

# Perception of Temporal and Spectral Information in French Vowels\*

Terry L. Gottfried<sup>†</sup> and Patrice Speeter Beddor<sup>††</sup>

*Perceptual effects of orthogonal variations in temporal and spectral information differentiating French /o/ and /ɔ/ were examined. Although both parameters contribute to acoustic differentiation of /o/ and /ɔ/, the phonetic and phonological structure of French suggests that duration might be a less important perceptual property in French than in languages like American English. Three 10-step /kot/-/kɔt/ continua were synthesized by systematically varying frequencies of the first two formants of the vowel nuclei. The three continua differed in vowel duration (140, 180, and 220 ms). Two perceptual tests, identification and 5-choice category rating, were presented to three listener groups: native French, native American English who had studied French, and native American English who did not know French. For both native American English groups, spectrally ambiguous vowels were identified and rated more often as /o/ when these vowels were long and as /ɔ/ when short, thereby showing a trading relation between temporal and spectral information. In contrast, native French listeners showed little effect of duration in either perceptual task. Despite this perceptual insensitivity to duration, acoustic measurements showed that these French subjects' productions of /o/ and /ɔ/ consistently maintained a duration difference. Results are interpreted to support the view that perceptual integration of the acoustic properties relevant to a phonemic contrast depends not only on covariation of the properties in the production of that contrast, but also on the prominence of this covariation in the language's phonological system.*

## INTRODUCTION

The temporal and spectral characteristics of vowels covary in the vowel contrasts of many languages (Lehiste, 1970; Straka, 1959). In some of these languages, phonetic measures and phonological patterning point toward vowel length as the distinctive property of the relevant contrasts (e.g., Czech, Hungarian, and Serbo-Croatian). In other languages in which temporal and spectral differences co-occur, vowel quality rather than quantity may be analyzed as the distinctive property. For example, in American English, the "tense-lax" pairs, such as /i, ɪ/, /æ, ɛ/, /u, ʊ/, and /ɑ, ʌ/ differ in both formant frequency and intrinsic duration: Tense vowels are more peripheral and longer than their lax counterparts (Peterson & Lehiste, 1960). These length

differences are traditionally analyzed as phonologically redundant (but see Delattre, 1962, for discussion).

Whether or not vowel length is contrastive in a language has been shown to influence certain perceptual judgments by speakers of that language. For example, Bastian and Abramson (1962) obtained steeper, more consistent labeling functions for Thai listeners than for American listeners on a length contrast in Thai. Similar differences were reported by Keating (1985) for Czech and American listeners on a length contrast phonemic in Czech.

Such findings, of course, do not mean that phonologically non-contrastive vowel duration is perceptually irrelevant. Two studies (Assmann, Nearey, & Hogan, 1982; Strange, Jenkins, & Johnson, 1983) found that when the nuclei of natural (Canadian and American English) vowels were truncated so that all vowels had the same length, listeners confused spectrally adjacent vowels that normally differ in length. Rakerd (1984) showed that a major factor in listeners' judgments of similarities between natural American English vowels was vowel duration. Indeed, under certain conditions, a duration difference alone is sufficient to specify a vowel contrast in English. Ainsworth (1972) showed that variations in duration of synthetic vowels affected the category name given these vowels by British English listeners. Mermelstein (1978) showed that synthetic vowels that were spectrally ambiguous for /ε/ and /æ/ (i.e., the frequency of the first formant fell between typical /ε/ and /æ/ values) were labeled according to vowel duration: Longer vowels were called /æ/ and shorter vowels /ε/ (see also Stevens, 1959).

There is substantial perceptual evidence, then, that vowel duration conveys information for vowel identification in English. Such findings might lead us to expect vowel duration to be perceptually relevant for any language in which, like English, phonologically nondistinctive length covaries with spectral properties. Alternatively, we might hypothesize that, for a given language, the perceptual relevance of duration for a particular vowel contrast depends not simply on the presence of duration as an acoustic correlate of that contrast, but rather on the overall prominence of vowel duration within the language's phonological system.

Bennett (1968) compared the perceptual role of vowel duration in English and German. Presented with an unfamiliar vowel contrast varying in both temporal and spectral characteristics ([u:]-[y]), British English listeners assigned more weight than did German listeners to the temporal variation. This perceptual difference may be related to phonological differences between English and German. In both languages, duration and spectral properties covary in several vowel contrasts. However, it is only in English that vowel duration provides information for voicing of a following obstruent, as well as for vowel identity.

French also differs from English in the phonological prominence of vowel duration, although the specific differences between French and English are not the same as those between German and English. Various factors suggest that vowel duration has a relatively minor role in Parisian French (Delattre, 1959; Fry, 1968). For example, temporal and spectral cues covary in only three vowel contrasts in French (Delattre, 1959): /a/-/a/, /ø/-/æ/, and /o/-/ɔ/. The phonemic status of the first two of these pairs is questionable. The /a/-/a/ contrast is not commonly observed in spoken French (Delattre, 1957), and the /ø/-/æ/ contrast is confined to two minimal pairs (Valdman, 1976, p. 56). The only common vowel distinction that consistently maintains spectral and temporal differences involves the mid back vowels /o/ and /ɔ/, as in the minimal pairs *paume-pomme* 'palm-apple,' *saute-sotte* 'leaps-foolish,' and *côte-cotte* 'rib-tunic.' These vowels are opposed only in syllables closed by a final consonant. (This is similar to the situation in English where the "tense-lax" vowel pairs also contrast only in closed syllables.)

Spectrally, French /o/ and /ɔ/ consistently differ in first and second formant ( $F_1$  and  $F_2$ ) frequencies (Chollet & Malécot, 1980). Temporally, measurements by Bricler-Labaeye (1970) indicate that /o/ is roughly 20-40% longer than /ɔ/ in closed syllables. Gottfried (1984) found that /o/ was, on the average, 25% longer than /ɔ/ in /t/-vowel-/t/ and vowel-/t/ syllables produced by native speakers of French. (Intrinsic vowel duration differences are larger for some vowel contrasts in English. See Peterson & Lehiste, 1960; Strange, Verbrugge, Shankweiler, & Edman, 1976.)

