

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI[®]

THE ACOUSTIC CORRELATES OF THE VOICING DISTINCTION

IN FRENCH COGNATE STOPS

by

Nassima Babaci Abdelli-Beruh

A dissertation submitted to the Graduate Faculty in Speech and Hearing Sciences

in partial fulfillment of the requirement for the degree of

Doctor of Philosophy,

The City University of New York

2002

UMI Number: 3047190

Copyright 2002 by
Abdelli-Beruh, Nassima Babaci

All rights reserved.

UMI[®]

UMI Microform 3047190

Copyright 2002 by ProQuest Information and Learning Company.
All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

© 2002

Nassima Babaci Abdelli-Beruh

All Rights reserved

This manuscript has been read and accepted for the Graduate Faculty in Speech and Hearing Sciences in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

April 28, 2002

Date

Lawrence J. Raphael

Lawrence J. Raphael, Ph.D.

Chair of the Examining Committee

4/29/02

Date

Robert Goldfarb

Robert Goldfarb, Ph. D.

Executive Officer

Supervisory Committee:

Winifred Strange, Ph.D.

The City University of New York

Valerie L. Shafer, Ph.D.

The City University of New York

Alice Faber, Ph.D.

Haskins Laboratories

THE CITY UNIVERSITY OF NEW YORK

Abstract

THE ACOUSTIC CORRELATES OF THE VOICING DISTINCTION
IN FRENCH COGNATE STOPS

by

Nassima Babaci Abdelli-Beruh

Advisor: Professor Lawrence, J. Raphael

Both English and French contain stops that contrast phonologically in voicing. However, the phonetic-acoustic realizations of the phonological contrast differ markedly across languages. Furthermore, these phonetic-acoustic realizations vary within each language as a function of contextual variations (e.g. prosody, phonetic contexts, speaking rate). This study examines how some acoustic parameters (aspiration duration, closure duration, phonation during stop closure, preceding vowel duration) vary with specific contextual variables in sentence-length utterances spoken by native French speakers. These contextual variables include: the voicing class of the sounds adjacent to the target stop (vowel vs. unphonated consonant); the position of the stop within a word (syllable-initial, -final); the place of articulation of the target stop (labial, alveolar, velar); and the identity of the vowel within the word in which a stop occurs (ϵ , \circ). Twenty-one CvC syllables were embedded in two different sentence contexts: In the /pa/ context, the

sentence was /il a di pa CvC a lui/. In the /pas/ context, the sentence was /il a di pas CvC sã lui/. Nine speakers were asked to place prominent stress on the CvC. Results show that in the /pa/ and /pas/ contexts, the duration of the vowel /a/ in /pa/ (/pa/:vowel), the Cvc:initial closure duration, the Cvc:aspiration duration, and the presence vs. absence of phonation in the Cvc:initial closure distinguish statistically word-initial /b, d, g/ from word-initial /p, t, k/. In the /pas/ context, speakers use one additional articulatory/acoustic feature differentially: the /s/ of /pas/ (/pas/:s/) is consistently un-phonated in anticipation of /p, t, k/ (produced acoustically as [s]). But /s/ is consistently phonated (produced acoustically as [z]) in anticipation of /b, d, g/ which is indicative of regressive voicing assimilation that occurs 150 ms before a phonologically voiced stop. Additional results show that the voicing distinction is maintained for word-final stops in the /pa/ context, whereas neutralization occurs in the /pas/ context. The present results also clearly demonstrate differential context sensitivity, where some measures are sensitive to the word-stop position while others are sensitive to the addition of a voiceless phone (/pas/ context). No voicing by place of articulation effect was evidenced.

Acknowledgements

None of the other degrees I have earned demanded as much time, emotional stamina and involved as many individuals as this Ph.D. After the initial pride and elation that usually accompanies the admission in a doctoral program, earning a doctorate becomes more than simply acquiring knowledge and expertise. Progressively, it necessitates faith, political astuteness, and above all, endurance; endurance on the part of the candidate whose enthusiasm erodes slowly; endurance on the part of the faculties whose duties gradually require more counseling; endurance on the part of the family members and friends who want to believe that the pursuit of a doctorate is worth the emotional drainage and the financial investment.

The dream of earning a doctorate in America brought me ashore, far away from my family in France. On the road, I have learned English, faced off with a brutal cultural shock, hopped over hurdles, foiled intrigues, conducted investigations, encountered rascals as well as shrewd and sincere administrators/faculties. En route, the indefectible support of my family and friends never dwindled and unceasingly strengthened. Under the auspices of Professor Irving Hochberg and Professor Joseph Glick, I journeyed from Psychology to Speech and Hearing Sciences; under the leadership of my beloved advisor Professor Lawrence J. Raphael, I stirred away from deception and looked forward to a brighter destiny. Under the shield of my loving family, I grew up to become stronger. The excursion into real-life has been traumatic at times and enriching as well.

My gratitude extends to Gary Chant, a technical wizard and a brownie Chef, and to Marc. S. Wein, President of Video Planning, who was my first encounter in New York City and initial sponsor.

Enfin, à mes parents à qui je dois énormément et qui ont su accepter mon éloignement avec sérénité. À mon père qui m'a inculqué l'amour des études; à ma mère qui m'a toujours réconfortée dans la tourmente; à mon frère et mes sœurs qui en dépit des petites querelles intestines ont toujours su m'apporter leur soutien. À mon mari et à mon fils qui en premières lignes ont du subir les exigences de ce doctorat.

Table of Contents

	Page
Abstract	Iv
Acknowledgments	vi
Tables of Contents	viii
Lists of Tables	xi
Lists of Figures	xii
Lists of Appendices	xiii
CHAPTER I: INTRODUCTION	1
PURPOSE	1
CHAPTRE II: LITERATURE REVIEW	4
Production of the stop consonant voicing distinction	4
<i>The Articulatory/Phonetic Realization of the Voicing Distinction in Syllable-Initial Position</i>	4
<i>Acoustic Features</i>	5
<i>Word-initial stops</i>	6
<i>Citation speech studies: Voice Onset Time (VOT) in various languages.</i>	6
<i>VOT studies in English and in French.</i>	7
<i>VOT and place of articulation.</i>	9
<i>Studies of speech in sentence context.</i>	10
<i>Consonant duration in English and in French.</i>	11
<i>Phonation during closure.</i>	12

Table of Contents

	Page
<i>The effect of place of articulation on closure duration.</i>	12
<i>Summary.</i>	13
<i>Word--final stops</i>	13
<i>Closure duration in English and in French.</i>	14
<i>Closure voicing in English and French.</i>	17
<i>Release in English and in French.</i>	17
<i>Preceding vowel duration in English and in French.</i>	19
<i>Other acoustic differences between English and in French.</i>	27
<i>Summary.</i>	28
Perception of the Stop Consonant Voicing Distinction	29
<i>Word-Initial Stops</i>	29
<i>Studies of English and French using Synthetic Speech</i>	30
<i>Word-Final Stops</i>	32
Conclusions	35
CHAPTER III: METHODOLOGY	37
Method	38
<i>Apparatus</i>	38
<i>Subjects</i>	39
<i>Design</i>	39
<i>Material</i>	40
<i>Procedure</i>	40

Table of Contents

	Page
<i>Dependent measures</i>	41
<i>Reliability measures.</i>	46
CHAPTER IV: RESULTS	47
<u>Durational measures for voicing.</u>	49
<u>Closure voicing measures.</u>	58
<u>Durational measures for place of articulation.</u>	62
CHAPTER V: DISCUSSION	64
Appendices	81
REFERENCES	83

Lists of Tables

	Page	
1	<i>Number of tokens per voicing category</i>	48
2	<i>Means and Standard Deviations for word-initial stops</i>	51
3	<i>Means and Standard Deviations for final-initial stops</i>	52
4	<i>Mean voicing-related durational differences and Standard Deviations for word-initial stops</i>	53
5	<i>Mean voicing-related durational differences and Standard Deviations for word-final stops</i>	54
6	<i>Number of times phonation is present or absent through the Cvc:closure</i>	59
7	<i>Number of times phonation is present or absent throughout /s/ of /pas/(/pas:/s/)</i>	59
8	<i>Number of times phonation is present or absent through the cvC:closure</i>	61
9	<i>List of articulatory/acoustic correlates of the voicing distinction that statistically differentiate /b, d, g/ from /p, t, k/ as a function of phonetic contexts</i>	65
10	<i>Summary of voicing-related differences as a function of phonetic context and word position</i>	69
11	<i>Percentage of phonated closures as a function of contexts and voicing categories.</i>	70
12	<i>Consonant closure durations from Laeuffer (1996a, 1992) and current findings</i>	75
13	<i>Vowel durations, vowel duration differences (Dif.) and vowel duration ratio for each study and each context. Durations are in ms.</i>	77

List of Figures

	Page
1 Waveform and corresponding spectrogram of /il a di pa <u>d</u> εt a lui/ (/pa/ context)	44
2 Waveform and corresponding spectrogram of /il a di pas <u>d</u> εt sã lui/ (/pas/ context)	45
3 Duration differences for word-initial stops	56
4 Duration differences for word-final stops	57

List of Appendices

	Page
Appendix A <i>Parameters of narrow-band and wide-band spectrograms, FFT analyses and windows' attributes</i>	81
Appendix B <i>CvC syllables</i>	82

CHAPTER I

INTRODUCTION

In English, the distinction between the phonologically voiced stops, /b, d, g/, and the phonologically voiceless stops, /p, t, k/ (as in the contrast between "back and pack") is conveyed by a set of physiological, acoustic-phonetic and perceptual features. This set of features differs from those conveying the distinction in other languages (e.g. French) which may have a similar number of phonological (cognate) oppositions. The differences between feature sets causes difficulty for native speakers of a language who are trying to produce and perceive the non-native phonemes of a second language. In addition to cross-linguistic differences, the acoustic-phonetic features that convey the oppositions within a language vary as function of the position of the stop within the word (syllable-initial, -medial, -final), the surrounding phonetic context (vocalic vs. consonantal; phonated vs. unphonated), the speaking rate (slow, "normal", fast), the style (overarticulated, more natural speech), the prosody (e.g. heavily stressed vs. weakly stressed syllables), the pattern of preferred syllabic structure ("syllabation ouverte," "syllabation fermée") (Laeufer 1992, 1996 a, b; Umeda, 1977, Delattre, 1953).

PURPOSE

The goal of the present study is to establish how variations in phonetic context affect the various acoustic correlates (i.e. aspiration duration, closure duration, closure voicing, vowel duration) which may serve to differentiate voiced from voiceless stops in French connected speech. We expect that the contribution of each parameter will vary

from one phonetic context to another. We also anticipate that some of the acoustic features that contribute to the phonological distinction will be different in word-initial vs. word-final position, while other features will be stable across contexts.

While the voicing contrasts between English stops have been studied extensively, less research has been conducted with French stops. Much of the published research in French has analyzed citation-forms read from lists read at relatively slow rates (Caramazza & Yeni-Komshian, 1974; Mack 1982) There are also few studies of stop voicing contrasts in continuous speech in French (Laeufer, 1992, 1996a, b). In her studies, Laeuffer used phonetically unbalanced contexts that might have affected the phonetic-realization of the phonological voicing contrast irrespective of the voicing class of the target stops: In all three studies, Laeuffer (1992, 1996, a,b) used /i/ before the word-initial stops and /a/ following the word-final stops. Furthermore, the durations she reported in all three studies were similar to citation speech values, which suggest that her speakers used over-articulated speech. Thus, this raises the possibility that her results are not generalizable to conversational speech in French. The current study was designed to improve on Laeuffer's studies: 1) by using a better controlled phonetic environment with the same vocalic context for word-initial and word final stops. 2) by ensuring that speakers use near normal average speaking volume and rate. This study was also designed to gain information concerning the influence of phonetic contexts other than a vocalic environment on the phonetic-acoustic realizations of the phonological voicing distinction. In this study and for the first time, the influences of vocalic and unphonated consonantal environments on the phonetic-acoustic features of the voicing distinction were compared to assess whether an unphonated phonetic environment would neutralize

the voicing-related differences. Finally, the interaction of place of articulation with voicing class was thoroughly examined.

CHAPTER II

LITERATURE REVIEW

Production of the Stop Consonant Voicing Distinction

The Articulatory/Phonetic Realization of the

Voicing Distinction in Syllable-Initial Position

For the phones [b, d, g] (phonated, unaspirated stops), the vocal folds are adducted before or shortly after the oral closure and they vibrate through the closure, the release and the vowel onset. For the phones [p⁻, t⁻, k⁻] (unphonated, unaspirated stops), the vocal folds are abducted (i.e., the glottis is wide open) at the time of the oral closure and they close around or very shortly after the oral release. For the phones [p^h, t^h, k^h] (unphonated, aspirated stops), the vocal folds are abducted throughout the oral closure and the oral release, and attain phonatory position only at the time of the vowel onset. No phonation occurs during the closure period for [p⁻, t⁻, k⁻] and [p^h, t^h, k^h] because the separation of the vocal folds offers no resistance to airflow and because there is no pressure difference between the supra- and sub-glottal cavities (Borden et al, 1994).

Lisker & Abramson (1964, 1967) have shown that the relative timing of laryngeal events (adduction and onset of vibration of the vocal folds) and supra-glottal events (oral and velopharyngeal closure and release) differentiates stops of different phonological voicing classes in initial pre-stressed position in English and in 10 other languages. The position and the vibratory state of the vocal folds at the time of the oral closure and oral release determine the temporal value of Voice Onset Time (VOT). VOT is classically defined as the time interval between the first sign of the release of oral occlusion and the

onset of phonation. By convention, when periodic pulsing precedes the release of the consonant, VOT is negative (voicing lead). When the onset of periodic pulsing occurs after stop release, VOT is positive (voicing lag).

Acoustic Features

The timing differences between the onset of phonation relative to the release of the oral closure generate acoustic features that distinguish voiced from voiceless stops. In English, the following features differentiate /p, t, k/ from /b, d, g/ in syllable-initial pre-stressed position (e.g. “back” vs. “pack”): degree of aspiration, degree of first formant attenuation and burst intensity (Lisker and Abramson, 1964, 1967; Liberman, Delattre & Cooper, 1958; Lisker, 1957). The duration of aspiration grows as a function of VOT: the longer the VOT (i.e. the more delayed the glottal closure and onset of vocal fold vibration relative to oral closure, the longer the duration of aspiration). No physiological data concerning these features are available for French; Caramazza & Yeni-Komshian (1974) and Flege (1987) among others have provided acoustic data.

For English /p, t, k/ realized as [p^h, t^h, k^h], when the oral closure is released, the high air pressure accumulated in the vocal tract is released suddenly and unimpeded turbulent air rushes from lungs through the open glottis and upper vocal tract. The sudden release of energy translates acoustically into a transient burst of noise (most often between 10 and 20 ms in duration) followed by the turbulence usually termed aspiration. For /b, d, g/, realized as [p[̄], t[̄], k[̄]], the release is weaker since less air pressure is allowed to build up (the glottis is closed shortly before or at the moment of consonant release). In addition, there is little or no aspiration as the closed glottis prevents any massive flow of air from rushing upwards from the trachea. For [p^h, t^h, k^h], the first formant energy is

attenuated until phonation begins. This first formant attenuation is ascribed to damping of energy in the F1 region by the trachea and lungs. The damping of energy is caused by an elongation of the back (pharyngeal) resonating cavity into the trachea, because of the open glottis: this articulatory configuration lowers F1 frequency because it requires a greater amount of energy than is supplied to maintain the F1 resonance at an elevated amplitude.

Word-initial stops

Speech is a complex phenomenon. To simplify its examination, seminal studies have been carried out on using synthetic or natural isolated syllables or words. Normal conversational speech is rapid and informally articulated, and thus marked by coarticulation that is relatively extensive compared to that found in most citation-form speech. Hence, results obtained from studies using citation-form speech and connected speech may lead to different conclusions with regard to the contribution of various acoustic features to [involved in] the phonemic distinction between [b, d, g] and [p, t, k].

Citation speech studies: Voice Onset Time (VOT) in various languages.

