National Library
of Canada

Bibliothèque nationale
du Canada

Canadian Theses Service

Services des thèses canadiennes

Ottawa, Canada
K1A 0N4

CANADIAN THESES

THÈSES CANADIENNES

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30. Veuillez prendre connaissance des formules d'autorisation qui accompagnent cette thèse.

THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED

LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE

NL 339 (r. 86/01)

Canada

238

0-315-26969-3

National Library of Canada | Bibliothèque nationale du Canada

Canadian Theses Division | Division des thèses canadiennes

Ottawa, Canada
K1A 0N4

PERMISSION TO MICROFILM — AUTORISATION DE MICROFILMER

• Please print or type — Écrire en lettres moulées ou dactylographier

Full Name of Author — Nom complet de l'auteur
Janet Daphne Mary Schwegel

Date of Birth — Date de naissance March 24, 1959	Country of Birth — Lieu de naissance Canada

Permanent Address — Résidence fixe
4319 - 38 Street
Edmonton, Alberta
T6L 5A6

Title of Thesis — Titre de la thèse
Acoustic Factors in the Perception of Voiced Labial and Dental Stop/Fricative Categories

University — Université
University of Alberta

Degree for which thesis was presented — Grade pour lequel cette thèse fut présentée
M. Sc.

Year this degree conferred — Année d'obtention de ce grade 1984	Name of Supervisor — Nom du directeur de thèse John T. Hogan

Permission is hereby granted to the NATIONAL LIBRARY OF CANADA to microfilm this thesis and to lend or sell copies of the film.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

L'autorisation est, par la présente, accordée à la BIBLIOTHÈQUE NATIONALE DU CANADA de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans l'autorisation écrite de l'auteur.

Date April 24/84	Signature J Schwegel

NL-91 (4/77)

THE UNIVERSITY OF ALBERTA

ACOUSTIC FACTORS IN THE PERCEPTION OF VOICED LABIAL AND DENTAL STOP/FRICATIVE CATEGORIES

by

JANET D. M. SCHWEGEL

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE
OF MASTER OF SCIENCE
IN
SPEECH PRODUCTION AND PERCEPTION

DEPARTMENT OF LINGUISTICS

EDMONTON, ALBERTA
SPRING 1984

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR JANET D. M. SCHWEGEL

TITLE OF THESIS ACOUSTIC FACTORS IN THE PERCEPTION OF VOICED LABIAL AND DENTAL STOP/FRICATIVE CATEGORIES

DEGREE FOR WHICH THESIS WAS PRESENTED MASTER OF SCIENCE

YEAR THIS DEGREE GRANTED SPRING 1984

Permission is hereby granted to THE UNIVERSITY OF ALBERTA LIBRARY to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

(SIGNED) ...JSchwegel........

PERMANENT ADDRESS:

....4319 - 38 Street.....
....Edmonton, Alberta....
T6L 5A6
....

DATED ..April 24,......1984

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled ACOUSTIC FACTORS IN THE PERCEPTION OF VOICED LABIAL AND DENTAL STOP/FRICATIVE CATEGORIES submitted by JANET D. M. SCHWEGEL in partial fulfilment of the requirements for the degree of MASTER OF SCIENCE.

John T. Hogan

Supervisor

A. Rozsypal

J. Baker

Gary Holdgrafer

Date April 24, 1984

Acknowledgements

I want to thank all dese people bery much:

Dr. John Hogan, my supervisor, for sharing so much time and so many ideas.

Dr. Wm. Baker, Dr. Anton Rozsypal, and Dr. Gary Holdgrafer, the members of my committee, for their provocative input.

Dr. Terry Nearey, Peter Assmann, and Al Opperthauser, for helping me control the computer.

Dr. Gary Prideaux, Mrs. Helen Hawkes, and Mrs. Florence Cunningham, for running the department well and showing so much consideration for me.

Ava Bandi, John Hogan, Bryan Larocque, Eric Vanderkruk, and Liz Vanderkruk, for their recordings.

My subjects, Peter Assmann, Ava Bandi, Donna Hawley, John Hogan, Murray Munro, Terry Nearey, and Tim Reeve, for their time.

Leslie Wickett, for help with the typing.

Mom and Dad, for being so proud of me.

Bryan Larocque, for being my all-hours, no-problem-too-large-or-small consultant and editor.

(As well, the Department of Linguistics and the Alberta Heritage Scholarship Fund, for financial support.)

Abstract

Although stops and fricatives are generally considered qualitatively different types of consonants, this study provides evidence that the distinction is quantitative. The primary factor in differentiating English voiced labial and dental fricatives and stops is the duration of the consonant noise.

Two experiments were designed to investigate how alterations in some acoustic attributes of stops and fricatives affect the perception of these consonants. The particular English consonants involved are /b/, /d/, /v/, and /ð/. These consonants appeared word-initially before the vowels /i/, /æ/, and /o/ in actual English words. Two Canadian English speakers recorded these words in carrier phrases. The intent of the experiments was to investigate the roles played by onset slope, intensity, and duration in distinguishing stops and fricatives. One experiment varied the onset, the amplitude, and the duration of the noise portion of the stops. The other experiment increased the amplitude and decreased the duration of the fricative frication.

The primary determinant of the manner of the obstruent - stop or fricative - is the consonant duration. The duration of consonant noise boundary between stops and fricatives occurred at approximately 25 ms. Signals shorter than this were identified as stops. Longer signals were identified as fricatives. Removing the abrupt onset of a

stop consonant without increasing duration had almost no effect on the consonant's categorization. Intensity influenced the recognition of some stimuli with manner-ambiguous durations.

Table of Contents

List of Tables

List of Figures

Figure Page

Figure Page

I. Introduction

What is the difference between a stop and a fricative? In terms of articulation, the difference is that stops involve a complete closure in the vocal tract while fricatives are produced by narrowing the vocal tract at some point. But what are the acoustic differences relevant to stop and fricative identification? The answer is not as clear. Perceptual studies of the distinction between stops and fricatives in English are not available. Many papers on the acoustic properties of fricatives exist (Carney & Moll, 1971; Delattre, Liberman, & Cooper, 1962; Fujisaki & Osamu, 1978; Harris, 1958; Heinz & Stevens, 1961; Hughes & Halle, 1956; Jassem, 1965; Lacerda, 1982; Lariviere, Winitz, & Herriman, 1975; McCasland, 1979; Strevens, 1960). There are also many studies of the acoustic properties of stops (for example, Blumstein & Stevens, 1980; Dorman, Studdert-Kennedy, & Raphael, 1977; Fischer-Jorgenson, 1954; Halle, Hughes, & Radley, 1957). As well, the differences between stops and glides (Miller & Liberman, 1979; Suzuki, 1970), and the differences between fricatives and affricates have been investigated (Dorman, Raphael, & Isenberg, 1980; Gerstman, 1957; Howell & Rosen, 1983; Repp, Liberman, Eccardt, & Pesetsky, 1978; and Van Heuven, 1979). But few studies have directly compared perceptual attributes of fricatives and stops (Baker, 1975; Isenberg, 1978; Malecot, 1968; Treon, 1970). This thesis investigates the perceptual effects of manipulating acoustic properties of some English fricatives and stops.

The particular English consonants under consideration are the word initial /b/, /d/, /v/, and /∂/. The acoustic properties of the noise portion of the consonants (stop burst or fricative frication) which were manipulated are duration, slope of onset (abrupt vs. gradual), and intensity. The frequency spectra of the consonants and the formant transitions were not altered. In the literature review, the acoustic properties of stops and fricatives will be discussed. At this point, the four phonemes and the reasons for studying these four are considered.

The English stops /b/ and /d/ and fricatives /∂/ and /v/ were chosen because they are the only English stop/fricative pairs that have approximately the same places of articulation: /b/ and /v/ are labial; /d/ and /∂/ are dental.

Voiced consonants were chosen intentionally. A comparison of the voiceless fricatives and stops would be complicated by the long voice onset time (VOT) after the release of voiceless stops which does not appear after voiceless fricatives. Gerstman (1957) conducted an experiment in which the overall duration and onset time of the voiceless affricate /tʃ/ was reduced. At short durations and abrupt onsets, considerably more voiced stops were heard than unvoiced stops, although the original stimulus consonant was unvoiced. Gerstman says that "the result is not surprising when we consider that voiceless stops usually show a considerable separation between friction and vowel ... while voiced stops do not" (p. 51). Since there was no delay between the consonant and vowel in Gerstman's stimuli, his consonants were heard as voiced when perceived as stops. Voiced consonants were selected as stimuli for the present experiments to avoid this complication associated with voicing of stops.

There is some question of the comparability of /b/ to /v/ and /d/ to /∂/. In the case of /b/-/v/, the question is primarily one of difference in place of articulation: /b/ is produced bilabially and /v/ is labiodental. But, since both are labial, their consonant-vowel formant transitions should be similar. The non-existence of an English bilabial fricative phoneme and a labiodental stop phoneme indicates that /b/ and /v/ are the phones to be compared, even though their places of articulation are not identical.

Some investigators regard the alveolar fricatives /s/ and /z/ as the counterparts of the stops /t/ and /d/. One reason for comparing /d/ and /∂/ instead of /d/ and /z/ is historical. According to Grimm's law, the "Indo-European bh, dh, and gh became, respectively, the Germanic sounds β, ∂, and γ, and later, in initial position at least, b, d, g" (Pyles and Algeo, p. 89). Also, Indo-European /p/, /t/, /k/ became Germanic /f/, /θ/, and /χ/. Further, Verner's law claims that these Proto-Germanic voiceless fricatives became voiced in a voiced environment before a stressed syllable. So, /f/, /θ/, /χ/ became /β/, /∂/, /γ/. These sound shifts, dh - /ð/ - /d/ and /t/ - /θ/ - /ð/ confirm the assumption of a relationship between the dental fricatives and stops.

Another justification for pairing /ð/ and /d/ is their similarity in production. The phonemes /θ/ and /ð/ are often called "interdental", but, though /θ/ may be interdental, /ð/ certainly is not. Umeda (1977) noted that many /ð/'s are articulated with a tap of the tongue against the alveolar ridge, an articulation very similar to that of /d/. Similarly, in preliminary investigation, the author noted small bursts in the oscillograms of some tokens of /ð/. Umeda described the articulation of /ð/ as glide-like because it does not "hold a constant articulatory manner" as do /s/ and /z/ (p. 849). The author would call the articulation of /ð/ stop-like because of the brief occlusion of the vocal tract sometimes present in its production. Either way, it is apparent that the articulation of /ð/ resembles that of /d/ more closely than the articulation of /z/ does, both in place and manner of production.

Substitutions made by children provide another argument for the comparability of /d/ and /ð/. In child acquisition of English fricatives, Moskowitz (1975) noted that /d/ is typically substituted for /ð/ at first. In her study of eight children, aged 1:1 to 3:5, Moskowitz found that seven out of the eight produced /d/ for /ð/ at least sometimes. (Only the 3:5 year old had complete control over his /ð/.) Moskowitz also notes that an older child may substitute /v/ for /ð/ in imitation when he has to match "the acoustic characteristics of the model he is to imitate. And /v/ serves the purpose more satisfactorily than /d/ does" (p. 147). However, "for the child who does not yet have total mastery over /ð/, and who also does not have /θ/ or /z/, the closest substitute in articulation is /d/. In his spontaneous speech, he seems to be attempting to learn how to change or modify /d/ appropriately to arrive at /ð/" (p. 147). This evidence supports the statement that /ð/ is highly similar to /d/.

A final piece of evidence supporting /d/-/ð/ similarity is their occurrence in morphophonemic alternations in some languages. Danish exhibits opposition of strong and weak consonants (Jakobson, Fant, & Halle, 1969). Strong consonants occur at the beginning of monosyllabic words, weak consonants occur at the end. For example, the phoneme /d/ occurs as [d] word initially as in dag, "day"; however, word finally /d/ is pronounced [ð] as in had [hað], "hate". Another language with such an alternation is Portuguese: intervocalically /d/, /b/, /g/ become [ð], [β], [γ]. Spanish also exhibits this allophonic variation. The historical sound changes of /ð/ to /d/, the articulatory similarity of these phones, the substitution of /d/ for /ð/ in language acquisition, and the occurrence of morphophonemic alternations between /d/ and /ð/ are all evidence that /d/ and /ð/ are similar and bear comparison.

Finally, the status of /ð/ in English must be discussed because of its peculiarity. The phoneme /ð/ is relatively rare in languages of the world. According to Ruhlen's 1976 inventory, it occurs in 36 out of the 700 languages surveyed - only 5 percent. In English, /ð/ is unusual in that it does not occur in initial position in "full content" words (Umeda, 1977). It could be considered a word initial function-word-allophone of /d/. But on the basis of the minimal pair test, it is an English phoneme. For example, the difference between the words "dough" and "though" is the word initial stop or fricative. In context, /d/ and /ð/ may be distinguished, word initially, solely by linguistic factors; nevertheless, it is interesting to study the perceptual properties of /ð/ because, although its word initial occurrence is restricted to function words, these function words occur very frequently.

In a dictionary frequency count done by Trnka (1935), /ð/ was rated least frequent out of 23 English consonants. Since /ð/ occurs in few words, especially few content words, it rates low in a count in which every word is considered only once. In this count /d/ was eighth, /b/ was tenth, and /v/ was fourteenth. French (1930) counted frequency of occurrence of phones in conversation, which meant that phones in frequently occurring words - such as function words - were counted as occurring more often. French's data was collected from extemporaneous telephone conversations. Tobias (1959) retranscribed French's data to a more standard transcription and analysed the relative occurrence of phonemes in the data. He found that /ð/ ranked sixth most frequent consonant, /d/ ranked third, /v/ thirteenth, and /b/ sixteenth in the retranscribed data. It is clear from a comparison of Trnka's and French's studies that although /ð/ occurs in relatively few words, these words occur frequently. The frequent occurrence of /ð/ can be attributed to the fact that many deictic words contain a /ð/: the, this, that, these, those, there, then, than.