Thus spectral and temporal information simultaneously vary in certain vowel contrasts in both French and English. However, the role of temporal information is less prominent in French in that the number of such contrasts is smaller in French (Delattre, 1959), and the frequency of the closed syllable context, where such contrasts occur, is less in French. Dauer (1983) reports that English closed syllables account for 56% of the total in spoken language, whereas French closed syllables account for only 26%.

The present study investigated whether this difference in vowel systems has perceptual manifestations, such that French listeners are less sensitive than American English listeners to temporal differences in vowels. Some previous research points toward the possibility of such perceptual differences. Gottfried (1984) examined the accuracy of identification of natural French vowels by native French and by French-speaking American listeners. Analysis of the errors by the Americans indicated inappropriate use of vowel duration in identifying the vowels. For example, when a natural token of French /ɔ/ was unusually long in duration, the Americans were more likely than the French to label that token /o/—that is, for the Americans, but not for the French, the temporal cue overrode the spectral cue.

This study used the trading relations paradigm (see Repp, 1982) to test the relative contribution of temporal and spectral information to vowel identification. In this paradigm, two acoustic parameters that naturally covary in some phonetic contrast are manipulated independently of each other. In instances where both parameters are perceptually relevant, a change in the value of one parameter can be "traded off" against an opposing change in the other parameter, thereby maintaining phonetic category identity. For example, voiceless and voiced initial stop consonants differ in voice onset time (VOT) and onset frequency of the  $F_1$  transition. To maintain phonetic equivalence of intermediate stimuli, VOT must be increased when  $F_1$  onset frequency is lowered (Summerfield & Haggard, 1977). Trading relations have also been found for cues signaling place of articulation in stops (e.g., Dorman, Studdert-Kennedy, & Raphael, 1977), manner of articulation in consonants (e.g., Repp, Liberman, Eccardt, & Pesetsky, 1978), and many other phonetic contrasts.

The focus of our study was whether French and American English listeners differed in the extent to which temporal information can be traded against spectral information in the perception of the French vowels /o/ and /ɔ/. American listeners were expected to use duration, as well as spectral information, as cues to the identity of these vowels. This expectation was based on the systematic role of duration in the perception of English vowels demonstrated by the studies cited above.

Given the systematic difference in the duration of /o/ and /ɔ/ in French, French listeners might also be expected to be sensitive to temporal information. Under this hypothesis, French listeners should identify long, spectrally ambiguous vowels as /o/ and short, spectrally ambiguous vowels as /ɔ/. That is, vowel identity should shift as a function of vowel duration. This prediction is consistent with trading relations studies that indicate integration of acoustic information when the sources of information cospecify a phonemic contrast (Repp, 1982, 1983).

Alternatively, the restricted role of duration in the French vowel system might lead us to predict that French listeners would be insensitive to temporal cues, even

for a vowel distinction in which duration differences are systematically maintained in French. The latter hypothesis claims that perceptual integration of acoustic properties relevant to a given vowel contrast depends in part on the extent to which these acoustic properties correlate within a phonological system.

### Method

#### Stimuli

The stimuli were five-formant synthetic versions of /o/ and /ɔ/ embedded in a /k/-vowel-/t/ context, corresponding to the French words *côte* and *cotte*. All stimuli were generated on a serial software formant synthesizer at Haskins Laboratories. The acoustic characteristics of the two endpoint synthetic stimuli were based on an analysis of natural tokens of the French words, spoken by a male native speaker of French. Intermediate versions of these syllables were synthesized by manipulating the spectral and temporal characteristics of the vowel nuclei.

The spectral manipulation was in  $F_1$  and  $F_2$ . Table 1 shows the initial  $F_1$  ( $F_{1i}$ ), steady-state target  $F_1$  ( $F_{1t}$ ), and final  $F_1$  ( $F_{1f}$ ) for each of the 10 steps in the /kot/-/kɔt/ continuum. Likewise, the initial ( $F_{2i}$ ), target ( $F_{2t}$ ), and final ( $F_{2f}$ ) values for  $F_2$  are shown. Initial, target, and final  $F_3$  were 2275 Hz, 2380 Hz, and 2800 Hz, respectively, for all stimuli. The  $F_4$  was fixed at 3300 Hz and  $F_5$  at 3850 Hz. The initial burst for /k/ was 20 ms, followed by a 20-ms voice-onset time. The initial  $F_1$  transition was 30 ms long; the initial  $F_2$  transition was 40 ms; and the initial  $F_3$  transition was 20 ms, starting 10 ms after the  $F_1$  and  $F_2$  transitions. All final transitions were 40 ms. After 40 ms of silence, there was a 20-ms final burst for /t/. The  $F_0$  was at 110 Hz for the initial transitions and steady-state, and fell linearly to 85 Hz in the last 80 ms of voicing.

TABLE 1

Initial (i), target (t), and final (f) synthesis values for  $F_1$  and  $F_2$  of the ten step/kot/-/kɔt/ continuum.

Stimulus Number	Formant values (in Hz)					
	$F_{1i}$	$F_{1t}$	$F_{1f}$	$F_{2i}$	$F_{2t}$	$F_{2f}$
1	315	400	260	1270	800	1340
2	320	415	265	1285	835	1355
3	325	430	270	1300	870	1370
4	330	445	275	1315	905	1385
5	335	460	280	1330	940	1400
6	340	475	285	1345	975	1415
7	345	490	290	1360	1010	1430
8	350	505	295	1375	1045	1445
9	355	520	300	1390	1080	1460
10	360	535	305	1405	1115	1475

Three temporal variations of each of these 10 stimuli were synthesized by manipulating the duration of the vowel steady-state. Vowel length (including formant transitions) was 220 ms for long vowels, 180 ms for medium vowels, and 140 ms for short vowels. The long vowel length was appropriate for French *côte* and the short vowel length for *cotte*. The ten spectral and three temporal variations provided a total of 30 stimuli.

