Journal of Phonetics (1984) 12, 91–114

Effects of consonant context on the perception of French vowels

Terry L. Gottfried

State University of New York at Purchase, Purchase, New York 10577, U.S.A.

Received 21*st October* 1983

Abstract: American English vowels have been found to be more accurately identified in consonantal context than in isolation. One might, however, expect different results for French because phonological and phonetic characteristics of French vowels are different from those of English. Native French listeners identified stressed oral vowels produced by native speakers in /t/-vowel and isolated vowel contexts more accurately than vowels produced in vowel-/t/ and /t/-vowel-/t/ contexts. In categorial discrimination, native French, non-native speakers of French, and non-speakers of French perceived isolated French vowels more accurately than vowels in TVT context, arguing against phonological appropriateness as the determinant of context effect differences between languages. American non-native speakers of French and non-speakers of French were significantly less accurate than native French in identification and discrimination of French vowels in all contexts, especially front rounded vowels which do not occur in English. Acoustical analyses of the French vowels showed less diphthongization than for English vowels in the same syllabic contexts, making vowel "targets" more reliable as sources of information for French vowel identity.

Introduction

Descriptions of the acoustic information for vowel identity have commonly focused on the center frequency of the first two or three speech formants (see e.g., Joos, 1948; Ladefoged, 1967). Each vowel may be described by a distinct set of formant frequencies, called its "target." The vowel is then perceived by listeners in relation to the position of the target in the acoustic "vowel space," a two-dimensional plot with the frequencies of the first two formants as coordinates.

Although the target formant frequencies are often adequate to differentiate vowels spoken by one person at a particular speaking rate, the same vowel spoken by different people may vary considerably in its target formant frequencies (Peterson & Barney, 1952). In addition, vowels in continuous speech often do not reach the target values when spoken in consonantal context (Gay, 1978; House & Fairbanks, 1953; Lindblom, 1963; Stevens & House, 1963). If target formant frequency information is the primary cue for vowel identity, then we would predict that vowel identification would be poorer for vowels coarticulated with consonants than for vowels alone because coarticulated vowels do not reach the target values.

In fact, several studies of American English vowels have shown that vowels in consonant-vowel-consonant (CVC) syllables, vowel-consonant (VC) syllables and some consonant-vowel

0095–4470/84/020091+24$03.00/0

(CV) syllables were easier to identify than isolated vowels (Gottfried, 1982; Gottfried & Strange, 1980; Strange *et al.*, 1976, 1979; Verbrugge *et al.*, 1976). From these studies, Strange and her colleagues concluded that dynamic spectral information in the formant transitions into and out of the syllable nucleus and relative syllable duration aided significantly in the more accurate identification of coarticulated vowels than isolated vowels. Further studies have shown that the target formant information is not necessary for good vowel identification. Jenkins *et al.* (1983) and Strange *et al.* (1983) excised the relatively steady-state portion (the target) from natural vowels spoken in CVC syllables. They obtained high degrees of accuracy when only the initial and final transitions and relative timing of these portions were left intact. Identification accuracy for these "vowel-less" syllables was equal to or greater than that for the "targets" extracted from the syllables.

Although a number of researchers (Assmann *et al.*, 1982; Diehl *et al.*, 1981; Macchi, 1980) have failed to obtain significant differences in vowel identification accuracy between isolated and coarticulated vowels, these failures to replicate the findings of Strange and her colleagues may be attributed to ceiling effects (see Gottfried, 1982; Strange & Gottfried, 1980). Thus, American English vowels in consonantal context are, in most natural circumstances, easier to perceive than isolated vowels.

There are phonological and phonetic differences between languages, however, that might lead one to postulate cross-language differences with respect to context effects on vowel identification accuracy. Specifically, the way that French vowels are articulated is quite different from the way American English vowels are articulated. Further, the phonotactic constraints on syllable structure are different for French and English. These differences in articulation (and consequently, the differences in acoustic information) and in phonological structure might lead to differences in how native speakers may perceive vowels in the two languages.

Delattre (1953) characterized French as a "tenser" language in terms of its articulation than English. The acoustic consequence of this articulatory tension includes less dipthongization of vowels. Delattre (1963) also tested this often-observed difference between French and English vowels in terms of their diphthongization and observed via cinefluororadiology that French speakers articulate their vowels with much less gliding and a much steadier articulation than American English speakers.

In addition, Delattre (1953) noted that vowels seem more prominent perceptually in French than in English because of syllable structure constraints. The syllable structure of French is such that CV syllables are much more preferred in French than in English. Juilland (1965) found that over 70% of the words in printed French ended phonetically with vowels, and Léon (1966) estimated that in spoken French nearly 80% of the syllables end in vowels. Estimates for syllables ending in vowels in English are much lower (44%; Dauer, 1983) and lax vowels /ɪ, ɛ, æ, ʌ, ʊ/ do not occur in syllable-final contexts.

Phonetic descriptions of French (see, e.g., Delattre, 1951; Léon, 1966) enumerate 12 oral and four nasal vowels /i, e, ɛ, a, ɑ, ɔ, o, u, y, ø, œ, ə, ɛ̃, œ̃, ɑ̃, ɔ̃/. However, the distinction between the more front /a/ and the more back /ɑ/ is not maintained by most French speakers (Delattre, 1957), and the /ə/ does not occur in stressed position. Of the stressed oral vowels, the high vowels /i, y, u/ and the low vowels /a, ɑ/ may occur in both open (CV, V) or closed (CVC, VC) syllables. The mid-vowels /e, ɛ, o, ɔ, ø, œ/ present an interesting pattern of distribution across syllabic contexts. In general the mid-vowels follow the "Law of Position," which states that the higher mid-vowels /e, o, ø/ occur in open syllables and the lower mid-vowels/ɛ, ɔ, œ/ occur in closed syllables. There are, however, exceptions to the Law of Position. The vowel /ɛ/ also occurs in open syllables, for example, in the minimal

pair *fée* /fe/ "fairy" versus *fait* /fɛ/ "fact". Most Parisians observe this distinction in their speech production (Monnot, 1977), although many speakers of other dialects do not (Martinet, 1945). The high-mid vowels /o, ø/ may occur in closed syllables, although the minimal pairs between /o/–/ɔ/ and between /ø/–/œ/ are relatively few and it is doubtful that many speakers maintain the distinction in rapid speech. The high-mid vowel /e/ never occurs in closed syllables.

One of the purposes of the present experiment was to determine whether, because of these differences in phonological and phonetic characteristics, native French listeners would show differences in the pattern of identification accuracy from that exhibited by native speakers of American English. If isolated vowels were identified less accurately by French listeners (as well as American English listeners) than vowels in consonantal context, this would be further evidence that dynamic acoustic information is important for vowel identification, regardless of the phonological and phonetic properties of the language. If, however, the results with French differed from those observed in English, one must then consider how the phonological and phonetic properties of French and English vowels might account for the different outcomes.

A second purpose of the present experiment was to determine whether the identification of French vowels differed for native vs non-native speakers of the language. Many studies indicate that native speakers of one language are often poor at perceiving the phonemic contrasts of another language (e.g., Goto, 1971; Marckwardt, 1944, 1946; Miyawaki *et al.*, 1975; Nemser, 1971; Sapon & Carroll, 1958). Learning a second language changes the speech perception of language learners (e.g., Caramazza *et al.*, 1973; Elman *et al.*, 1977; MacKain *et al.*, 1981; Politzer, 1961; Streeter & Landauer, 1976; Williams, 1977, 1979), although non-native speakers of a language are not exactly like native speakers in their speech perception (Elman *et al.*, 1977; Goto, 1971; Marckwardt, 1944, 1946; Politzer, 1961; Sheldon & Strange, 1982; Williams, 1979).

In order to test the influence of native language on the perception of vowels in a second language, native speakers of French and American English were tested on identification and discrimination of French vowels. Two groups of American listeners were studied: Americans who had several years of French study and Americans who had never studied French. From the results of the cross-language studies cited above, one would expect that American English listeners who knew French would not be as accurate in vowel identification and discrimination as native French listeners, but more accurate than Americans who had never studied French. Moreover, if we observe native speakers of American English who are less familiar with French to differ from native French listeners in their relative accuracy on vowels in CVC syllables vs isolated vowels, we might conclude that the different phonotactic constraints of English and French could account for any observed cross-language differences in context effects. On the other hand, if Americans who do not know French show the same context effects as French listeners, then cross-language differences in context effects might be attributed to the acoustic properties of the French syllables themselves.

Experiment 1 examined the vowel identification ability of native French speakers and native American English speakers who spoke French, using a key word task similar to one used in experiments with American English vowels (Gottfried & Strange, 1980; Strange *et al.*, 1976, 1979). Because American English speakers who did not know French would not be familiar with the French orthography used on the answer sheets, they were not tested in this experiment. Listeners used response sheets that were compatible with the syllable context heard because several experiments have indicated that these are optimal for vowel identification in American English (see Gottfried, 1982; Strange & Gottfried, 1980).