Other studies of phoneme usage also rank /ð/ high in frequency, usually just slightly lower than /d/. Denes (1963) found /ð/ to be the seventh most frequently spoken consonant. According to Denes, /d/ is fourth most frequent, /b/ is twelfth, and /v/ is thirteenth. Word initially, however, /ð/ is the most frequent consonant, /b/ is the seventh most frequent consonant, /d/ is ninth, and /v/ is nineteenth. Wang and Crawford (1960) compiled the results of ten frequency counts of English phonemes, including the Trnka and French surveys mentioned above. In nine of the ten, the tenth being Trnka's dictionary count, /ð/ was between twelfth and sixth most frequent: in four studies it was sixth, in three it was seventh. In these nine studies /d/ was ranked between tenth and third, usually third, fourth, or fifth; /b/ always ranked between 12 and 17; and /v/ ranked between 10 and 17. The phonemes /b/ and /v/ occur approximately equally often in English. Likewise, /ð/ and /d/ occur with similar frequency in English conversation. These equalities support the comparison of the labial and dental stops and fricatives.

Much is known about English stops and fricatives. But little is known about the perceptual relationship between the two types of consonants. The many aspects in which they are related encourages an investigation of the perception of these consonants. The intent of this thesis is to contribute to an understanding of the bases of recognition of the English voiced labial and dental stops and fricatives. Although the question posited at the beginning of this chapter asked for the differences between stops and fricatives, the intent here is to illuminate not only the differences but also the similarities between /b/ and /v/ and /d/ and /∂/.

II. Literature Review

A. Spectral Properties

Many spectrographic studies have been done on the properties of fricatives and stops. (Such studies include Baker, 1975; Blumstein & Stevens, 1980; Delattre, Liberman, & Cooper, 1955; Delattre et al., 1962; Halle et al., 1957; Heinz & Stevens, 1961; Hughes & Halle, 1956; Jassem, 1965; McCasland, 1979; and Strevens, 1960.) The spectral qualities of stops and fricatives have not often been compared, although some conjectures have been made. Strevens (1960) supposes that the spectrum of the burst of a voiceless stop would be like the spectrum of a closely homorganic fricative. Harris (1958) also suggests that the frication of a fricative is comparable spectrally to the burst of a stop. In comparing the spectrograms of affricates and fricatives, Howell and Rosen (1983) found no spectral differences. These studies imply comparability between fricatives and stops, at least with respect to the noise spectra.

Another spectral consideration in the comparison of consonants, other than that of the noise portion, is the consonant to vowel transition. Most studies of formant transitions and formant transition loci have considered stop consonants, and ignored fricatives. Carney and Moll (1971), however, suggests that Öhman's (1966) observations apply to fricatives as well as stops: VCV articulations can be represented as vowel to vowel diphthongal gestures with an independent consonant superimposed on those gestures.

Delattre et al. (1955) discuss second formant transition loci and postulate that since they reflect the place of articulation of a consonant, all homorganic consonants should have the same second formant locus. For example, the second formant transitions from /b/, /p/, and /m/ into a following vowel would have the same frequency patterns since formant transitions reflect the movement of articulators from the consonant to the following vowel, and /b/, /p/, and /m/ all have the same place of articulation. In a later paper, Delattre et al. (1962) present the results of an investigation of second formant transition loci for English fricatives. As in Delattre's earlier stop experiments, they synthesized stimuli with flat (horizontal) second and third formants at various frequencies. The locus found for the second formant of /ð/ was 1400 Hz; the locus for /v/ was 700 Hz. The placement of the F2 locus for /ð/ is between those found earlier for /b/ (700 Hz) and /d/ (1800 Hz). This is the expected result since the second formant transition varies according to place of articulation, and the place of articulation of /ð/ is close to that of /d/, but slightly more anterior. The /v/ and /b/ loci were found to be the same, indicating that the place of production of these two consonants is the same.

This brief review of spectral properties of stops and fricatives indicates similarities between homorganic stops and fricatives with respect to (1) the frequency spectra of the noise portions and (2) the consonant-vowel formant transitions. That is, the noise of the burst of the stop /b/ and the noise of the frication of the fricative /v/ are comparable in noise frequency. Likewise, the stop /d/ and the fricative /ð/ should be similar. As well, the consonant-vowel formant transitions for the labials and those for the dentals should be the same.

B. Onset

While studies comparing perceptually relevant qualities of stops and fricatives are rare, many studies have been done comparing fricatives and affricates. (For instance, Dorman et al., 1980; Gerstman, 1957; Howell & Rosen, 1983; Repp et al., 1978; and Van Heuven, 1979.) In studies of English consonants, the particular fricative and affricate usually considered are /ʃ/ and /tʃ/. The cues for the fricative/affricate distinction are duration and rise time (Gerstman, 1957; Cutting & Rosner, 1974; and Howell & Rosen, 1983).

In initial position, Gerstman (1957) showed that rapid rise times and brief durations of fricative noise lead to the perception of affricates, while slower rise times and longer frication lead to the perception of fricatives. Gerstman claimed that fricatives are distinguished from affricates principally on the basis of rise time. Cutting and Rosner (1974) studied rise times in both speech and non-speech stimuli and found that musical stimuli (sawtooth and sine waves) with onset rise times of less than 40 ms were identified as "plucked" 92 percent of the time. Stimuli with rise times of 50 to 80 ms were called "bowed" 87 percent of the time. This indicates, generally, that listeners can distinguish signal quality based simply on duration of rise time. Using speech stimuli, Cutting and Rosner found results similar to those of the musical stimuli. Obstruent consonant noise signals with rise times from zero to 30 ms long were identified as /tʃ/ 88 percent of the time; those with rise times 60, 70, or 80 ms long were called /ʃ/ 80 percent of the time; and stimuli with 40 or 50 ms rise times were ambiguous. From these results, Cutting and Rosner concluded that the auditory mechanism has a natural sensitivity at approximately 40 ms.

Howell and Rosen (1983) dispute Cutting and Rosner's results. They found that the mean rise times for affricates and fricatives were 33 and 76 ms in running speech and 49 and 123 ms in isolation for /tʃ/ and /ʃ/, respectively. In nonsense syllables they were 61 and 120 ms long. Howell and Rosen also attempted to measure the rise time of the voiced affricate /dʒ/ and the voiced fricative /ʒ/. At the same time, they remeasured /tʃ/ and /ʃ/ using a different procedure. They found mean rise times for /dʒ/ of 49 ms and for /ʒ/ of 90 ms, for /tʃ/ 53 ms and for /ʃ/ 99 ms. The comparison indicates only a small difference between the rise times of voiced and voiceless consonants.

Howell and Rosen use their results to dispute Cutting and Rosner's claim of a "natural auditory sensitivity" at 40 ms. The rise time boundary must not be fixed at 40 ms because rise time varies according to context - running speech versus isolated words. And, even if there were no context variation in rise time, the boundary between /tʃ/ and /ʃ/ could not be at 40 ms because this would result in almost all of Howell and Rosen's measured stimuli being classified as fricatives.

Howell and Rosen noted that their own measurements of rise times are notably longer than Gerstman's (1957) measurements for /tʃ/ and /ʃ/. Gerstman found affricates with rise times as short as 5 ms. When Howell and Rosen synthesized /tʃ/ and /ʃ/, they found natural sounding affricates with 30 to 50 ms long rise times. Although they have yet to determine the cause of the discrepancies, Howell and Rosen do concur with Gerstman that rise time distinguishes (voiceless) fricatives from affricates.

Howell and Rosen's findings, and the findings of others comparing stops with affricates and affricates with fricatives, indicate categorical perception of the manner classes based on differences which vary along acoustic continua. One of these continua is duration of onset rise time. The present study investigated the role of the duration of onset in the identification of stops and fricatives.

C. Duration

Rise time is one factor in the fricative-affricate distinction; another separate but related factor is consonant duration. Van Heuven (1979) reanalysed Gerstman's (1957) data, factoring out variances due to duration and to rise time in order to verify Gerstman's analysis. Van Heuven found that duration accounts for 75 percent of the response variance in the data, and rise time and steady time each accounted for only an additional seven percent of the variance. This contradicts Gerstman's analysis which attributes the fricative/affricate distinction mainly to rise time, not to overall duration. Other studies in which duration is critical merit mentioning to show the variety of situations in which duration is important in distinguishing speech segments. Liberman, Delattre, Gerstman, and Cooper (1956) found that the duration of the consonant to vowel formant transition was a sufficient cue for distinguishing stops from semivowels. As the transition duration increased the stimuli were increasingly perceived as semivowels. Duration also seems to be a factor in the fricative-stop distinction. Grimm (1966) noted that truncated fricatives - fricatives with part of their frication removed - were generally heard as stops. Carden, Levitt, Jusczyk, and Walley (1980) also found that the "primary cue to the stop-fricative manner constrast in CV syllables is the duration of the aperiodic noise that precedes the formant transitions" (p. 62).

A few studies have measured the duration of stop and fricative consonants and have come up with somewhat varying measurements. Umeda (1977) found the mean duration of /v/, word initially before a stressed vowel, was 78 ms. He found that /∂/ is usually shorter in the same position: 52 ms. Umeda attributed this difference to the fact that the only English words that begin with /∂/ are function words which tend to be spoken relatively quickly. Malecot (1968) measured the duration of the word initial /v/ before a stressed vowel as 202 ms. (Malecot did not include /θ/ and /∂/ in his investigations.) Abbs and Minifie (1969) gave the following mean durations for /v/ and /∂/ in initial position in CV syllables: /v/ - 153 ms; /∂/ - 123 ms. The variation found in the duration of /v/ for these studies is remarkable: Umeda - 78 ms, Abbs and Minifie - 153 ms, and Malecot - 202 ms. The variation may be explained by the fact that word initial stressed consonants show a relatively large variance compared to word-medial and word-initial unstressed consonants (Umeda, 1977).

Another possible reason for the variation in the measured duration of the fricatives is speaker variation. Variation in the articulation of fricatives has been noted by Hughes and Halle (1956), Klatt and Cooper (1975), Malecot (1968), and Nartey (1983). Hughes and Halle found great discrepancies among the spectra of a fricative spoken by different speakers in various contexts. Klatt noted that the duration of the fricative /ʃ/ was variable across all the speakers studied. Malecot found differences existing among subjects in the force of articulation of consonants, including fricative consonants. He also noted variability in the amplitude and duration of the consonant. As well as expecting variation in the duration of consonants produced, we should expect variation in the spectra and intensity of the consonants.

Many studies measure the duration of pre-burst closure for stops. In comparison to fricatives, it is the duration of the noise portion of the stop, rather than the pre-burst closure, which is of interest because it is the burst noise that is comparable to the frication noise. Klatt (1975) measured the VOT of stop consonants (from the release of the plosive to the onset of vertical striations in the second and higher formants of the vowel). He noted that for voiced stops, burst duration equals VOT by definition. He made the following measurements of voiced stops: The burst duration for /b/ was on average 11 ms; for /d/, 17 ms. The standard deviation for his measurements was approximately 5 ms for both /b/ and /d/.

Klatt explains the short duration of /b/ as a result of a rapid labial release. Further, the rapid release generates a burst spectrum that is weak in intensity because there is no resonating cavity in front of the lips. The short duration and the weak intensity of the /b/ release cause the burst to be judged as not very loud, since the loudness of the burst is proportional to both its intensity and its duration. This attribute of /b/ bursts may account for identifications of truncated, or otherwise altered syllable initial consonants, as /b/: The burst of /b/ is very low in intensity; therefore, a consonant with no burst or little burst might be identified as a /b/.

Although there is a large difference between the durations of stop bursts and fricative frication, there may be a durational continuum between the two. The experiments conducted here altered noise durations of stops and fricatives to test for such a relationship. The large differences in the original durations of these manner classes indicate that duration may be a major factor in distinguishing stops and fricatives.

D. Intensity

Not only does intensity play a role in distinguishing fricatives from each other, there is evidence that intensity influences the perception of voicing and place of articulation of speech sounds. It is well accepted that intensity distinguishes the fricatives /s/ and /ʃ/ from /f/ and /θ/. McCasland (1979) found that /s/ and /ʃ/ are identified as /f/ and /θ/ if they are spliced to replace /f/ or /θ/ in a CV syllable and attenuated in amplitude. When /s/ and /ʃ/ reach the level of intensity of /f/ and /θ/, they are identified as /f/ or /θ/, depending on the original consonant.

Abbs and Minifie (1969) measured the intensity of the fricatives /s/, /z/, /f/, /θ/, and /ð/ in CV and VC syllables. They found that /s/ and /z/ were significantly more intense than the other fricatives. This study supports McCasland's evidence that intensity is a cue used to distinguish fricatives.

The evidence of the effect of intensity on perception of voicing is not as consistent as that of its role in distinguishing fricatives. Minifie (1973) claims voiced fricatives are less intense than voiceless fricatives because of the reduced intraoral pressure in the production of voiced fricatives, even though Abbs and Minifie (1969) found no consistent differences in intensity between voiced and unvoiced fricatives. So, it is not clear whether or not intensity is a reliable cue for the perception of voicing in fricatives.

Repp (1979) found that amplitude of aspiration noise is a cue for the distinction between voiced and voiceless syllable-initial stop consonants in English. Repp identified a trading relationship between the amplitude of the aspiration noise and VOT, in the identification of synthetic /da/ and /ta/ syllables which varied along a VOT continuum. An increase in the aspiration amplitude resulted in shift of the voiced-voiceless boundary towards shorter VOT.

Ohde and Stevens (1983) studied the effect of the relative amplitude of the release burst of a stop on its place identification. They varied the amplitude of the first 10 to 15 ms of voiced and voiceless synthetic stop stimuli which varied spectrally along a labial-dental place continuum. Ohde and Stevens found that "the relative amplitude of the burst significantly affected the perception of the place of articulation of both voiceless and voiced stops" (p. 706). An increase in amplitude biased a response toward the dental consonant over the labial one. This effect was noted especially in stimuli with spectra intermediate between labial and dental end-stimuli.

Malecot (1968) measured the amplitude of English stops and fricatives in various positions within a nonsense disyllable. His study is one which compares stops and fricatives, but it does not include /ð/ and /θ/. In initial position, Malecot found that the voiced stops /b, d, and g/ have greater amplitudes than the voiced fricatives /v/ and /z/.