Three audio tapes were constructed using these stimuli: a familiarization set, an identification test, and a rating test. The familiarization set consisted of 15

---

randomly selected members of the stimulus set. The identification test contained 10 instances of each of the 30 stimuli, for a total of 300 test items arranged in random order and grouped in blocks of 20. The interstimulus interval within blocks was 1.5 s; the interblock interval was 5 s.

The 2-choice identification task, which provided one measure of phonetic categorization, was supplemented by a 5-choice category rating task where the endpoint categories were for good exemplars of /o/ and /ɔ/, and the middle categories were for vowels judged to be intermediate to the good exemplars. We speculated that this rating task might be more sensitive to the effects of vowel duration than the identification task. That is, even though vowel duration might not influence *phonemic* categorization, it might affect finer judgments of categorization, for the French as well as the American listeners. The rating test consisted of 3 instances of each of the 30 stimuli, yielding 90 randomly arranged test items that were grouped into blocks of 20. The rating test interstimulus interval within blocks was 3 s; the interblock interval was 5 s.

### Subjects

Three groups of 12 listeners each participated in the experiment: native speakers of French, native speakers of American English who had studied French (that is, "French learners"), and native American English speakers with no knowledge of French (that is, "naive English"). These three groups enabled us to look at the effect of duration on vowel identification across various degrees of competence in a particular phonological system: native knowledge versus second language knowledge versus no knowledge of the vowel system within which the /o/-/ɔ/ contrast functioned.

The native American English speakers were students at Yale University or the State University of New York at Purchase. The French learners had studied French at least five years or had spent one or more years in France. Most of the native French speakers were students or faculty at Yale or SUNY-Purchase. All subjects were paid for their participation.

### Procedure

Subjects were tested in small groups of one to three in sound-attenuated rooms at Haskins Laboratories and at SUNY-Purchase. Stimuli were presented binaurally over headphones in a 1-hour session. Instructions were given in English for all listeners. The same procedure was used for the three groups, except that, prior to testing, the naive English listeners heard a brief tape-recorded training series consisting of natural-speech tokens of *côte* and *cotte* produced by a native French speaker. They were told that these were natural speech examples of the French vowels they were going to identify in the perceptual tests. Otherwise, the experiment consisted of the familiarization set, the identification test, and the rating test, in that order.

Subjects were told they would hear computer-generated tokens of *côte* and *cotte*. The familiarization set was presented to acquaint listeners with the stimuli and the speed of the identification test. In the identification test, subjects were instructed to listen to each syllable and record the vowel in each syllable on printed score sheets. Native French and French learners recorded "ô" or "o"; naive English subjects (who were unfamiliar with French orthography) recorded "o" or "ɔ," symbols that were explained to them during their training session.

After the identification test and a short break, listeners were informed that the identification test had contained several different versions of *côte* and *cotte*, but that some versions were more natural, or more native-like, than others. A rating test required them to categorize vowels on a 5-point scale, where 1 represented a good exemplar of /o/, 3 represented an ambiguous vowel midway between a good /o/ and

/ɔ/, and 5 a good /ɔ/. Listeners circled their ratings on score sheets. Thus 10 identification judgments and three ratings per subject were obtained for each stimulus.

### Results and Discussion

The pooled results of the identification test for the 12 native French speakers are shown in Figure 1. Percent /o/ responses are plotted as a function of stimulus number (corresponding to the ten spectral changes) for the three vowel durations. The vertical line drawn through the 50% point in the curves indicates the /o/-/ɔ/ boundary for the functions. As the figure shows, the cross-over point was not different for the three durations. An analysis of variance was performed on the number of /o/ responses for each vowel duration, summed over the 10 spectral patterns. For the native French speakers, the number of /o/ responses did not significantly change as a function of vowel duration,  $F(2,22)=0.03$ .

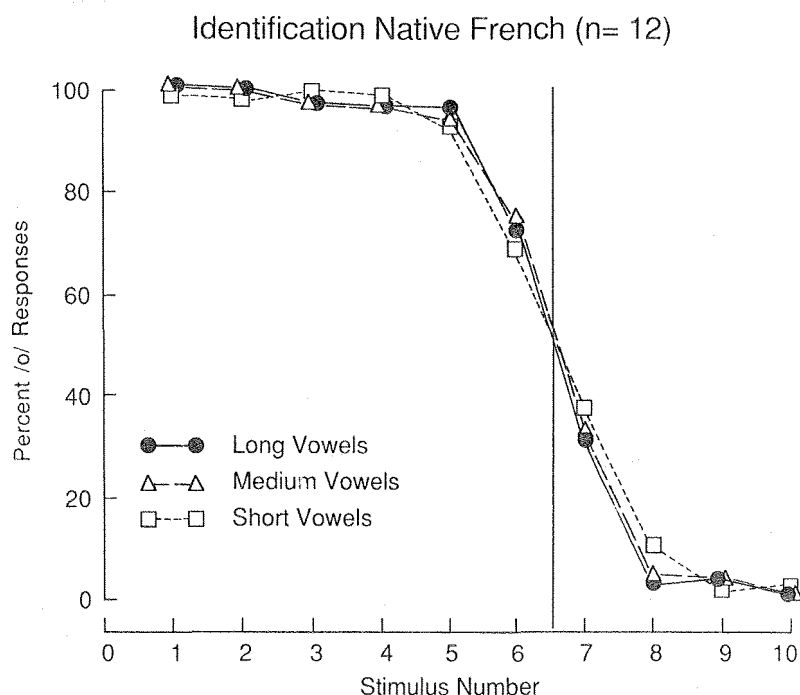


Figure 1. Identification of synthetic vowel series by native French listeners.

Figure 2 gives the identification test results for the 12 French learners. Unlike the native French, the learners' performance showed an effect of vowel duration, as shown by the diverging (rather than overlapping) identification curves for the three vowel durations and by the different 50% crossover points (indicated by the vertical lines) for the three durations. The number of /o/ responses by the French learners increased significantly as duration increased,  $F(2,22)=16.93$ ,  $p<.001$ .