1. Experiment 1

A. Method

(i) *Stimulus materials*

Two adult male and two adult female speakers who were native speakers of Parisian French were recorded individually in a sound-attenuated room, using an Electro-Voice microphone (660SP) and Revox tape recorder (A77). Speakers produced two tokens each of 11 French oral vowels /i, e, ɛ, a, ɑ, ɔ, o, u, y, ø, œ/ in each of four different syllabic forms: /t/-vowel-/t/ (TVT), vowel-/t/ (VT), /t/-vowel (TV), and the vowels in isolation (#V#).

Speakers produced the stimuli by reading aloud from individual cards presented by the experimenter. Cards were arranged in random order; all tokens of a given syllabic form were produced consecutively. All tokens were printed on the cards as TVT words, using normal French orthography to represent the vowels. Exemplary words for the vowel were also printed under the target syllable on the card. The speakers were instructed to produce only the vowel (#V#), the initial consonant and the vowel (TV), the final consonant and the vowel (VT), or the entire syllable (TVT), as appropriate.

Speakers were told to produce each token briskly, "as you would say it in a sentence". Tokens were discarded if (a) a speaker expressed dissatisfaction with his/her utterance, (b) the experimenter detected a clear phonemic error, or (c) the peak recording amplitude fell above or below a predetermined range. In addition, a phonetician who is a native speaker of French listened to all of the vowel tokens in each condition. She determined that all tokens were acceptable versions of the intended vowel, although she did note that speakers were not very consistent in differentiating the back low vowel /ɑ/ from the front low vowel /a/.

The tokens from each speaker were low-pass filtered (4800 Hz) and digitized at 10 000 samples/s by a PDP 11/34 computer. Four different listening tests were then constructed by digital to analog conversion and recording onto audio tape via a Crown tape recorder (CX822). Each test series consisted of 88 different tokens (4 speakers × 11 vowels × 2 tokens of each vowel) of a given syllabic form; the sequence of vowels and speakers was randomly determined, with the constraints that consecutive syllables were produced by different speakers, and the same vowel did not occur more than twice consecutively. The sequence of speakers and vowels was different for each of the syllable conditions. The interstimulus interval was 4 seconds, with 10 seconds between blocks of 10 stimuli.

(ii) *Subjects and procedure*

Groups of subjects were tested on all four of the syllable context conditions (TVT, VT, TV, #V#) and were assigned to one of four counter-balanced orders (#V#, TVT, VT, TV; TVT, #V#, TV, VT; VT, TV, #V#, TVT; and TV, VT, TVT, #V#). There was a total of 24 subjects; 16 were native speakers of Parisian French, and eight were native speakers of American English who had had extensive experience with speaking French. The native French listeners were high-school students (ages 15–17 years) visiting Minnesota in the summer of 1981 who volunteered to participate in the experiments. They were all from the area around Paris and spoke English with varying degrees of competence. The American English speakers who also spoke French had all studied French for several years (mean years of study was 4.8 with a standard deviation of 2.8). Many of these were French majors or minors at the University of Minnesota, and many also spoke languages other than English and French.

Listening tests were presented to small groups of subjects in a quiet experimental room via a Revox tape recorder (A77), MacIntosh amplifier (MC 400), and an acoustic suspension loudspeaker (AR 2ax). Subjects identified vowels by circling the appropriate syllable on

response forms provided for them. For TVT test syllables, the response alternatives were in the form of TVT syllables (i.e., tite, tét(er), tette, tatte, tâte, totte, tôte, toute, tute, teut(on), tœute). For TV syllables, response alternatives appeared as TV syllables (i.e., ti, thé, tais, ta, tâ, to(que), tôt, tout, tu, teu(ton), teu(f)). For VT syllables, response alternatives were VT syllables (i.e., ite, ét(é), haite, hatte, hâte, hotte, haute, oute, ut, (m)eute, hœute). For isolated vowels, response alternatives were isolated vowels (i.e., y, et, hais, a, â(me), o(r), eau, où, eu, eux, œu(f)).

Before each test began, subjects performed a familiarization exercise in which they identified vowels spoken in the syllable context they were subsequently tested on. The tokens in familiarization were spoken by a different speaker from the speakers used in the test. Subjects responded on the same response forms that they used in the test phase of the experiment. Forty items were presented for familiarization; subjects were given feedback after each block of ten items. Each subject then responded to the 88-item listening test without feedback for a total of eight identifications of each intended vowel in that syllable context. After completion of the first test, subjects then performed the familiarization and test for the second, third, and fourth syllable context conditions.

B. Results

Figure 1 presents the overall percent errors for the TVT, VT, TV, and isolated vowel (#V#) conditions for the native French listeners and the French-speaking American listeners. An error was defined as a vowel response other than the one intended by the speaker; errors were summed over all 11 vowels and were expressed as the percentage of total number of responses. Let us consider first the performance by the native French listeners.

(i) *Native French listeners*

Overall error rates varied significantly as a function of consonantal context. An analysis of variance of subjects' number of errors revealed a significant difference across the TVT, VT, TV, and isolated vowel conditions, $F(3,36) = 13.49$, $MSE = 26.27$, $p < 0.01$. Vowels were identified significantly more accurately when they were presented in TV (17.5% errors)

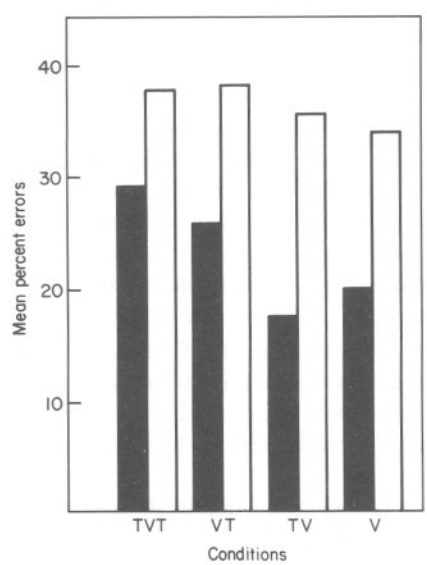

Figure 1 Mean percent identification errors as a function of syllable context for native French (■) and French-speaking American (□) listeners.

context than when they were in TVT (29.2% errors) and VT (25.8% errors) contexts, $p < 0.05$ (by Tukey's test of honestly significant differences). Isolated vowels (20.0% errors) were also significantly easier to identify than vowels in TVT syllables, $p < 0.05$, but not significantly easier than vowels in VT syllables, $p > 0.10$.

As discussed in the introduction, the distinction between /a/ and /ɑ/ is not maintained by many French speakers (Delattre, 1957), and the French phonetician who evaluated the French vowel tokens in this study commented that the /a/ and /ɑ/ sounds were not well differentiated. Confusions of the /a/ and /ɑ/ were quite common among the native French listeners, accounting for 22.4% of the errors in all context conditions. When the confusions between /a/ and /ɑ/ were removed from the analysis, the error rates were reduced considerably, but the effect of syllabic context on vowel identification accuracy was still significant, $F(1,12) = 14.56$, $MSE = 25.54$, $p < 0.01$. The TVT and VT conditions still had higher error rates (24.1% and 20.6%, respectively) than the TV and isolated vowel conditions (11.9% and 15.0%, respectively).

These results are almost exactly opposite to the results with American English speakers identifying American English vowels. The presence of the final consonant /t/ did not aid, and indeed, seemed to hinder accurate identification of French vowels. These results differ somewhat from the results of a study by Gottfried (1979). In that study, the error rates, like those in the present experiment, were considerably higher than those observed for American English. However, syllable context in the early study (also comparing vowels in isolation and in combination with /t/) did not significantly affect vowel identification accuracy. Several factors may account for differences between these studies. The Gottfried (1979) study did not include the vowel /ɑ/, had fewer speakers, fewer listeners, and considerable dialect heterogeneity among the listeners (who were from several regions of France). It is nonetheless striking that neither study showed an advantage of a consonant context for vowel identification in French.

There are several possible explanations for these results. First, it is possible that phonological unfamiliarity hindered vowel identification in the different syllabic contexts. For example, the Law of Position of French vowels states that the high-mid vowels are usually found in open (e.g., TV and #V#) syllables; /e/ is found only in such syllables. Low-mid vowels are usually found in closed (e.g., VT and TVT) syllables; /ɔ/ and /œ/ are found only in such syllables. Thus, the vowels /ɛ, ɔ, œ/ are usually inappropriate for TV and #V# syllables, and the vowels /e, o, ø/ are usually inappropriate for VT and TVT syllables. We would thus expect more confusions of low-mid vowels for high-mid vowels (e.g., /ɛ/→/e/, /ɔ/→/o/, /œ/→/ø/) in the TV and #V# context conditions.

