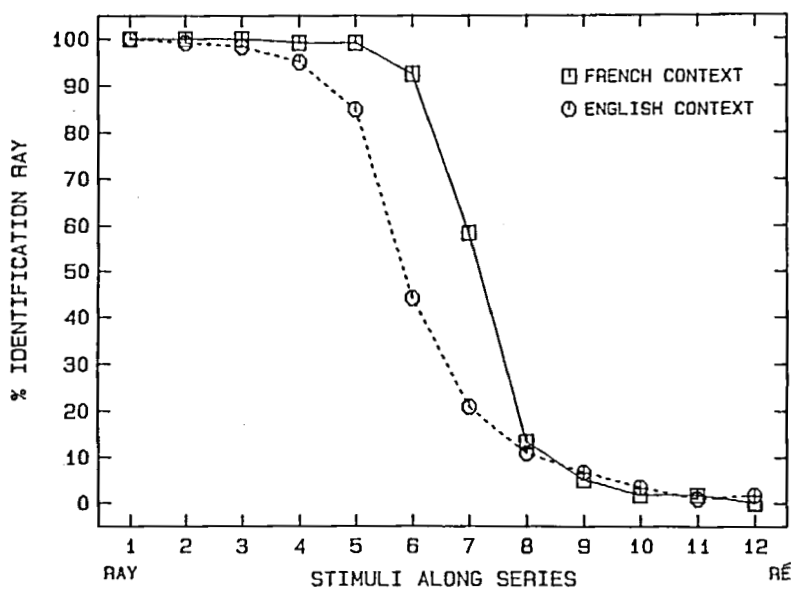

Fig. 6. Group results for the identification as "ray" of the "ray–ré" series in English (circles) and French (squares) context.

subjects divided the language-selective series into two well-defined categories. For this series, however, there was a significant effect of the base language. In English context, the mean category boundary was located at stimulus value 6.18 (which corresponds to a stimulus of approximately 19% French followed by 81% English), whereas in French context, it was located at 7.14 (which corresponds to 23% French followed by 77% English). These results show that in accordance with our predictions, the middle stimuli were categorized as the English endpoint in the French condition and as the French endpoint in the English condition (t (11) = 2.56, $p < 0.03$). In other words, the base language had a contrastive effect on the categorization of the language-selective series.

In summary, although we did not reproduce the assimilative effect of the base language with the categorical perception paradigm, we did find affirmative answers to both our experimental questions. We found that under certain conditions, the base-language context does affect the identification of a code-switched word and, moreover, that this base-language effect interacts with the acoustic-phonetic nature of the code-switched word.

GENERAL DISCUSSION

Experiment 1 demonstrates that bilingual subjects identify and discriminate a stimulus series ranging from a word in one of their languages to a word in their other language very much like monolingual listeners perceive a series within one language. That is, they divide the series into two discrete categories and discriminate best the stimuli they have categorized differently. The identification of a between-language series may thus serve as a tool to investigate context effects in the perception of code-switched words, that is, words that bilinguals bring in from their alternate language when talking to another bilingual.

In Experiment 2 we used this tool to find corroborating evidence for an effect of the base language on the perception of a code-switched word. Evidence for a base-language effect had been previously found in the form of delayed comprehension and lexical decision for code-switched words (Macnamara and Kushnir, 1971; Soares and Grosjean, 1984) and a preponderance of base-language candidates at the initial gates of both code-switched and base-language words (Grosjean, 1988). In addition, we investigated the finding from Grosjean's gating study that the nature of this base-language effect might be influenced by the acoustic-phonetic make-up of the code-switched items, in particular, their compatibility with the phonetic structure of the base language. To this end, we tested the effect of the base language on the identification of both a language-neutral and a language-selective series, that is, a series that was more or less compatible with both languages and a series that was compatible with only one language. And, indeed, we found an effect of the base language, but only on the language-selective series. As expected for this series, the effect was of a contrastive nature, that is, the middle stimuli of the series were identified as the alternate-language endpoint. The expected assimilative effect of the base language on the language-neutral series was not found; the middle stimuli of this series were identified the same in both conditions.

Several possible explanations come to mind for the lack of an assimilative effect of the base language on the perception of the language-neutral series. The first is that a larger step size between the stimuli of the language-neutral compared to the language-selective series may have masked an assimilative effect. In other words, the change from English to French across the series may be too abrupt for the effect of the base language to be evident in the identification data. It is not possible to compare the step sizes of the between-language continua directly, because the stimulus dimensions of the two continua are not commensurate. However, there is an indirect way to address the issue, namely, by comparing the number of stimuli categorized inconsistently in the two series. As can be seen in Figures 5 and 6, the number of stimuli categorized inconsistently was virtually the same in the two series. Under the assumption that only inconsistently categorized stimuli are subject to base-language effects, the base language had the same number of stimuli to act upon in the two series. In other words, the category boundary had the same leeway in both series, so that the step size between the critical stimuli was at least functionally equated.