However, these French-speaking American listeners had quite diverse backgrounds in French. It is possible that those learners with near-native competence in French performed more like native French speakers than did the less competent learners. To address this question, we divided the 12 French learners into two groups of six: advanced graduate students in French (all of whom reported native or near-native fluency) versus all other learners. As seen in Figure 3, the identification curves of the three vowel durations overlapped more for the advanced learners (upper graph) than they did for the other learners (lower graph). In other words, the advanced learners responded more like the native French than did the less

advanced learners. Although both learner groups showed an effect of duration, a significant interaction of duration with group indicated that the duration effect was stronger in the less advanced group,  $F(2,20)=3.80, p<.05$ .

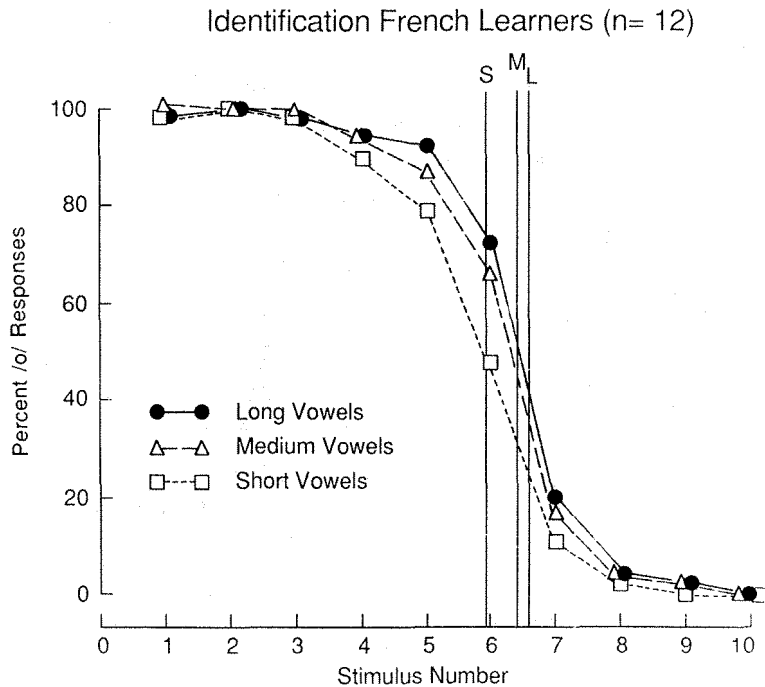


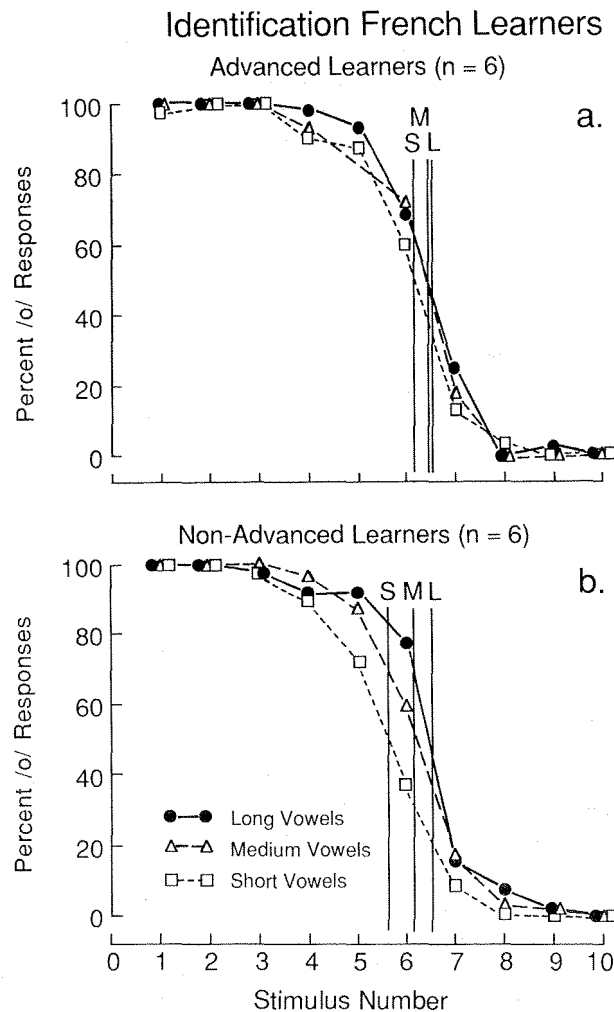
Figure 2. Identification of synthetic vowel series by French learners.

The identification data for the naive English speakers are given in Figure 4. The naive English listeners were generally less consistent than the French learners (especially the more advanced learners) in labelling the stimuli as /o/ or /ɔ/, as shown by the relatively gradual slope of all three identification functions. However, like the learners, the naive English listeners showed a significant effect of vowel duration [ $F(2,22)=14.78, p<.001$ ] and this effect was not significantly different from that of the learners [ $F(2,22)=0.78$ ].

The 5-choice category ratings are summarized in Figure 5. The figure plots the mean rating (averaged over spectral differences) given to vowels as a function of duration for the three listener groups. Given our 5-point scale, the mean rating would be 3.0 if the vowels of a particular duration were rated /o/-like and /ɔ/-like in equal proportions. The more vowels that are rated /o/-like, the lower the mean rating would be. The ordinate of Figure 6 is inverted so that more /o/-like rating is plotted toward the top of the graph.

The line representing the responses of the native French listeners has nearly a zero slope, which shows that the 5-choice category rating (like the 2-choice vowel identification) was not affected by vowel duration for these listeners. In contrast, the sloping functions of both native English groups indicate an effect of duration such that more of the long stimuli were rated as close to /kot/ and more of the short stimuli as close to /kɔt/. However, statistical analysis did not reveal a significant interaction of duration effects with language group performance,  $F(3,66)=2.07$ . This might be attributed to the high variability of the naive English group. These subjects often assigned different ratings to the three presentations of each stimulus, and they generally reported this task to be much more difficult than the identification test. Assuming that the large variation in these responses reflected a relatively poor ability to rate unfamiliar stimuli consistently, we omitted the naive English

listeners from the overall analysis. The revised analysis revealed a significant interaction of duration and linguistic background—that is, vowel duration affected categorization ratings by the learners, but not the native French listeners,  $F(2,44)=7.75, p<.01$ .