Lisker & Abramson (1964) showed that VOT distinguished phonetic categories of stops in initial prevocalic position in all but one of 11 languages whether the languages had 2 or 3 phonological categories at each point of articulation. In the two-category languages, /b, d, g/ and /p, t, k/ could be, respectively, voiced unaspirated and voiceless unaspirated (e.g. Puerto Rican Spanish) or voiceless unaspirated and voiceless aspirated (e.g. American English). In the three-category languages, the phonemic distinction among homorganic stops could be phonetically realized as voiced unaspirated, voiceless unaspirated and voiceless aspirated stops (e.g. Thai) or as voiceless unaspirated, voiceless

slightly aspirated and voiceless aspirated stops (e.g. Korean). Finally, four-category languages, such as Hindi, employed the full combination of voiced unaspirated and aspirated and voiceless unaspirated and aspirated stops. In words spoken in isolation, VOT differentiated the phonetic categories in all but one of the four-category languages. None of the 2-category language used the endpoints of the VOT spectrum ([e.g.] i.e., voicing lead versus long lag) and the modes of the VOT distributions representing each phonetic category were separated by no more than 150 ms. Despite differences in VOT means and ranges from language to language, VOT categories displayed 3 median values: -100 ms (voicing lead), +10 ms (short lag), +75 ms (long lag). In all languages, there were few data points displaying low negative values.

VOT studies in English and in French.

In English initial pre-stressed /b, d, g/, voicing lead is infrequent and short voicing lag is predominant in citation form (Lisker & Abramson, 1964, 1967). Zlatin (1974) also found that the primary mode of production for /b, d, g/ was short voicing lag. She reported VOT values that ranged from +8 ms to +23 ms while Lisker & Abramson (1964, 1967) reported a broader range (0 ms to 21 ms). But Zlatin (1974) noted that 9 of 20 speakers used voicing lead for all /b, d, g/ "with percentages varying between 3.3 and 100" (p. 986). She found that 33% of the single-word productions of /b, d, g/ were characterized by voicing lead with a large variability from -20 ms to -150 ms, centered around a mean lead time of -81 ms. Lisker & Abramson (1964) also reported that the distributions of /b, d, g/ were bimodal for English speakers with modes at about -100 ms [b, d, g] and at 0 ms [p̄, t̄, k̄]. Lisker & Abramson (1967) estimated that physiological measures were necessary in order to explain why /b, d, g/ had two ranges of VOT values.

VOT distributions for [p^h, t^h, k^h] were unimodal with VOT values greater than 40 ms (Lisker and Abramson, 1964). Finally, the range of VOT values [distributions of] for [p⁻, t⁻, k⁻] was smaller than those for [p^h, t^h, k^h].

In French, the same phonemic oppositions have different phonetic realizations. Caramazza & Yeni-Komshian (1974) requested monolingual speakers of Metropolitan French (UF), of Canadian French (UCF) and of Canadian English (UCE) to produce stop + [a] syllables. They used nine words in each language for each of the six stop consonants. No minimal pairs were used. A total of 54 French and 54 English words were used. Caramazza & Yeni-Komshian (1974) found that monolingual French (UF) speakers from the Western part of France "produced voiced consonants with VOT values almost exclusively in the voicing lead (negative) region" (Caramazza & Yeni-Komshian, 1974, p. 242.) (mean and mode values not reported). Initial /b, d, g/ uttered by UF speakers were produced with voicing lead 94% of the time and with voicing lag 5% of the time. The frequencies of voicing lead were: /b/= 26, /d/= 29, /g/= 29; the frequencies of voicing lag were: /b/= 3, /d/= 1, /g/= 1. Syllable-initial /p, t, k/ uttered by UF speakers were produced with short voicing lag (mean and mode values unknown) that were significantly smaller than the values produced by either monolingual Canadian French (UCF) or monolingual Canadian English (UCE) speakers. Results indicated that both Metropolitan French (UF) and Canadian English speakers (UCE) produced VOT distributions that were non-overlapping. However, monolingual French Canadian speakers (UCF) produced VOT values that were intermediate to those of the monolingual French (UF) and English (UCE) speakers for all stops and the VOT distributions of the voiced and voiceless stops overlapped.

In summary, in English, phonological voicing does not mean that phonation constitutes the necessary distinctive feature that speakers employ to make a phonemic distinction with phonologically voiceless cognates. It only means that /b, d, g/ and /p, t, k/ belong to "opposing" linguistic/phonemic categories. However in French, phonation is an articulatory feature that distinguishes the two phonological categories. French and English differ in their linguistic usage of the 3 phonetic categories ([b, d, g], [p̄, t̄, k̄] and [p^h, t^h, k^h]). Both employ the short voicing lag (i.e., voiceless, unaspirated) category, but English speakers use it for /b, d, g/ and French speakers use it for /p, t, k/.

VOT and place of articulation.

Lisker & Abramson (1964) showed that in English and in citation speech the members of each contrasting pair (/p-b, t-d, k-g/) have distinctive and non-overlapping VOT values. The mean VOT values of stops with VOT in the lag range varied systematically with the place of occlusion/articulation: VOT values increased as the place of articulation moved back into the oral cavity (/b/=1 ms, /d/=5 ms, /g/=21 ms and /p/=58 ms, /t/=70 ms, /k/= 80 ms) (Lisker & Abramson, 1964). Zlatin (1974) found a similar relationship between place of articulation and VOT, but reported that the significant differences were primarily attributable to comparisons between labials and velars while the labial/alveolar difference was not significant. Zlatin (1974) added that place of articulation also affected production of stops with voicing lead: mean lead values grew gradually shorter as the place of articulation moved back in the oral cavity from labials to velars. "These means were most likely offset not by the alterations in the use of lead time..., but by the use of longer mean lag times for voiced stops" (Zlatin, 1974, p. 986). This influence of place of articulation on VOT is maintained across a wide range of

speaking rates (Lisker & Abramson, 1967; Miller & Volaitis, 1989; Volaitis & Miller, 1992). Caramazza & Yeni-Komshian (1974) did not examine the effect of place of articulation on VOT among Canadian or French speakers.

Studies of speech in sentence context.

Lisker & Abramson (1967) asked the same speakers who had produced the isolated words in their 1964 study to say the same words in sentences at different speaking rates. It was difficult to measure VOT unless the word-initial stop was also in sentence-initial position as phonation often continued unbroken during the closure intervals of sentence-medial stops. For a stop in medial phonetic context, VOT does not reliably distinguish the members of each cognate pair (i.e. VOT distributions overlapped). The overlap of the distributions always reflected a decrease in the $[p^h, t^h, k^h]$ VOT distributions. The overlap could not be attributed to greater variability since VOT values in sentences were less variable and shorter. Lisker & Abramson (1967) suggested that this reduction in VOT could be taken as evidence for alternative acoustic features whose contribution could be redundant in careful speech (citation speech) but indispensable at more normal speaking style.

In French, there is no study that reports the descriptive values for each voicing category and hence it is impossible to establish a direct comparison with the English data. Furthermore, studies differ with respect to the position of the syllable within the sentence. Laeufer (1996a) used phrase-medial and phrase-final position and Durand (1985) used sentence-final position. In addition, results indicate that durational features other than VOT (e.g. consonant closure duration, preceding vowel duration) are sensitive to the

position of the word in the sentence (Laeufer, 1992, 1996a,b; Luce & Charles-Luce, 1985; Stathopoulos & Weismer, 1984; Suen & Beddoes, 1974).

Laeufer (1996a) used monosyllabic words embedded in short sentences and found that 23% of initial French /b, d, g/ were on average produced with short-voicing lag (about 12 ms) and 77% were produced with voicing lead (average unreported). She found that initial French /p, t, k/ were produced with short voicing lag (mean = 24 ms (/p/=14 ms, /t/=18 ms, /k/=30 ms) and with small speaker variability (value unreported).

Consonant closure duration in English and in French.

The closure interval or “silent” interval observed on a spectrogram represents the duration of the oral/velopharyngeal closures required for the production of oral stops. Closure duration varies as a function of a number of factors, one of which is the voicing classification of the consonant. Luce & Charles-Luce (1985) found that English /p, t, k/ were produced with longer closure durations than English /b, d, g/ by about 18 ms in sentences (mean /b, d, g/ = 68 ms, mean /p, t, k/ = 86 ms). In word-initial position followed by a stressed vowel, however, many studies reveal that English /b, d, g/ are slightly longer than English /p, t, k/ (Laeufer, 1996a; Lisker, 1977; Stathopolous & Weismer, 1983; Umeda, 1977). Laeufer (1996a) reported a small difference of 8 ms (mean English /p, t, k/= 97 ms; mean English /b, d, g/ = mean 105 ms).

In French, Laeufer (1996a), reported that word-initial /p, t, k/ closures were longer (139 ms) than those of word-initial /b, d, g/ (118 ms) by about 21 ms. O'Shaughnessy (1981) reported a similar difference using a French Canadian speaker. Laeufer (1996a) added that her findings were consistent with those of previous studies (Wajkop, 1978; Durand, 1936; Simon, 1967).

Overall, a slight difference exists between the closure durations of word-initial /p, t, k/ and word-initial /b, d, g/ in both languages. Also, the direction of the difference is less consistent in English than in French, which raises the question of whether closure duration can serve to differentiate the two voicing categories at the production level and whether it carries perceptual significance. In their production study, Luce & Charles-Luce (1985) have contended that, in English, closure duration did not reliably distinguish the two categories in half of the sentences they analyzed.

Phonation during closure.

French and English speakers differ dramatically in the percentage of stops that are phonetically voiced (i.e., phonated): Laeufer (1992) found that 77% of the total number of /b, d, g/ closures were phonated by French speakers while only 18% of the /b, d, g/ closures were phonated by English speakers. Laeufer (1996a) also found that French speakers fully phonated /b, d, g/ closures (i.e. that phonation ran through the entire closure interval) in 34% of the cases while English speakers did so in only 3% of the instances.

The effect of place of articulation on closure duration.

In agreement with Umeda (1977), Stathopoulos & Weismer (1983) and Suen & Beddoes (1974), Luce & Charles-Luce (1985) found that the mean values for alveolar stop closure durations were the shortest and that those for labials the longest. Velar stop closures had intermediate values. Suen & Beddoes (1974) have posited that the differences in duration can be attributed to differences in muscular effort required by the articulators to reach their final position: "since lips are less mobile than the back of the tongue, which is less mobile than the tip of the tongue (Hudgins and Stetson, in Lehiste

1970), it may be expected that shorter SI (i.e. silent interval) [closure duration] is achieved by the most mobile articulator to reach its final articulatory position (Suen & Beddoes, 1974, p. 131). Suen & Beddoes (1974) found this place of articulation effect in all stop positions in a word and for all combinations of voicing and stress. Furthermore, the difference between labials and stops that are produced in other places of articulation was largest in word-final position. The difference in closure duration between cognates, however, was about the same across place of articulation: Stathopoulos & Weismer (1983) reported that /p, t, k/ closures were longer than /b, d, g/ ones and unaffected by place of articulation.

Summary:

Many phonetic features reliably distinguish word-initial /b, d, g/ from /word-initial /p, t, k/ in very specific contexts: VOT reliably differentiates the two phonological categories in both English and French, but only in the context of words read in isolation. Closure duration differences related to the voicing classification of the stops are small in both languages, suggesting that closure duration may not be a salient perceptual cue. Phonation during closure is predominant in French /b, d, g/ (Laeufer, 1992) while it is less common in English /b, d, g/ (Zlatin, 1974, Lisker & Abramson, 1967). Finally, it is likely that context-dependent spectral measures (e.g. F1 onset frequency) participate to mark the voicing distinction. Overall, for French there is much less information available than for English. The proposed study is designed to provide this information.

Word--final stops

In English, various acoustic features mark the voicing distinction between stops following a stressed syllable: preceding vowel duration, formant transitions into the

articulatory closure, phonation during the closure interval (closure voicing), duration of the closure interval, F1 offset frequency, and initial (post-release) formant transitions. As in initial stops, the articulatory-acoustic realizations of stops in word-final position vary with context.

Closure duration in English and in French.

There is a consensus that English word-final /p, t, k/ closures are longer than those of word-final /b, d, g/ (Stathopoulos & Weismer, 1983). Suen & Beddoes (1974) reported that in citation speech the difference was of about 33 ms (30%) (mean /p, t, k/ = 111 ms and mean /b, d, g/ = 78 ms), while Revoile, Pickett, & Holden (1982) reported that /p, t, k/ closures were about 41% longer than closures of /b, d, g/ in sentences. The difference was larger in word-final position than in word-initial (Suen & Beddoes, 1974, Stathopoulos & Weismer, 1983, Luce & Charles-Luce, 1985).

Luce & Charles-Luce (1985) systematically manipulated the local phonetic context (vowel or stop) following the target stop, the sentence position (phrase-final or non-final) and the voicing category of the stop. The local phonetic environment, defined as the phoneme following the test word, was either a stop or a vowel. Four possible carrier phrases were used in experiment 1: a) the test word occurred in phrase-final position and the test word was followed by a unstressed mid-central vowel (e.g. "When Mark read cvC, Elaine made a checkmark"); b) the test word occurred in phrase-final position and was followed by the voiceless alveolar stop /t/ (e.g. "If Ted says cvC, Tom will leave the room"); c) one word occurred between the test word and the phrase boundary and the test word was followed by unstressed mid-central vowel (e.g. "When Mark read cvC aloud, Elaine made a checkmark"); d) one word occurred between the test

word and the phrase boundary and the test word was followed by voiceless alveolar stop /t/ (e.g. "If Ted says cvC today, Tom will leave the room"). Luce & Charles-Luce (1985) estimated that the closure duration difference between /b, d, g/ and /p, t, k/ was not statistically significant in about 53% of the utterances analyzed. The differences failed to reach significance in all cases where the target stop was followed by a voiceless sound (/t/) in non phrase-final position or in phrase-final position and for most test words ending in /d/ or /t/. These results suggest that closure duration carries little perceptual weight.

Laeufer (1992) manipulated the position of the CvC within the sentence as well the location of sentence stress. There were three conditions: in the condition "phrase-medial position under focus/stress," English speakers said "Yell cvC at Marie" while French speakers said "Dis cvC à Marie." In the condition "phrase-final position under focus," speakers read "Yell cvC" or "Dis cvC." Finally, in the condition "phrase-medial in the non-stress/non-focused condition," the sentence frame was "Yell cvC at MARY" or "Dis cvC à MARIE." The sentence focus was placed on the word written in upper-case letters. Laeufer (1992) differentiated voiceless releases from vocalic releases. She defined "voiceless releases" as "these strong release bursts [which] were measured from the instant of release noise to cessation of the high amplitude noise" (p. 417). Vocalic releases called "détente vocalique" (Delattre, 1951) are "heard as a schwa and visible in the waveform as a low-amplitude vowel of variable duration" (p. 417).

Laeufer (1992) reported that in phrase-final position under focus, French /p, t, k/ closure duration was longer (mean = 165 ms) than French /b, d, g/ closure duration (mean = 100 ms) by about 65 ms; in English the difference was only 43 ms (mean /p, t, k/ = 133 ms, mean /b, d, g/ = 90 ms.). In sentence-medial context under focus (i.e with sentence

stress on the target word), the difference between the mean closure duration of French /b, d, g/ and /p, t, k/ was reduced to 32 ms (/p, t, k/ = 101 ms, /b, d, g/ = 71 ms.). There was a similar effect in English: the mean closure duration for /p, t, k/ was longer than that for /b, d, g/, but the difference was smaller when the word was in sentence-medial than in sentence-final position.

Considering only non-vocally released stops, Laeuffer (1992) reported that the closure durations of French stops were significantly longer than those of English stops in both phrase-final and phrase-medial position under focus. In phrase-final position under focus, French /b, d, g/ closures were longer than English /b, d, g/ closures by 10 ms (mean French = 100 ms and mean English = 90 ms) and French /p, t, k/ closures were longer than English /p, t, k/ closures by 42 ms (mean French = 165 ms, mean English = 133 ms). Interestingly, in sentence-medial context under focus, the difference between French and English /b, d, g/ grew to 16 ms (mean French = 71 ms, mean English = 55 ms) and the difference between French and English /p, t, k/ closures decreased to 30 ms (mean French = 101 ms, mean English = 71 ms). There were significant main effects of language and voicing in both stressed conditions (sentence-medial and sentence-final position). Differences in closure durations between French and English could not be attributed to differences in tempo or emphasis: Laeuffer assessed possible differences in tempo and emphasis. She calculated the mean duration of two of the sentence types for each speaker and for each group and found that the durations of English sentences read by the bilingual French speakers and by the monolingual English speakers were of similar duration. The degree of emphasis was assessed by asking trained listeners to

evaluate emphasis on a “scale from 1 (focus but no particular emphasis) to 4 (emphatic stress)” (Laeufer, 1996a, p. 91).