Since only voiced consonants are being considered here, studies which compare intensity with respect to voicing are relevant only in showing that intensity affects identification. The same holds for the effect of intensity on fricative identification, since we are interested in the comparison of stops and fricatives. From Malecot's evidence we would expect increased intensity of an ambiguous consonant stimulus to bias response towards stop, rather than fricative. And Ohde and Stevens' study suggests higher intensity will increase the proportion of dental responses.

It is to determine the validity of these conjectures that intensity is included as a factor in the present investigation.

E. Consonant Confusions

Consonant confusion statistics are relevant to the current study for the following reasons: 1) They specify errors in consonant identifications to expect. 2) They indicate what proportion of confusions between two consonants is normal. And, 3) they give an estimation of the distance between consonants in perceptual space. This allows us to speculate on the salient distinctions between consonants. In order to get an idea what identification errors are common among the four consonants under consideration, the results from four consonant confusion studies have been tabulated. The studies included in the table are Kent, Wiley, and Strennen (1979); Miller and Nicely (1955); Tolhurst (1954); and Wang and Bilger (1973). Reporting more than one study allows for the comparison of a variety of results; this is important because there are some discrepancies in the proportion of misperceived stimuli and the usual choice of substituted consonant. Also, the results of consonant confusion studies are not easily replicated, suggesting that any claims must be considered in light of other findings.

The table of consonant confusions studied, Table 2.1, shows the most frequent misperceptions are between /b/ and

Table 2.1 Results of consonant confusion studies.

	Study			
Confusion	Kent (1979)	Miller & Nicely (1955)	Tolhurst (1954)	Wang & Bilger (1973)
/b/-/d/	none	rare	rare	rare
/b/-/v/	some	many	many	many
/b/-/ð/	few	few	some	some
/v/-/d/	none	rare	few	rare
/v/-/ð/	some	many	many	many
/d/-/ð/	some	few	some	few

<u>Note</u>. "rare" indicates approximately 1% occurrence
"few" indicates approximately 5% occurrence
"some" indicates approximately 10% occurrence
"many" indicates approximately 20% occurrence

/v/ and between /v/ and /ð/. Confusions between /d/ and /ð/ and between /b/ and /ð/ occur occasionally, but between /b/ and /d/ and between /v/ and /d/ they are rare. In general, /d/ is well-perceived, even under poor acoustic conditions (such as high signal to noise ratios and narrow frequency ranges). On the other hand, /ð/ is frequently misperceived. Kent et al. (1979) found that at 40 dB SL all consonants are perceived correctly more than 80 percent of the time, except /θ/ and /ð/.

The prevalence of /v/-/ð/ confusions is interesting because it involves an error in place identification. The phonemes /v/ and /ð/ are often confused, even under ideal conditions. Miller and Nicely (1955) found that /v/-/ð/ and /f/-/θ/ distinctions were among the most difficult for listeners to make. Wang and Bilger (1973), who varied signal to noise ratios and presentation levels, found that /ð/ was identified as /v/ twice as often as it was correctly identified. And Kent noted that even at 60 dB SL /v/ and /ð/ were identified correctly only 93 percent of the time. The high rate of confusion between the labial and dental fricatives suggests consonant-vowel formant transitions are not always sufficient to establish place of articulation.

More surprising than the /v/-/ð/ confusion is the /b/-/ð/ one, since the latter differ in both manner and place. Despite their differing articulatory features, these phonemes apparently are somewhat similar acoustically. The /b/-/ð/ errors do not occur very frequently, but they do occur in all the studies reported. The labial/dental misidentification is another demonstration of the relative difficulty of discerning place of articulation, as apparent in /v/-/ð/ errors. The manner confusion suggests similarity between voiced stops and fricatives. Manner confusion also appears in the /v/-/b/ substitutions, which are frequent, and the /ð/-/d/ substitutions, which occur occasionally. The relatively high occurrence of stop-fricative confusions indicates strongly that these consonants are close together in the auditory system's perceptual space. Under less than optimum listening conditions stops and fricatives are confused.

Carden et al. (1980) conducted an experiment designed to explore not just consonant confusions but particularly stop-fricative confusions. As their aim is similar to the aims of the present study and since Carden et al. achieve interesting results, their study will be discussed in some detail.

The experiments of Carden et al. investigated the relationship between manner identification and place identification. Manner-ambiguous stimuli were created by truncating the frication from natural fricatives. From subject categorizations of /f/, /v/, /θ/, and /ð/, they concluded that perception of place is dependent on the manner perceived. In one experiment, they found that the proportion of labial responses was greater when a stimulus was identified as a stop than when the same stimulus was identified as a fricative. For example, the truncated /θ/ was identified as /b/ 27 percent of the time and as /θ/ 20 percent of the time. It was identified as a labial 75 percent of the time it was identified as a stop, but as a dental 84 percent of the time it was identified as a fricative. Carden et al. stated the phenomenon thus: "...perceived manner affects the perception of the formant-transition cue" (p. 79). They suggest that listeners compare incoming transitions to one of two labial-dental boundaries, one boundary dividing the stops, the other separating the fricatives. Two boundaries are proposed because of the difference in place of articulation between the labial stops and fricatives and between the dental stops and fricatives. Since the stops and fricatives are articulated at slightly different places, their formant transitions would be different, and the optimal boundary between the labial and dental stops would not coincide with the optimal boundary between the labial and dental fricatives. Carden et al. further suggest that the fricative transitions, for /f/ and /θ/ at least, lie on the labial side of the labial-dental stop boundary. This explains why truncated /θ/ is perceived as labial when identified as a stop.

One difficulty with this explanation is its inability to account for the identifications of truncated /ð/, which was mainly called dental, not labial, whether identified as a stop or a fricative. Carden et al. suggest /ð/ is articulated further back, towards /∂/ than /θ/ is. Although this may be true, one wonders whether this slight difference is sufficient to cause a noticeable difference in the formant transitions and is enough to separate /θ/ from /∂/ over the stop labial/dental boundary.

Carden et al. recognize alternative hypotheses to explain their results, hypotheses based on acoustic attributes of the ambiguous stimuli. One acoustic hypothesis they propose is that lack of burst is a labial cue. Labial stops typically have smaller amplitude, shorter duration bursts than dental stops. For this reason, manner-ambiguous stimuli lacking bursts may be identified as labials when identified as stops. The other cue proposed is length of frication distinguishing between /v/ and /∂/. For some speakers, the frication of /∂/ is very short, thus fricative stimuli with zero or short frication might be identified as dental more than as labial. These two cues are as predictive of Carden's results as the two-boundary hypothesis.

Although direct comparisons of stops and fricatives have not often been done, from the nature of stops, fricatives, and affricates, we can expect certain similarities and differences to be evident. We expect the spectral aspects of the consonants and the consonant-vowel transitions will be similar when the place of articulation of the consonants is similar. We expect duration of onset to distinguish stops (with abrupt burst onsets) from affricates (with slower onsets) and fricatives (with very gradual onsets). Since natural stops are much shorter than natural fricatives, short duration of noise should indicate a stop while long duration of noise should indicate a fricative. Finally, since stops are higher in intensity than fricatives, higher consonant amplitude should influence perceptions of consonants toward stops. To verify these conjectures, duration, intensity, and slope of onset are investigated in the experiments described in the next chapter.

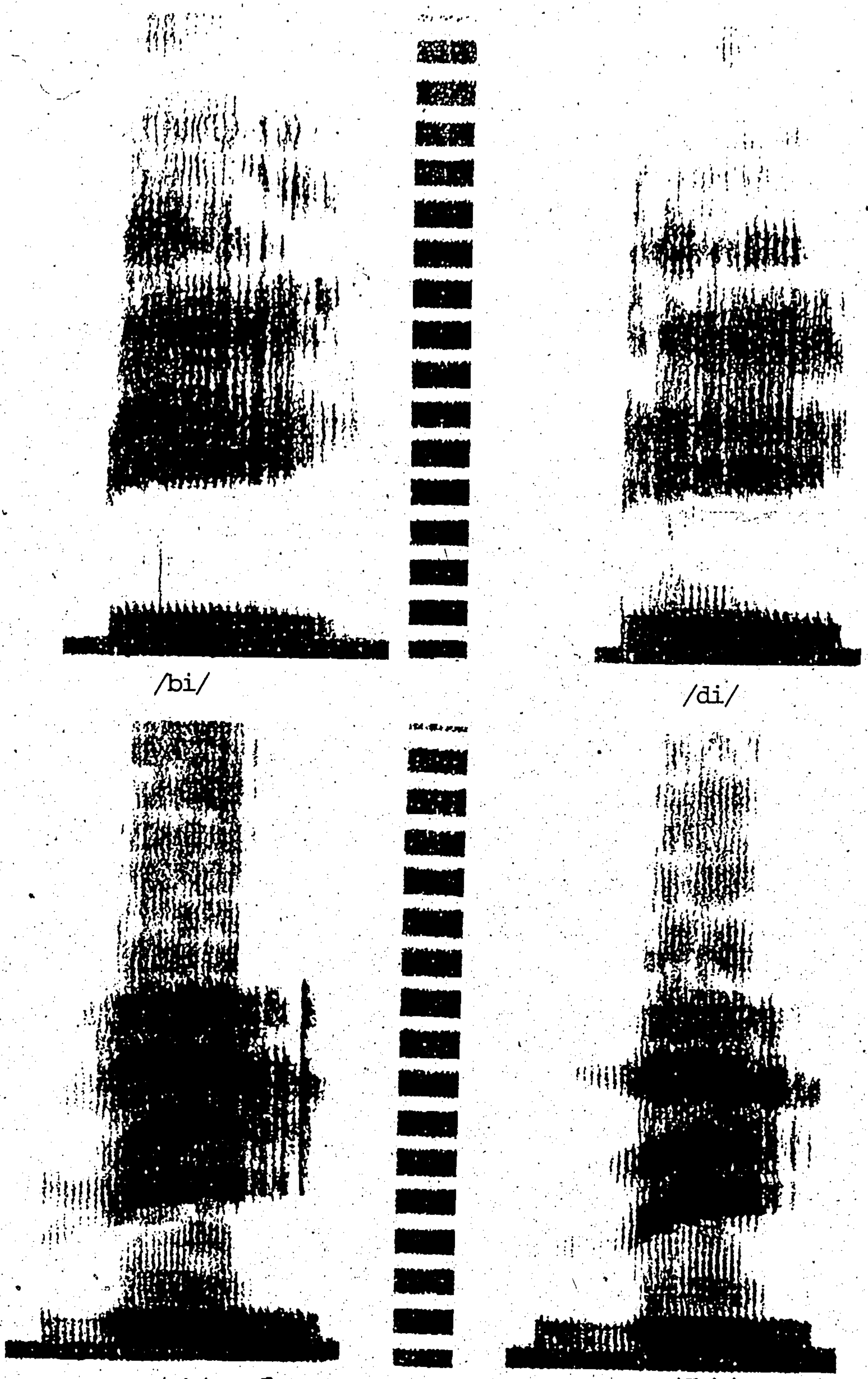
III. Experimental Method

Experiment I involved manipulating stop consonants to have them perceived as fricatives. Experiment II involved altering fricative consonants to be perceived as stops. The consonants studied are /b/, /d/, /v/, and /ð/. A recording was made of two male Canadian English speakers saying words starting with these consonants in the sentence frames, "I like that . . ." and "I block . . . ." The frames produce speaking rates in line with conversational speech. One speaker recorded the stop-initial words for Experiment I. The other speaker recorded the fricative-initial words for Experiment II. These two speakers produced consonants which were typical in duration and shape of onset. The two speakers recorded were chosen for their stable, clear enunciation and intonation of the words involved. The speakers recorded each stop and fricative initial word in the sentence frames a number of times. From these recordings one token was chosen as the original syllable for the creation of the experimental stimuli.

The words recorded were chosen so that the consonants occurred before a high front vowel, /i/, (high F2); a low vowel, /æ/, (mid F2); and a high back vowel, /o/, (low F2). The vowel /o/ was chosen instead of /u/ because of wide speaker variation in pronunciation of /u/ in this area of North America, and to avoid the dialect which pronounces /u/ as /yu/. The attempt to study only real English words, rather than nonsense syllables, also favoured the choice of the vowel /o/ over /u/ because of the greater variety of monosyllabic words with the vowel /o/.

The words chosen for this study are partners in minimal pairs - one word in the minimal pair begins with a stop, and the other word begins with a fricative. These words are "bee", "vee" (as in the letter V); "bat", "vat"; "boat", "vote"; "dee" (as in the letter D), "thee"; "Dan", "than"; "dough", and "though." Figures 3.1, 3.2, and 3.3 show spectrograms of these words. The spectrograms are grouped by vowel to facilitate comparison of the four consonants before the same vowel. The comparison shows the formant transitions from the English labial stop and fricative and dental stop and fricative into the vowels /i/, /æ/, and /o/ are similar in frequency and direction of change, indicating that CV formant transitions do not distinguish homorganic stops and fricatives. The main difference between the stops and fricatives appears to be the burst onset of the stops and the periodic noise of the fricatives.

It is important to use real words when the consonant /ð/ is being studied because of the limited occurrence of this consonant in the language. It is also especially difficult to find minimal pairs involving the consonant /ð/. In some instances, such as "Dan" and "than", the available real words are suspect in their status as a word. "Than" is problematic, as are all English words starting with /ð/, because it is a function word, not a content word. Function words are pronounced quickly, usually with reduced vowels.



/bi/ /di/

/vi/ /dhi/

Figure 3.1 Spectrograms of /bi/, /di/, /vi/, and /dhi/.