A second possible explanation is that the sentence preceding the target words was constant, so that subjects knew when and where a code-switched word could occur. However, this was also the case for the language-selective series, on which the base language did have an effect, albeit a contrastive one. In addition, in Grosjean's gating experiments (1988), target words were also always preceded by an identical string of words, and subjects still proposed a preponderance of base-language words for the code-switched words with base-language counterparts even at rather late gates, that is, subjects did show an assimilative effect of the base language.

A third possible reason for the loss of the assimilative effect of the base language is the forced choice between the two endpoints of the series in the categorical perception experiment. That is, on top of the base-language context remaining constant, it is also the case that the base-language word or code-switched word remained constant. This contrasts with Grosjean's experiment, in which the target could be one of any number of words belonging to either English or French. Maybe the restricted choice in a categorical perception experiment allowed subjects to concentrate their attention selectively on the acoustic information contained in the stimuli while disengaging from effects of the base language. In view of this possibility, it is all the more interesting that we did find the predicted contrastive effect of the base language on the language-selective series. The finding of the contrastive effect and not the assimilative effect when using the categorical perception paradigm, but both effects when using the gating paradigm, might be a first indication that these two effects arise at different stages of language processing, with different underlying mechanisms (see, e.g., Miller, Green, and Schermer, 1984). Further experiments with different series and different languages are needed to clarify these questions. As part of this process, Grosjean is currently investigating the effect of the base language on the identification of a language-neutral series ranging from a German to a French word.

In summary, the present data, stemming from a novel use of the categorical perception paradigm, support evidence from previous experiments with different procedures that the perception of code-switched words by bilinguals does not occur solely on the basis of the

acoustic-phonetic information in the code-switched word, but is influenced by the base language of the conversation. Moreover, this effect varies with the acoustic-phonetic characteristics of the code-switched words, in particular, with whether they are phonetically compatible with the base language.

REFERENCES

BLANK, M.A. (1980). Measuring lexical access during sentence processing. *Perception & Psychophysics*, 28, 1–8.

ELMAN, J.L., DIEHL, R.L., and BUCHWALD, S.E. (1977). Perceptual switching in bilinguals. *Journal of the Acoustical Society of America*, 62, 971–974.

GROSJEAN, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception & Psychophysics*, 28, 267–283.

GROSJEAN, F. (1982). *Life with Two Languages: An Introduction to Bilingualism.* Cambridge, MA: Harvard University Press.

GROSJEAN, F., and SOARES, C. (1986). Processing mixed language: some preliminary findings. In J. Vaid (ed.), *Language Processing in Bilinguals: Psycholinguistic and Neuropsychological Perspectives* (pp. 145–179). Hillsdale: Lawrence Erlbaum Associates.

GROSJEAN, F. (1988). Exploring the recognition processes of guest words in bilingual speech. *Language and Cognitive Processes*, 3, 233–274.

LADEFOGED, P. (1975). *A Course in Phonetics.* New York: Harcourt Brace Jovanovich.

LIBERMAN, A.M., HARRIS, K.S., HOFFMAN, H.S., and GRIFFITH, B.C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, 54, 358–368.

MACNAMARA, J., and KUSHNIR, S. (1971). Linguistic independence of bilinguals: the input switch. *Journal of Verbal Learning and Verbal Behavior*, 10, 480–487.

MARSLEN-WILSON, W. (1984). Function and process in spoken word recognition – a tutorial review. In H. Bouma and D.G. Bouwhuis (eds.), *Attention and Performance X: Control of Language Processes* (pp. 125–150). Hillsdale: Lawrence Erlbaum Associates.

MILLER, J.L., GREEN, K., and SCHERMER, T.M. (1984). A distinction between the effect of sentential speaking rate and semantic congruity on word identification. *Perception & Psychophysics*, 36, 329–337.

REPP, B.H. (1984). Categorical perception: Issues, methods, findings. In N.J. Lass (ed.), *Speech and Language: Advances in Basic Research and Practice, Vol. 10* (pp. 243–335). New York: Academic Press.

REPP, B.H., and LIBERMAN, A.M. (1987). Phonetic category boundaries are flexible. In S.N. Harnad (ed.), *Categorical Perception* (pp. 89–112). New York: Cambridge University Press.

SOARES, C., and GROSJEAN, F. (1984). Bilinguals in a monolingual and a bilingual speech mode: the effect on lexical access. *Memory & Cognition*, 12, 380–386.