**Figure 3.** a. Identification of synthetic vowel series by advanced French learners. b. Identification of synthetic vowel series by non-advanced French learners.

Our findings suggest that, although native English listeners perceptually integrated duration and spectral information in categorizing /kot/ and /kɔt/, the native French were insensitive to temporal information in this vowel contrast. This is somewhat surprising given that temporal (as well as spectral) differences have been reported in French productions of this vowel pair (Brichler-Labaeye, 1970; Delattre, 1959; Gottfried, 1984). To determine whether this reported duration distinction between /o/ and /ɔ/ was indeed maintained in the productions by our native French subjects, we collected speech samples from five female native speakers of Parisian French, all of whom had already participated in the perceptual part of the study.



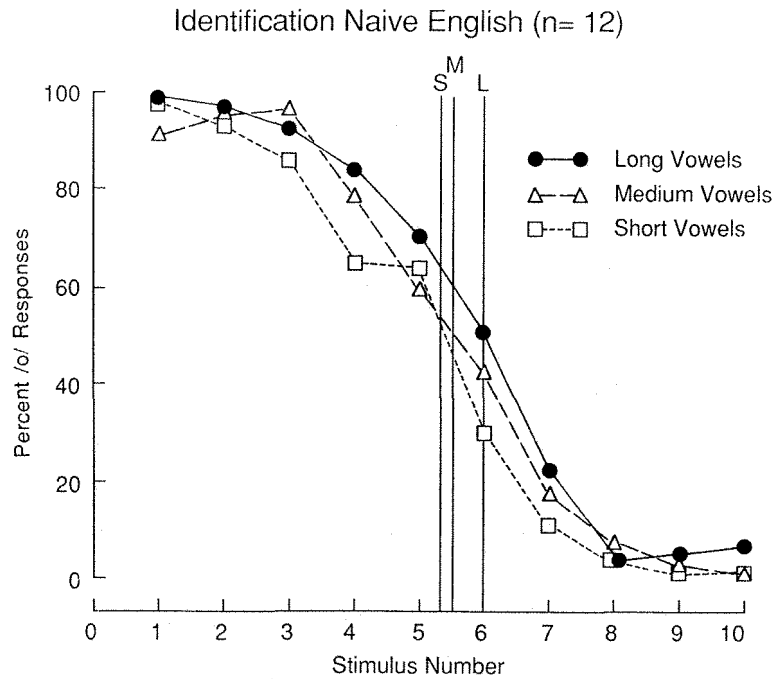


Figure 4. Identification of synthetic vowel series by naive English listeners.

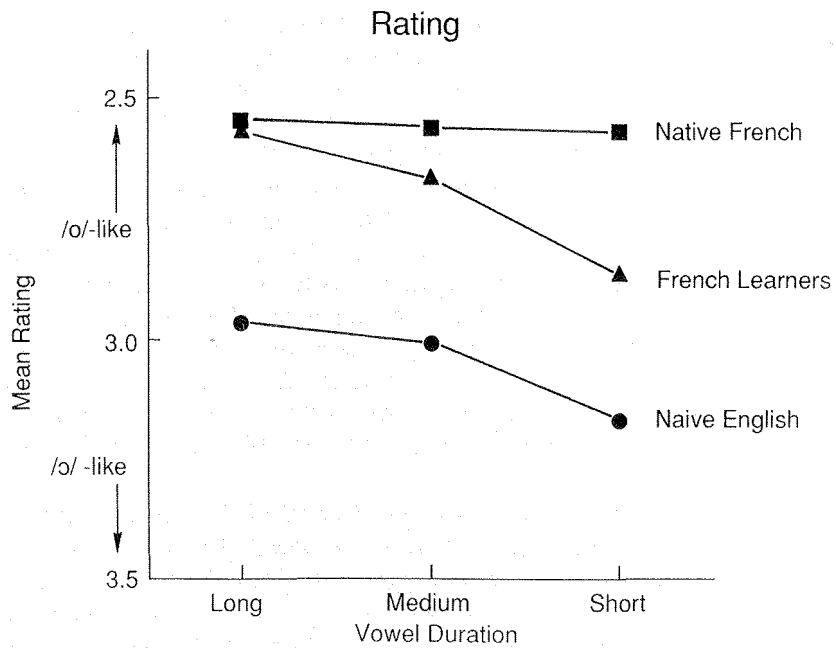


Figure 5. Mean rating of synthetic vowels as a function of duration by native French, French learners, and naive English.

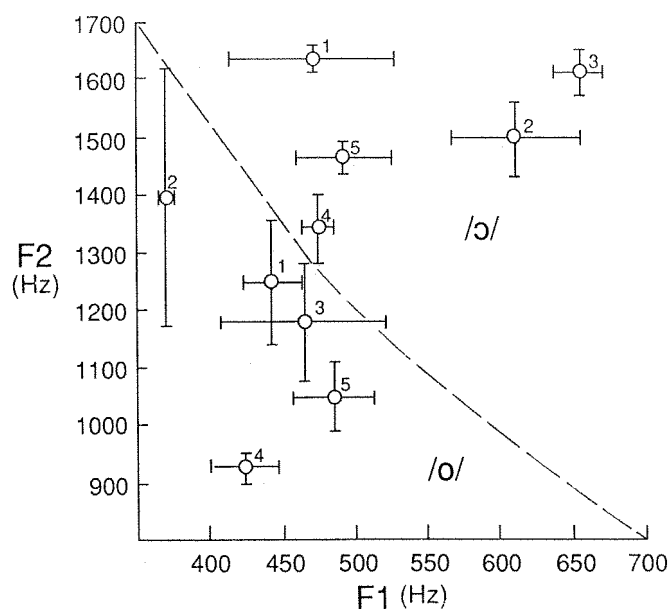


Figure 6. Mean first ( $F_1$ ) and second ( $F_2$ ) formant frequencies (in Hz) for /o/ and /ɔ/ spoken by five talkers.