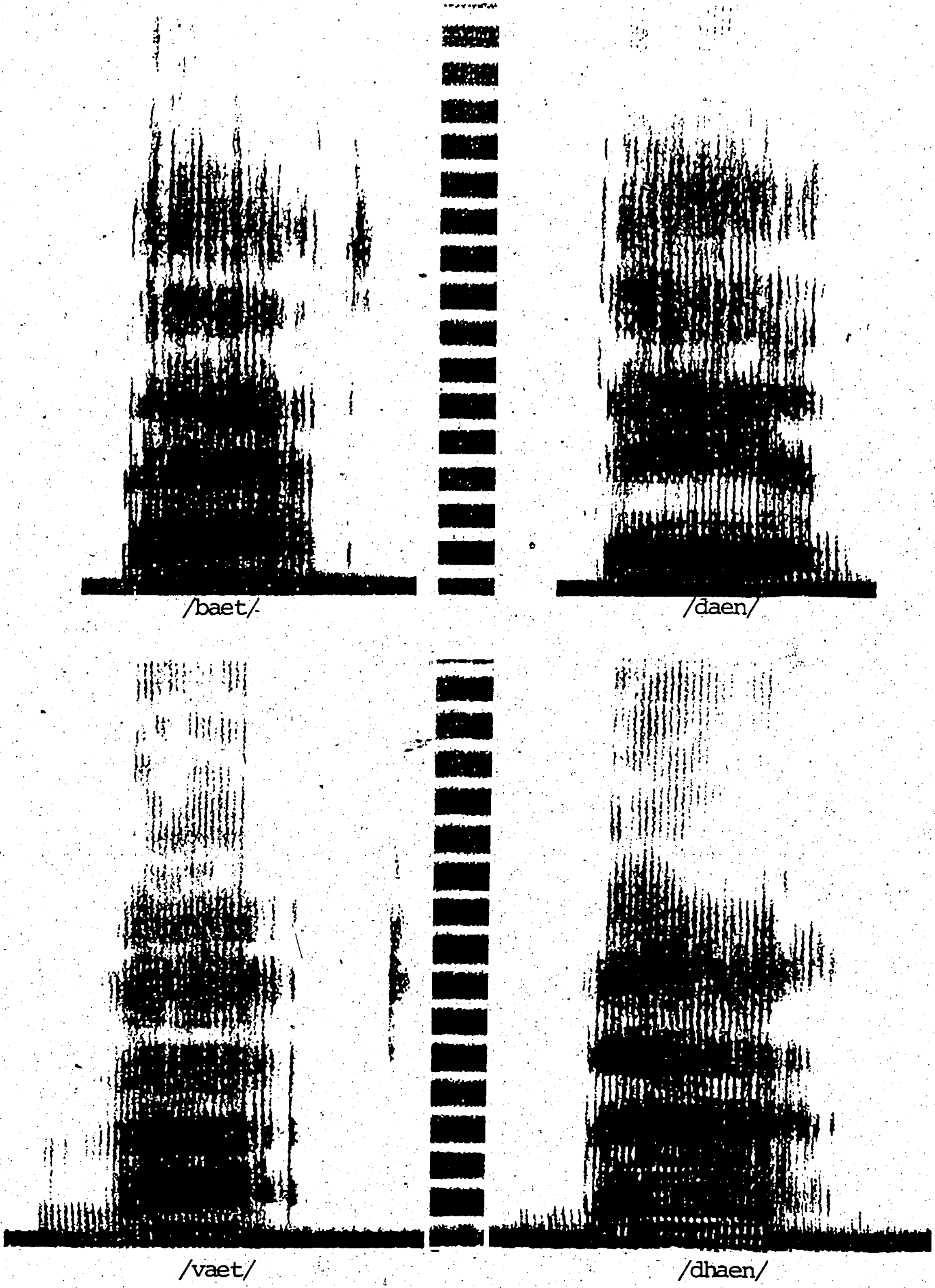
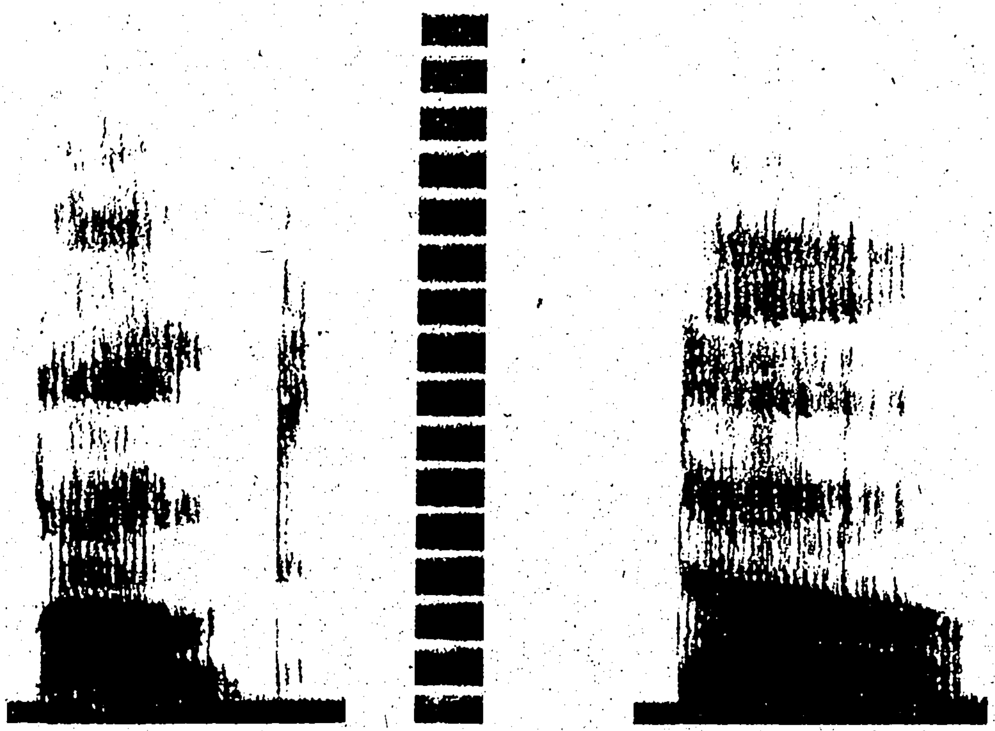
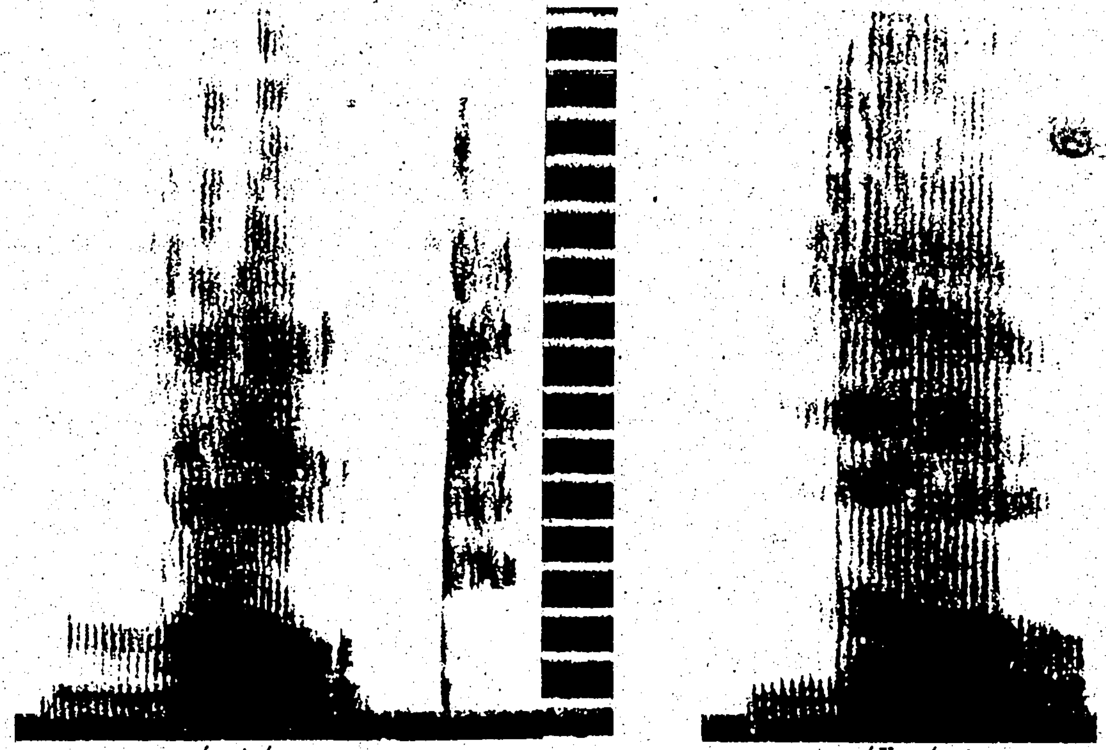


Figure 3.2 Spectrograms of /baet/, /daen/, /vaet/, and /dhaen/.



/bot/ /do/



/vot/ /dho/

Figure 3.3 Spectrograms of /bot/, /do/, /vot/, and /dho/.

"Dan" raises difficulties because the vowel pronunciation varies depending on the speaker. In some instances in running speech, and consistently for some speakers, the vowels in "Dan" and "than" are not the same. Articulating the stimulus words in sentence frames helped to alleviate these problems by making all the words have the same position and function in a sentence. A word at the end of either sentence frame could be regarded as a noun, given an appropriate semantic context. While placing the stimulus words in a sentence frame may gloss over differences in linguistic roles, it is desirable in a perceptual study because it allows us to concentrate on the acoustic differences between the speech sounds.

The recorded sentences were band pass filtered from 68 Hz to 6800 Hz to eliminate low-frequency hum and prevent aliasing. Next the sentences passed through an analog-digital converter where they were sampled at a rate of 16 kHz. The particular word to be modified was extracted from the carrier phrase and stored in a file in a DEC PDP-12 minicomputer. Up to this point the procedure for developing the stimuli for Experiments I and II was identical.

A. Experiment I

Stimuli

The stimuli for Experiment I varied in duration, abruptness of onset, and amplitude of the word-initial stop.

In brief, duration was altered by repeating the word-initial consonant. Abruptness of onset was modified by windowing the original syllable to remove the burst of the stop and by windowing to create a very gradual onset. Amplitude was varied by multiplying the consonant portion of the syllable by two, effectively doubling the amplitude of the consonant. In constructing the experimental stimuli care was taken to create tokens with oscillograms which resembled natural speech and which sounded like natural speech.

The durations of the word-initial stops to be studied in Experiment I were found to vary greatly. The duration of the word-initial consonant burst was approximately 10 ms for the words "bat", "Dan", "dee", and "dough." The corresponding duration for "bee" and "boat" was, however, only 6 ms and 4 ms, respectively. The noise of these two consonants was extended to enable all the stimuli to be comparable and modified similarly. Extending the consonant noise was also necessary for the changes in duration; repetition of short /b/'s created signals which did not sound like speech. To create a 10 ms long /b/ from the 6 ms one of /bi/, the last 4 ms of the 6 ms burst was attached at the end of the 6 ms. To create a 10 ms /b/ from the 4 ms one of /bot/, the last 3 ms of the original was repeated and attached to the original. These extended-consonant syllables were not used as "original" words. The 6 and 4 ms burst consonants were present in the experiment as the original consonants. But in situations in which the consonants were modified in any way, the extended-burst consonants were used.

One modification of the stop consonants for this experiment was changes in abruptness of onset. This involved two separate alterations - one removed the burst of the stop, the other windowed the stop by a ramp to create a gradual onset. Removing the burst was done by multiplying the original signal by a 3 ms long cosine squared window. By doing this, "burstless" consonants were created. The modification used to produce stimuli with "gradual" onsets was one of multiplying the original stimuli by a linear ramp window rising from zero to one in 10 ms. This windowing had the effect of removing the burst of the consonant and decreasing the amplitude of the signal immediately after the burst. With these two onset-altering modifications, two distinct sets of stimuli were produced from the stops: syllables with initial burstless consonants, and syllables with gradual onset consonants.

Figure 3.4 shows all the windows applied to the Experiment I stimuli. The cosine-squared window used to remove stop bursts appears first in the figure. Next is the 10 ms long linear ramp used to create gradual onset consonants. Below that appear linear ramps 20, 30, 40, 50, and 60 ms long. The latter are applied to the increased duration stimuli of Experiment I as described below. Oscillograms of the stimuli discussed thus far can be found in Figures 3.5 to 3.9. Figures 3.5 and 3.6 show the entire,

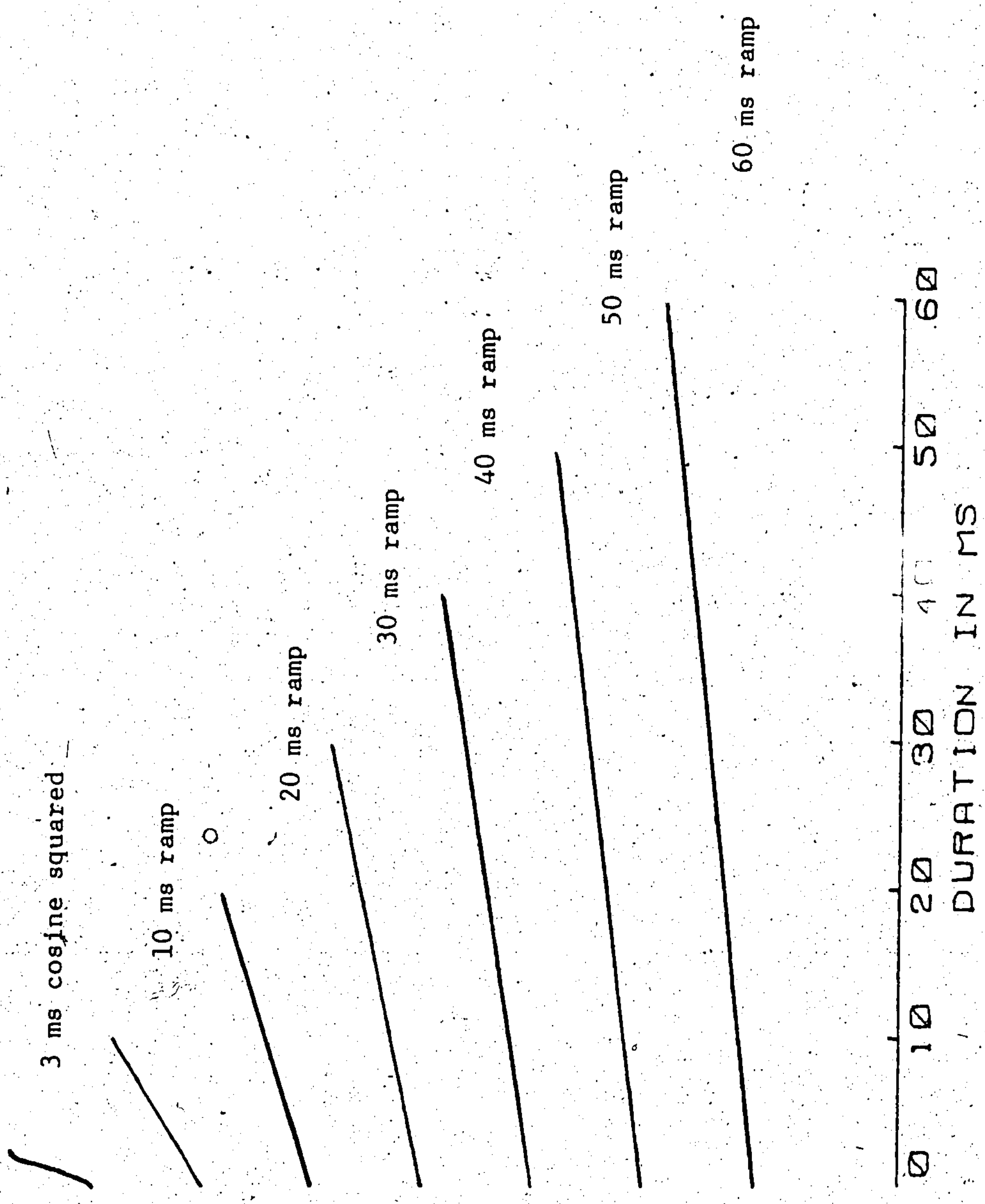


Figure 3.4 Oscillograms of windows which modify Experiment I words.

unmodified words used in this experiment. Figure 3.7 contains oscillograms of the first 20 ms of the words, showing the original unaltered stops in detail. Figure 3.8 shows the six stops after windowing by the 3 ms long window that removes the burst of the consonants. Oscillograms of the consonants after being windowed by the 10 ms long ramp are displayed in Figure 3.9.