Closure voicing in English and French.

The presence of glottal pulsing in the closure interval (closure voicing) may be prompted by specific instructions, at least in citation speech. Revoile et al (1982) and Wolf (1978) both requested the English speakers to release all final consonants and both noticed on the spectrograms a low frequency voicing bar in the occlusions of all /b, d, g/ stops. It was of low-intensity and decreased in amplitude throughout the occlusion interval. This low-frequency murmur occupied, on average, 87% of the closure (Revoile et al, 1982). There was no murmur in the occlusions of the /p, t, k/ stops (Wolf, 1978; Revoile et al, 1982). Hillenbrand, Ingrisano, Smith & Flege (1984) and Raphael (1981), none of whom gave particular instructions to the speakers, found closure voicing during the production of /b, d, g/ or /p, t, k/ final stops. Laeufer (1992) reported that in the unfocussed context (i.e. sentence stress is not on the target word), English /b, d, g/ had glottal pulsing throughout most of their closures. But she did not present any descriptive statistics for either French or English. The current research will quantify the number of occurrences of stops with closure voicing for each voicing category in French.

Release in English and in French.

In studies where English speakers were instructed to release final stops, bursts were most often higher in frequency, greater in amplitude and of greater duration for voiceless than for voiced stops (Revoile et al, 1982; Wolf, 1978). These differences were small and did not always occur. Rositzke (1943) reported that 63% of English final stops

were not released. Mack (1981) has reported that prepausal voiced stops are commonly followed by vocalic release in French but not in English.

Laeufer (1992) reported that French word-final /p, t, k/ stops were strongly released with both longer release duration and greater release amplitude compared to English. In both languages, /b, d, g/ releases were less intense than those of /p, t, k/. In English and in all three contexts (i.e. sentence-medial unfocussed and focussed, and sentence-final focussed) English stops were followed at times by a "pseudo" release "indicative of a more passive separation of the articulators than that of aspirated releases... or of less oral pressure during closure due to glottalization" (Laeufer, 1992, p.417). Laeufer (1992) reported that pseudo-releases were indicated in the sound-wave by a small spike, marking the instant of the burst release, and by the absence of the following interval of high intensity noise typical of fully released stops.

The average duration of French stop releases (51 ms) was found to be significantly greater than those for English stops (25 ms) only in sentence-final context (sentence-final context = 52 ms difference for /p, t, k/ and 25 ms difference for /b, d, g/; sentence-medial context = 15 ms difference for /p, t, k/ and 9 ms difference for /b, d, g/). Overall for both languages, release durations were significantly longer in sentence-final context (57 ms) than in sentence-medial focus context (20 ms), and /p, t, k/ releases were significantly longer than /b, d, g/ in sentence-final context (French = 62 ms-difference, English = 35 ms-difference) but not in sentence-medial focus context (in French = 13 ms-difference; in English = 7 ms-difference). In medial-focus context, absolute durations and duration differences decreased in both languages.

Overall, /p, t, k/ releases were significantly longer than /b, d, g/ releases in both languages and French releases were about twice as long as the English ones. The longer French releases are presumably related to a difference in stop closure duration. There is a greater supra-glottal pressure build-up during the longer French stop closures than during the shorter English ones. In voiceless stops, the transglottal airflow might be further increased in French by a more active opening of the glottis. In contrast to English stops involve some glottal narrowing or even glottal closure in some American Dialects (Benguerel, Hirose, Sawashima & Ushijima, 1978). Altogether, these factors predict greater and longer aspiration following the release of French than of English final stops.

Preceding vowel duration in English and in French.

There is general agreement that the durations of vowels preceding English final /b, d, g/ are about 33% longer than the duration of the vowel before English /p, t, k/ in citation speech (House & Fairbanks, 1953; Peterson & Lehiste, 1960; Chen, 1970; Raphael, 1972; Revoile et al, 1982; Wolf, 1978). This voicing-related vowel duration effect has been attributed to a longer sustention of muscular activity for the vowels preceding voiced stops, at least in CVC and VCV contexts in citation speech (Raphael, 1975). The durational difference in the muscular activity underlying the vowel articulation is close to the difference in vowel durations measured acoustically.

In sentences, however, researchers have demonstrated the importance of preceding vowel duration to be context-dependent: vowel duration differences are smaller in non phrase-final and non prepausal positions than in phrase-final and prepausal positions (Klatt, 1976, Umeda, 1975). Klatt (1976) pointed that "a large difference in vowel duration is only seen in phrase-final environments, so it is only in these cases that

the durational cue has primary importance.” (p. 1219). But Luce & Charles-Luce (1985) found that vowel duration, unlike closure duration, consistently distinguished English voicing categories at or beyond significance level of 0.003 in all sentence contexts. In agreement with Luce & Charles-Luce (1985), Van Summers (1987) reported significantly longer vowel durations before /b, d, g/ than before /p, t, k/ in non phrase-final and non prepausal positions. Compared to duration differences obtained in citation speech, Van Summers (1987) found that durations of vowels preceding English final /b, d, g/ are only 21% longer than the durations of the vowel before English /p, t, k/, while Luce & Charles-Luce (1985) found a 31% difference. Laeuffer (1992) reported that vowels preceding /b, d, g/ were significantly longer than vowels before /p, t, k/ in sentence context. Laeuffer (1992) also found that the vowel durations differed by 36% in phrase-final position (ratio = 0.64, a 111 ms difference), by 28% in phrase-medial context under focus (ratio = .72, a 60 ms difference) and by 18% in phrase-medial non-focused context (English ratio = .82, a 29 ms difference). In all three contexts in English, she found that these vowel duration differences associated with the consonant voicing were statistically significant at the 0.01 level.

In a cross-linguistic study, Chen (1970) reported that the durations of the vowels preceding /b, d, g/ were longer than the durations of the vowels preceding /p, t, k/ by an average 92 ms in English. The ratio of vowel durations (voiceless/voiced contexts) was 0.61. Luce & Charles-Luce (1985) reported a smaller difference (55 ms), with a ratio of 0.69. The difference between studies may stem from the fact that Luce & Charles-Luce (1985) used a sentence context whereas Chen's (1970) data were taken from isolated syllables. Chen (1970) also reported that the voicing-conditioned vowel duration effect

was considerably smaller in French, Russian and Korean. He reported that in French, the voiceless/voiced context ratio was closer to 1 (.87) than in English (0.61): the voicing-related vowel duration difference was reduced from 33% (a mean difference of 92 ms) in English to 13% in French (a mean difference of 53 ms). In Russian and in Korean the mean differences were, respectively, 29 and 28 ms with ratios of .82 and .78. In view of these observations, Chen (1970) concluded that voicing-conditioned vowel duration is "a language-universal phenomenon" and that the extent to which it is exploited varies with the "language-specific phonological structure"(p.139.) This statement that voicing-conditioned vowel duration is "a language-universal phenomenon" has been disputed by Flege & Port (1981) who reported that in a dialect of Saudi Arabia, such differences in preceding vowel durations were non-existent when speakers were asked to produce utterances embedded in short carrier sentences. The fact that in some languages vowel duration does not vary systematically according to the voicing category of the following consonants suggests that voicing-conditioned vowel-duration variation is not a phonetic universal. At the same time, the fact that no language shows longer vowel durations before all instances of /p, t, k/ and shorter vowel durations before all instances of /b, d, g/ indicates that languages may exploit a physiologically based mechanism that conditions vowels before /b, d, g/ to be longer than before /p, t, k/. Hence, a phonetic universal may exist even though it is not fully exploited in all languages. A language that does not fully exploit the "universal phonetic pattern "must create a phonetic rule (Keating, 1985).

Mack (1982) argued that in Chen's study (1970), the syllabic structure of the words was not held constant across languages. To control for syllable structure she used homophonous sets of minimal pairs in English and French so that the durational

percentages could be compared. The French speakers were residents of Southern France and the English speakers were from Southern New England. Like Umeda (1977) and Chen (1970), Mack (1982) did not include the aspirated portion of the vowel following the preceding syllable-initial voiceless stops in the measurement of the vowel. There were large and consistent differences in the use of voicing-dependent vowel duration by English and French monolingual speakers. The average durational ratio was lower in English (.54) than in French (.73). Interestingly, the durations of the vowels produced by the English monolinguals in the voiceless context were only moderately longer than the vowels produced in the same context by French monolinguals (154.1 vs. 133.1 ms). The vowels produced by the monolingual English speakers in the voiced context were considerably longer than the vowels produced by the French monolinguals in the same context (285.3 vs. 182.1 ms). The average durational range was 82.2 ms larger for the English speakers, but the standard deviations were smaller than for the French speakers.

Overall, stop consonant voicing affects the preceding vowel duration in both English and French, but the effect is smaller in French than in English. "Based on these studies (Chen, 1970, Zimmerman & Sapon, 1958) and other single-language studies (e.g. House, 1961), it is commonly assumed that obstruent voicing is a universally important factor influencing the duration of preceding vowels and that English, but not the other languages examined, exploits this natural tendency by using it as the basis for a low-level phonological rule of vowel lengthening before voiced obstruents." (Laeufer, 1992, 412). The fact that the English monolingual speakers have lower standard deviations reflects the greater stability of the durational ratio in English.

Laeufer (1992) contended that most cross-linguistic studies did not take into account confounding variables such as the length of the test words, the immediate segmental environment or the presence in some of the languages of a stop devoicing rule which tends to neutralize the voicing distinction word-finally (Delattre, 1962; Mack 1982). She added that comparisons across languages have usually been based on separate studies of each language, ignoring factors such as the positions of sounds in words and speaking rates that can have large effects on vowel and segment durations. Laeufer noted that Mack's (1982) comparison of words read in isolation took these factors into account. She added, however, that Mack's conclusions of a learned vowel lengthening rule in English, which she based on the difference between the English ratio (.53) and the French (.74), was premature because the effect of stop-consonant voicing on preceding vowel duration is very sensitive to contextual variables in both languages. In English, the preceding vowel duration difference is greatest in word- or phrase-final context, while it is significantly smaller in medial position (Laeufer, 1992; Luce & Charles-Luce, 1985; Umeda, 1975; Klatt, 1975, 1976). Luce & Charles-Luce (1985) reported that vowel durations were about 69 ms longer in sentence-final than in non-sentence final-position (mean vowel duration for phrase-final = 184 ms; mean vowel duration for non-phrase-final 115 ms). It is greater before English bilabials than before English alveolars and velars in stops only (Luce & Charles-Luce, 1985) and it is greater in stressed than in unstressed syllables (Luce & Charles-Luce, 1985; Van Summers, 1987). In contrast, it is smaller in polysyllabic words than in monosyllables (Klatt, 1973; Harris & Umeda, 1974; Port, 1981) and in spontaneous connected speech than in real word or phrase lists to the point of being almost non-existent in non phrase-final and prepausal position. The

voicing effect on vowel duration also decreases as speaking rate increases (Harris & Umeda, 1974). The effect is more pronounced in the context of fricatives than stops (House & Fairbanks, 1953; Peterson & Lehiste, 1960), while flapping of postvocalic alveolars in American English has been found to lessen the effect. In French, similar contextual variability was reported. The effect on preceding vowel duration is greater in word-, sentence- and in phrase-final position than in medial positions. The voicing effect is smaller in connected speech than in word or phrase list reading (O'Shaughnessy, 1981). There is also a reduction in differential vowel duration from French mono- to polysyllables (O'Shaughnessy, 1981).

Laeufer also contended that although Mack's (1982) studies remedy some shortcomings of Chen's (1970) study, it failed to consider some "other language specific conditioning factors" such as intrinsic vowel duration: most English stressed vowels are of greater duration than their closest French equivalents (Delattre, 1951; 1965). The greater vowel duration of English vowels has been shown to correlate with greater vowel duration differences, and thus with lower ratios of the durations in the voiced/voiceless contexts. The effect of voicing has been found to be smaller for inherently short vowels such as /ɪ/ than for inherently long ones such as /a/ (Peterson & Lehiste, 1960; Crystal & House, 1982). But Luce & Charles-Luce (1985) found that preceding vowel duration differences were greatest in words containing /i/, least in words containing /I/ and intermediate in words containing /a/. Because the number of (near)-homophonous pairs in French and English with short/lax English vowels is small, the cross-linguistic comparisons are limited.

Laeufer (1992) extended the Mack (1982) study to contexts in which sentence focus was not on the target word. In the condition "phrase-medial position under focus/stress," English speakers said "Yell X at Marie" and French speakers said "Dis X à Marie." In the condition "phrase-final position under focus," speakers read "Yell X" or "Dis X." Finally, in the condition "phrase-medial in the non-stress/non-focused condition," the sentence frame was "Yell x at MARY" or "Dis x à MARIE." Results indicated that vowels preceding /b, d, g/ were longer than vowels before /p, t, k/ in both languages. Also, there was a larger dissimilarity between the durations of vowels in both French and in English in phrase-final position (French ratio = .74, a 53 ms difference, English ratio = 0.64, a 111 ms difference) than in phrase-medial context [both in stressed syllables (French ratio = .78, a 38 ms difference, English ratio = .72, a 60 ms difference) and in unstressed syllables (French ratio = .91, a 11 ms difference, English ratio = .82, a 29 ms difference)]. The ratio of .74, for French stops in phrase-final position, however, combines a higher ratio of .77 for stops with vocalic release and a lower ratio for stops with voiceless releases. Laeufer (1992) reported that vowel durations were shorter and the voicing related differences were smaller before consonants (stops and fricatives) with vocalic (i.e. phonated) releases than before consonants with unvoiced releases. A weak but significant negative correlation existed between durations of the preceding vowel and the durations of the schwa releases.

Overall, ratios of the durations of vowels preceding voiceless vs. voiced English stops were always smaller than ratios of the durations of vowels preceding French stops. In addition, the voicing-related vowel duration differences were statistically significant within each language and between languages in phrase-final position only. The vowel

duration differences were influenced by the word-position in the sentence and by the stress/focus condition. Shifting the location of a word from phrase-final position to phrase-medial position led to a decrease in vowel duration. Additional reduction in the vowel duration differences took place when sentence stress/focus was shifted away from the target syllables. A shift in the position of the word within the sentence had a greater effect on English than on French vowel duration differences. A shift in stress/focus also had a greater effect on English than on French vowel duration.. There is no explanation for the former effect, except perhaps that English uses phrase-final lengthening to exaggerate this effect The latter effect is expected since English is usually classified as a stress-based language and French is not.

In phrase-final position, the ratio of .77 for vowels preceding French voiceless/voiced stops is close to the value found by Mack (1982) in citation speech (.74). Laeufer (1992) suggested that it "is presumably also due to the presence of final schwa releases, as is expected of speakers from the Southern France, where historic word-final schwas (still preserved in spelling) have not been lost and quasi-bisyllabic renditions are the norms" (p. 427). This may, however, be because the speakers over-articulated.

Overall, similarities between the two languages can be seen: in that vowel duration lengthening is greatest in syllables in sentence-final position in both languages and is less pronounced for syllables in sentence-medial position. English vowel duration differences are, however, apparently more susceptible to contextual variations. In English, as in French, a shift in sentence stress/focus had less effect than a shift in word position within the sentence: changes in stress/focus diminished vowel duration differences by 31 ms (in sentence-medial context) while a change in the position of the

syllable in the sentence yielded a 66 ms decrease. "In the unfocussed context, where English vowels are prosodically less prominent and phonologically voiced consonants have glottal pulsing throughout most of the closure, English vowel duration differences are much less marked and thus much more French-like" (Laeufer, 1992, p. 433). Laeufer (1992) concluded that "similar prosodic factors determine similar variations in the size of the voicing effects" (p.431). To account for the greater differences in vowel duration in English than in French (Chen, 1970, Laeufer, 1992, 1996b), Laeufer (1992) noted that "a low-level allophonic rule of vowel-lengthening before a voiced obstruent has been claimed to be operational in English but not in French or in any other language studied"(p. 430).