## ACOUSTICAL ANALYSIS OF NATURAL SPEECH

### Method

The key words in the natural speech samples were six minimal pairs differing in /o/ and /ɔ/: *côte-cotte* 'rib-tunic,' *saute-sotte* 'shift-foolish,' *paume-pomme* 'palm-apple,' *rauque-roc* 'hoarse-rock,' *nôtre-notre* 'ours-our,' and *rôde-rodé* 'prowl-grind.' Eight additional words (viz., *belle* 'beautiful,' *flamme* 'flame,' *peur* 'fear,' *sèche* 'dry,' *rage* 'rage,' *sûr* 'sure,' *dîtes* 'say,' and *foule* 'crowd') were included so that the speakers would be less likely to be aware of the vowel contrast of interest.

Each of these 20 words was embedded in two frames: a variable sentence (see Appendix) and a fixed carrier sentence. The variable frame provided relatively natural and meaningful conditions for the production of the key words. These sentences were constructed to have approximately the same number of syllables, with each member of a minimal pair occupying a similar sentential position. The fixed carrier sentence, *On dit le mot \_\_\_\_\_ souvent* ('They say the word \_\_\_\_\_ often'), provided control over the acoustic context of the key words. Speakers read each randomly-arranged list of sentences twice, so that there were four instances (per speaker) of each key word, two produced in the variable frame and two in the fixed carrier frame.

The 20 tokens (4 repetitions x 5 speakers) of each of the 12 key words containing /o/ or /ɔ/ were digitized at a sampling rate of 10 kHz, after low-pass filtering at 4.9 kHz. Measurements of vowel duration were obtained from the digitized speech samples (using a waveform display program at Haskins Laboratories). Vowel onset was determined by the beginning of the significant periodic portion of the waveform for syllables with initial /p/, /k/, and /s/. Vowel onset of syllables with initial /n/ and /r/ was determined from the point at which the waveform changed its characteristic pattern, which usually corresponded with a noticeable increase in overall amplitude. Vowel offset was determined by the cessation of significant periodic pulsing for syllables ending in stop consonants. For the syllable ending in /m/, offset was determined by the change in characteristic waveform pattern and the

sharp reduction in overall amplitude. Thus, the vowel durations included formant transitions as well as the relatively steady-state portion.

Formant frequencies were measured for all /k/-vowel-/t/ words, using a linear predictive coding (LPC) analysis. A Hamming window of 24 ms width was placed on the part of the signal that had the most extreme formant excursion (as determined by visual inspection) from the initial and final transition portions. This was usually in the first half of the vowel and corresponded to a point in the vowel shortly (about 50 ms) after the rapid increase in amplitude at vowel onset. Formant frequency measurements were accurate within a range of  $\pm 20$  Hz.

### Results and Discussion

Table 2 gives the mean durations for the /o/ and /ɔ/ portions extracted from the six minimal pairs. As shown by the values in the table, vowel durations varied as a function of sentence frame and syllabic context (cf. O'Shaughnessy, 1981). There was also considerable variation in duration from speaker to speaker. However, averaging across sentence frame, all five speakers produced a longer /o/ than /ɔ/ in each of the vowel contrasts, except /rod/-/rɔd/.<sup>1</sup>

TABLE 2

Mean duration in ms (averaged over two tokens from five speakers) for /o/ and /ɔ/ and ratio of long/short vowel duration for six minimal pairs in variable and fixed sentence frame.

	Sentence Frame		Mean
	Fixed	Variable	
/kot/	143	78	111
/kɔt/	97	73	85
Ratio	1.48	1.06	1.30
/sot/	114	78	96
/sɔt/	81	75	78
Ratio	1.41	1.05	1.24
/pom/	169	136	152
/pɔm/	121	82	102
Ratio	1.40	1.65	1.50
/rok/	152	161	157
/rɔk/	107	81	94
Ratio	1.43	1.99	1.67
/notr/	161	106	133
/nɔtr/	138	65	101
Ratio	1.17	1.62	1.31
/rod/	205	164	184
/rɔd/	188	168	178
Ratio	1.09	0.97	1.03
Mean /o/	157	121	139
Mean /ɔ/	122	91	106
Ratio	1.29	1.33	1.31

The effect of sentence frame on vowel duration was not consistent across syllabic contexts. The duration difference between /o/ and /ɔ/ was larger in the variable sentence frame than in the fixed sentence frame for /pom/-/pɔm/, /rok/-/rɔk/, and /notr/-/nɔtr/. However, /kot/-/kɔt/ and /sot/-/sɔt/ were better differentiated in the fixed frame. Most speakers in the variable sentence frame did not maintain a difference in duration between /kot/ and /kɔt/ or /sot/ and /sɔt/.<sup>2</sup>

By contrast, the spectral difference between /kot/ and /kɔt/ was very large. Mean  $F_1$  values for /kot/ and /kɔt/ were 438 Hz and 541 Hz, respectively; mean  $F_2$  values were 1160 Hz and 1512 Hz, and mean  $F_3$  values were 2608 Hz and 2654 Hz, respectively. Figure 6 plots the mean  $F_1$  and  $F_2$  for each of the five speakers for /kot/ and /kɔt/. Error bars represent the standard error of the formant value for each individual speaker. The figure shows that all five speakers made a clear spectral distinction between /kot/ and /kɔt/.<sup>3</sup>

### GENERAL DISCUSSION

The results of the perceptual tests of native French listeners show no evidence of perceptual integration of spectral and temporal information in the /kot/-/kɔt/ contrast. These listeners were insensitive to temporal information not only in categorizing the vowels as /o/ or /ɔ/, but also in a 5-choice category rating task. Acoustical analysis revealed that five of these same native French, who did not use duration as a perceptual cue, consistently maintained a distinction in duration (at least in stressed position) in their productions of these vowels.