Table I shows the average number per subject of high-mid → low-mid and low-mid → high-mid vowel confusions in each of the syllable context conditions. There were only a few more low-mid → high-mid vowel confusions than high-mid → low-mid vowel confustions in all syllable context conditions. None of the differences between these two kinds of confusions were significant. Thus, it was not the case that French listeners were "regularizing" the vowel stimuli to match the expectations according to the Law of Position. Rather, they were no more likely to identify a mid vowel in open syllables as a high-mid vowel than as a low-mid vowel; likewise, they were no more likely to identify a mid vowel as low-mid than as high-mid in closed syllables.

Of the mid vowel errors, the largest proportion in all context conditions was for the front rounded mid vowels /ø, œ/. Confusions of these vowels accounted for 20.8% of all errors made by French listeners. This is not surprising because the two sounds are rarely contrasted in French words and are usually spelled the same way ("eu"). The /œ/sound, if it

Table I Mean high-mid/low-mid vowel confusions for French listeners

Confusion type	Syllable context TVT	VT	TV	#V#
High-mid → Low-mid	4.13	4.44	3.00	3.06
Low-mid → High-mid	6.38	5.13	3.06	3.63
Difference	−2.25	−0.69	−0.06	−0.56

is considered to be different from /ø/, never occurs in syllable-final position. The /ø/ usually occurs in syllable-final position, but also occurs in syllables with certain consonants following (e.g., /z, t, ʒ/). There are only two contexts where the /ø/ and the /œ/ may contrast: when followed by /n/ and when followed by /l/. This yields the minimal pairs *jeûne-jeune* /ʒø:n/-/ʒœn/ "fast [noun]-young" and *veule-veulent* /vø:l/–/vœl/ "flabby-want [third person plural]". However, these are virtually the only such minimal pairs, and many French speakers do not make such a contrast. For them, the sounds /ø/ and /œ/ are in complementary distribution.

For this reason, it may be reasonable to remove /ø/–/œ/ confusions from the analysis, as was done for the /a/-/ɑ/ confusions above. When confusions of /ø/ and /œ/ were also removed, the error rates were reduced, but the significant differences between consonant context conditions were not altered, $F(3,36) = 13.67$, $MSE = 19.11$, $p < 0.001$. The average percent errors in the TVT (18.1%) and VT (15.1%) conditions were higher than in TV (7.4%) and #V# (11.9%) conditions. Thus, the poorer identification of vowels in TVT and VT syllables cannot be attributed to the confusions of /a/ and /ɑ/ or /ø/ and /œ/, vowel contrasts whose phonemic status in French is unstable.

(ii) *Comparison of native and non-native listeners*

Figure 1 also shows the overall percent erors for the TVT, VT, TV, and isolated vowel conditions by native speakers of American English who speak French. An analysis of variance (for unweighted means) revealed that there was a significant difference between native French and French-speaking Americans in their overall accuracy in identifying vowels, $F(1,22) = 13.72$, $MSE = 215.31$, $p < 0.01$. Syllable context significantly affected vowel identification accuracy for the two subject groups combined, $F(3,66) = 6.93$, $MSE = 27.61$, $p < 0.001$. Vowels in TV context were significantly easier to identify than vowels in TVT and VT contexts, $p < 0.01$. Vowels in isolation were significantly easier to identify than vowels in TVT syllables, $p < 0.05$, but not significantly easier than vowels in VT syllables, $p > 0.10$.

The subject groups did not significantly differ in their relative performance on the syllable context conditions. That is, the group × condition interaction was not significant, $F(3,66) = 2.44$, $p > 0.05$. (Although the group × condition interaction was not significant, the American group did not show significant differences between context conditions when the American data were analyzed separately, $F(3,21) = 0.67$, $MSE = 25.11$. This may have been due to the small number of American subjects and the higher error variance, which reduced the power of the statistical test to detect a difference.)

Like native speakers of French, the Americans made many confusions of /a/ and /ɑ/. These confusions, however, did not account for as many of the American English listeners' errors (15.0% of all errors) as was the case for French listeners (22.4%). When the /a/-/ɑ/ confusions were removed from the analysis for all listeners, there still were significant effects of native languege [$F(1,22) = 12.58$, $MSE = 224.26$, $p < 0.01$] and syllable context

Table II Mean high-mid/low-mid vowel confusions for French-speaking American listeners

Confusion type	Syllable context TVT	VT	TV	#V#
High-mid → Low-mid	5.88	3.88	6.00	5.63
Low-mid → High-mid	6.75	6.63	8.88	8.50
Difference	−0.88	−2.75	−2.88	−2.88

[$F(3,66) = 9.40$, MSE = 27.34, $p < 0.001$] on the number of identification errors. The interaction of language group and syllable context was again not significant, $F(3,66) = 1.49$.

American listeners made many mid vowel confusions, similar to those of native French listeners. Table II presents the average per subject of high-mid → low-mid and low-mid → high-mid vowel confusions for each syllable context condition. There were significantly more low-mid → high-mid vowel confusions in the VT, TV, and isolated vowel context conditions ($t(7) = 3.27$ for VT syllables; $t(7) = 3.02$ for TV syllables; and $t(7) = 2.54$ for isolated vowels, $p < 0.05$ for all of these conditions). The high number of low-mid → high-mid vowel confusions would be predicted by the Law of Position for the TV and isolated vowel contexts (i.e., the open syllables).

The results for the VT condition are puzzling until it is noted that most of the low-mid → high-mid vowel confusions are confusions of /œ/→/ø/. Confusions of these vowels were very common for native listeners (20.8%) as well as non-natives (20.5% of all errors). Because of the unstable status of this vowel contrast, confusions of /ø/→/œ/ and /œ/→/ø/ were removed from the analysis of mid vowel confusions among the American listeners. Table III presents the average number per subject of high-mid → low-mid and low-mid → high-mid vowel confusions, disregarding confusions of /ø/ and /œ/. This table shows that Americans made relatively more low-mid → high-mid vowel confusion in only the two contexts predicted by the Law of Position, TV and #V# [$t(7) = 2.85$ and $t(7) = 2.12$, respectively, $p < 0.05$].

C. Discussion

Native speakers of French identified the vowels of their language significantly more accurately than native speakers of American English who speak French. The result is similar to findings in earlier cross-language studies (e.g., Goto, 1971) and replicates the results of Gottfried (1979), using more speakers and listeners. Unlike the Gottfried (1979) study, which found no significant differences between syllable context conditions, this experiment showed that French listeners made significantly fewer errors in the TV and isolated vowel conditions than in the TVT and VT conditions. These results contrast with those of previous studies using American English vowels (Gottfried, 1982; Gottfried & Strange, 1980; Strange

Table III Mean high-mid/low-mid vowel confusions, excluding /ø/-/œ/ confusions, for French-speaking American listeners

Confusion type	Syllable context TVT	VT	TV	#V#
High-mid → Low-mid	3.63	1.13	2.38	1.44
Low-mid → High-mid	3.13	2.25	5.38	5.00
Difference	+0.50	−1.13	−3.00	−3.56

et al., 1976, 1979), in which American English listeners made significantly more errors in identifying American English vowels in TV syllables and in isolation. In addition, the error rates for native French listeners were much higher than for the American listeners who identified American English vowels in a comparable task. Even with /a/-/ɑ/ and /ø/-/œ/ confusions removed, the error rates for French vowel identification by native French listeners were considerably higher than error rates for the identification of American English vowels by American listeners.

These results may be due to (1) French vowels being intrinsically more difficult to identify, especially in TVT and VT syllables; (2) more of the French vowel tokens being ambiguously produced than the English vowel tokens of the previous studies, especially in the TVT and VT syllables; (3) poorer attention to (or poorer understanding of) the task by French listeners than by American English listeners. This third possibility might have some validity because the French listeners were younger, and possibly less inclined to pay attention, than the Americans. In addition, the instructions to the French were given in French, but by a non-native speaker, so the French listeners may not have understood the task as well as the American listeners. The fact that Americans who knew French made a similar pattern of errors as the French, however, renders this explanation somewhat less plausible. Listeners were well practised, and certainly by the last three tests understood the task very well. The possibility of intrinsic differences in vowel identification difficulty and the role of misproduced tokens on identification errors are discussed in the Acoustical Analysis section below.

It is interesting that the non-native French speakers tended to make identification errors in accordance with the Law of Position, whereas native French speakers did not. As stated in the introduction, this "Law" is not an absolute law, but rather a description of general distributional tendencies of French vowels. Native French listeners, especially those who speak Parisian French, would be aware of the many exceptions to the rule, perhaps more than Americans who had learned the language. In addition, Valdman (1976) discussed the use of the Law of Position as a pedagogical device to simplify the teaching of French phonetics. Absolute application of the rule (high-mid vowels in open syllables; low-mid vowels in closed) would result in a non-standard, southern French dialect, but an acceptable dialect nonetheless. Having learned this general rule, and being less familiar with the exceptions, Americans may have been more influenced in their vowel perception by the Law of Position than native French speakers.