The variation in duration of the word-initial consonant was created by repeating the original consonant noise from two to six times. If the abrupt onset were left on the consonant that was repeated, the result was not recognizable as a speech sound because of the periodicity created by the repetition of the burst. To ameliorate this condition, the repeated sections had their bursts removed by application of a 3 ms cosine squared window. The duration repetition was done by concatenating the burstless consonant to itself, followed by the vowel. The concatenation generated consonants two, three, four, five, and six times the original consonant length. Because of a buzz still heard due to the periodic fluctuation in the noise amplitude of the consonant, these tokens were further altered by multiplying them by a linear rise ramp of approximately the same duration as the consonant. A 20 ms ramp was applied to the stimuli created by doubling the consonant length; a 30 ms ramp was applied to the stimuli that had consonants three times the duration of the originals; and so on. The ramped onset stimuli were more natural in structure, as well

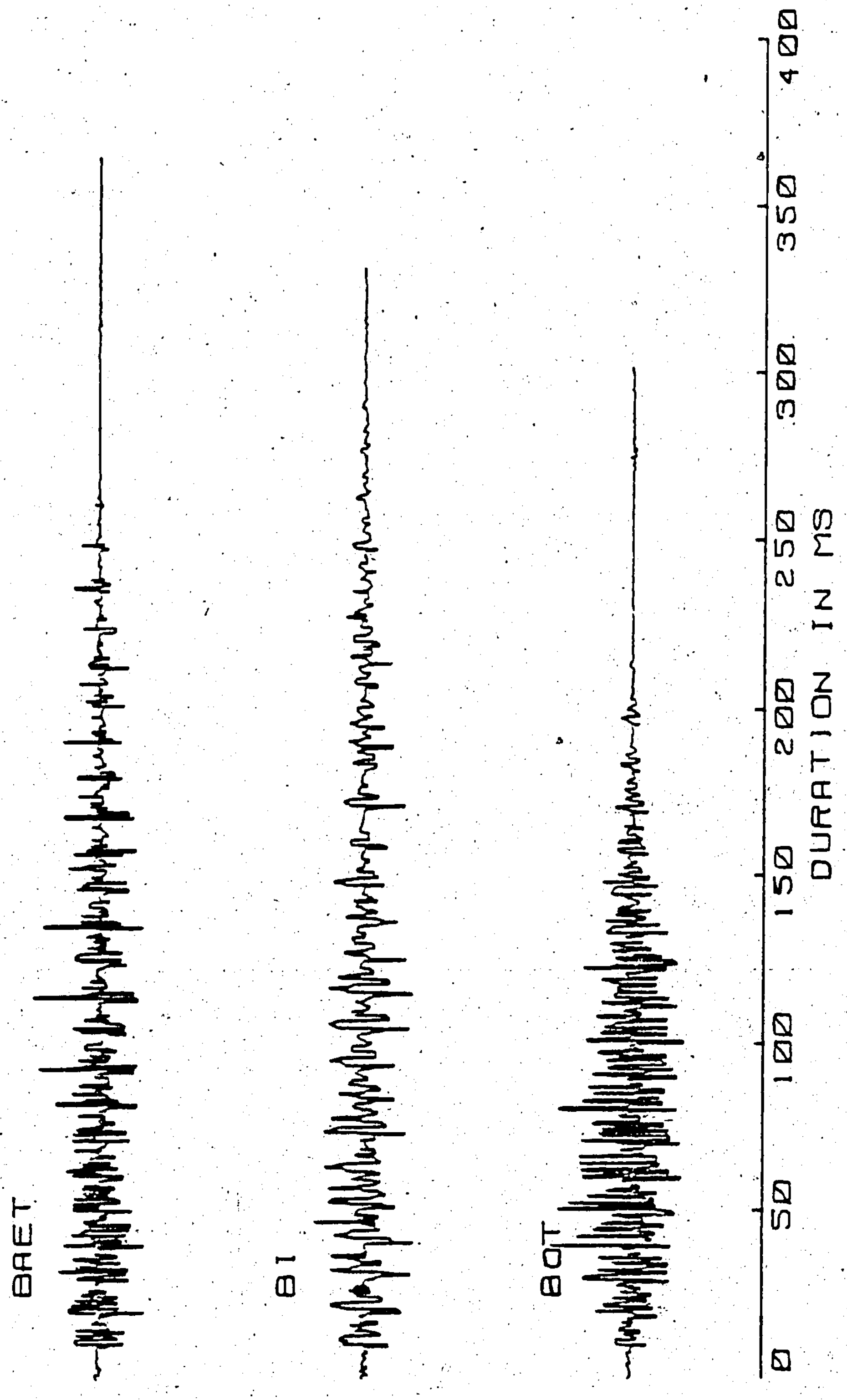


Figure 3.5 Oscillograms of unmodified words beginning with labial stops.

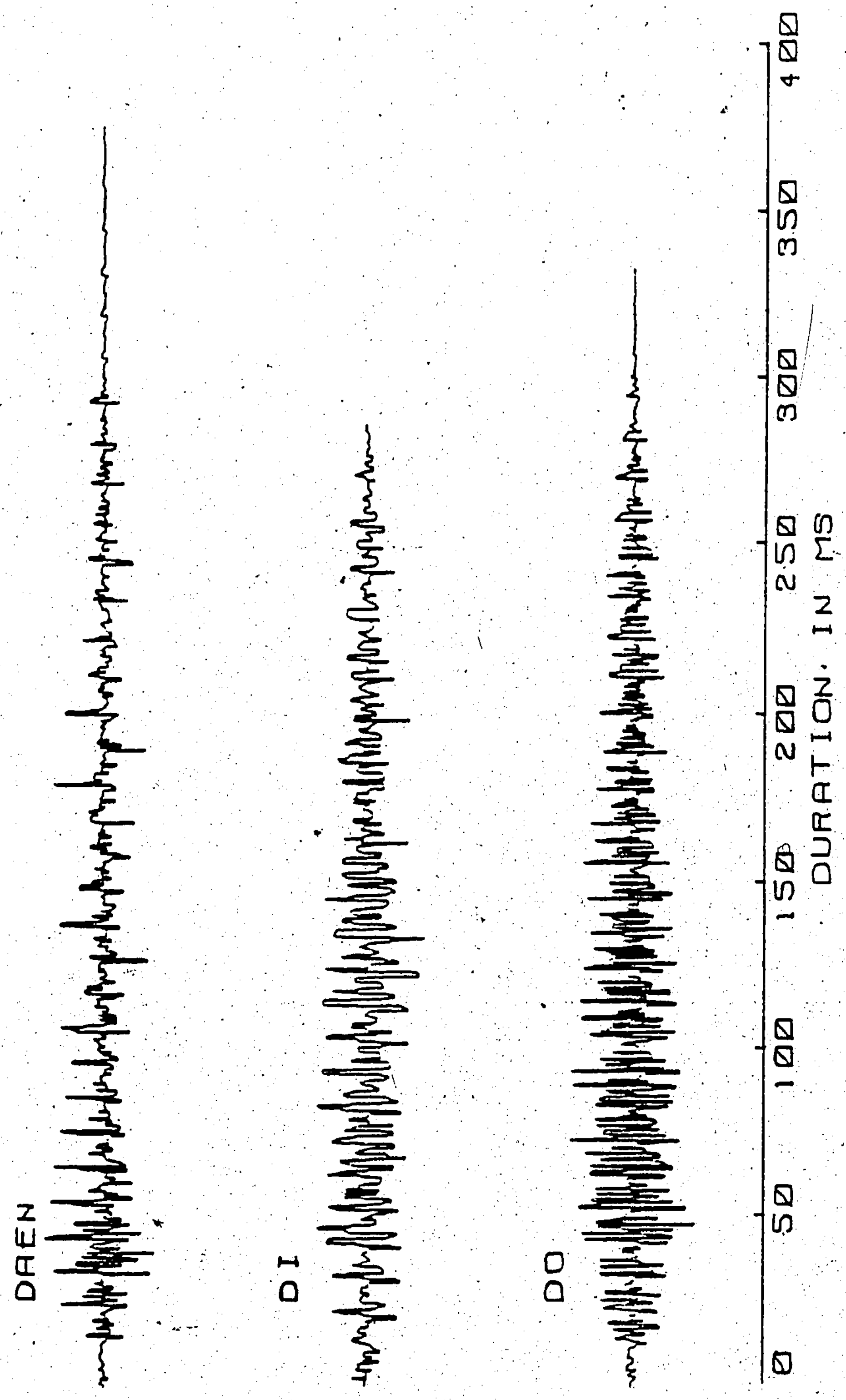


Figure 3.6 Oscillograms of unmodified words beginning with dental stops.

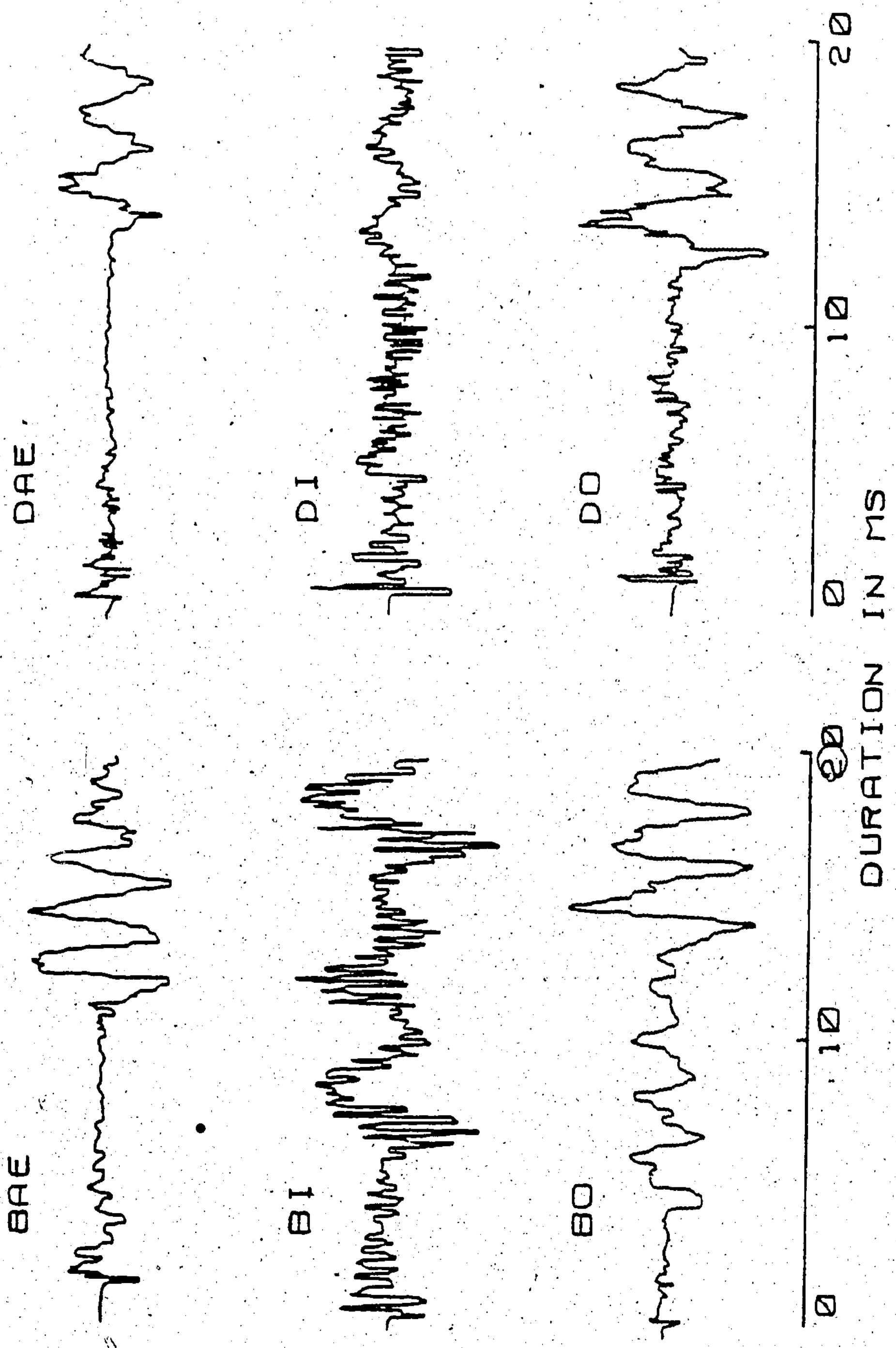


Figure 3.7 Oscillograms of unmodified stop consonants.