Other voicing-related acoustic differences between English and French.

Several researchers who recorded isolated English monosyllables containing the vowel /ae/ found that the vowel formant offsets preceding voiceless stops were lower in intensity and were also higher in frequency than the vowel offsets before voiced stops (Revoile et al, 1982; Wolf, 1978; Van Summers, 1987). The vowel transition differences appeared primarily in the final five pitch periods preceding the consonant occlusion. Wolf (1978) also found that for /aeC/ syllables there were differences in the initial portion of the vowel: For syllables ending in /b, d, g/, the mean F1 (extracted with LPC and averaged over the first 50 ms) was lower than for syllables ending in /p, t, k/ and the mean F2 over the first 50 ms was higher in syllables ending in /b, d, g/ than in syllables ending in /p, t, k/. Similar differences in initial formant onset frequency were also observed when the target syllables were produced in sentences and the same vocalic context (/ae/) was used (Van Summers, 1987).

Van Summers (1987) examined changes in the formant structure (F1 initial transition slope, F1 initial transition duration, steady-state durations of F1, final transition duration, F1 offset frequency) of the vowel (/æ/) in sentences. He found that the voicing classification of the final consonant did not have any effect on the F1 initial transition slope or on F1 initial transition duration. Syllables ending in voiced stops, however, had higher steady state formant frequencies, longer F1 steady state duration, and longer final transition durations than syllables ending in voiceless stops.

Summary.

Vowel duration differences between syllables ending in /b, d, g/ and syllables ending in /p, t, k/ are greater in English than in French. Release and closure duration differences are larger in French than in English. Laeufer (1992, 1996b) suggested that the significantly longer closure durations and smaller voicing-related vowel duration differences (lower ratio of non vocalically-released final voiced over voiceless consonants) in French compared to English could be affecting the pattern of voicing-related vowel duration. In English, shorter consonants with a smaller durational range followed longer vowels with a greater durational range while in French the opposite is true (i.e. closure durations are longer). She hypothesized that these patterns could be related to the prosodic systems of the two languages (i.e. English is a stress-based language while French is not) and to the different status assigned to the voicing feature in each language. "In French, the voicing contrast is carried more by the characteristics of the consonant and, in particular, [by] the presence/absence of pulsing during the closure" (Delattre 1951, p. 417). Hence, the consonant needs to be long enough to convey the voicing status perceptually. The relative shortness of the voiced consonant relative to the

voiceless is presumably linked to the difficulty of maintaining strong glottal pulsing for too long (Lisker, 1977; Westbury, 1983). In English on the other hand, final stops are often partially devoiced (Flege and Brown, 1982) and the contrast is carried more by the vowel, in particular, by the relative durations of vowels preceding phonologically voiced/voiceless consonants (Denes, 1955). Laeufer (1996a) maintains that French grants more emphasis to the postvocalic consonants and English to the vowels. This explains the comparatively greater magnitude of the voicing effect on vowel duration in English (Laeufer, 1992).

Perception of the Stop Consonant Voicing Distinction

Word-Initial Stops

Some explanations of the strategies used in the perception of the stop consonant voicing distinction are based on theories which incorporate speech production as part of the perceptual mechanism. Other explanations are based on the theoretical stance which does not include reference to speech production as part of the perceptual system for speech. From either theoretical point of view, it is of interest to understand the nature of the perceptual processes involved and, ultimately, to explore the possibility that those processes are related to the processes of speech production.

The term VOT in the acoustics of production refers to the time interval between the release of the stop (burst) and the onset of quasi-periodic vocal fold pulsing. In studies using synthetically produced speech, the same term refers to the interval between the onset of stimuli (which may or may not include a burst) and the onset of a periodic source. VOT is sometimes called the separation cue (Summerfield & Haggard, 1977). To

create /p, t, k/ stimuli, the upper formants (F2 and F3) are excited by noise and the F1 transitions are attenuated until the higher formants are periodically excited.

Synthetic VOT continua in the lag range are created by delaying the onset of the periodic source while simultaneously attenuating the F1 transition. This manipulation mirrors natural speech processes and results in a similar co-variation of other acoustic parameters. The delay in the onset of periodicity and the attenuation (often realized as the absence) of F1 are synthetic analogues of the effect of the open glottis at the moment of stop release; this glottal configuration inhibits phonation while extending the pharyngeal cavity, weakening the intensity of F1. The onset of the unattenuated portion of the F1 transition is thus both raised in frequency and delayed in time, relative to the F2 and F3 onsets. The duration and extent of frequency change of the F1 transition are thus reduced. Synthesizing the F2 and F3 transitions with an aperiodic (noise) source until the onset of the periodic source (and of unattenuated F1 transition) simulates the aspiration found in natural speech, which is another result of the glottis being open at the moment of stop release.

Synthetic VOT continua in the voicing lead range are created by inserting varying durations of a periodic source, simulating the first harmonic (fundamental frequency), during the closure period preceding stop release. A number of studies using synthetic speech stimuli have attempted to clarify the contribution of the various potential acoustic cues to the perception of voicing of initial pre-stressed stops. At times this has been done at the expense of the naturalness of the tokens (Summerfield & Haggard, 1977).

Studies of English and French using Synthetic Speech

An increase in F1 onset frequency or in F1 extent or a delay in time at which F1 begins relative to upper formants each produced a significant increase in voiceless responses (Liberman et al, 1958; Summerfield & Haggard, 1977). Delaying F1 onset, while keeping F1 onset frequency constant, was found to be a strong cue to voicelessness: Liberman et al (1958) effected this experimental separation of the two cues by using the vowel /o/ in which there is no transition and hence no change in the frequency of F1 onset. Results showed that relative onset time had perceptual significance independent of F1 onset frequency, but they did not assess whether F1 onset frequency had independent perceptual significance. Summerfield & Haggard (1977) have also shown that F1 transition extent becomes salient in the case where F1 onset frequency is low.

Liberman et al (1958) found that synthesizing the F2 and F3 transitions as noise without cutting back F1 (delaying F1 onset) was ineffective as a cue to voicelessness, whereas adding such simulated aspiration in the upper formants when F1 was cut back strengthened the perception of voicelessness. This phenomenon is a trading relationship, where the perceptual value of one cue increases as the perceptual value of another decreases and vice versa. These compensatory changes are the consequences of the multiplicity of cues for a given phonetic contrast. For example, as F1 steady state increased in frequency, shorter VOTs were necessary to effect a shift from voiced to voiceless judgments and the slope of the identification function became steeper (Summerfield & Haggard, 1977). Similarly, the VOT boundary shifted to longer values as formant transition durations increased (Lisker et al, 1977). Systematic changes in the

voiced/voiceless boundary were attributable to the manipulation of F1 but not to variations in F2 and F3 (Lisker et al, 1975; Summerfield & Haggard, 1977).

F1 onset frequency at the onset of periodicity is a powerful cue, but results point to its limited value in the context of high vowels for which there is no or minimal F1 transition and hence no F1 onset co-variation with changes in VOT (Summerfield & Haggard, 1974). Overall, in the context of minimal F1 transition, aspiration becomes a more salient cue. When there are relatively extensive potential F1 transitions (as in low vowels), a high frequency F1 onset cues a voiceless percept and the extent of frequency change becomes less relevant perceptually. The transition extent becomes more relevant perceptually when the F1 transition has a low onset and cues a voiced percept. These results emphasize the importance of the context sensitivity of the cues to the voicing distinction in initial-word position. They suggest that cues are incorporated in a context-dependent perceptual processing system that allows listeners to differentiate patterns of cues as a function of contextual variations (e.g. vocalic environment). The voicing contrast has been extensively studied in English, but much less research has been conducted on the trading relationships between possible cues to the voicing distinction in French speakers.

Word-Final Stops

Preceding vowel duration is a sufficient cue to the perception of the voicing contrasts in synthetic speech where other potential cues are neutralized (Raphael, 1972; Raphael, Dorman, Freeman & Tobin, 1975). However, in naturally-spoken words where the spectral and temporal characteristics of acoustic signals co-vary as a function of contextual variables, vowel duration is not the only or the most effective cue (Wardrip-Fruin, 1982; Revoile et al, 1982; Hogan and Roszypal, 1980; Raphael, 1981).

As in the case of initial voicing, the perceptual salience of acoustic features for voicing in final stops is context-dependent. Results by Wolf (1978) and Revoile et al (1982) point to the greater importance of F2 final transition offset frequency than of closure voicing (i.e. phonation during closure); in the context of vowels with high F1 steady-state frequency, frequency changes in F1 transitions into the closure are extensive. Phonation during closure may be a more powerful cue in the context of syllables in which the vowels have low first formant steady state (i.e. high vowels) (Wardrip-Fruin, 1982; Raphael, 1981). In addition, Raphael (1981) found that the salience of the transitions seems to vary with the context, weak in certain pre-vocalic contexts and strong in some pre-consonantal contexts.

As in the perception of voicing contrasts in word-initial position in English, there seems to exist a trading relationship between some cues: a long duration of closure voicing and a short preceding vowel or a short duration of closure voicing and a long preceding vowel yielded more voiced responses (Hogan and Roszypal, 1980). Hogan and Roszypal (1980) found that varying closure voicing duration produced a relatively stronger perceptual effect than varying the duration of the vowel nucleus for speakers of English. This is in agreement with other findings (Wardrip-Fruin, 1982; Raphael, 1981).

The fact that, in English, deletion of the final release burst does not induce a change in the perception of the voicing classification of stops (Wolf, 1978; Revoile et al, 1980; Hillenbrand et al, 1984; Raphael, 1981) is consistent with the finding that 63% of the final stops are not released (Rositzke, 1943). Other findings, however, point to the differential effects of the burst deletion. The deletion of the burst affected the voiceless percept more than the voiced one (Wolf, 1978, Revoile et al, 1982). Flege & Hillenbrand

(1987) examined monolingual native speakers of French and English. Using naturally produced syllables, they created two continua from /pɛk/ to /pɛg/ and from /plk/ to /plg/. Only tokens with a fully voiced closure of about 75 ms with an audible final burst were used for the perception study. The continua were created by editing varying amounts of phonation from the final closure interval. To minimize the contribution of vowel duration to the voicing decisions, the vowel duration of /plg/ was reduced from 149 ms to 110 ms by removing alternate glottal pulses. Closure voicing and portions of the preceding vowel were removed in 10 ms-steps. Up to 120 ms of the signal were removed, including part of the vowel transitions. Half of the stimuli contained a burst (with-burst stimuli) and half did not (no-burst stimuli). Listeners were informed in their native language that they would be presented with English words that had been edited. They were told to indicate whether they heard a final /g/ or /k/ and to guess when uncertain. French speakers found the task more difficult than the English speakers, presumably because the words were English. Results indicated that removal of glottal pulsing from the final closure interval induced a perceptual shift for both English and French speaking listeners, suggesting that phonation is perceptually salient in both languages. But the responses of the English speakers, and not of the French, were affected by the vowel in the /ɛ, l/ continua.

The removal of the release bursts affected the voicing judgments of the French speakers whereas it did not affect the voicing responses of the English speakers. For French speakers, the percentages of /k/ responses were greater than the percentages of /g/ responses when 59 ms of closure voicing was removed from stimuli without a burst and when 73 ms of closure voicing was deleted from the stimuli with a burst. For English speakers there were no differences in perception between the continua containing the

same amount of voicing with or without a burst. These results confirm previous findings using English speakers (Hillenbrand et al 1984). Flege & Hillenbrand (1987) have hypothesized that the salience of bursts to French speakers may stem from the fact that word-final stops may be more consistently released in French than in English. Clearly, this study confirms expected differences between speakers of English and French. Because English velar stops are always released, it is likely that the English speakers were hearing unnatural and unfamiliar stimuli so that they were probably relying on alternative cues

Wolf (1978) and Raphael, Dorman & Liberman (1980) also demonstrated the contribution of the initial first formant transitions to the perception of the voicing class of word-final alveolar stops. Overall, the initial formant transitions provide some information about the vowel. This information is integrated into the perceived vowel duration. The question remains as to whether it is the transition or the increase in vowel duration that plays a role in the voicing perception of final stops.

Conclusions

Perceptual studies offer a complicated picture in which cues trade with one another to convey the voicing distinction to the listeners. These studies indicate that listeners attend to multiple cues whose relationships vary according to the context, which mirrors a similar sensitivity in the production of speakers. Also, findings from production and perception studies suggest that the acoustic features and cues to the voicing distinction differ between French and English. In both the perception and the production literature, there is an abundance of studies of English, but not of French. The aim of the

proposal is to remedy this situation by examining the acoustic correlates of the voicing distinction in French in various phonetic contexts.

CHAPTER III

METHODOLOGY

Overview

The aim of this study is to document how differences between phonetic contexts correlate with the temporal parameters (i.e. aspiration duration, closure duration, closure voicing duration) that can serve as an index of the voiced-voiceless opposition of French stops. It is anticipated that some of the acoustic features that contribute to the phonological distinction will differ from word-initial to word-final position, while other features will be distinctive in both contexts. It is expected that in the /pa/ context, /b, d, g/ closures will be longer than /p, t, k/ closures; /b, d, g/ closures will more consistently be phonated than /p, t, k/ closures; /b, d, g/ will be preceded by longer vowel durations than /p, t, k/. In the /pas/ context, it is expected that these duration differences will be attenuated and that a smaller number of /b, d, g/ closures will be phonated. In addition, it is hypothesized that the amount of closure voicing in the inter-vocalic context (i.e. the sentence frame: /il a di pa CvC a lui/) will be greater than in the voiceless fricative sentence context (i.e. the sentence frame: /il a di pas CvC sã lui/. This study will provide relative proportions of vowel durations between voiceless and voiced contexts (ratios) and determine the extent to which the values of the temporal parameters differ between sentence context and citation form speech.

Overall, this study will provide a better controlled phonetic context, using the same phonetic segment ([s] or [a]) immediately before and after the target stop (i.e. /il a

di pas CvC să lui/, il a di pa CvC a lui). The effect of phonetic context within the word was examined by using two vowels (ε, ɔ) in the target CVC syllable. These two vowels were selected because it was possible to construct minimal pairs. The contexts of the target syllable were also manipulated (voiceless consonant or vowel preceding and following) as well as the place of articulation of the stops (labial, alveolar, velar). Productions that are reasonable approximations to casual conversational speaking style were used as the basis for assessing which temporal measure(s) differentiate /b, d, g/ from /p, t, k/ in particular phonetic contexts. All target consonants were spoken in syllables under primary stress.

Method

Apparatus

The tokens were recorded on the two channels of a DAT-tape. The recording level of the left channel was higher than that of the right channel, so that in case clipping occurred on the left channel, signals on the right channel could be used for final acoustic analyses. A high definition microphone (Behringer model 2000) connected to a DAT-tape recorder (Sony, model DTC-Ze700) was positioned on a stand in front of the speaker. A pre-amplifier "ultra-grain" (Sony) was used. All files were then imported from the DAT-tape onto a PC via the Computer Speech Lab (CSL) model 4300 (version 5.05) or Multispeech, a windows version of CSL. The utterances recorded on the left channel were then imported at a sampling rate of 16,000 points per second. The CSL software (KAY Elemetrics-Windows version) was used to make the acoustic measurements. The parameters for narrow-band and wide-band-spectrograms and for Fast Fourier Transforms (FFTs) are listed in Appendix A.

Subjects

Twelve participants were French speakers raised and educated in France. Their ages ranged from 21 to 25 years old. They were students at the Institut de Linguistique et Phonétique Générales et Appliquées in Paris. All went through the typical foreign language training before entering college (7 years of a second language and 3 years of a third language). None had spent more than 2 months in a non-French speaking country for at least one year prior to recording. No participant spent vacations abroad on a regular basis. None believed they could express themselves fluently and without accent in any of the languages learned in school. Finally, none had any hearing or speech-language related problems according to self reports.

Design

There were four independent variables: (1) identity of the vowel in the target syllable (ϵ , ɔ), (2) position of the stop within the word (initial, final), (3) place of articulation (labial, alveolar, velar) and (4) preceding/following phonetic context (voiceless fricative or /pas/ context versus vowel or /pa/ context). In the /pas/ context, the following carrier sentence was used /il a di pas CvC sã lui/; in the /pa/ context, the following carrier sentence was used /il a di pa CvC a lui/. For example, in the word-initial stop condition speakers were asked to say /il a di pas D ϵ t sã lui/ (/pas/ context, i.e. voiceless fricative environment) or /il a di pa D ϵ t a lui/ (/pa/ context, i.e. vowel environment). In the word-final stop condition, speakers were asked to say /il a di pa b ɔ B a lui/ (/pa/ context) or /il a di pas b ɔ P sã lui/ (/pas/ context).