It should also be noted that the Law of Position predicted the Americans' errors occurring only for open syllables. This could be due to the low-mid vowel tokens in the TV and #V# conditions being unusually long in duration. The vowels /ɔ/ and /œ/ are ordinarily short (relative to /o/ and /ø/) when they are in the phonologically appropriate closed syllable context. Americans, perhaps more than French listeners, may have been misled by inaccurate duration information in the TV and isolated vowel conditions (see Acoustical Analysis section below).

It might be argued that the large number of errors for native French and especially for French-speaking Americans is attributable to infelicitous or even misleading response alternatives. The TV and #V# response alternatives were necessarily closer to "natural" French orthography and included more real French words than the TVT and VT response alternatives. Thus, the appropriateness of response alternatives presents a problem, as it also does for American English vowel perception (see Assmann *et al.*, 1982; Diehl *et al.*, 1981; Gottfried, 1982; Strange & Gottfried, 1980). In order to eliminate this problem, Experiment 2 used a task that was successfully employed with American English vowel perception

(Gottfried, 1982). This particular task retained the categorial aspects of the identification task but did not require the listeners to be familiar with French orthography. Therefore, native speakers of American English could be tested who did not know French, as well as native French speakers and native American English speakers who knew French.

2. Experiment 2

A. Method

(i) *Stimulus materials*

The syllable tokens used in the ABX discrimination task were the same as those used in the TVT and isolated vowel conditions of Experiment 1. Trials consisted of three tokens spoken by three different speakers. The intended vowel in the first two tokens of a trial were different from each other; the third token was of the same vowel category as the first token or the second token. For example, a trial might consist of /u/ spoken by speaker 1, /y/ spoken by speaker 2, and the /u/ spoken by speaker 3. In this case, the first and third syllable are the "same" in that they belong to the same vowel category.

Eight pairs of vowels were tested using this categorial discrimination task: /i, e/, /e, ɛ/, /ɛ, a/, /a, ɑ/, /ɑ, ɔ/, /u, y/, /y, ø/, /ø, œ/. These pairs were selected because the vowels of each pair were often confused for each other in identification, as determined in a previous study (Gottfried, 1979). There were four possible discrimination trial orders for each pair /A, B/: ABA, ABB, BAB, and BAA. There were 24 possible orders in which three out of four speakers could be arranged for each trial (e.g., speaker 1/speaker 2/speaker 3; speaker 1/speaker 3/speaker 4; speaker 2/speaker 3/speaker 4; etc.). Vowel pairs were randomly assigned to the 96 speaker/order triples so that each vowel pair appeared 12 times on a test.

The stimuli within each trial were separated by a 1-second interval. Trials were separated by 3 seconds. Trials were blocked in groups of ten with a 6-second interval separating each block. The order of trials was randomly determined. Four different listening tests were constructed by reconverting the digitized tokens, filtering the analog signals, and recording them onto audio tape via a Crown tape recorder (CX 822). The stimuli for one test were the first TVT tokens of each speaker (TVT); another test consisted of the first #V# tokens spoken by each speaker (#V#).

(ii) *Subjects and procedure*

Groups of subjects were tested in both syllable context conditions in two possible orders: TVT, #V#; and #V#, TVT. There was a total of 40 subjects, ten native speakers of French, ten native speakers of American English who spoke French, and 20 American English speakers who did not speak any French. The native speakers of French came from several areas of France, including southern France. Most were graduate students at the University of Minnesota, all were over 21 years old, and were therefore older than the native French speakers of Experiment 1. These French speakers were also all proficient in English. The ten French-speaking Americans had studied French for several years; many were teachers of French at the college or high school level (mean number of years of study was 7.7 years, with a standard deviation of 4.9). The last group of subjects were volunteers from introductory psychology classes at the University of Minnesota and received extra class credit for participation. All of this last group were native speakers of American English, and most were from the Upper Midwest region of the United States.

Subjects were instructed to put a "1" next to the trial number if the third stimulus of the triad was the same as the first stimulus; a "2" if the third stimulus was the same as the second stimulus. Examples of the trials and five practice trials with feedback were presented

before each condition. After the instructions and practice trials, subjects were presented the first 96-trial ABX discrimination test. When the first test was completed, subjects were given instructions, practice, and the second 96-trial discrimination test.

B. Results

The overall percent errors for each of the context conditions (TVT, #V#) are presented in Fig. 2 for each of the subject groups – native French (F), French-speaking American (FA), non-French-speaking American (A) – collapsed over the two presentation orders. An error was defined as responding "1" when the trial was ABB or BAA, responding "2" when the trial was ABA or BAB, or making no response. Errors were summed over all trials of the context condition and expressed as the percentage of the total number of trials. Let us consider first the performance of native French listeners.

(i) *Native French listeners*

The overall percent errors for French listeners was significantly different as a function of syllable context. An analysis of variance revealed that isolated vowels were easier to discriminate (12.7% errors) than vowels in TVT syllables (21.9% errors), $F(1,18) = 43.83$, $MSE = 8.33$, $p < 0.001$. The order of condition presentation did not significantly affect overall performance, $F(1,18) = 1.96$, $MSE = 17.28$. Order also did not significantly affect performance on particular conditions; e.g., performance on a particular condition was not better when presented second, $F(1,18) = 0.57$.

The error rates for some of the vowel comparisons were greater than for others. Table IV presents the average number of errors (of 12 possible) per subject on each of the vowel pair comparisons in the isolated vowel and the TVT conditions. Confusions of /a/ and /ɑ/ were by far the most common in both syllable contexts (accounting for 39.3% of the errors in the isolated vowel condition and 27.3% of the errors in the TVT condition). There were also

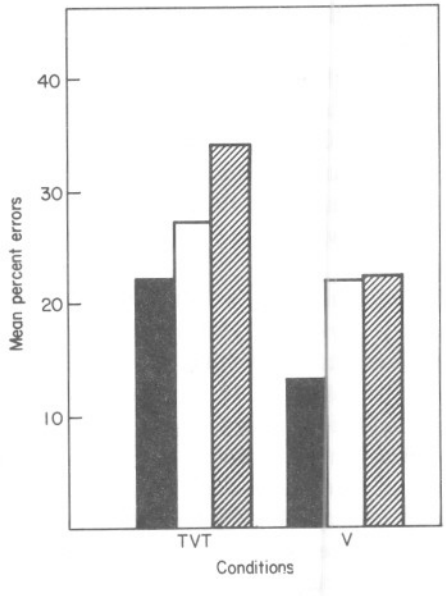

Figure 2 Mean percent discrimination errors as a function of syllable context for native French (■), French-speaking American (□), and non-French-speaking American (▨) listeners.

Table IV Mean discrimination errors (out of 12 possible) for vowel pair comparisons by native French listeners

Vowel pair	Syllable context	
	TVT	#V#
/i, e/	4.40	1.10
/e, ɛ/	1.40	0.70
/ɛ, a/	0.10	0.20
/a, ɑ/	5.70	4.80
/ɑ, ɔ/	3.30	1.60
/u, y/	0.60	0.50
/y, ø/	2.30	0.20
/ø, œ/	3.20	3.10

many confusions of /ø/ and /œ/ in the isolated vowel condition. The TVT condition had many /ø/-/œ/ confusions, but also /ɑ/-/ɔ/ and especially /i/-/e/ confusions. That this last comparison pair was confusing is somewhat surprising because French listeners did not make very many /i/-/e/ confusions in discriminating #V# syllables. One might attribute the discrimination confusions to the phonological inappropriateness of /e/ in closed syllable context. However, the majority of these discrimination confusions involved a particular token of /i/ that apparently sounded like /e/ to these listeners. It appears that the formant frequencies of this token were like those for /e/, leading listeners to call it the same as one of the /e/ tokens (see Acoustical Analysis below).

(ii) *Comparison of the three subject groups*

Figure 2 also allows one to compare the performance of native French, French-speaking American, and non-French-speaking American listeners on the isolated vowel and TVT discrimination tests. Because order of test presentation (TVT first vs #V# first) did not significantly affect vowel discrimination in any of the groups, the results for the two orders were combined for the analyses.

The subject groups differed significantly in their average percent errors. An analysis of variance (of unweighted means) revealed a significant difference between the subject groups, $F(2,37) = 12.10$, $MSE = 55.77$, $p < 0.001$. As expected, native French listeners were able to discriminate French vowels significantly better (17.3% errors) than non-native speakers of French (24.6% errors) and non-speakers of French (28.2% errors), $p < 0.05$. The overall differences between the Americans who spoke French and the Americans who did not speak French was not significant.