We believe that the explanation for this discrepancy between perception and production measures lies in the phonological structure of the French vowel system. The /o/-/ɔ/ distinction is the only common vowel contrast in French in which duration and formant frequencies covary. We suggest, then, that our native French listeners were insensitive to duration as a cue to vowel identity in this contrast because duration is not a reliable cue in other French contrasts.

Support for a phonological explanation of the native French data is provided by the results of the native English listeners, inasmuch as the same stimuli were perceived differently by the two language groups. In contrast to the native French listeners, the native English listeners were influenced by the duration of French /o/ and /ɔ/ in their categorization of these vowels. The English listeners' sensitivity to duration as a cue to the French distinction is consistent with the prominent role of duration in the English vowel system (Ainsworth, 1972; Peterson & Lehiste, 1960; Rakerd, 1984). Our American English data are also similar to the British English results reported by Bennett (1968): Both subject groups used duration in categorizing the vowels of an unfamiliar contrast.

If the effect of duration shown by the native English speakers is a function of the systematic role of duration in differentiating spectrally adjacent vowels in English, it is not surprising that both groups of English speakers—the naive English and the French learners—demonstrated this effect. The naive English listeners apparently used the temporal information appropriate to English vowel contrasts in labeling the unfamiliar /o/-/ɔ/ contrast. (Many American listeners said that the vowels sounded like the English /oʊ/-/ʌ/ contrast, as in 'coat' vs. 'cut.')

Similarly, the French learners also used this information, since French /o/ and /ɔ/ (which differ in duration as well as spectral shape) provide no phonetic impetus for the learners to discard vowel duration as a relevant cue. Yet we do see a suggestion of a developmental trend among the learners: The advanced learners were more like the native French in their perception of these vowels than were the less advanced learners.

In conclusion, our findings showed that vowel duration was perceptually relevant for the native English listeners, but not the native French, as a cue to the /kɔt/-/kɔt/ distinction. While it is of course possible that the French might demonstrate a trading relation under more stringent conditions (e.g., smaller spectral steps might lead to a greater duration effect), comparison with the native English listeners indicates a relative insensitivity of the French to temporal information. We have argued that this insensitivity is a function of the highly restricted role of vowel duration in the French vowel system, but it is possible that our perceptual data also reflect a change in progress involving the French /o/-/ɔ/ contrast. That is, it may be that a phonological contrast for which a particular phonetic parameter is an acoustic, but not a perceptual, correlate is an unstable contrast. If so, then the current perceptual irrelevance of temporal information may be followed at some later time by loss of temporal differences in the production of /o/ and /ɔ/. Regardless of the future of the /o/-/ɔ/ contrast in French, our perceptual data lead us to the following conclusion: Perceptual integration of the acoustic properties relevant to a given vowel contrast does not simply follow from experience with that contrast. Rather, perceptual integration depends on the extent to which the acoustic properties correlate within the broader context of a phonological system.

### ACKNOWLEDGMENT

This research was supported by NIH Contract N01-HD-2910 and Grant HD-01994 to Haskins Laboratories, the State University of New York Research Foundation (to the first author while at SUNY at Purchase), and NIH Postdoctoral Fellowship Grant NS-07196 (to the second author). We wish to thank Dr. Bruno H. Repp and two anonymous reviewers for their helpful comments on earlier versions of the manuscript.

### APPENDIX

Sentences used as the variable sentence frames.

1. Sa voix était rauque après qu'elle avait chanté.  
'Her voice was hoarse after she had sung.'
2. J'ai vu une belle fleur dans le jardin.  
'I saw a beautiful flower in the garden.'
3. Il y avait une saute de vent ce matin.  
'There was a gust of wind this morning.'
4. Nos amis n'aiment pas notre histoire.  
'Our friends didn't like our story.'
5. Son souffle fait trembloter la flamme de la chandelle.  
'His breath makes the candle's flame tremble.'
6. Il mange la pomme jaune que je lui ai donnée.  
'He is eating the yellow apple that I gave him.'
7. L'enfant avait peur quand il se couchait.  
'The child was afraid when he went to bed.'
8. Elle a mangé une côte de boeuf hier soir.  
'She ate a beef rib yesterday evening.'
9. On dit qu'il rôde les rues en quête de victimes.  
'They say that he prowls the streets in search of victims.'

10. Le charpentier sèche le chêne vert.  
'The carpenter is drying the green oak.'
11. Elle n'est pas assez sott e pour le croire.  
'She isn't silly enough to believe it.'
12. Ça lui faisait rage de voir ça.  
'It enraged him to see that.'
13. La gitane qui a lu dans ma paume avait tort.  
'The gypsy who read my palm was wrong.'
14. Il a mis la cote de mailles avant le combat.  
'He put on the coat of mail before the combat.'
15. Je suis sûr qu'elle viendra demain.  
'I am sure that she'll come tomorrow.'
16. Le mécanicien rode la soupape d'échappement.  
'The mechanic is grinding the escape valve.'
17. Que dites-vous de son tableau récent?  
'What do you say about her recent picture?'
18. Ils ont fait foule autour du candidat.  
'They made a crowd around the candidate.'
19. Votre cousine sera là et la nôtre aussi.  
'Your cousin will be there and ours also.'
20. On peut voir un roc massif près du sentier.  
'You can see a massive rock near the path.'