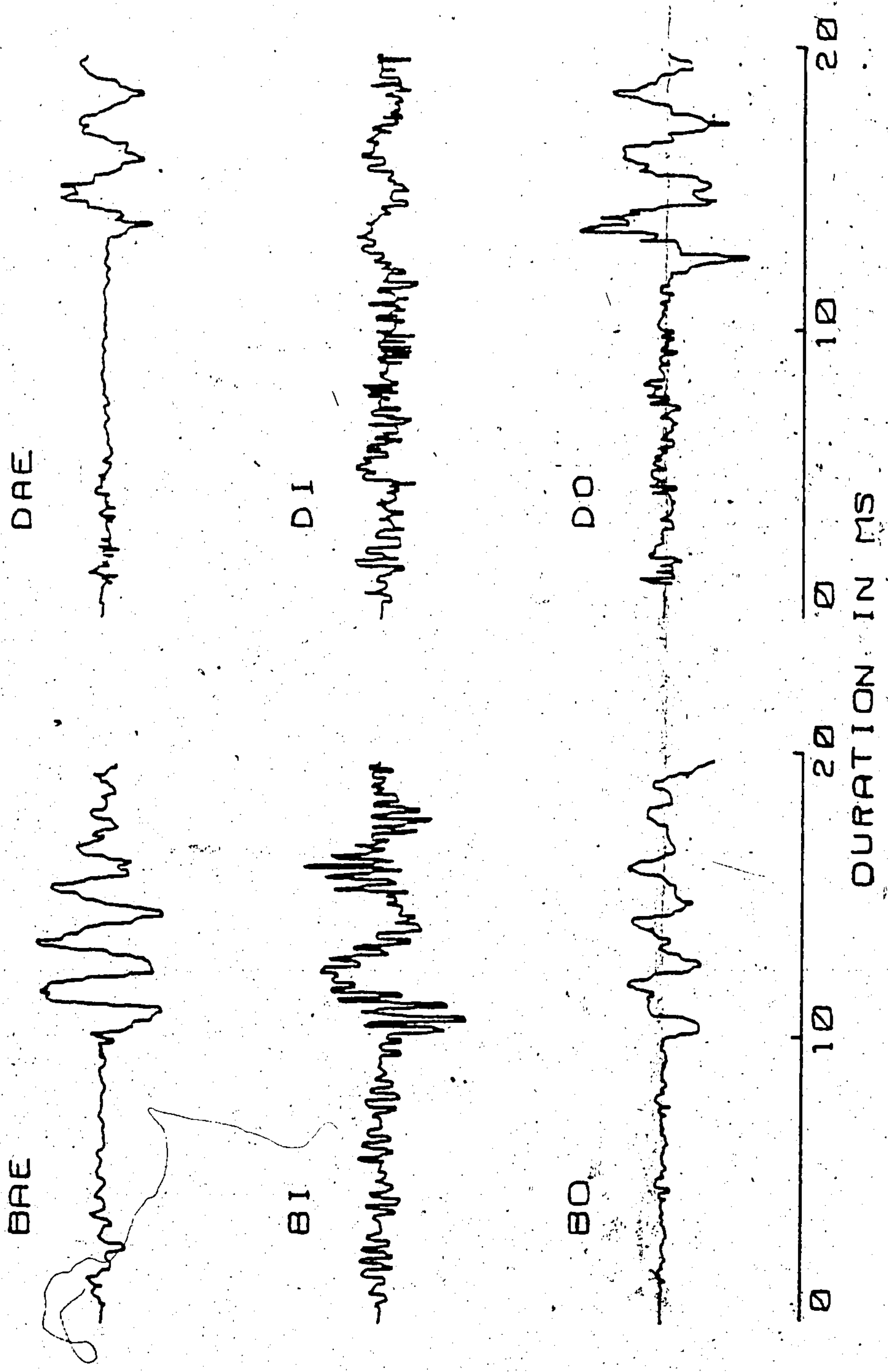


Figure 3.8 Oscillograms of burstless stop consonants.

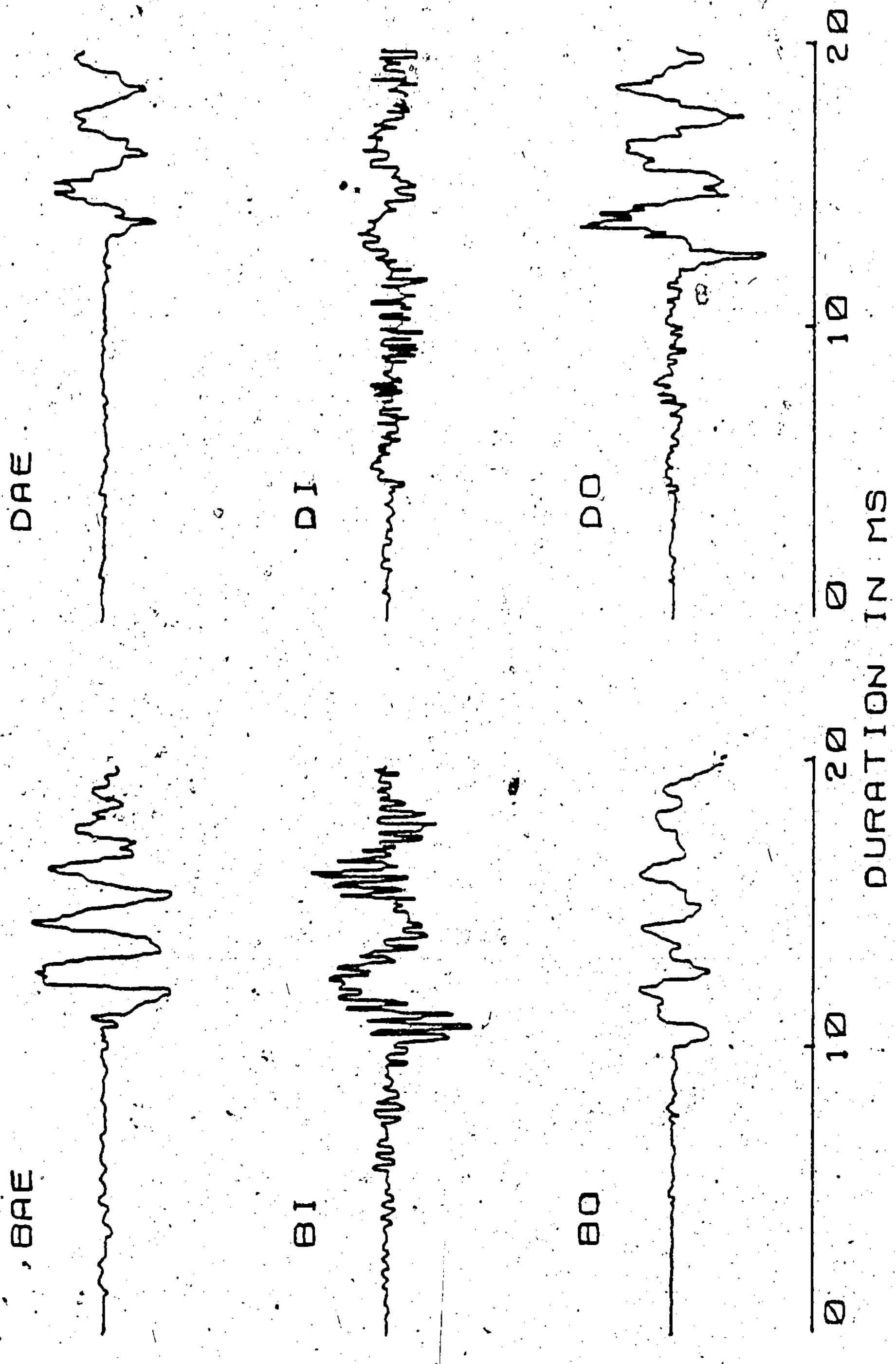


Figure 3.9 Oscillograms of gradual onset stop consonants.

as sound, than the unramped repeated consonants. Structurally the stimuli resemble fricatives, which have gradual increases in the noise amplitude from consonant onset to vowel onset. The concatenations resulted in consonants with duration ranging from approximately 10 ms long to approximately 60 ms long in 10 ms steps. To illustrate the effect of repeating the consonant to increase the duration of the consonant and of applying the ramps to these consonants, Figure 3.10 contains the /di/ stimuli of all six durations.

The three modifications discussed so far, removal of burst, application of ramp window, and change in duration of consonant, resulted in 48 stimuli: 6 originals (2 consonants with 3 vowels), 6 without a burst, 6 with gradual onsets (ramped), and 30 with various consonant durations. Each of these tokens was used as the base for the amplitude factor. To vary the amplitude, the consonant section of a stimulus was doubled in amplitude by multiplying each point by two. This resulted in 48 stimuli with increased-amplitude consonants. Experiment I involved 96 stimuli in total. These tokens embody variations in presence or absence of stop burst, gradual or abrupt onset, duration of initial consonant, and amplitude of the initial consonant.

Figure 3.11 provides more sample oscillograms to illustrate the construction of Experiment I stimuli. Again, these traces are based on the syllable "dee." Included in this figure are oscillograms of the first 40 ms of the

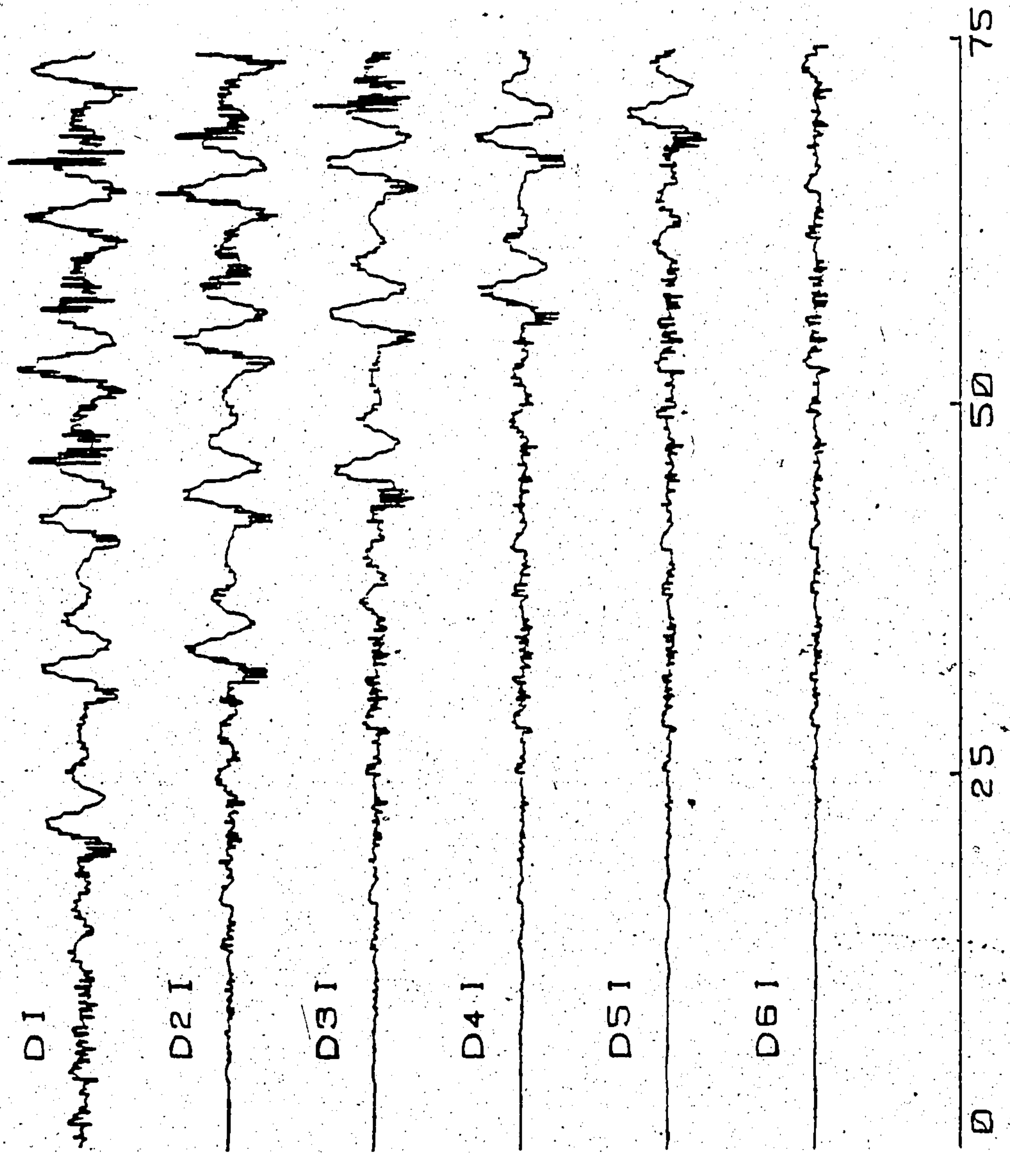


Figure 3.10 Sample oscillograms of Experiment I stimuli that vary in duration.