Overall error rates also varied significantly as a function of consonant context, $F(1,37) = 80.56$, $MSE = 16.27$, $p < 0.001$. Vowels in TVT syllables were discriminated more poorly on the average (mean percent errors of 29.1%) than vowels in isolation (19.7% errors). There was also a significant interaction of subject group and syllable context, $F(2,37) = 3.80$, $p < 0.05$. The source of this interaction lies in the relative performance of the two American groups on the different syllable context conditions. The French-speaking Americans were not significantly better than non-French-speaking Americans in discriminating the isolated vowels (21.9% vs 22.1% errors, respectively). On the other hand, French-speaking Americans discriminated vowels in TVT syllables significantly better than Americans who spoke no French (27.3% vs 34.2% errors, respectively).

A more detailed comparison of the subject groups may be accomplished by examining errors for particular vowel comparisons. Table V presents the average number of errors (of

Table V Mean discrimination errors (out of 12 possible) for TVT vowel pairs by all listeners

Vowel pair	Subject group† F	FA	A
/i, e/	4.40	3.50	5.60
/e, ɛ/	1.40	2.10	3.05
/ɛ, a/	0.10	0.30	0.25
/a, ɑ/	5.70	4.90	5.25
/ɑ, ɔ/	3.30	3.90	4.20
/u, y/	0.60	2.50	3.85
/y, ø/	2.30	4.20	6.30
/ø, œ/	3.20	4.80	4.30

†F = native French; FA = French-speaking American; A = non-French-speaking American.

12 possible) in the TVT condition on each of the vowel comparisons for each of the subject groups. Examination of this table shows a marked tendency by non-French-speaking Americans to confuse the /u/-/y/ and /y/-/ø/ vowel comparisons. These confusions were comparatively uncommon among French listeners. Americans who had studied French were between native French and non-French-speaking Americans in the rate at which they confused these front rounded vowels.

Table VI presents the vowel-by-vowel data for the isolated vowel condition. All groups made many errors on comparisons of /a/-/ɑ/ and /ø/-/œ/. Most of the errors by French listeners were on these vowel comparisons (see above). French-speaking Americans also confused /e/-/ɛ/ pairs. Non-French-speaking Americans tended to make many errors on the /i/-/e/ and /y/-/ø/ pairs.

The confusion of the front rounded vowels is understandable since these are unfamiliar contrasts to American English listeners unless they had studied French or another language in which these vowels are contrasted (e.g., German). About half of the non-French-speaking Americans had studied some German, but most of these people reported that they could not speak any foreign language well.

Table VI Mean discrimination errors (out of 12 possible) for #V# vowel pairs by all listeners

Vowel pair	Subject group† F	FA	A
/i, e/	1.10	3.30	4.80
/e, ɛ/	0.70	5.00	2.50
/ɛ, a/	0.20	0.50	0.15
/a, ɑ/	4.80	4.40	4.55
/ɑ, ɔ/	1.60	1.70	1.55
/u, y/	0.50	0.80	0.80
/y, ø/	0.20	1.60	4.00
/ø, œ/	3.10	3.70	2.85

†F = native French; FA = French-speaking American; A = non-French-speaking American.

Table VII Mean discrimination errors for TVT vowel pairs including and not including front-rounded vowels

Pair type	Subject group† F	FA	A
Front-rounded	6.10	11.50	14.45
Non-front-rounded	14.90	14.70	18.35

†F = native French; FA = French-speaking American; A = non-French-speaking American.

Table VII shows the average number of errors for each subject group in the TVT condition on vowel comparisons that included a front rounded vowel (i.e., /u, y/, /y, ø/, /ø, œ/) versus all other vowel comparison. A two-way analysis of variance (vowel comparison type × subject group) revealed that there was a significant interaction of subject group and vowel comparison type, $F(2,37) = 7.43$, $MSE = 7.53$, $p < 0.01$. That is, all American English listeners made relatively more errors on front rounded vowel comparisons than native French listeners, $p < 0.05$. Non-French-speaking Americans did not make significantly more of these errors than the French-speaking Americans. However, French-speaking Americans did make fewer errors in the non-front-rounded vowel comparisons than the other American group, $p < 0.05$. Indeed, when the number of errors on the non-front-rounded vowels were examined, French-speaking Americans made no more errors than native French speakers.

Table VIII shows the average number of errors for each subject group in the #V# condition on vowel comparisons that included a front-rounded vowel versus all other vowel comparisons. There was a tendency for non-French-speaking Americans to make relatively more front-rounded vowel errors in the isolated vowel condition, although this trend was not significant. (The interaction of subject group and vowel comparison type did not quite reach significance, $F(2,37) = 3.11$, $MSE = 6.65$ $p < 0.06$).

C. Discussion

This detailed vowel-by-vowel analysis has shown that the pattern of errors, as well as the number of errors, differed significantly as a function of linguistic experience. Native French listeners made relatively many /a/-/ɑ/ and /ø/-/œ/ confusions, as would be predicted from their identification errors and from phonetic descriptions of modern French (e.g., Valdman, 1976). Americans made significantly more errors on the front rounded vowel comparisons in the TVT context than French listeners. These confusions could account for the difference between native French and French-speaking American performance in the TVT context.

Table VIII Mean discrimination errors for #V# vowel pairs including and not including front-rounded vowels

Pair type	Subject group† F	FA	A
Front-rounded	3.80	6.10	7.65
Non-front-rounded	8.40	14.90	13.55

†F = native French; FA = French-speaking American; A = non-French-speaking American.

French-speaking Americans were better at discriminating the non-front-rounded vowels than non-French-speaking Americans. The pattern of results for the isolated vowels was not as clear, although there was a slight tendency for non-French-speaking Americans to make more front rounded vowel confusions than native French in this context also.

The result that is most striking from this experiment is that vowels in isolation were easier for listeners to discriminate than vowels in TVT syllables. These data are all the more surprising in that they were obtained from both native and non-native speakers of French, as well as from non-speakers of French. As mentioned in the introduction, there is reason to believe that native American English listeners should perceive vowels in CVC syllables *more* easily than isolated vowels because of phonological expectations from English. Because non-French-speaking Americans discriminated French isolated vowels better than vowels in TVT syllables, we cannot attribute the difference between this study and the experiments with American English vowels to the phonological expectations of the listeners. These discrimination data, taken with the identification test results of Experiment 1, point to possible differences in the acoustic information that is important for vowel perception in English and French. To examine these differences and to investigate the influence that misarticulation might have had on vowel perception, acoustical measurements (formant frequencies and duration) were made of each vowel token in each syllable context.

3. Acoustical analysis

A. Method

Formant frequencies and syllable durations were measured from digitized waveforms that were stored on the PDP 11/34 computer. Syllable durations were determined from visual displays of the digitized waveforms. Vowel onset was determined as an abrupt amplitude increase. Offset for the TVT and VT syllables was marked at the point of final stop closure, indicated by a reduction in overall amplitude. Offset for the TV and #V# syllables was defined as the cessation of significant energy (i.e., at least twice the background noise level).

A fast Fourier transform (FFT) of each token was made from the center of the periodic portion of the syllable; the FFT sampled the waveform via a 24-ms Hamming window. Center frequencies were ascertained for the first two speech formants by examining the spectral peaks of the spectral section that resulted from the FFT, taking an average of peaks if there was more than one harmonic associated with a formant.

B. Results

(i) *Average syllable duration*

Table IX presents the average durations of the tokens in each of the syllable contexts for each of the vowels. As in English, vowels in isolation were usually longer than vowels in other syllable contexts. However, TVT and VT syllables were on the average longer than vowels in TV syllables, a result that runs counter to what occurred in English (see Gottfried, 1982). The relative durations of the vowels were not the same across different contexts. The vowels /ɑ/, /o/, and /ø/ are often transcribed with diacritics for length, and French phoneticians (e.g., Delattre, 1959) have noted that these vowels are considerably shorter in the open syllable contexts than in the closed syllable contexts, where they contrast with the shorter vowels /a/, /ɔ/, and /œ/. These vowels, as expected, were relatively long in the TVT and VT syllables compared to TV syllables; the other vowels in those syllable contexts were of about the same duration as in the TV context.

Table IX Mean absolute durations (in ms) for vowels in each syllable context

Vowel	Syllable context TVT	VT	TV	#V#	Mean
/i/	223	259	234	307	256
/e/	282	276	281	343	296
/ɛ/	256	271	274	338	285
/a/	266	251	254	297	267
/ɑ/	342	336	301	349	332
/ɔ/	273	251	263	306	273
/o/	309	346	261	336	313
/u/	243	309	260	334	287
/y/	269	290	275	324	290
/ø/	311	339	289	382	330
/œ/	285	307	299	370	315
Mean	278	294	272	335	

The two vowels that were inappropriate to open syllables (TV syllables and isolated vowels) were the low-mid vowels /ɔ/ and /œ/. The durations for these vowels in TV and isolated vowel syllables were not different from those of the high-mid vowels /o/ and /ø/. Apparently because these vowels do not occur in syllable-final position, speakers had difficulty making the distinction in duration that they made between high- and low-mid vowels in non-final position. French-speaking Americans showed confusions of low-mid → high-mid vowels, especially confusions of /ɔ/→/o/, although native French listeners did not.