## REFERENCES

- Ainsworth, W. A. (1972). Duration as a cue in the recognition of synthetic vowels. *Journal of the Acoustical Society of America*, 51, 648-651.
- 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.
- Bastian, J., & Abramson, A. S. (1962). Identification and discrimination of phonemic vowel duration. *Journal of the Acoustical Society of America*, 34, 743-744. (Abstract)
- Bennett, D. C. (1968). Spectral form and duration as cues in the recognition of English and German vowels. *Language and Speech*, 11, 65-85.
- Brichler-Labaeye, C. (1970). *Les voyelles françaises: Mouvements et positions articulatoires à la lumière de la radiocinématographie*. Paris: Klincksieck.
- Chollet, G., & Malécot, A. (1980). An acoustic study of /o/ and /ɔ/ in French in selected contexts. *Travaux de l'Institut de Phonétique d'Aix*, 7, 59-73.
- Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics*, 11, 51-62.
- 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*, 23, 547-552.
- Delattre, P. (1962). Some factors of vowel duration and their cross-linguistic validity. *Journal of the Acoustical Society of America*, 34, 1141-1142.
- Dorman, M. F., Studdert-Kennedy, M., & Raphael, L. J. (1977). Stop-consonant recognition: Release bursts and formant transitions as functionally equivalent, context-dependent cues. *Perception & Psychophysics*, 22, 109-122.
- Durand, M. (1946). *Voyelles longues et voyelles brèves*. Paris: Klincksieck.
- Fry, D. B. (1968). Prosodic phenomena. In B. Malmberg (Ed.), *Manual of phonetics* (pp. 365-410). New York: American Elsevier.
- Gottfried, T. L. (1984). Effects of consonant context on the perception of French vowels. *Journal of Phonetics*, 12, 91-114.

- Jones, D. (1950). *The phoneme: Its nature and use*. Cambridge: W. Heffner and Sons.
- Keating, P. (1985). Linguistic and nonlinguistic effects on the perception of vowel duration. *UCLA Working Papers in Phonetics*, 60, 20-39.
- Lehiste, I. (1970). *Suprasegmentals*. Cambridge, MA: MIT Press.
- Mermelstein, P. (1978). On the relation between vowel and consonant identification when cued by the same acoustic information. *Perception & Psychophysics*, 23, 331-336.
- O'Shaughnessy, D. (1981). A study of French vowel and consonant durations. *Journal of Phonetics*, 9, 385-406.
- Peterson, G. E., & Lehiste, I. (1960). Duration of syllable nuclei in English. *Journal of the Acoustical Society of America*, 30, 693-703.
- Rakerd, B. (1984). Vowels in consonantal context are perceived more linguistically than are isolated vowels: Evidence from an individual differences scaling study. *Perception & Psychophysics*, 35, 123-136.
- Repp, B. H. (1982). Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception. *Psychological Bulletin*, 92, 81-110.
- Repp, B. H. (1983). Trading relations among acoustic cues in speech perception are largely a result of phonetic categorization. *Speech Communication*, 2, 341-362.
- Repp, B. H., Liberman, A. M., Eccardt, T., & Pesetsky, D. (1978). Perceptual integration of acoustic cues for stop, fricative and affricate manner. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 621-637.
- Smith, M. R. (1983). *The integration of temporal and phonetic cues to word stress*. Doctoral dissertation, University of Connecticut, Storrs, CT.
- Stevens, K. N. (1959). Effect of duration on identification. *Journal of the Acoustical Society of America*, 31, 109. (Abstract)
- Straka, G. (1959). Durée et timbre vocalique. *Zeitschrift für Phonetik und allgemeine Sprachwissenschaft*, 12, 276-300.
- 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.
- Summerfield, Q., & Haggard, M. (1977). On the dissociation of spectral and temporal cues to the voicing distinction in initial stop consonants. *Journal of the Acoustical Society of America*, 62, 435-448.
- Valdman, A. (1976). *Introduction to French phonology and morphology*. Rowley, MA: Newbury House.

## FOOTNOTES

\**Language and Speech*, in press.

†Lawrence University, Appleton, WI.

††Yale University. (Currently at the University of Michigan Program in Linguistics, Ann Arbor, MI).

<sup>1</sup>Table 2 shows that the durations for /rod/ were only marginally longer than /rɔd/ in the fixed frame, and were slightly shorter than /rɔd/ in the variable frame. Spectral analysis revealed that the vowels in the /r/-vowel-/d/ context had very similar formant frequencies. The mean values of F<sub>1</sub>, F<sub>2</sub>, and F<sub>3</sub> for /rod/ were 420, 947, and 2876 Hz, respectively; for /rɔd/ these values were 468, 1023, and 2878 Hz. Only one of the five speakers made a clear and consistent spectral distinction between these vowels. The similarity of both duration and formant frequency measures suggests that the /rod/-/rɔd/ distinction was not phonemic for most of our speakers. (When asked about this putative minimal pair, these speakers usually responded that they did not contrast these words, despite the orthography, which represents a difference in vowel duration and timbre in several other minimal pairs.)

<sup>2</sup>The lack of temporal differentiation in these pairs might be due to the prosodic characteristics of the sentences in which they were embedded. Although we attempted to balance these sentences for key word position, differences in stress may have attenuated differences in duration for the minimal pair /sot/- /sɔt/. An analogous explanation of the slight duration differences between /kot/ and /kɔt/ in the variable sentence frame is *not* possible, because the key words are in

---

unstressed position in both sentences (see Appendix). It may be, however, that in unstressed position these words do not ordinarily differ in duration. Jones (1950, pp. 124-125) asserts that intrinsically long and short French vowels differ in duration only in phrase-final position. Descriptions of American English have stated that the duration of long vowels, but not that of short vowels, is greatly affected by stress position (see Durand, 1946, p. 29). Smith (1983) showed that the intrinsic duration difference between /i/ and /ɪ/ is smaller in unstressed sentence positions (mean ratio of long/short is 1.38) than in stressed position (ratio is 1.69). Our data do not include /kɔt/ and /kɒt/ in stressed position in a variable sentence frame, but the vowel of /kɔt/ was nearly 50% longer than that of /kɒt/ when these words were in stressed position in the fixed sentence frame (see Table 2).

<sup>3</sup>There were fairly large context effects on the spectral characteristics of /kɔt/ and /kɒt/. In general,  $F_1$  was higher in the fixed context, and  $F_2$  was lower. However, the spectral differentiation between /o/ and /ɔ/ was large in both conditions. Because there were only two tokens by each speaker in each context,  $F_1$  and  $F_2$  values were averaged over context conditions.