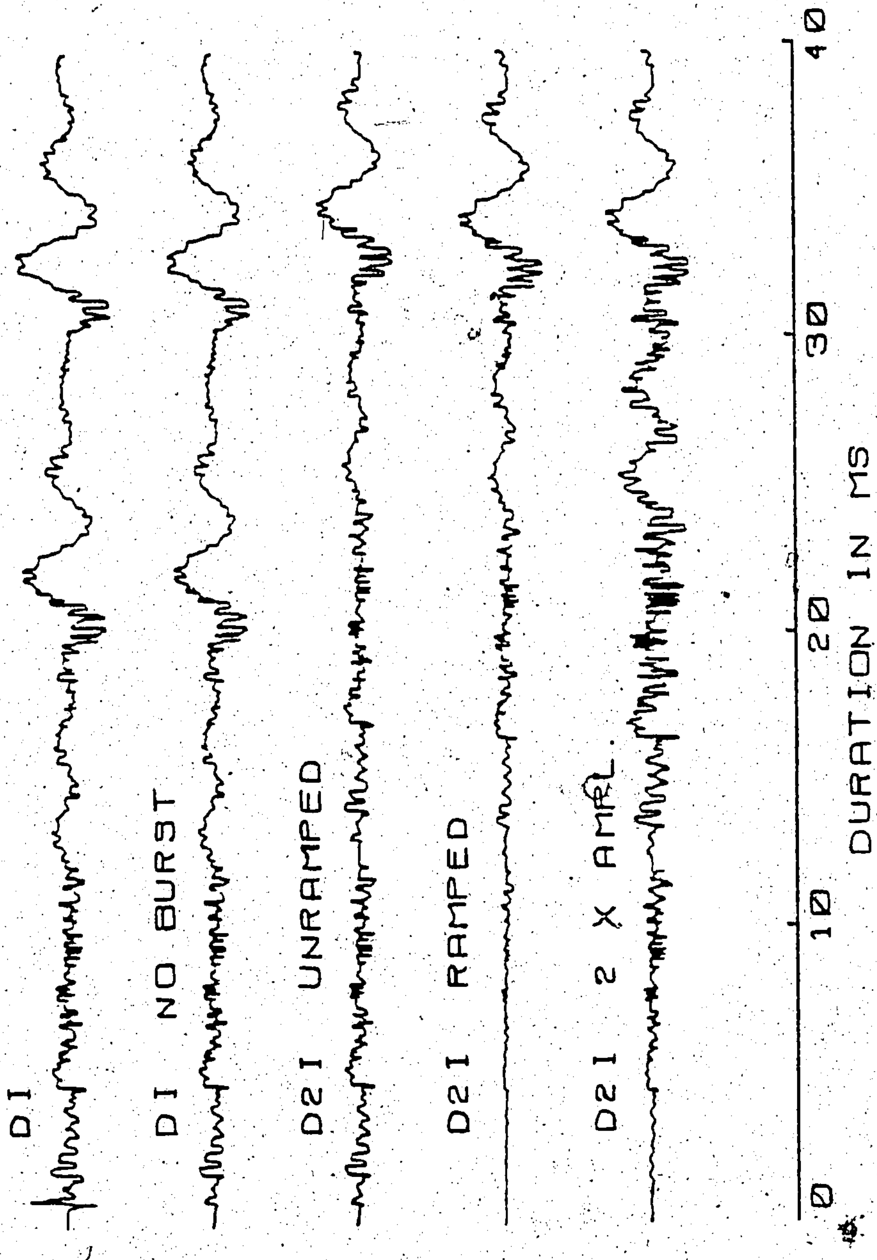


Figure 3.11 Oscillograms depicting the construction of Experiment I stimuli.

original /di/ syllable, this syllable with the burst removed, the first increase in duration due to repetition of the burstless consonant, the increased-duration syllable ramped, and finally, the ramped trace modified by an increase in the amplitude of the consonant.

Presentation

The methods of presentation used in the two experiments were different. For Experiment I, the stimuli were randomized and three replications of different randomizations were recorded onto tape. The stimulus words were placed at the end of the carrier phrase "Please say the word . . . ." Ten practice sentences preceded the 288 (96 stimuli x 3 replications) items and two filler sentences followed to round the number of items to 300. A 3 ms pause was inserted after every tenth sentence. The entire tape was 20 minutes long. It was played back on a TEAC A-7030 GSL stereo tape deck over Telephonics TDH-49 headphones in a sound treated room. A sine wave which covered the minicomputer's range of amplitude output was used to calibrate the recording level so that the stimuli would not be clipped. The same sine wave was used to set the the playback level on the TEAC tape deck.

Subjects were given an answer sheet which included instructions, a confidence rating scale, and numbered choices of two words corresponding to the stimuli on the tape. (A copy of this answer sheet is included in Appendix

B.) Subjects were asked to choose which word they heard at the end of the sentence, "Please say the word . . . ." They were given a choice of a word starting with either a stop or a fricative. The vowel and optional consonant at the end of the word corresponded to the word heard on the tape at that item number. A word beginning with a stop and a word beginning with a fricative appeared in alternate columns for each item. The alternation was designed to remove any bias towards words in either column. Subjects were also asked to indicate their confidence in their response on a scale of one to three, three representing high confidence and one representing low confidence. They indicated these confidence ratings by circling the appropriate number on the response sheet.

This design prevented subjects from confusing the place of articulation of the word-initial consonant. For instance, if a subject was presented with the word "thee", he or she had as a response choice "dee" or "thee." The subject could not respond "vee", even if the item was perceived as starting with a /v/. (Two of subjects did, however, indicate when they heard a /ð/ when given only the choice of /b/ or /v/. These responses will be discussed briefly in Chapters IV and V.)

Every subject listened to the tape at least four times. Subjects 1 and 5 listened six times. This gave twelve replications of each stimulus syllable for Subjects 2, 3, and 4, and eighteen replications for Subjects 1 and 5.

Subjects

A small number of subjects were used, but each subject heard each stimulus many times. The subjects for both experiments were English speakers from various regions of Canada and the United States who are currently residing in Edmonton, Alberta. The dialect of English the subject spoke was not expected to affect his or her perception of English consonants. The subjects reported having no known hearing impairments.

Five people participated in Experiment I, four males and one female. All of the subjects were university students or staff, and all had some knowledge of phonetics. Subjects 1 and 3 were familiar with the design of the stimuli.

B. Experiment II

Stimuli

Experiment II was designed to alter word-initial fricatives to determine which factors would cause a fricative to be perceived as a stop. Modifications were made to the duration and amplitude of the fricative. In this experiment the amplitude factor had three levels. Since the original fricative waveforms were of such low amplitude, it was possible to multiply each consonant by two twice. The first level was normal amplitude; the second level was doubled amplitude; and the third level was quadrupled amplitude.

Each amplitude increase added 6 dB to the intensity level of the frication.

The second factor in Experiment II was duration. The duration of the fricative was decreased to resemble a stop by gating out middle sections of the fricative frication. Sections were cut out leaving three, two, and one period of the original voiced fricative. The periods that were kept for the three-period stimuli were the first two and the last period (before the vowel) of the original fricative. The two-period long stimuli were made of the first and last period; the one-period ones consisted of only the last period of frication before the vowel transition. The original fricatives varied in duration, and because the fundamental frequency of the recorded words vary slightly, the durations of the shortened fricatives vary slightly around 30, 20, and 10 ms.

By creating three new durations of consonant from each fricative, and by creating two new consonant amplitude sizes, 72 items were prepared for the second experiment. These items consisted of two consonants (/v/ and /ð/) with three vowels (/i/, /æ/, and /o/), four durations (original, three-periods of frication, two-periods, and one period), and three amplitudes (original, double, and quadruple consonant amplitude). Figures 3.12 and 3.13 show oscillograms of the entire syllables used in Experiment II. Figures 3.14 and 3.15 contain oscillograms of the consonant portion of the original Experiment II syllables. Figure 3.16 is a

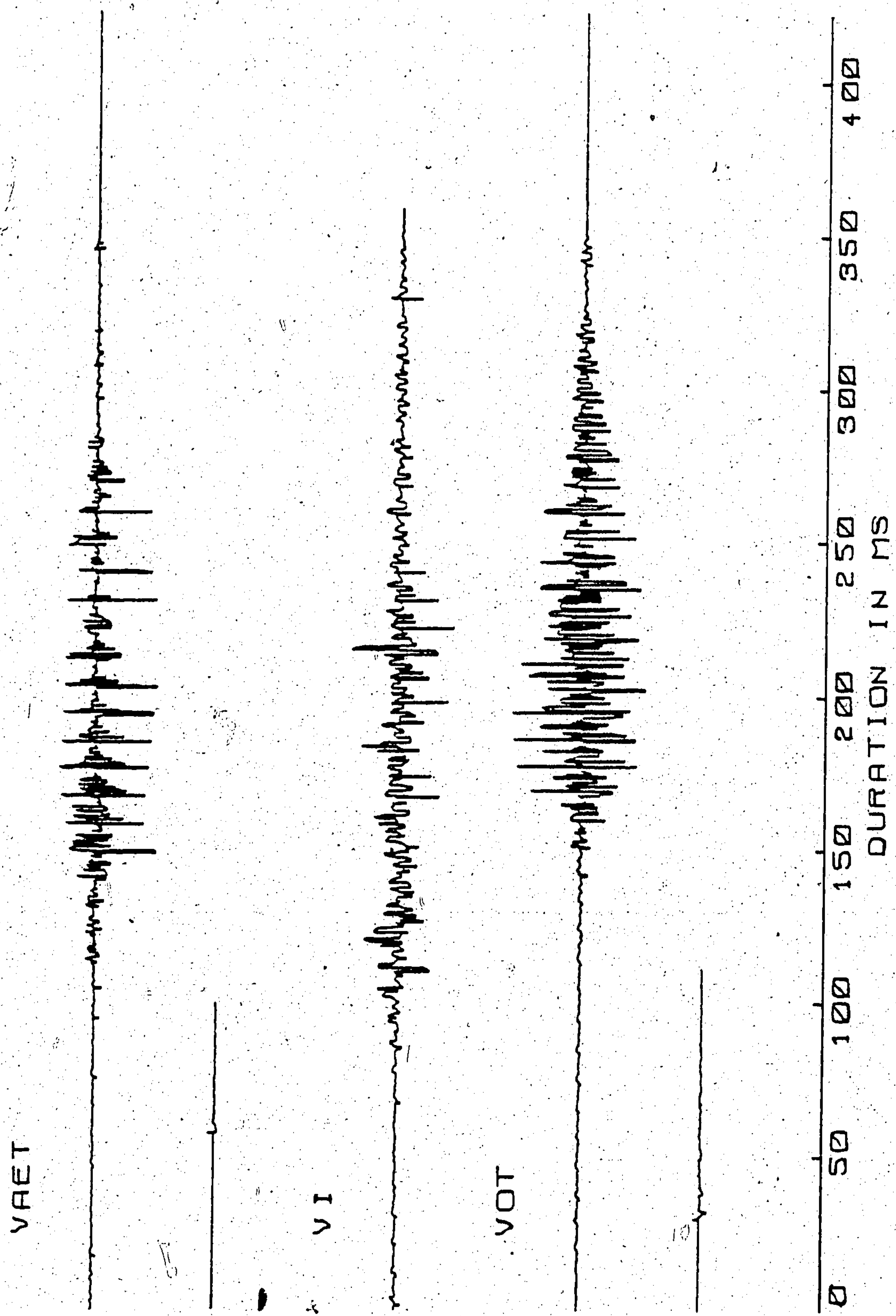


Figure 3.12 Oscillograms of Experiment II words starting with /v/.

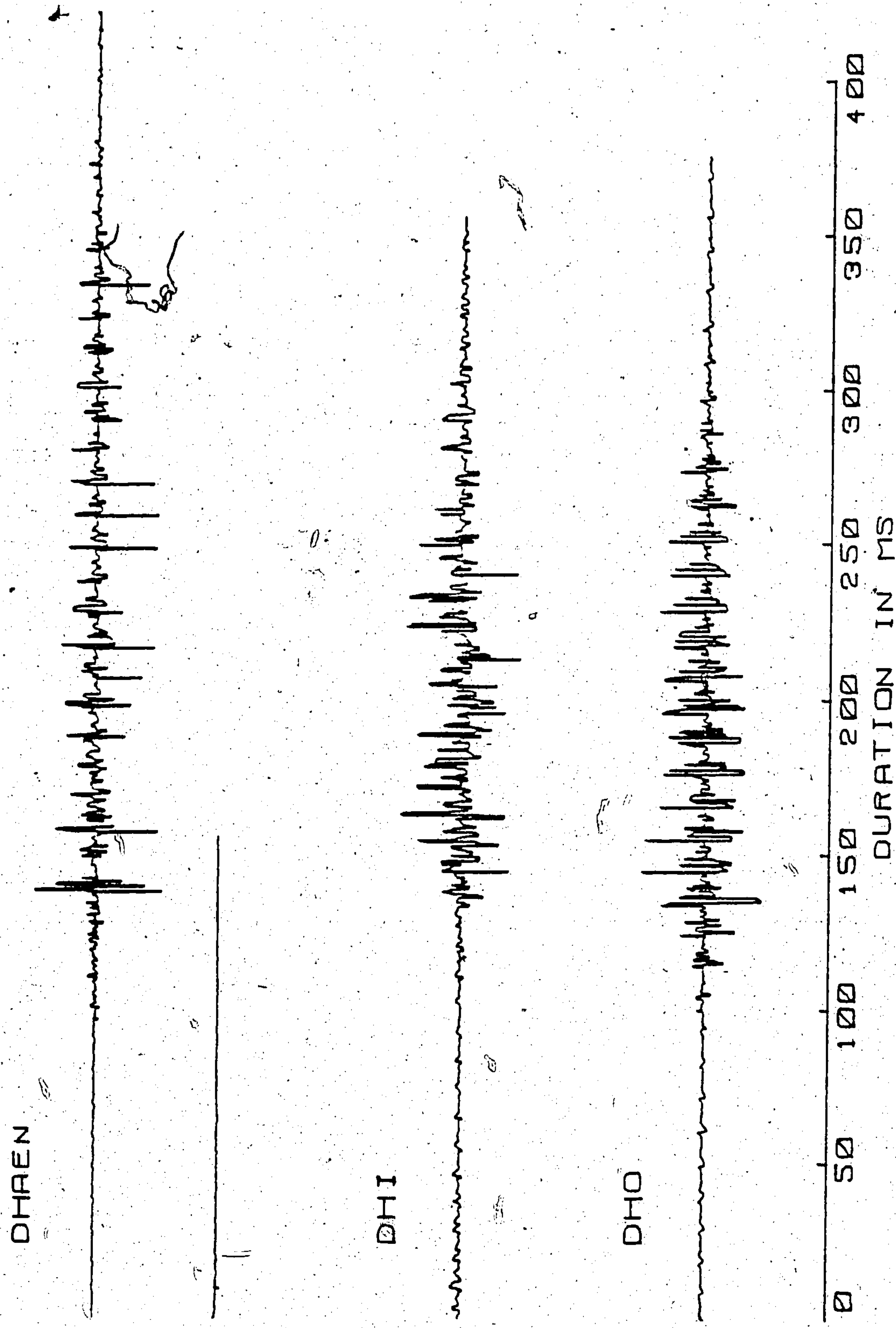


Figure 3.13 Oscillograms of Experiment II words starting with /dh/.