The mid vowels /e/ and /ɛ/ are not differentiated by duration ordinarily, so the lack of duration differences between these vowels did not affect native listeners' errors. However, French-speaking Americans did seem to make /ɛ/→/e/ confusions more often in the TV and #V# conditions, as if they were influenced by the Law of Position (see Results, Experiment 1). It is also possible that they were treating the high-mid vowel /e/ like the tense English vowel, which in English is longer than the vowel /ɛ/. If this were true, then the lack of difference between these vowels in duration might have led the Americans to make /ɛ/→/e/ confusions in the open syllable contexts, just as Americans make lax → tense (or short → long) confusions on American English vowels.

(ii) *Average formant frequencies*

Figure 3 plots the vowel targets of the syllables used in each context condition, according to the frequency of the first formant (F1) versus the frequency of the second formant (F2), averaged over the tokens of all four speakers. As the figure shows, the four average vowel spaces for the different syllable contexts are quite similar. The formant frequencies for vowels in /t/ contexts tended to be less dispersed in the vowel space than isolated vowels. This is probably attributable to the effects of articulation with the consonant. The effect of initial /t/ on the second formant frequencies of /u/ and /o/ was less marked than for American English vowels (see Gottfried, 1982; Stevens & House, 1963). This is possibly due to French vowels being less affected by consonant articulation, as noted by Delattre (1953) in his discussion of the phonetic characteristics of French.

The average /a/ and /ɑ/ formant frequencies were very close to each other. The poor identification and discrimination of these vowels by listeners is thus attributable to the similarity of their spectral targets. The average /ɑ/ and /ɔ/ formant frequencies were also close,

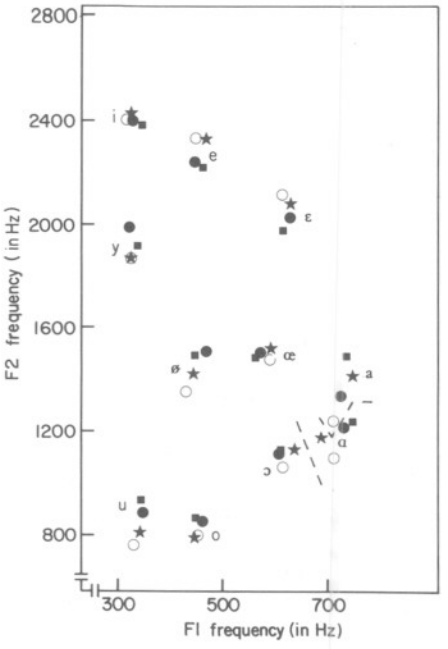

Figure 3 Mean first formant (F1) and second formant (F2) frequencies (four speakers) in four syllable contexts. ■, TVT; ● TV; ★ VT; ○ V.

especially in the VT and isolated vowel conditions. For example, the average formant frequencies for /ɑt/ were displaced in the direction of the formant frequencies for /ɔt/. It could be argued that this displacement is reflected in the higher number of /ɑ/→/ɔ/ confusions in the VT and isolated vowel contexts (37.5% error rate in the VT conditions; 26.6% in the isolated vowel condition) than in the other contexts (7.8% for TV; 14.8% for TVT).

The average /ø/ and /œ/ formant frequencies, on the other hand, were relatively distinct from each other in the vowel space. Thus, the many confusions of /ø/ and /œ/ cannot be attributed to poor differentiation of the productions as reflected in the average formant frequencies (but see individual token analysis below).

(iii) *Individual token analysis*

Average target values may not reveal the effect of individual misarticulated tokens on the vowel identification and discrimination performance. Figure 4 shows the first and second formant frequencies for each vowel token, for each speaker and syllable context. As is the case for American English vowels (Gottfried, 1982: Gottfried & Strange, 1980; Peterson & Barney, 1952; Strange *et al.*, 1979), there is considerable overlap in the formant frequencies of French vowels in different categories, and the second formant frequencies of French front vowels were higher for the female speakers' tokens. When the males' tokens are considered separately from the females', there still is overlap in the vowel space for /ɑ/ and /a/, for /ɑ/ and /ɔ/, and for /ø/ and /œ/. The acoustical measurements were examined for the particular tokens that were consistently identified by a majority of the native French listeners as a vowel other than the intended vowel. For some specific confusions, vowel durations could account for higher error rates (e.g., /œ/→/ø/ for one male's VT token; /ɔ/→/o/ on TVT tokens by another speaker. Several other common confusions by both French and American listeners might also be attributable to misleading or ambiguous vowel durations. On the

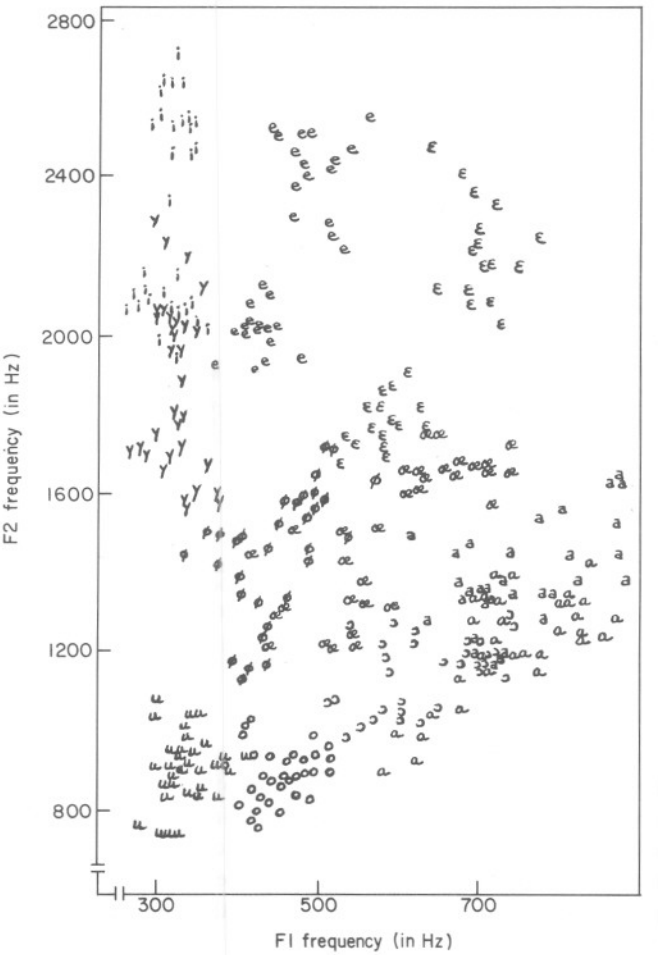

Figure 4 First formant (F1) and second formant (F2) frequencies for all tokens.

other hand, target formant frequencies usually could not account for common confusions. For example, a token of /tyt/ was misidentified as /tøt/ by 63% of the French listeners, but its formant frequencies were not greatly different from the other token of /tyt/ spoken by the same speaker, which was usually identified correctly.†

†It was suggested in discussing the results of Experiment 2 that the large number of /tit/-/tet/ confusions may have been due to a particular token of /tit/. The first and second formant frequencies of this token were 351 Hz and 2021 Hz, compared to an average of 300 Hz and 2087 Hz for this speaker's /e/ tokens in the other syllable contexts. This difference is probably the effect of coarticulation on the vowel, and may be observed in the formant frequencies for /tit/ produced by the other speakers. On the other hand, the first formant seemed to be raised more and the second formant lowered more in this particular token than in the other /tit/ token produced by this speaker (F1 = 332 Hz; F2 = 2099). It appears that the target formants for the first token of /tit/ were displaced in the direction of /tet/. It is also possible that other acoustic factors, such as diphthongization, affected the perception of this token.

D. Conclusions

It appears from the analysis of acoustic measurements that many of the perceptual confusions might have been due to poor differentiation of vowel targets and duration in speakers' articulation. Native French listeners seemed affected by the closeness of /a/ and /ɑ/ targets in the vowel space. American listeners were more affected than French by the poor differentiation of the duration of mid vowels in the TV and isolated vowel contexts.

In addition, the individual token analysis revealed that some of the common vowel confusions could be explained in terms of inappropriate vowel durations and target formant frequencies on specific tokens. However, not all highly confused tokens had aberrant target formant frequencies or durations. Other factors must account for the poor identification and discrimination of these vowels.

4. General discussion

One of the motivations for conducting cross-language speech perception research is the desire to make general statements about the nature of speech perception. The results of these experiments with French vowels are therefore all the more interesting because they contrast so greatly with the results of studies of American English vowels. Even with the highly confusable tokens removed, French vowels in TVT and VT syllables were *more* difficult to identify than vowels in TV and isolated vowels, by all listener groups, a result exactly opposite that found for American English vowels.