Material

Twenty-one CVC syllables in which V was (/ɛ/ or /ɔ/) were embedded in each of the two carrier sentences. The target syllables were either proper nouns, common nouns or non-sense words. The carrier sentences were /il a di pas CvC sã lui/ (/pas/ context); /l a di pa CvC a lui/ (/pa/ context). Speakers were requested to place prominent stress on the CvC, which appeared in boldface in the text. The list of CvC syllables are listed in Appendix B. Even though there are 12 minimal pairs, there are only 21 different CVC's. Three words served for the word-initial and the word-final contexts: /tɛt/, /kɔd/ and /pɛk/. Note that the sentences are grammatically awkward. But they were constructed primarily with the aim to maintain the phonetic context constant around the target stop consonants. The /pas/ carrier phrase could be translated as "He said give CvC without him" and the /pa/ carrier sentence could be translated as "He said no CvC to him".

Procedure

Speakers were requested to fill out a language background questionnaire prior to the recordings. Their hearing was not directly measured. Some speakers reported having a recent screening revealing no hearing deficit.

Speakers were then presented with lists of sentences written in phonetic alphabet. They were asked to read the sentences at a comfortable speaking rate. Even though they were asked to read, they were urged to speak as if they were talking to another native French speaker. After some practice trials, speakers were asked whether they felt their pace and speaking rate to be near their normal average speaking level and rate. When confirmed, the recording began when the speaker was ready. If the speaker was not

ready, he/she was asked to practice more and attempt to approximate "near normal" speaking behavior.

Twelve speakers were asked to produce the articulatory liaison normally used by speakers of French. Each sentence was preceded by a number that the subject read before saying the sentence. Subjects were told that they could pause between the number and the beginning of the phrase. When subjects omitted a word or mispronounced a word, they were requested to repeat the entire sentence.

Lists of sentences were blocked by sentence type. The subjects were asked to speak all the CvCs in the same carrier sentence. After finishing reading the list, speakers were invited to rest. They were then asked to say the second series of CvC's in the other carrier sentence. The order of the carrier-sentence condition was counterbalanced across subjects. In addition, no cognate pairs of sounds occurred in consecutive utterances.

Dependent measures

Figures 1 & 2 display spectrograms illustrating the measurements criteria for each context. An upper case letter has been associated with each measurement and affixed to the spectrograms.

Aspiration duration (A).

To measure the onset of aspiration duration, the cursor was placed on the first deflection on the waveform corresponding to the release burst and right before the first periodic pulse at its trough. On the wide-band spectrogram, this landmark generally corresponded to the first periodically excited upper formant energy (vertical striation).

"Aspiration +Vocalic nuclei": "Asp+Vow" in [pas] and [CvC] (A+V).

On the waveform, for the onset of "Asp+Vow" in [pas], the cursor was placed on the first burst and for its offset on the first sign of frication on the waveform (i.e. intense high-frequency energy (4000-8000 Hz) visible on the wide-band spectrogram). On the spectrogram this measurement represents the sum of A and V, where V stands for the vocalic interval.

"Aspiration+Vowel" in [pa] and [CvC] (A+V).

On the waveform, the cursor was placed on the first release burst and on the point where there was a reduction in amplitude which corresponded to a dramatic change in the wave. This also corresponded to the last visible periodically excited high formant energy (vertical striation) on the wide-band spectrogram.

In all cases, the trail of energy beyond the last periodically excited upper-formant energy (striation) visible on the wide-band spectrogram was ignored.

Closure duration (C).

The cursor was placed on the last periodically excited cycle of second formant energy and then moved to the first release burst. In the closure interval, if there were a spike followed by silence, it was ignored and the cursor was placed on the following burst that was immediately followed by aspiration/frication.

Closure voicing (CV).

Categorical measures were made: absence (when less than 25% of the closure duration contained periodic energy) and presence (when more than 75% of the closure duration contained periodic energy). Evidence of periodicity (phonation) was represented in the spectrogram by regularly spaced vertical striations at the level of the baseline.

Frication period of /s/ of [pas](F).

The cursor was placed on the first and last signs of intense, high-frequency frication [4000-8000 Hz]. Trails were ignored.

Phonated frication (PF)

Whether the /s/ was fully or partially voiced was determined using the criteria applied to determine the presences vs. absence of closure voicing.

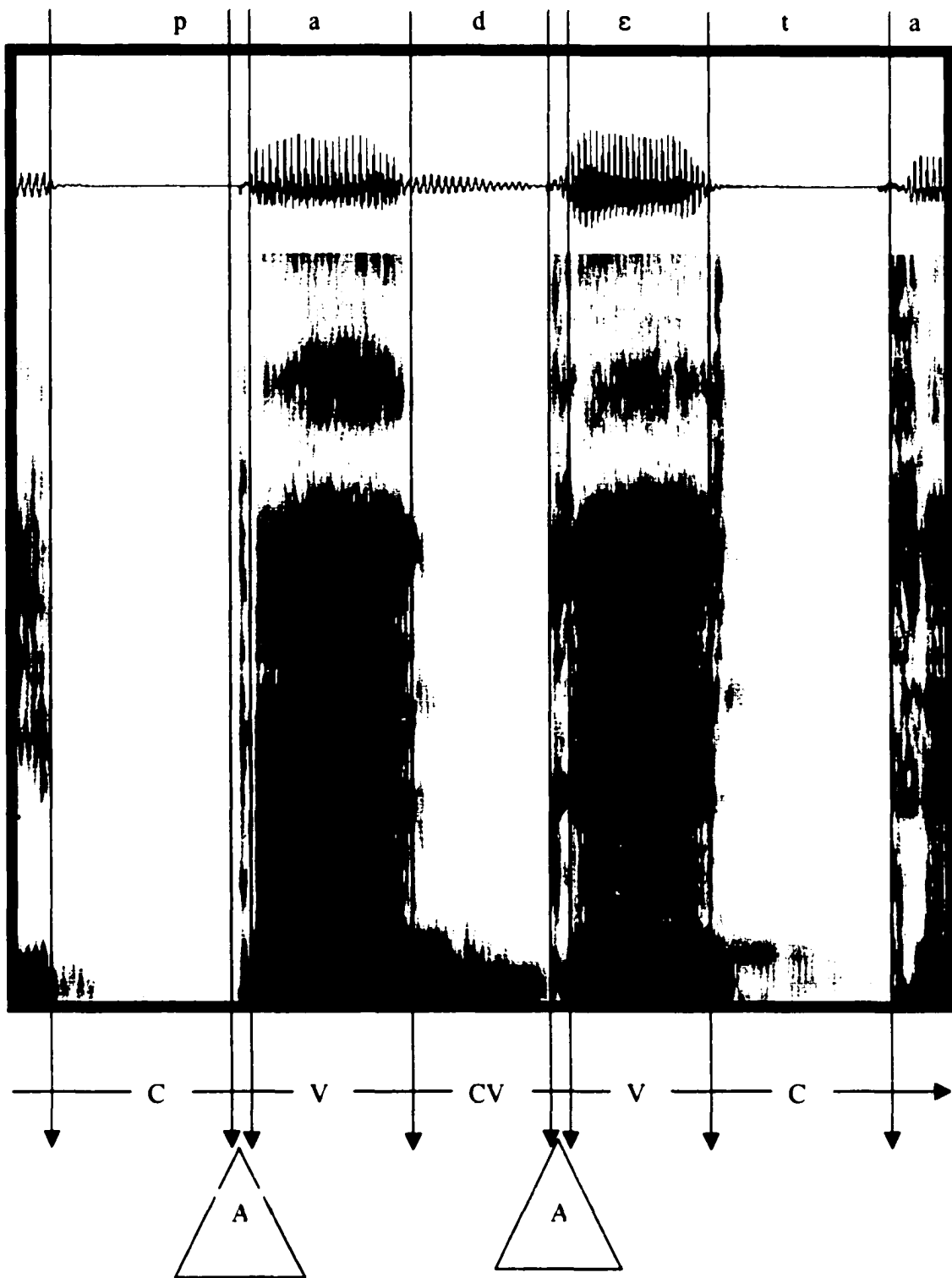


Figure 1. Waveform and corresponding spectrogram of /il a di pa dεt a lui/ (/pa/ context)

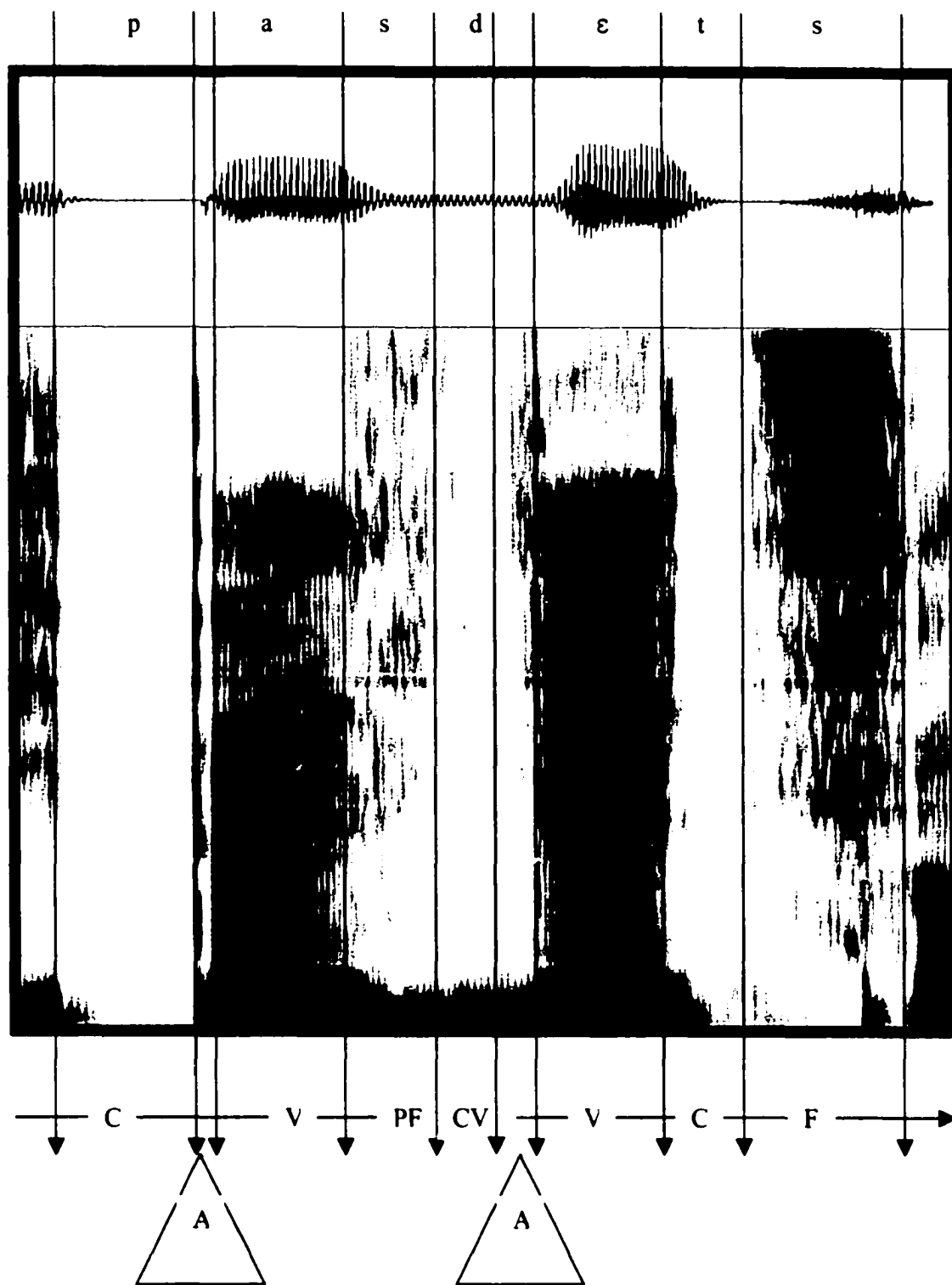


Figure 2. Waveform and corresponding spectrogram of /il a di pas dεt sã lui/ (/pas/ context)

The utterances of nine out of twelve speakers were analyzed. Three speakers were thus eliminated from the subject pool: One speaker's speech had a heavy Southern French accent from Marseille. Another speaker acknowledged speaking another language while a third speaker remembered having received speech therapy at a very young age. The first token of each type was not analyzed. In total, seven repetitions of twenty-one words produced by nine subjects were analyzed for the /pa/ and /pas/ contexts. A total of 2623 utterances was analyzed, of which 8 % were excluded from the final analyses because they were obscured by noise, flapped, or produced with incomplete closure.

Reliability measures.

Reliability measures were performed by a second judge on 10% of the tokens. Agreement was higher than 90% on all measures. Agreement was deemed to occur when values were within +/- 5 ms for measurements under 50 ms. For measurements greater than 50 ms, agreement was deemed to occur when values were within +/-10% of the longer of the two measurements.

CHAPTER IV

RESULTS

Separate correlated t-tests were performed on all mean durations. Means were computed for each phonological voicing category and for each subject. T-tests were computed on the average differences across subjects for each phonological voicing group. Partial Bonferroni correction was applied to adjust the alpha level to $p < 0.0035$. Chi-squares were performed on the frequency of phonated and unphonated closures. Acoustic measurements were made on a total of 2414 utterances, distributed as a function of phonetic context (See Table 1).

Table 1

Number of tokens per voicing category for each phonetic context

	<i>/pa/</i> context		<i>/pas/</i> context	
	<i>/b, d, g/</i>	<i>/p, t, k/</i>	<i>/b, d, g/</i>	<i>/p, t, k/</i>
Word-initial	310	322	283	295
Word-final	294	318	305	287
Total	604	640	588	582
	1244 utterances		1170 utterances	

Durational measures for voicing

In this study, measures of vowel duration included the period of aspiration following the release of a preceding stop (aspiration +vowel). To allow for comparisons between the present results and those of other research in which vowel duration did not include aspiration, a measure of vowel duration was obtained by subtracting the aspiration durations from the (aspiration +vowel) durations. Means and standard deviations for each voicing category in each context are displayed in Tables 2 and 3. Mean voicing related duration differences and standard deviations are displayed in Table 4 and 5. Differences were calculated by subtracting /p, t, k/ means from /b, d, g/ means. These duration differences have been plotted in figures 3 and 4.

Statistical analyses of the durational differences indicate that for word-initial stops in the /pa/ and /pas/ contexts as well as for word-final stops in the /pa/ context, vowels preceding /b, d, g/ are longer than vowels preceding /p, t, k/, and /b, d, g/ closures are shorter and less aspirated than /p, t, k/ closures. Note that the significant differences observed for the (aspiration+vowel) acoustic correlate are entirely attributable to vowel duration differences and not to the aspiration duration differences of the vocalic nuclei. Even though the three acoustic correlates (i.e. vowel duration, closure and consonant aspiration) reliably distinguish statistically /b, d, g/ from /p, t, k/ in the three contexts mentioned above, there is a noticeable reduction of closure and aspiration duration differences for word-initial stops in the /pas/ context. The reduction in closure duration differences stems from a “shortening” of /p, t, k/ closures while the reduction of the aspiration duration differences is due to /b, d, g/ being more aspirated. Additional statistical analyses indicate that for word-final stops in the /pas/ context, there remains only one significant distinguishing acoustic

correlate: vowel durations before /b, d, g/ are longer than before /p, t, k/. Note that all other duration differences have been reduced with closure duration differences incurring the greatest reduction, followed by aspiration duration differences. Further data analyses show that closure voicing is also influenced by variations in the phonetic contexts. Closure voicing data are reported in Tables 6, 7 & 8.

Table 2

Means and Standard Deviations (in parentheses) for the duration of each acoustic correlate for word-initial stops are displayed as a function of contexts. Durations are in ms.