One might attribute the higher error rates on French vowels in TVT and VT syllables to their infrequency in these contexts in French relative to English. This explanation implies that the different distributional generalizations and preferred syllable structures of the two languages would lead native French listeners to make more vowel errors on the TVT and VT syllables in French and native English speakers to make more errors on the TV syllables and isolated vowels in English. The errors that French listeners made, however, were not of the kind predicted by such an explanation of phonological appropriateness. Furthermore, non-French-speaking American listeners, while they made more errors overall, showed the same consonantal context effects as the other two groups who had some experience with French phonology. For these reasons, one might consider the acoustic parameters that could account for the differences between the identification of English and French vowels. For example, in the French tokens, the lag time between burst and voiced formant transitions (VOT) was shorter, and the burst had less energy; the final stop release was weaker and tended to occur much longer after final stop closure than in English. It is not certain whether the lack of aspiration of initial stop consonants and weak, delayed final stop release would affect vowel perception. While consonant transitions have been shown to supply information about American English vowel identity (Jenkins *et al.*, 1983; Strange *et al.*, 1983), the specific nature and the relative importance for vowel perception of the acoustic parameters associated with the consonant have yet to be determined.

The influence of the French final stops on the acoustic characteristics of the vowels was mentioned briefly in the introduction. Delattre (1953) stated that French speakers tend to leave as many syllables open as possible in their speech. This observation seems to be borne out by at least some of the speakers: the delay between the final stop closure and the final stop release occasionally gave the impression that another syllable was starting. Acoustically, a final stop consonant in French did not seem to affect the formant transitions at the end of the syllable as greatly as in English.

The formant frequencies of French vowels were much more stable than those of English vowels. Delattre (1963) made cineradiographic measurements of articulation that showed

considerably less diphthongization in French vowels than in English vowels. The spectrographic measurements of the stimuli used in the present study support the same conclusion. In English TVT syllables, the formant frequencies rarely are stable across an appreciable length of time. For lax vowels especially, the formant frequencies shift throughout the syllable. The tense vowels are relatively more stable at the beginning of the syllable, but they are usually marked by considerable shifts in formant frequencies at the end of the syllable. In contrast most French TVT syllables had relatively short transitions and reached a "plateau" during which the formant frequencies did not shift appreciably. The transitions at the end of the syllable were also rapid, and the spectrogram of the syllable often looked remarkably symmetrical.

Because the French vowels are not diphthongized and are relatively stable in formant frequencies across the syllable, one might say that they are more like steady-state synthetic vowels than English vowels. The dynamic information in an English syllable is considerable – formant frequencies change continuously throughout the syllable, not just during a "consonant transition" portion. Certainly, a French syllable also has changes in acoustic parameters across time, but these changes are relatively small compared to those of English. The target formant frequencies therefore provide a more complete description of the acoustic properties of French vowels than of English vowels.

The results of these experiments with French might be interpreted to suggest that, although a "target" theory of vowel perception (Joos, 1948; Ladefoged, 1967) is inadequate to explain American English vowel identification and discrimination, such a theory might account for French vowel perception. The most accurate French vowel perception occurs in CV and isolated vowel contexts, where the target formant information should be better for differentiating vowels. French vowels, by virtue of their relatively "pure" (undiphthongized) quality may therefore be specified more accurately by the target formant information than English vowels.

On the other hand, by returning to the target formant account of vowel perception, one is left with the problem of the lack of invariance of target formant frequencies for vowels across speakers. Assmann *et al.* (1982) were able to categorize Canadian English vowels according to transformations of the raw formant frequency values, but their program was more successful when they incorporated more than one spectral section of the formant frequencies. Cursory inspection of the F1–F2 plot of French vowels from the present study shows considerable overlap in vowel categories. However, it is possible that attempts to determine the best transformations of raw spectral section data for differentiating the vowels would be more successful in a language like French in which the target formants are more clearly specified in the syllable.

Yet another difference between English vowels and French vowels is the use of duration in distinguishing spectrally adjacent vowels. Duration is an important source of information for identifying English vowels: the pairs /i/-/ɪ/, /æ/-/ɛ/, /ɑ/-/ʌ/, and /u/-/ʊ/ are differentiated acoustically (in closed syllables) not only by their characteristic formant frequencies, but also by the duration of the vocalic nucleus.

There are French vowels that are contrasted acoustically in a similar manner (i.e., the pairs /ɑ/-/a/, /o/-/ɔ/, and /ø/-/œ/). The duration measurements of VT and TVT syllables showed that the ratio of /ɑ/ to /a/ was 1.31; the ratio of /o/ to /ɔ/ was 1.25; and the ratio of /ø/ to /œ/ was 1.10. These ratios did not show as great a contrast as the ratio of long to short vowels in English (e.g., a ratio of 1.33 in the Gottfried, 1982, study). Klatt (1976) found that increments of 20% were necessary for a just noticeable difference in duration (a ratio of 1.20) for American English listeners. Hence, the observed differences in duration

may not have aided French or American listeners in differentiating /ø/ and /œ/. In addition, closed syllables, in which vowel duration provides distinctive information, are less frequent in French than in English (Dauer, 1983).

For these reasons, syllable duration may be less valuable as a source of information about French vowel identity. Perhaps speakers of "syllable-timed" languages like French (Ladefoged, 1975; Dauer, 1983) rely more on spectral information than on duration because the spectral information is more reliable than in languages with stress-timing, like English, which show vowel reduction in unstressed syllables. It would be necessary to test this idea with other syllable-timed languages (e.g., Spanish and Japanese) and other stress-timed languages (e.g., German and Russian).

All of these acoustic factors may have enabled listeners to perceive French vowels in isolation better than coarticulated vowels. It should be noted, however, that after the confusing tokens were removed from the analysis, French isolated vowels were not perceived by native French listeners any more accurately than English isolated vowels were by American English listeners (7.7% error rate for French; 6.7% error rate for American English listeners in Gottfried, 1982). It is possible that with a different group of French listeners or with different French vowel tokens one could observe lower error rates for French isolated vowels. Nevertheless, the French vowel identification data of this study are better described as showing significantly less accuracy for coarticulated vowels, *not* more accuracy for isolated vowels, relative to results on English vowels.

Thus, it seems that the presence of a final /t/ in the syllable is the determining factor in the difference between the French and English vowel studies. Final /t/ contributes information for the identification of English vowels, possibly in both the formant transitions associated with the consonant and in the clear demarcation of the syllable's duration. The final /t/ in French syllables is relatively unusual, there is not as much movement of the formants toward the end of such syllables, and vowel duration is not as important in differentiating French vowels as in differentiating English vowels. Thus, the presence of the final /t/ in French syllables did not aid vowel identification, and may have, in fact, interfered with the identification of vowels in TVT and VT syllables.

The presence of an initial /t/ in French syllables, however, did not interfere with accurate vowel identification. Indeed, the accuracy of vowel identification in TV syllables was numerically greater than the accuracy of isolated vowel identification. The difference was not statistically significant in this study, but suggests that the initial consonant might carry some useful information for French vowel identity in addition to target formant frequency information. The importance of initial consonants to French vowel identification is particularly likely if there are coarticulation effects of the vowel on the preceding consonant, as Delattre (1953) suggested. Data from studies by Benguerel and his colleagues (Benguerel & Adelman, 1976; Benguerel & Cowan, 1974) indicate that coarticulation effects of the vowel on the preceding consonants may be strong in French. Although listeners may not ordinarily rely on the acoustic information about vowel identity available in the consonants preceding the vowel, those consonants in French syllables carry information that can be used in vowel identification. Further study of French vowels in other CV syllables might reveal whether there is a reliable difference in identification accuracy between isolated vowels and vowels in syllables with initial consonants.

The results of these experiments also indicate that native speakers of French are in general better able to perceive the vowels of their language than non-native speakers of French and non-speakers of French. Several cross-language studies mentioned in the introduction suggest that such a relation between linguistic competence and speech percep-

tion often exists (e.g., Elman *et al.*, 1977; Goto, 1971; Politzer, 1961; Streeter & Landauer, 1976; Williams, 1979).

The fact that the non-native speakers were not as accurate in their perception might have been due, in part, to phonological expectations set up by second language phonetic instructions. That is, native French listeners had relatively little trouble ignoring violations of the general Law of Position, whereas French-speaking Americans made more of the expected errors based on this distributional rule. However, the difficulty experienced by the Americans might also be attributed to the relative importance of duration in differentiating the vowels of English (see Gottfried, 1982; Jenkins *et al*, 1983; Peterson & Lehiste, 1960; Strange *et al.*, 1983). By treating the high-mid vowel /e/ like the English tense vowel /ej/, the Americans would have expected a difference in French vowel duration (as in the English /ej/-/ɛ/ contrast) that did not occur between the French /e/ and /ɛ/. Hence, they made many /e/-/ɛ/ confusions.

The effect of the Americans' native language could also be seen in their performance on the discrimination task. The relative number of front rounded vowel confusions was greater for Americans (French-speaking and non-French-speaking) than for French listeners. The front rounded vowels do not occur in English, so they were relatively unfamiliar for the American listeners. When front rounded vowel comparisons were removed from the analysis of the discrimination test, error rates for native French and French-speaking Americans were about the same.

These results are similar to those of Politzer (1961), who found differences in vowel discrimination accuracy among learners of French. He reported that students with more years of instruction showed more accurate discrimination. Although the number of subjects tested in the present experiments was small, the possible correlation between years of instruction and perceptual performance was investigated. The Pearson product-moment correlation of error scores on the identification test (Experiment 1) and years of French study was not significant, $r = -0.25$, $t(7) = -0.63$. However, the correlation of error scores for the discrimination test (Experiment 2) and years of French study was significant, $r = -0.62$, $t(7) = -2.24$, $p < 0.05$. Although this correlation is suggestive of a relation between years of linguistic experience and perceptual test performance, the number of subjects was too low and the measure of linguistic experience was too approximate to make any broad generalizations. It would be of interest to examine this relationship more systematically (e.g., by obtaining independent measures of bilingualism from the French-speaking Americans).

The research reported here was part of a PhD dissertation for which I received support from the Graduate School of the University of Minnesota, from a grant awarded to J. J. Jenkins and W. Strange from the National Institute of Mental Health (MH-21153), and from research grants to the Center for Research in Human Learning from the National Institute of Child Health and Human Development (HD-01136, HD-07151) and the National Science Foundation (BNS 77-22075). I wish to thank Dr Winifred Strange for helpful comments on earlier drafts of this manuscript. I gratefully acknowledge Drs. James J. Jenkins, Winifred Strange, and Kathleen Houlihan for their guidance of the dissertation work; Patrice S. Beddor, Lenief Heimstead, and James Nead for their help in the collection and analysis of data; and Dr Genevieve Escure for her expert evaluation of the French vowel tokens.

Reprint requests may be sent to Terry L. Gottfried, Division of Natural Sciences, State University of New York, College at Purchase, Purchase, NY 10577.

References

Assmann, P. F., Nearey, T. M. & Hogan, J. T. (1982). Vowel identification: Orthographic, perceptual, and acoustic aspects. *Journal of the Acoustical Society of America,* **71**, 975–989.

Benguerel, A.-P. & Adelman, S. (1976). Perception of coarticulated lip rounding. *Phonetica,* **33**, 113–126.
Benguerel, A.-P. & Cowan, H. A. (1974). Coarticulation of upper lip protrusion in French. *Phonetica,* **30**, 41–55.
Caramazza, A., Yeni-Komshian, G., Zurif, E. B. & Carbone, E. (1973). The acquisition of a new phonological constrast: The case of stop consonants in French-English bilinguals. *Journal of the Acoustical Society of America*, **54**, 421–428.
Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics,* **11**, 51–62.
Delattre, P. (1951). *Principes de phonétique française à l'usage des étudiants anglo-américains.* Middlebury, Vermont: Middlebury College Press.
Delattre, P. (1953). Les modes phonétiques du français. *The French Review,* **27**, 59–63.
Delattre, P. (1957). La question des deux 'A' en français. *The French Review,* **31**, 141–148.
Delattre, P. (1959). Rapports entre la durée vocalique, le timbre et la structure syllabique en français. *The French Review,* **32**, 547–552.
Delattre, P. (1963). Voyelles diphtonguées et voyelles pures. *The French Review,* **37**, 64–76.
Diehl, R. L., McCusker, S. B. & Chapman, L. S. (1981). Perceiving vowels in isolation and in consonantal context. *Journal of the Acoustical Society of America,* **69**, 239–248.
Elman, J. L., Diehl, R. L. & Buchwald, S. E. (1977). Perceptual switching in bilinguals. *Journal of the Acoustical Society of America,* **62**, 971–974.
Gay, T. (1978). Effect of speaking rate on vowel formant movements. *Journal of the Acoustical Society of America,* **63**, 223–230.
Gottfried, T. L. (1979). Identification of French vowels. In: *Speech Communication Papers.* (Wolf, J. J. & Klatt, D. H. eds). New York: Acoustical Society of America.
Gottfried, T. L. (1982). *Perception of French and American English vowels: A cross-language study.* PhD dissertation, University of Minnesota.
Gottfried, T. L. & Strange, W. (1980). Identification of coarticulated vowels. *Journal of the Acoustical Society of America,* **68**, 1626–1635.
Goto, H. (1971). Auditory perception of normal Japanese adults of the sounds 'L' and 'R'. *Neuropsychologia*, **9**, 317–323.
House, A. S. & Fairbanks, G. (1953). The influence of consonant environment upon the secondary acoustical characteristics of vowels. *Journal of the Acoustical Society of America,* **25**, 105–113.
Jenkins, J. J., Strange, W. & Edman, T. R. (1983). Identification of vowels in 'vowel-less' syllables. *Perception & Psychophysics*, **34**, 441–450.
Joos, M. A. (1948). Acoustic phonetics. *Language Supplement,* **24**, 1–136.
Juilland, A. (1965). *Dictionnaire inverse de la langue française.* The Hague: Mouton.
Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America,* **59**, 1208–1221.
Ladefoged, P. (1967). *Three Areas of Experimental Phonetics.* New York: Oxford University Press.
Ladefoged, P. (1975). *A Course in Phonetics.* New York: Harcourt, Brace, Jovanovich.
Léon, P. R. (1966). *Prononciation du française standard.* Paris: Didier.
Lindblom, B. E. F. (1963). Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America*, **35**, 1773–1781.
Macchi, M. H. (1980). Identification of vowels spoken in isolation versus vowels spoken in consonantal context. *Journal of the Acoustical Society of America,* **68**, 1636–1642.
MacKain, K. S., Best, C. T. & Strange, W. (1981). Categorical perception of English /r/ and /l/ by Japanese bilinguals. *Applied Psycholinguistics,* **2**, 369–390.
Marckwardt, A. H. (1944). An experiment in aural perception. *The English Journal,* **33**, 212–214.
Marckwardt, A. H. (1946). Phonemic structure and aural perception. *American Speech*, **21**, 106–111.
Martinet, A. (1945). *La prononciation du français contemporain.* Paris: Librairie E. Droz.
Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J. J. & Fujimura, O. (1975). An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception & Psychophysics,* **18**, 331–340.
Monnot, M. (1977). The [e]-[ɛ] distinction in modern French: Phonetic and phonological implications. In: *Papers from the 1977 Mid-America Linguistics Conference.* (Lance, D. M. & Gulstad, D. E; eds). Columbia, Missouri: University of Missouri Press.
Nemser, W. (1971). *An Experimental Study of Phonological Interference in the English of Hungarians.* Bloomington, Indiana: Indiana University Press.
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.
Peterson, G. E. & Lehiste, I. (1960). Duration of syllable nuclei in English. *Journal of the Acoustical Society of America,* **32**, 693–703.
Politzer, R. L. (1961). Auditory discrimination of French vowels by English speakers. *Canadian Journal of Linguistics,* **7**, 32–44.
Sapon, S. M. & Carroll, J. B. (1958). Discriminative perception of speech sounds as a function of native language. *General Linguistics,* **3**, 62–72.

Sheldon, A. & Strange, W. (1982). The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics*, **3** 243–261.
Stevens, K. N. & House, A. S. (1963). Perturbations of vowel articulations by consonantal context: An acoustical study. *Journal of Speech and Hearing Research*, **6**, 111–128.
Strange, W., Edman, T. R. & Jenkins, J. J. (1979). Acoustic and phonological factors in vowel identification. *Journal of Experimental Psychology: Human Perception & Performance*, **5**, 643–656.
Strange, W. & Gottfried, T. L. (1980). Task variables in the study of vowel identification. *Journal of the Acoustical Society of America*, **68**, 1622–1625.
Strange, W., Jenkins, J. J. & Johnson, T. L. (1983). Dynamic specification of coarticulated vowels. *Journal of the Acoustical Society of America*, **74**, 695–705.
Strange, W., Verbrugge, R. R., Shankweiler, D. P. & Edman, T. R. (1976). Consonant environment specifies vowel identity. *Journal of the Acoustical Society of America*, **60**, 213–224.
Streeter, L. A. & Landauer, T. K. (1976). Effects of learning English as a second language on the acquisition of a new phonemic contrast. *Journal of the Acoustical Society of America*, **59**, 448–451.
Valdman, A. (1976). *Introduction to French Phonology and Morphology*. Rowley, Massachusetts: Newbury House.
Verbrugge, R. R., Strange, W., Shankweiler, D. P. & Edman, T. R. (1976). What information enables a listener to map a talker's vowel space? *Journal of the Acoustical Society of America*, **60**, 198–212.
Williams, L. (1977). The perception of stop consonant voicing by Spanish-English bilinguals. *Perception & Psychophysics*, **21**, 289–297.
Williams, L. (1979). The modification of speech perception and production in second-language learning. *Perception & Psychophysics*, **26**, 95–104.