	/pa/ context		/pas/ context	
	/b, d, g/	/p, t/ k	/b, d, g/	/p, t, k/
/pa/:aspiration	13 (2)	13 (2)	/pas/:aspiration	14 (3) 14 (3)
/pa/:vowel	78 (14)	68 (10)	/pas/:vowel	86 (13) 77 (13)
/pas/:aspiration +vowel	91(13)	80 (10)	/pas/:aspiration +vowel	100 (13) 91 (12)
			/pas/:/s/	70 (14) 77 (15)
Cvc:closure	65 (11)	87 (16)	Cvc:closure	66 (11) 73 (11)
Cvc:aspiration	9 (4)	23 (5)	Cvc:aspiration	14 (5) 24 (4)

Table 3

Means and Standard Deviations (in parentheses) for the duration of each acoustic correlate for word-final stops are displayed as a function of contexts. Durations are in ms

	/pa/ context		/pas/ context		
	/b, d, g/	/p, t, k/	/b, d, g/	/p, t, k/	
cvC:aspiration	16 (3)	15 (3)	cvC:aspiration	17 (4)	17 (4)
cvC:vowel	87 (16)	74 (11)	cvC:vowel	72 (12)	66 (11)
cvC:aspiration +vowel	103 (16)	89 (12)	cvC:aspiration +vowel	88 (11)	83 (9)
cvC:closure	53 (6)	74 (12)	cvC:closure	60 (11)	63 (13)
/a/:aspiration	4 (3)	25 (5)	/sā/:/s/	106 (11)	106 (12)

Table 4

Mean voicing related durational differences and Standard Deviations (in parentheses) for each acoustic correlate for word-initial stops are displayed in each context. Differences were computed by subtracting /b, d, g/ values from /p, t, k/ values. Duration differences are in ms.

	/pa/ context		/pas/ context
	/b, d, g/ - /p, t, k/		/b, d, g/ - /p, t, k/
/pa/:aspiration	0 (1)	/pas/:aspiration	0 (1)
/pas/:vowel	10 (5)**	/pas/:vowel	9 (2)***
/pas/:aspiration +vowel	11 (5)**	/pas/:aspiration +vowel	8 (2)***
		/pas/:/s/	-7 (6)
Cvc:closure	-21 (7)***	Cvc:closure	-7 (4)**
Cvc:aspiration	-15 (5)***	Cvc:aspiration	-10 (5)**

Note. * $-4.2 < t < 4.2$, $p < 0.0035$, ** $-5.3 < t < 5.3$, $p < 0.0004$,

*** $-8.9 < t < 8.9$, $p < 0.00003$

Table 5

Mean voicing-related durational differences and Standard Deviations (in parentheses)

for each acoustic correlate for word-final stops are displayed in each context.

Differences were computed by subtracting /b, d, g/ values from /p, t, k/ values. Duration differences are in ms.

	/pa/ context		/pas/ context
	/b, d, g/-/p, t/ k		/b, d, g/ - /p, t, k/
cvC:aspiration	1 (2)	cvC:aspiration	0 (2)
cvC:vowel	13 (6)*	cvC:vowel	6 (3)*
cvC:aspiration +vowel	14 (6)**	cvC:aspiration + vowel	6 (4)
cvC:closure	-22 (7)***	cvC:closure	-3 (4)
/a/:aspiration	-21 (6)***	sā:/s/	0 (4)

Note. * $-4.2 < t < 4.2$, $p < 0.0035$; ** $-5.3 < t < 5.3$, $p < 0.0004$;

*** $-8.9 < t < 8.9$, $p < 0.00003$

Voicing related duration differences have been plotted in figures 3 and 4. These figures are displayed on the next pages. For positive differences, bars are above the abscissa whereas for negative differences, bars are below the abscissa.

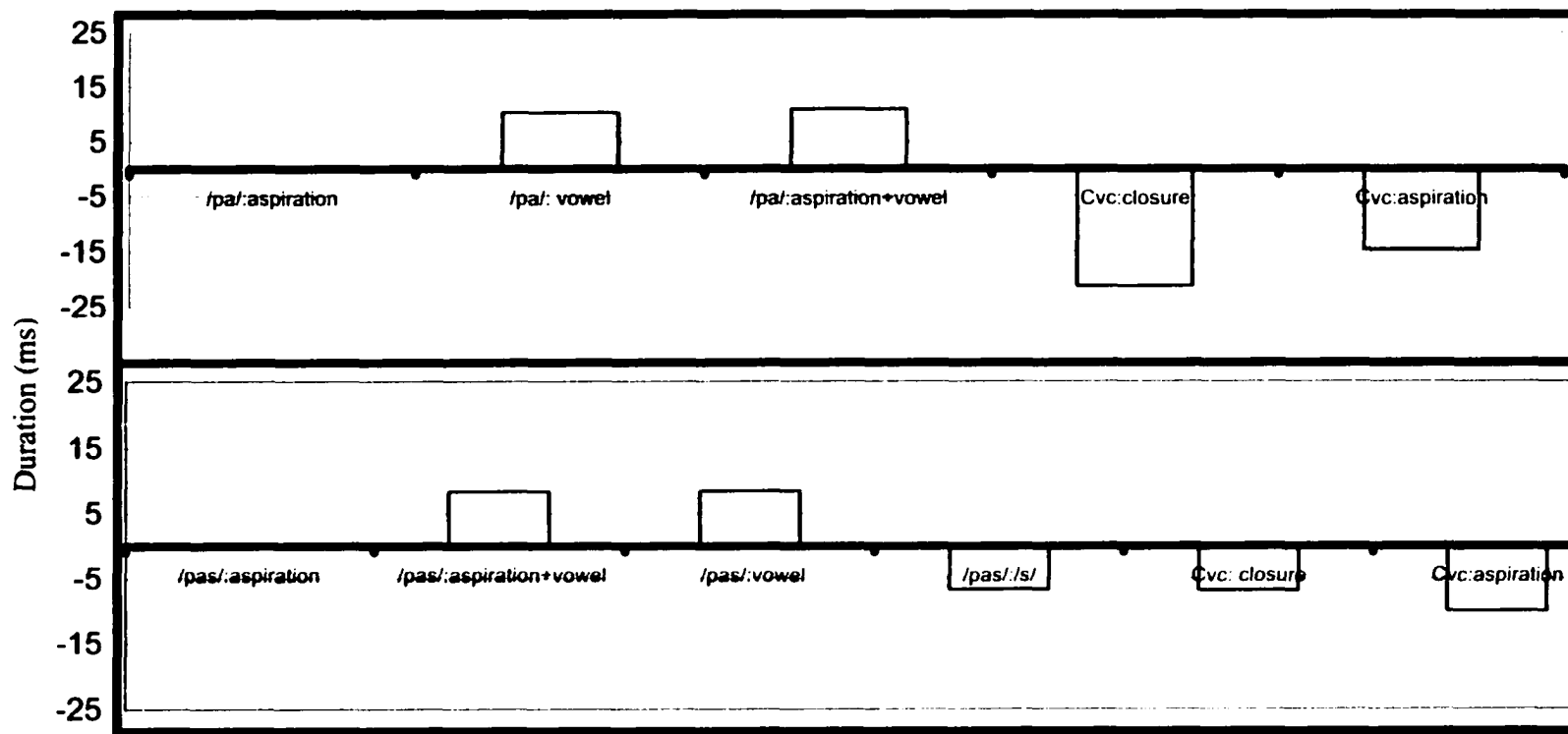


Figure 3. Duration differences of each acoustic correlate for word-initial stops in the /pa/ (top graph) and the /pas/ contexts (bottom graph). Differences were computed by subtracting /p, t, k/ from /b, d, g/ durations. A positive difference indicates that /b, d, g/ durations are greater than /p, t, k/ durations. A negative difference indicates that /b, d, g/ durations are smaller than /p, t, k/ durations.

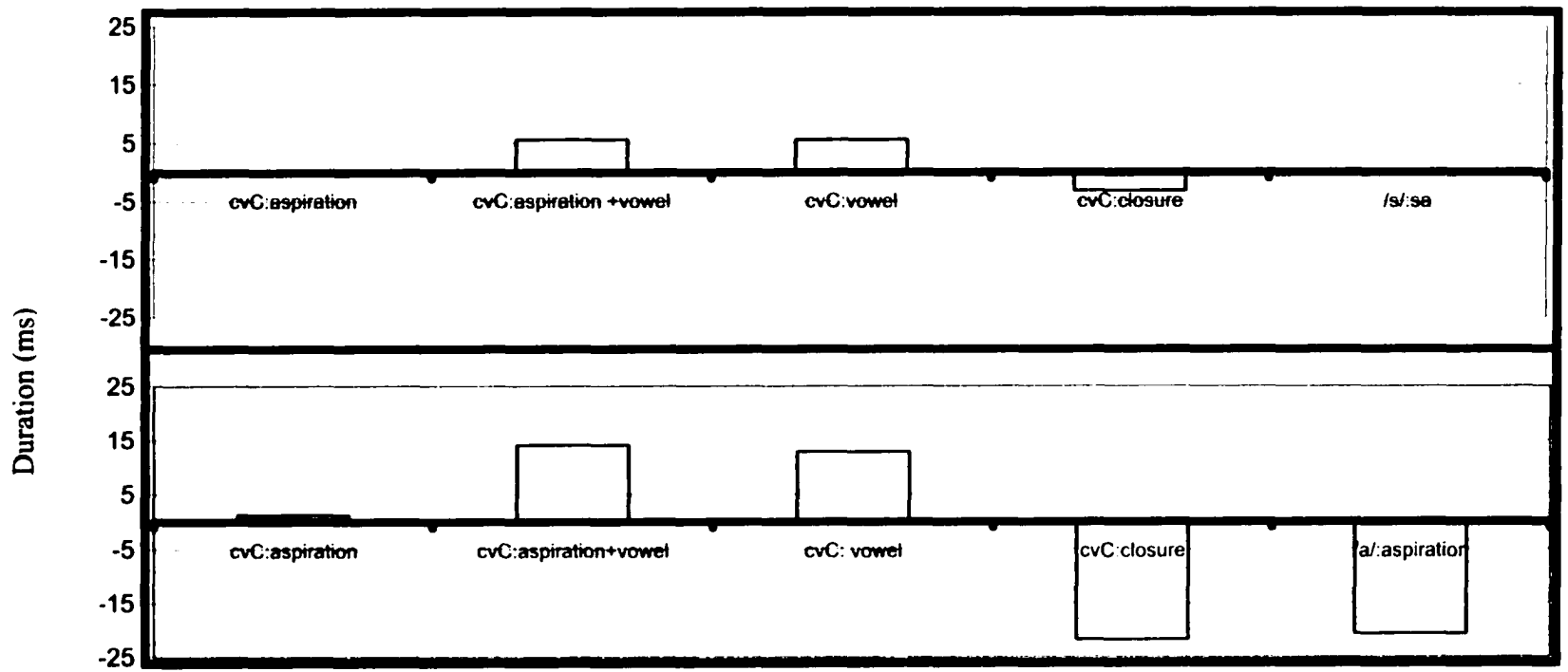


Figure 4. Duration differences of each acoustic correlate for word-final stops in the /pa/ (top graph) and the /pas/ contexts (bottom graph). Differences were computed by subtracting /p, t, k/ from /b, d, g/ durations. A positive difference indicates that /b, d, g/ durations are greater than /p, t, k/ durations. A negative difference indicates that /b, d, g/ durations are smaller than /p, t, k/ durations.

Closure voicing measures.

For word-initial stops in the /pa/ context, there is a significantly greater number of cases in which phonation runs throughout (75% or more of) the syllable-initial /b, d, g/ closures than the /p, t, k/ closures ($\chi^2(1) = 640$, $p < .0005$) (See Table 6). For word-initial stops in the /pas/ context, the number of times phonation runs through the /s/ of /pas/ (see Table 7) and through the initial closure of the Cvc (see Table 6) is significantly greater for initial /b, d, g/ than for initial /p, t, k/ (for /s/, $\chi^2(1) = 400$, $p < .0005$, for Cvc initial closure, $\chi^2(1) = 450$, $p < .0005$), for /pas/:/s/). Note that, acoustically, the /s/ of /pas/ that preceded syllable-initial /b, d, g/ is realized acoustically as [z]. This is an instance of regressive voicing assimilation.

Table 6

Number of times phonation is present or absent through the Cvc closure

	/pa/ context		/pas/ context	
	/b, d, g/	/p, t, k/	/p, t, k/	/p, t, k/
No phonation	1	321	29	293
Phonation Throughout	308	1	254	0

Table 7

Number of times phonation is present or absent throughout the /s/ of /pas/ (/pas/:/s/)

	/b, d, g/	/p, t, k/
No phonation	48	292
Phonation Throughout	235	3

The number of times phonation ran throughout word-final stop closures in the /pa/ context is significantly greater for /b, d, g/ than for /p, t, k/ ($\chi^2(1) = 580$, $p < .0005$). The number of times phonation ran throughout word-final stop closure in the /pas/ context is significantly greater for word-final /b, d, g/ than for word-final /p, t, k/ ($\chi^2(1) = 40$, $p < .001$) (See Table 8).

Overall, syllable-initial and syllable-final /b, d, g/ closures are more often phonated than syllable-initial and syllable-final /p, t, k/ closures in both the /pa/ and /pas/ contexts. Note that there is a smaller number of phonated /b, d, g/ closures in the /pas/ context. However, the number of phonated /b, d, g/ closures in the /pas/ context is smaller for word-final stops than for word-initial stops to such an extent that the number of unphonated word-final /b, d, g/ closures exceeds the number of phonated /b, d, g/ closures. This results attests to the phonetic/acoustic neutralization of the distinction between /b, d, g/ and /p, t, k/.

Table 8

Number of times phonation is present or absent through the cvC closure

	/pa/ context		/pas/ context	
	/b, d, g/	/p, t, k/	/p, t, k/	/p, t, k/
No phonation	3	316	193	268
Phonation Throughout	288	1	93	9

In summary, four acoustic correlates of the voicing distinction (preceding vowel duration, closure duration, closure voicing and aspiration) reliably differentiate /b, d, g/ from /p, t, k/ in three contexts (word-initial stops in /pa/ and /pas/ contexts and word-final stops in the /pa/ context). A striking case of regressive voicing assimilation emerges: /s/ of /pas/ is phonated throughout and is produced acoustically as a [z]. One striking effect on word-final stops in the /pas/ context was that voicing-related duration differences were dramatically reduced, suggesting that neutralization has taken place.

Durational measures for place of articulation

Statistical analyses performed on the mean durations and Standard Deviation for each place of articulation indicate that Cvc:aspiration reliably increases as place of articulation moves back in the mouth for word-initial stops in the /pa/ (\underline{M} /b/ = 3 ms, \underline{SD} = 2; \underline{M} /d/ = 10 ms, \underline{SD} = 6; \underline{M} /g/ = 13 ms, \underline{SD} = 5; \underline{M} /p/ = 16 ms, \underline{SD} = 4; \underline{M} /t/ = 21 ms, \underline{SD} = 5; \underline{M} /k/ = 33 ms, \underline{SD} = 7) (/b/ vs. /g/: $t(8) = -7.57$, $p < 0.00005$; /p/ vs. /t/: $t(8) = -4.41$, $p < 0.002$; /t/ vs. /k/: $t(8) = -8.59$, $p < 0.00004$; /k/ vs. /p/: $t(8) = -10.73$, $p < 0.00005$) and /pas/ contexts (\underline{M} /b/ = 7 ms, \underline{SD} = 4; \underline{M} /d/ = 16 ms, \underline{SD} = 6; \underline{M} /g/ = 21 ms, \underline{SD} = 7; \underline{M} /p/ = 14 ms, \underline{SD} = 4; \underline{M} /t/ = 25 ms, \underline{SD} = 6; \underline{M} /k/ = 32 ms, \underline{SD} = 6) (/b/ vs. /g/: $t(8) = -7.4$, $p < 0.0001$; /d/ vs. /b/: $t(8) = 5.10$, $p < 0.00097$; /p/ vs. /t/: $t(8) = -5.69$, $p < 0.00052$; /k/ vs. /p/: $t(8) = 8.72$, $p < 0.000038$). The aspiration of the word /a/ that follows cvC's (i.e. word-final stops) in the /pa/ context increases as place of articulation moves back in the mouth for /p, t, k/ only (\underline{M} /p/ = 15 ms, \underline{SD} = 4; \underline{M} /t/ = 24 ms, \underline{SD} = 6; (\underline{M} /k/ = 36 ms, \underline{SD} = 7) (/p/ vs. /k/: $t(8) = 11.18$, $p < 0.000006$; /t/ vs. /p/: $t(8) = 6.8$, $p < 0.0002$; /k/ vs. /t/: $t(8) = 9.07$, $p < 0.000006$). Word-initial labial closures in the /pas/ context (\underline{M} /b/ = 76 ms, \underline{SD} = 8; \underline{M} /p/ = 87 ms, \underline{SD} = 11) are significantly longer than either alveolar (\underline{M} /d/ = 62 ms, \underline{SD} = 11; \underline{M}

/t/ = 69 ms, SD = 13) or velar closures (M /g/ = 62 ms, SD = 10; ; M /k/ = 66 ms, SD = 11) (/b/ vs. /g/: $t(8) = 4.19$, $p < 0.0027$; /p/ vs. /t/: $t(8) = 5.52$, $p < 0.00062$; /k/ vs. /p/: $t(8) = -19.2$, $p < 0.00000019$). Word-final labial closures in the /pa/ context (M /b/ = 62 ms, SD = 9; M /p/ = 82 ms, SD = 14) are also significantly longer than either word-final alveolars (M /d/ = 47 ms, SD = 8; M /t/ = 71 ms, SD = 13) or word-final velars (M /g/ = 49 ms, SD = 4; M /k/ = 70 ms, SD = 11) (/b/ vs. /g/: $t(8) = 4.5$, $p < 0.001$; /d/ vs /b/: $t(8) = -9.2$, $p < 0.000027$; /p/ vs. /k/: $t(8) = 4.49$, $p < 0.001$; /t/ vs. /p/: $t(8) = 6.4$. 2, $p < 0.0002$). Finally, place-related vowel duration differences are not consistent across context: for word-initial stops in the /pa/ context, vowels before velars (M /g/ = 82 ms, SD = 14; M /k/ = 70 ms, SD = 10) and labials (M /b/ = 78 ms, SD = 13; M /p/ = 67 ms, SD = 10) are significantly longer than vowels before alveolars (M /d/ = 74 ms, SD = 12; M /t/ = 65 ms, SD = 10). In word-final stops in the /pa/ context, vowels before velars (M /g/ = 92 ms, SD = 16; M /k/ = 77 ms, SD = 11) and labials (M /b/ = 92 ms, SD = 16; M /p/ = 80 ms, SD = 14) are significantly longer than vowels before alveolars (M /d/ = 78 ms, SD = 16; M /t/ = 65 ms, SD = 11) while for word final- stops in the /pas/ context, vowels before /p/ (M = 72 ms, SD = 10) and /k/ (M = 67 ms, SD = 12) are longer than vowels before /t/ (M = 59 ms, SD = 116).

Overall, systematic and unsystematic results have been found for place of articulation. Among the systematic results are: aspiration duration increases as place of articulation moves back in the mouth in three contexts (word-initial stops in /pa/ and /pas/ contexts and word-final stops in the /pa/ context); /b/ closures are longer than either /d/ or /g/ closures for word-initial stops in /pas/ context and for word-final stops in the /pa/ context. Place-related vowel duration difference, however, were not systematic.

CHAPTER V

DISCUSSION

In the /pa/ context, results showed that the duration of the vowel /a/ in /pa/ is of greater duration before word-initial /b, d, g/ than before word-initial /p, t, k/. Closure durations of the first consonant in the target Cvc and the aspiration durations of these consonants are shorter for word-initial /b, d, g/ than for word-initial /p, t, k/. Finally, speakers consistently phonate during the closure of word-initial /b, d, g/ while they consistently produce word-initial /p, t, k/ closures without phonation.

In the /pas/ context, speakers performed in a similar manner. That is, /pas/:vowel duration, Cvc:closure duration, Cvc:aspiration duration, and the presence vs. absence of phonation in the Cvc:initial closure statistically distinguish word-initial /b, d, g/ from word-initial /p, t, k/ in the same ways as they did in the /pa/ context. In the /pas/ context, speakers used one additional articulatory/acoustic feature differentially: /s/ of /pas/ is consistently unphonated in anticipation of /p, t, k/ and so produced as [s] whereas /s/ is consistently phonated (produced acoustically as [z]) in anticipation of /b, d, g/. The latter constitutes a case of regressive voicing assimilation. Table 9 summarizes these results.

Table 9

List of articulatory/acoustic correlates of the voicing distinction that statistically differentiate /b, d, g/ from /p, t, k/ as a function of phonetic contexts

	Word-initial stops	Word-final stops
/pa/ context	/pa/:vowel duration Cvc:closure duration Cvc:closure voicing Cvc:aspiration duration	cvC:vowel duration cvC: closure duration cvC:closure voicing /a/:aspiration duration
pas/ context	/pas/:vowel duration Cvc:closure duration Cvc:closure voicing Phonated /s/ of /pas/ Cvc:aspiration duration	cvC:vowel duration cvC:closure voicing

In summary, speakers used many of the same acoustic correlates to distinguish word-initial /b, d, g/ from word-initial /p, t, k/ in both the /pa/ and /pas/ contexts. To signal the voicing distinction, speakers produced an articulatory/acoustic gestalt in which longer vowel durations are accompanied by shorter phonated closures and minimally aspirated consonants (for /b, d, g/), whereas shorter vowel durations are associated with longer, unphonated closures and maximally aspirated stops (for /p, t, k/). These combinations reflect the complementarity of the articulatory gestures: vowel duration before /b, d, g/ requires longer sustention of muscular activity (Raphael, 1975), which may shorten the articulation of the stop consonant in its vicinity. In addition, when closures are phonated, there is a minimal build up of air pressure in the oral cavity and hence minimal aspiration following the release. Lisker & Abramson (1964) pointed out that in languages with two phonological voicing categories, such as French or English, phonation and aspiration are mutually exclusive and VOT measures clearly separate the categories. This is not true in the case of languages with four phonological voicing categories such as Hindi, which contain voiced aspirates.

Clearly, the present results show that for word-initial stops there is an intricate relationship between gestures and that this relationship is maintained across variations in context: However, the weight of each component within this relationship varies by different amounts as a function of contextual variations. For instance, speakers accommodate for the presence of the /s/ in /pas/ by reducing the durations of other articulatory/acoustic segments. Voicing-related duration differences in the /pa/ and /pas/ contexts are listed in Table 10. The differences were computed by subtracting segment durations in the /p, t, k/ contexts from those in the /b, d, g/ contexts. Negative values

indicate that the durational measurements in the /p, t, k/ contexts are greater than those in the /b, d, g/ contexts. cursory inspection of the left-hand column of this table shows that the voicing related differences between the vowels of /pa/ and /pas/ are small in both /pa/ and /pas/ contexts (10-msec voicing difference in the /pa/ context; 9-msec voicing difference in the /pas/ context). Similarly, the voicing-related differences between Cvc:aspiration are about the same in the /pa/ and /pas/ contexts (-14-msec voicing difference in the /pa/ context; -10-msec voicing difference in the /pas/ context). But, the voicing related differences between Cvc:closures are larger in the /pa/ context than in the /pas/ context (-22-msec voicing difference in the /pa/ context; -7-msec voicing difference in the /pas/ context). Overall, voicing related Cvc:closure duration differences are more sensitive to a change from the /pa/ to the /pas/ context than are vowel and aspiration measures. It may be the case that this reduction of the Cvc: closure durations in the /pas/ context is due to the additional phoneme /s/, irrespective of its voicing identity. Because the introduction of a voiceless fricative phone is confounded with the addition of an extra phoneme, it is not possible to ascribe the decrease in Cvc:closure solely to the presence of a voiceless phone before the target stop consonant. However, a study by Luce & Charles-Luce (1985) suggest that this reduction in Cvc:closure may indeed be a result of a voiceless phone in the vicinity of /b, d, g/. They systematically manipulated the local phonetic context (vowel or stop) following the target stop, the voicing category of the stop and the sentence position (phrase-final or non-final). They found that closure duration differences failed to reach significance in all cases where the target stop was followed by a voiceless sound (/t/) in non-phrase-final position, in phrase-final position, and for most test words ending in /d/ or /t/. These results led them to conclude that

closure duration carries little perceptual weight. Finally, the percentage of phonated /b, d, g/ closures is also affected by the introduction of the /s/ of /pas/. Percentages of phonated closures are listed in Table 11. This percentage drops from 99% in the /pa/ context to 89% in the /pas/ context.

Table 10

Summary of voicing-related differences as a function of phonetic context and word position. The differences were computed by subtracting segment durations in the /p, t, k/ context from those in the /b, d, g/ context. Negative values indicate that the durational measurements in the /p, t, k/ contexts are greater than those in the /b, d, g/ contexts. Duration differences are in ms.

Word-initial stops			Word-final stops		
/b, d, g/-/p, t, k/			/b, d, g/-/p, t, k/		
Vowel duration	/pa/:vowel	78 - 68 = 10	/pa/ context	cvC:vowel	87 - 74 = 13
	/pas/:vowel	86 - 77 = 9	/pas/ context	cvC:vowel	72 - 66 = 6
CVC:closure duration	/pa/ context	65 - 87 = -22	/pa/ context		53 - 74 = -21
	/pas/ context	66 - 73 = -7	/pas/ context		60 - 63 = -3
Cvc:aspiration	/pa/ context	9 - 23 = -14			
	/pas/ context	14 - 24 = -10			

Table 11

Percentage of phonated closures as a function of contexts and voicing categories.

	<i>/pa/</i> context		<i>/pas/</i> context	
	<i>/b, d, g/</i>	<i>/p, t, k/</i>	<i>/b, d, g/</i>	<i>/p, t, k/</i>
Word- initial	99%	.003%	89%	0%
Word- final	98%	.003%	30%	.03%
Average	98.5%	.003%	59.5%	.015%

Overall, these results indicate that for word-initial stops, the distinction tends to be reduced in the /pas/ context: the introduction of a voiceless fricative phone (/s/) before the target stop affects closure durations of /p, t, k/ more dramatically than closure durations of /b, d, g/ (see Table 10). Furthermore, the number of fully phonated closures is smaller in the /pas/ context than in the /pa/ context.

For word-final stops, findings show that in the context of /pa/, 4 acoustic parameters (cvC:vowel duration, cvC:final closure duration, phonation in the cvC:final closure, /a/:aspiration duration) statistically distinguish word-final /b, d, g/ from word-final /p, t, k/: cvC:vowel durations are longer before word-final /b, d, g/ than before word-final /p, t, k/. Word-final /b, d, g/ closures are shorter than word-final /p, t, k/ closures. The aspiration durations of the word "à" (/a/) following word-final /b, d, g/ are shorter than those following word-final /p, t, k/. Finally, word-final /b, d, g/ closures are consistently phonated whereas word-final /p, t, k/ closures are not.

For word-final stops in the /pas/ context, 2 acoustic parameters (cvC:vowel duration, phonation in the cvC:closure) statistically differentiate word-final /b, d, g/ from word-final /p, t, k/ in the same way as in the /pa/ context: cvC:vowel durations are significantly longer before word-final /b, d, g/ than before word final /p, t, k/ and /b, d, g/ closures are significantly more often phonated than /p, t, k/ closures. However, in this condition (word-final stops in the /pas/ context), the number of unphonated /b, d, g/ closures (70%) exceeds the number of phonated /b, d, g/ (30%) closures. The percentage of cvC:closures that are phonated decreases from 98% for word-final /b, d, g/ in the /pa/ context to 30% in the /pas/ context (where /s/ follows the target cvC). Hence, in this condition, neutralization of the voicing distinction is phonetically realized by a combination of articulatory maneuvers:

speakers reduce closure duration differences by decreasing both word-final /b, d, g/ and /p, t, k/ closures (see Table 10) and they do not any longer systematically phonate /b, d, g/ closures (see Table 11). Neutralization also affects vowel duration: vowel duration differences for word-final stops decrease from 13 ms in the /pa/ context to 6 ms in the /pas/ context (see Table 10).

Overall, similarities in the patterns of results emerge for word-initial and word-final stops in both /pa/ and /pas/ contexts. The same acoustic correlates combine to produce patterns of differences that are qualitatively similar but that differ in magnitude. For example, vowels preceding the target stop are longer before word-initial and word-final /b, d, g/ than before word-initial and word-final /p, t, k/ regardless of the phonetic context (/pa/ or /pas/). But these voicing-related differences tend to be greater in the context of /pa/ than in the context of /pas/. The /b, d, g/ closure durations are shorter than the /p, t, k/ closure durations in both phonetic contexts and regardless of word-stop position. They are, however, always shorter in the /pas/ than in the /pa/ context whether the target stop is in word-initial or in word-final position. Similarly, Cvc:aspiration durations are shorter for word-initial /b, d, g/ than for word-initial /p, t, k/ but the difference is larger in the /pa/ context than in the /pas/ context. Finally, word-initial and word-final /b, d, g/ closures are more often phonated than word-initial or word-final /p, t, k/ closures; but word-final /b, d, g/ closures become less frequently phonated in the /pas/ context. The percentage of phonated closures is minimally affected by a change in stop position within the word, but it is greatly affected by the interaction of word-stop position with the addition of a voiceless fricative following the word-final stop (cvC): The percentage of phonated word-final stop closures decreases by

about 50% from 89% for word-initial stops in the /pa/ context to 30% for word-final stops in the /pas/ context (see Table 11).

In summary, all voicing-related acoustic correlates (vowel duration, closure duration, CvC:aspiration, closure voicing), tend to be reduced in the /pas/ context with the reduction being greater for word-final stops than for word-initial stops. Even though the presence of phonation during word-final /b, d, g/ closures is a statistically significant acoustic correlate of voicing class in the /pas/ context (30%), closure phonation is much less common than in the /pa/ context (98%). This dramatic difference in the occurrence of phonation in word-final /b, d, g/ may be caused by the difficulty in maintaining vocal fold pulsing in anticipation of a voiceless sound (the /vCs/ context), although in the case of word-initial stops, speakers are able to phonate through the frication of /s/ of /pas/ in anticipation of /b, d, g/ (the /VsC/ context). The order of the phones may account for the difference: That is, regressive voicing assimilation may produce a stronger effect than carryover assimilation. Note that the closure durations are more affected by the difference between the /pa/-/pas/ contexts than are any of the other acoustic correlates. Vowel duration remains relatively immune to the variation in the context from /pa/ to /pas/ for word-initial stops but it shows a mild sensitivity to the same contextual variation for word-final stops.

Clearly, in the /pas/ context, speakers tend to use smaller voicing-related durational differences, as well as less frequently phonated closures. It is not known if and to what extent smaller voicing-related durational differences affect perception. Given that for word-initial stops, closure voicing and vowel duration remain virtually unaffected by the difference between the /pa/ and /pas/ contexts, it seems unlikely that context plays an important role perceptually. But for word-final stops, it is expected that perception of the voicing contrast

will be strongly affected by these contextual differences. In light of the reduction in voicing-related word-final closure durations and in the percentages of phonated word-final /b, d, g/ closures in the /pas/ context, one would expect identification levels of word-final /b, d, g/ in /pas/ context to be low. The basis for neutralization at the production level and possibly at the perceptual level might be the differential manner in which word-endings are treated relative to word-beginnings.

There are some similarities between the findings of this study and those of Laeuffer's (1996a;b 1992). Consonant closure durations for each study are listed separately for each context in Table 12. Laeuffer (1996a,b ;1992) used a vocalic context before and following target consonants which allows us to compare her data with the results of this study collected for the /pa/ context. There is no available comparison with /pas/ context. Note that the differences listed on the right hand-side of table 12 are all in the same direction: /b, d, g/ closures are shorter than /p, t, k/ closures in all studies. Voicing related differences are larger in two of Laeuffer's studies relative to our study. Note also that word-initial stops in Laeuffer's study (1996a) are about 50% longer than word initial stops in the present study. Perhaps Laeuffer's speakers were asked to produce all word-initial stops in a blocked session at the beginning of the recording period which would account for the much longer durations observed in this condition only. Hence, sentences might have been produced using careful speech in the initial-stop condition, whereas speakers may have become less careful by the time they produced sentences in the final-stop condition. The consonant durations found in this study are shorter than those of Laeuffer's in all conditions, which suggests that our speakers did not use careful speech.

Table 12

Consonant closure durations from Laeuffer (1996a, 1992) and current findings. The difference (Dif.) was calculated by subtracting /p, t, k/ consonant closure durations from /b, d, g/ consonant closure durations. Durations are in ms.

		/b, d, g/	/p, t, k/	Dif
Laeuffer (1996a)				
Medial context	Word-initial	118	139	-21
Under focus				
Laeuffer (1996b)				
Table 6	Word-final	73	105	-32
Laeuffer (1992)				
Medial context	Word-final	71	102	-31
Under focus				
Table V				
Current study				
/pa/ context	Word-initial	65	87	-22
	Word-final	53	74	-21
Current study				
/pas/ context	Word-initial	66	73	-7
	Word-final	60	63	-3

Similarly, Laeuffer's vowel durations are twice as long as the vowel durations found in the present study, which reinforces the hypothesis that her speakers produced over-articulated speech. Vowel durations and vowel duration ratios are listed in Table 13 separately for each study and each context. The vowel duration ratios were obtained by dividing durations of vowels preceding /p, t, k/ by durations of vowels preceding /b, d, g/. Laeuffer's vowel duration ratios (1992, 1996b) are the smallest, which indicates that her speakers maximized the voicing-related vowel durational differences, probably because they used careful speech. In the present study, vowel durations are shorter, voicing-related vowel durational differences are smaller and the ratios are closer to 1, reinforcing the conclusion that speakers did not use careful speech.

Table 13

Vowel durations, vowel duration differences (Dif.) and vowel duration ratio for each study and each context. Durations are in ms.

	/bdg/	/ptk/	Dif.	Ratio
Laeufer (1992) medial context under focus non vocally released stops (table X)	171	134	37	.78
Laeufer (1996b) (table 5)	154	120	34	.78
Current study /pa/ context only				
Word-initial /pa/-vow	78	68	10	.87
Word-final CVC-vow	88	74	14	.84
Current study /pas/ context only				
Word-initial /pas/-vow	86	77	9	.90
Word-final CVC-vow	72	66	6	.92

Comparing Laeufer's percentages of phonetically voiced closures with percentages obtained in the present study is difficult since she used a different criteria than the one used in the current study. She recorded the percentages of closures voiced throughout (100%); the percentages of tokens with phonation over more than half of the closure duration (50+%); the percentages of stops with phonation over half or less of the closure interval (50-%); and finally the percentages of totally unphonated /b, d, g/ closures. In addition, we rarely found closures devoid of phonation for either /p, t, k/ or /b, d, g/. In spite of differences in criteria used in the present study, Laeufer (1996b) percentages of /b, d, g/ closures phonated over more than half their durations (95%) is close to the percentage found in the current study (98%).

Overall, Laeufer's findings and those of the present study concur. But her longer durations and larger durational ratios, added to the fact that her values were not much different from those that Mack (1982) found in citation-form material, suggest that her speakers were producing sentences using over-articulated or careful speech. In agreement with Luce & Charles-Luce (1985), the present findings suggest that closure duration is not a distinguishing acoustic correlate when the target sound is followed by a voiceless sound. Luce & Charles-Luce (1985) found that English speakers fail to produce a significant voicing-conditioned difference in closure duration in the case where the word-final stop was followed by the voiceless sound /t/. The present study reveals that vowel duration statistically distinguishes the voicing categories in all contexts, which agrees with the findings of Luce & Charles-Luce (1985) that voicing-related vowel duration

differences were significant in all contexts, whether the phoneme following the word-final stop was a voiceless stop or a vowel.

We also examined the potential influence of place of articulation on the voicing distinction. In the /pa/ context, vowel duration (/pa':vowel) is longer for word-initial /g/ than for word-initial /b/ or /d/. Cvc:aspiration durations increase as place of articulation moves back in the mouth for both word-initial /b, d, g/ and word-initial /p, t, k/. In the /pas/ context, Cvc:aspiration durations also increase as place of articulation moves back in the mouth and Cvc:closure durations for labials are longer than those for alveolars or velars for both word-initial /b, d, g/ and word-initial /p, t, k/.

For word-final stops in the /pa/ context, cvC:vowel durations are shorter for alveolars than for labials or velars for both voicing groups; in addition, the aspiration durations (/a':aspiration) of the word "à" is longer following /k/ than following /t/ or /p/. Finally, the cvC:closure duration of /b/ is longer than those of either /d/ and /g/. For word-final stops in the /pas/ context, cvC:vowel durations are shorter for /t/ than for /p/ and /k/.

In agreement with Laeufer (1996a), labial closures are consistently longer and in some cases, closure durations decrease as place of articulation moves back in the mouth. Laeufer (1992) contends that "it is related to the variation in VOT values: from front to back, closure duration decreases as VOT increases" (p.97). Note that in our study, aspiration duration also increases as place of articulation moves back in the mouth.

Conclusions

Overall, the present research highlights the critical influence of phonetic context on the production of the stop consonant voicing distinction in French. The results indicate that phonation in /b, d, g/ closures is the most consistently reliable articulatory/acoustic correlate

of the voicing distinction, whereas consonant closure duration is the least reliable correlate. It is in agreement with the conclusions reached by Luce & Charles-Luce (1985). Perceptual studies are needed to test this as a working hypothesis.

The results of the present study raise a number of other questions that need to be answered by subsequent research. For instance, the question remains as to whether word-final /b, d, g/ can be distinguished from word-final /p, t, k/ in a voiceless context. Word-final consonants may, in general, be more difficult to perceive than word-initial consonants. Future perception studies should attempt to test this hypothesis and to see if a voiceless context has specific effects on the perception of stop consonant voicing. Further, work needs to be done to determine which acoustic correlates and clusters of correlates are relied upon by listeners to perceive phonological voicing distinctions in all contexts/word-positions.

Appendix A

Parameters for narrow-band and wide-band spectrograms, for FFT analyses and windows' attributes are listed.

Narrow-band spectrogram	Wide-band spectrogram	FFT	Windows' attributes
512 points (45.78 Hz)	100 points (234.38 Hz)	256 points (15.625 ms window)	Not normalize

Appendix B

CvC syllables are displayed as a function of the voicing category and the place of articulation

	Word-initial stops	Word-final stops
Labials	<u>B</u> εk- <u>P</u> εk	gε <u>B</u> [*] -gε <u>P</u>
	<u>B</u> ɔt- <u>P</u> ɔt	bɔ <u>B</u> -bɔ <u>P</u>
Alveolars	<u>D</u> εt- <u>T</u> εt	tε <u>D</u> -tε <u>T</u>
	<u>D</u> ɔp- <u>T</u> ɔp	kɔ <u>D</u> -kɔ <u>T</u>
Velars	<u>G</u> εt- <u>K</u> εt	pε <u>G</u> -pε <u>K</u>
	<u>G</u> ɔd [*] - <u>K</u> ɔd	dɔ <u>G</u> -dɔ <u>K</u>

* non-sense words

References

- Belasco, S. (1953). The influence of force of articulation of consonants on vowel duration. Journal of the Acoustical Society of America, 1015-1016.
- Benguerel, A. P., Hirose, H., Sawashima, M. & Ushijima, T. (1978). Laryngeal control in French stop production: a fiberoptic, acoustic and electromyographic study. Folia Phonetica, 30, 175-198.
- Borden, G, Harris, K. S. & Raphael L. (1994). Speech Science Primer (3rd ed.). William & Wilkins, Baltimore, Maryland.
- Caramazza, A., & Yeni-Komshian, G. H. (1974). Voice onset time in two French dialects. Journal of Phonetics, 2, 239-245.
- Chen, M. (1970). Vowel length variation as a function of the voicing of the consonant environment. Phonetica, 22, 129-159.
- Crystal, T. H., & House, A. S. (1982). Segmental durations in connected speech signals: Preliminary results. Journal of the Acoustical Society of America, 72, 705-706.
- Delattre, P. (1951). Principes de phonétique Française à l'usage des étudiants anglo-américains. Middlebury, VT: Middlebury College store.
- Delattre, P. (1953). Les modes phonétiques du Français. The French Review, XXVIII, 1, 59-63.
- Delattre, P. (1962). Some factors of vowel duration and their cross-linguistic validity. Journal of the Acoustical Society of America, 34, 1141-1143.
- Delattre, P. (1965). Comparing the phonetic features of English, French, German: An

- interim report. New York: Chilton Books/Heidelberg: Julius Groos Verlag
- Denes, P. (1955). Effect of duration on the perception of voicing. Journal of the Acoustical Society of America, 27(4), 761-764.
- Durand. M. (1936). Etudes expérimentale sur la durée des consonnes parisiennes. Bibliothèque du Français Moderne, Paris.
- Durand. P. (1985). Variabilité acoustique et invariance en Français: Consonnes occlusives et voyelles. Editions du Centre National de la Recherche Scientifique, Paris.
- Flege, J.E. & Brown, W. S. (1982). The voicing contrast between English /p/ and /b/ as a function of stress and position-in-utterance. Journal of Phonetics, 10, 335-345.
- Flege, J. E. & Hillenbrand. J. (1987). A differential effect of release burst on the stop voicing judgment of native French and English listeners. Journal of Phonetics, 15, 203-208.
- Flege, J.E. & Port, R. (1981). Cross-language phonetic interference: Arabic to English. Language and Speech, 24, 125-146.
- Gottfried, T., & Beddor, P. (1988). Perception of temporal and spectral information in French vowels. Language and Speech, 31, 57-75.
- Harris, M. & Umeda, N. (1974). Effect of speaking mode on temporal factors in speech: vowel duration. Journal of the Acoustical Society of America, 56, 1016-1018.
- Hillenbrand, J., Ingrisano, D., Smith, H.L. & Flege, J.E. (1984). Perception of the voiced-voiceless contrast in syllable-final stops. Journal of the Acoustical Society of America, 76, 18-26.
- House, A. S., & Fairbanks, G. (1953). The influence of consonantal environment upon

- the secondary acoustical characteristics of vowels. Journal of the Acoustical Society of America, 25, 105-113.
- House A. S. (1961). On vowel duration in English. Journal of Acoustical Society of America, 33, 1174-1178.
- Klatt, D. H. (1975). Vowel lengthening is syntactically determined in connected discourse. Journal of Phonetics, 3, 129-140.
- 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.
- Keating, P. (1985). Universal phonetics and the organization of grammars. In Phonetic Linguistics: Essays in honor of Peter Ladefoged. Edited by Victoria A. Fromkin. Academic Press.
- Laeufer, C. (1992). Patterns of voicing-dependent vowel duration in French and English. Journal of Phonetics, 20, 411-440.
- Laeufer, C. (1996a). The acquisition of a complex phonological contrast: Voice timing patterns of English initial stops by native French speakers. Phonetica, 53, 86-110.
- Laeufer, C. (1996b). The acquisition of a complex phonological contrast: Voice timing patterns of English final stops by native French speakers. Phonetica, 53, 1175-142.
- Liberman, A. M., Delattre, P. C., & Cooper, F. S. (1958). Some cues for the distinction between voiced and voiceless stops in initial position. Language and Speech, 1, 153-167.
- Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing initial stops:

- Acoustical measurements. Word, 20, 384- 422.
- Lisker, L., & Abramson, A. S. (1967). Some effects of context on voice onset time in English stops. Language and Speech, 10, 1- 28.
- Lisker, L., Liberman, A. M., Erickson, D. M., Dechovitz, D. and Mandler, R. (1977). On pushing the voice onset time (VOT) boundary about. Language and Speech, 20, 209-216.
- Lisker, L. (1986). Voicing in English: a catalogue of acoustic features signaling /b/ versus /p/ in trochees. Language and Speech, 29, 3-11.
- Luce, P. A., & Charles-Luce, J. (1985). Contextual effects on vowel duration, closure duration, and the consonant/vowel ratio in speech production. Journal of the Acoustical Society of America, 78, 1949-1957.
- Mack, M. (1981). French and English word-final stop consonants: Monolingual and bilingual production. Brown University Working Papers in Linguistics, 66-72
- Mack, M. (1982). Voicing-dependent vowel duration in English and French: monolingual and bilingual production. Journal of the Acoustical Society of America, 71, 173-178.
- Miller, J. L., & Volaitis L.E. (1989). Effects of speaking rate on the perceived internal structure of phonetic categories. Perception and Psychophysics, 46 , 505-512.
- O'Shaughnessy, D. (1981). A study of French vowel and consonant durations. Journal of Phonetics, 9, 383-408.
- Peterson, G., & Lehiste, I. (1960). Duration of syllable nuclei in English. Journal of the Acoustical Society of America, 32, 693-703.
- Port, R. F. (1981). Linguistic timing factors in combination. Journal of the

Acoustical Society of America, 69, 262-274.

- Raphael, L. J. (1972). Preceding vowel duration as a cue to the perception of voicing of word-final consonants in American English. Journal of the Acoustical Society of America, 51, 1296-1301.
- Raphael, L. J. (1975). The physiological control of durational differences between vowels preceding voiced and voiceless consonants in English. Journal of Phonetics, 3, 25-33.
- Raphael, L. J. (1981). Duration and contexts as cues to word-final cognates opposition in English. Phonetica, 38, 126-147.
- Raphael, L. J., and Dorman, M. F. (1980). The contribution of CV transition duration to the perception of final consonant voicing in natural speech. Journal of the Acoustical Society of America, 67(1), 51.
- Raphael, L. J., and Dorman, M. F. (1980). Silence as a cue to the perception of syllable-initial and syllable-final stop consonants. Journal of Phonetics, 8, 269-275.
- Raphael, L. J., Dorman, M. F., & Geffner, D. (1980). Voicing-conditioned durational differences in vowels and consonants in the speech of three and four-year old children. Journal of Phonetics, 8, 335-341.
- Raphael, L. J., Dorman, M. F., & Liberman, A. L. (1980). On defining the vowel duration that cues voicing in final position, Language and Speech, 23(3), 297-307.
- Raphael, L. J., Dorman, M. F., Freeman, F., & Tobin, C. (1975). Vowel and nasal duration as cue to voicing in word-final stop consonants in English; Spectrographic and perceptual studies. Journal of Speech and Hearing Research, 18, 389-400.

- Revoile, S., Pickett, J. M., & Holden L. D. (1982). Acoustic cues to final voicing for Hearing-impaired and normal-hearing listeners. Journal of the Acoustical Society of America, 72, 1145-1154.
- Rositzke, H. A. (1943). The articulation of final stop in General American English. American Speech, 18, 39-42
- Simon, P. (1967). Les consonnes Françaises. Klincksieck, Paris.
- Stathopoulos, E. T., & Weismer, G. (1983). Closure duration of stop consonants. Journal of Phonetics, 11, 395-400.
- Suen, C-Y. & Beddoes, M. P. (1974). The silent interval of stop consonants. Language and Speech, 126-134.
- 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.
- Umeda, N. (1975). Vowel duration in American English. Journal of the Acoustical Society of America, 58, 434-445.
- Umeda, N. (1977). Consonant duration in American English. Journal of the Acoustical Society of America, 61, 846-858.
- Van Summers, W. (1987). Effect of stress and final-consonant voicing on vowel production: Articulatory and acoustic analyses. Journal of the Acoustical Society of America, 82(3), 847-862.
- Volaitis, L. E., & Miller, J. L. (1992). Phonetic prototypes: influence of place of articulation and speaking rate on the internal structure of voicing categories. Journal of the Acoustical Society of America, 92, 723-735.

Wajskop, M. (1978). Indices temporels des occlusives intervocaliques en Français.

Rapport des Activités de l'Institut Phonétique de l'Université de Bruxelles, 12,

71-98.

Wardrip-Fruin, C. (1982). On the status of temporal cues to phonetic categories:

Preceding vowel duration as a cue to final voicing in final stop consonants.

Journal of the Acoustical Society of America , 71, 187-195.

Westbury, J. (1983). Enlargement of the supraglottal cavity and its relation to stop

consonant voicing. Journal of the Acoustical Society of America , 73, 1322-1336.

Wolf, C. G. (1978). Voicing cues in English final stops. Journal of Phonetics, 6, 299-309.

Zlatin, M. A. (1974). Voicing contrast perceptual and productive voice onset time

characteristics of adults. Journal of the Acoustical Society of America, 56, 981-

994.