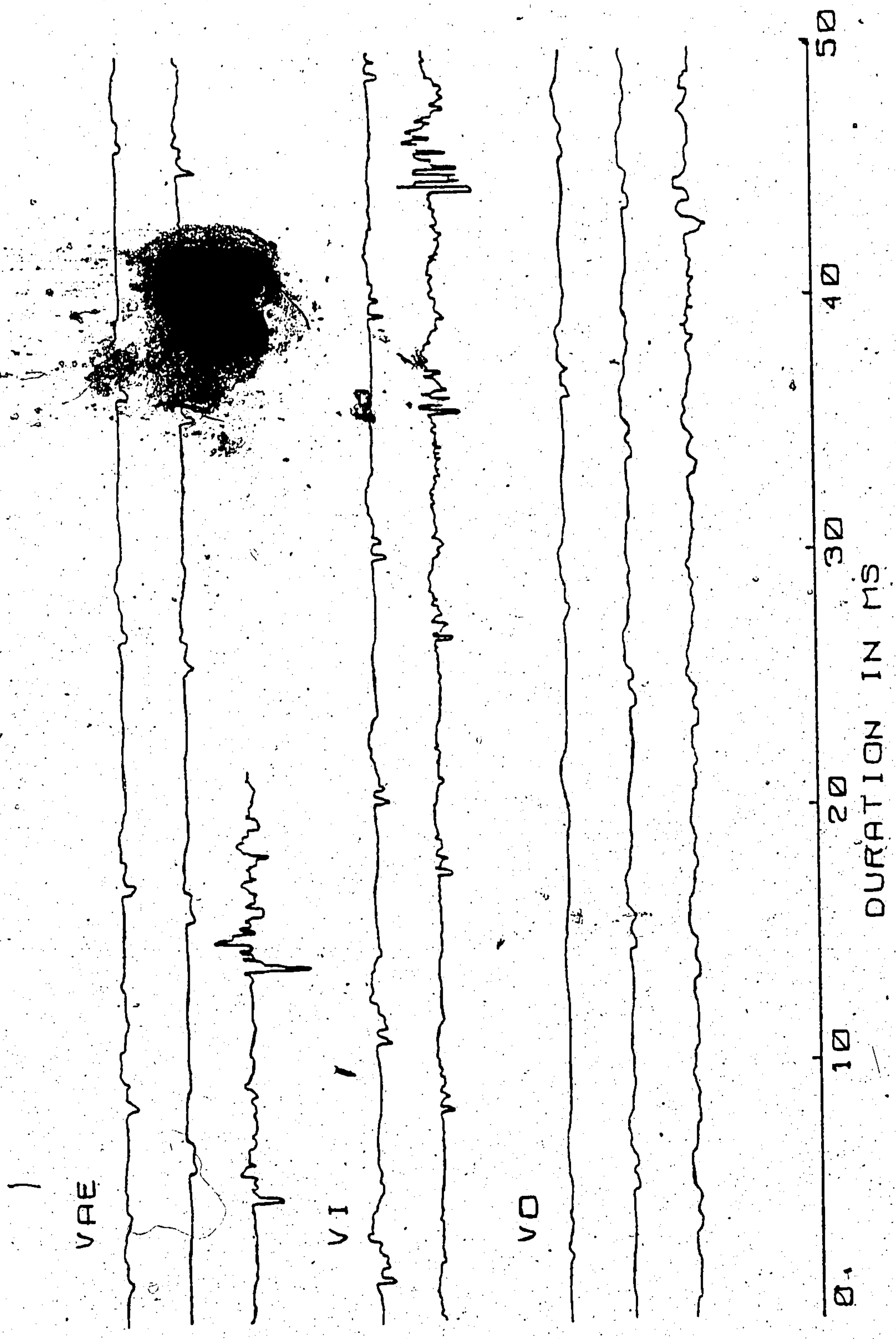


Figure 3.14 Oscillograms of the word-initial /v/.

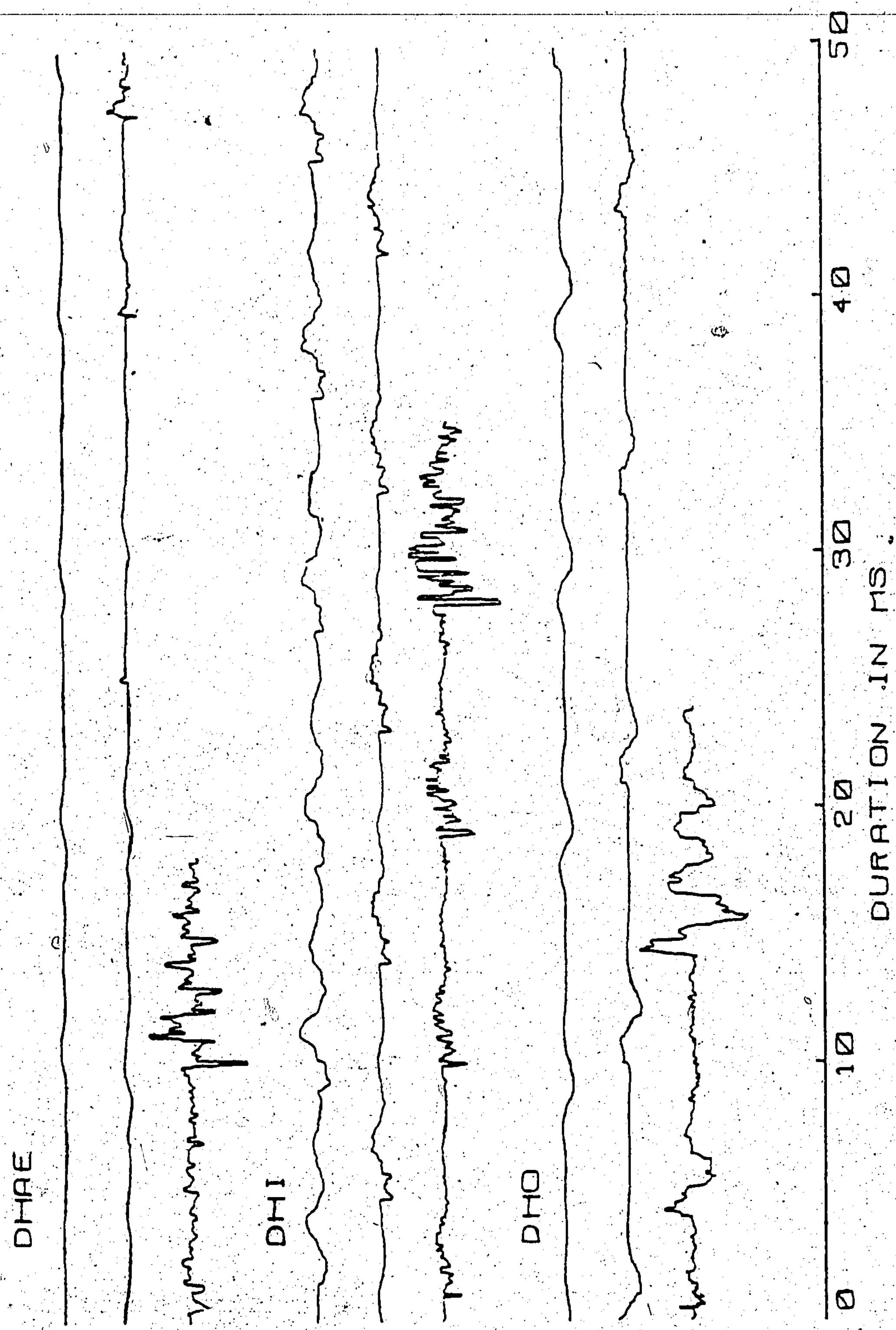


Figure 3.15 Oscillograms of the word-initial /dh/.

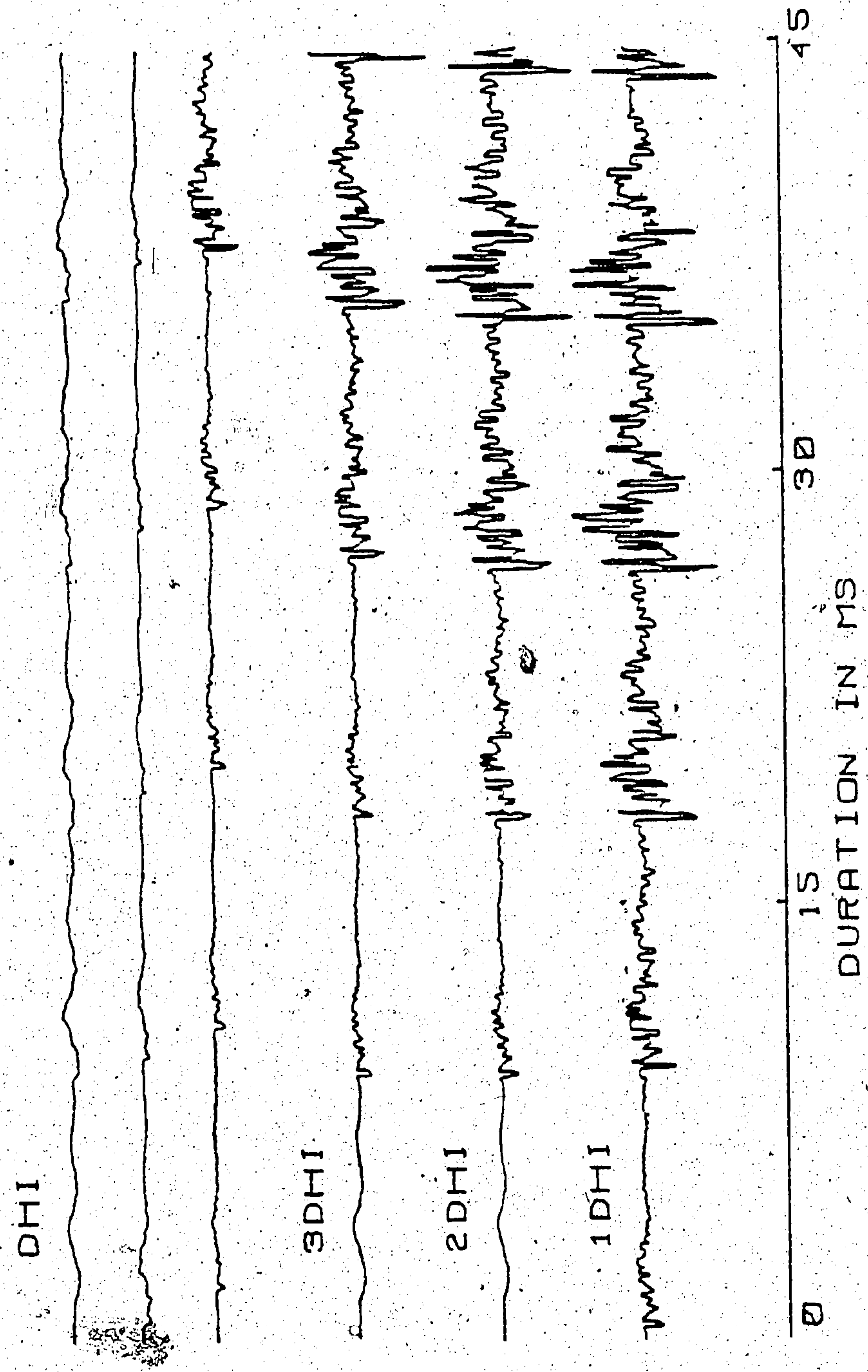


Figure 3.16 Oscillograms of a sample of the stimuli that vary in duration in Experiment II.

sample of the durational stimuli used in Experiment II. The last figure, 3.16, shows the entire original consonant from the syllable "thee", the stimulus constructed from three periods of this syllable, the two-period stimulus, and the one-period stimulus from the syllable "thee."

Presentation

The method of presentation of stimuli for Experiment II was quite different from that for Experiment I. After the presentation is described, the motivation for the differences will be discussed. The 72 Experiment II stimuli were randomized and presented to subjects in sets of three replications (216 items). These items were presented as isolated words over a Heco Sound-Master 15 loudspeaker at a comfortable listening level in a sound-treated room. The presentation was computer controlled, rather than recorded. The computer played the stimulus repeatedly until it received a response to that item. Then the computer played the next item on the list of randomized stimuli.

For each item, subjects were given the choice of responding /b/, /v/, /d/, or /ð/. Subjects responded on a four column by three row touchpad. Each column represented one of the four word-initial consonant choices. The order of the consonants on the touchpad was rearranged every trial. The rows on the touchpad represented the confidence rating the subject associated with his or her response. The top row signified a confidence of 3 in Experiment I, i.e., high confidence. The middle row indicated moderate confidence, and the last row meant low confidence. A subject responded by deciding which consonant the test word started with and how confident he or she was of the answer, and he or she pressed the appropriate key. Usually, subjects listened to a word two or three times before responding. The PDP 12 minicomputer stored each response in a separate file for each trial for each subject. Each subject performed the three-replication trial six times, giving 18 responses to each stimulus item for each subject. A trial took approximately 15 to 20 minutes to perform, depending on the subject's familiarity with the procedure and the stimuli.

The presentation of Experiment II stimuli was different from the presentation of Experiment I for a number of reasons. Mainly, Experiment II differed to allow subjects more choices in their responses, to allow them to respond with either a labial or dental place of articulation regardless of the place of articulation of the original consonant. Comments from subjects in Experiment I indicated perception errors in place of articulation. The change to responding with any of the four voiced labial and dental stops and fricatives would allow subjects to express these perceptions. In addition, indications of a dependency of place of articulation on manner of articulation, as proposed by Carden et al. (1980), could be examined from the Experiment II data.

The change from headphones to a speaker was made because some Experiment I subjects complained that the headphones were uncomfortable and because headphones are not usually involved in speech perception. The headphones had to fit tightly to reduce interference from outside noise, especially since subjects had only one chance to perceive each stimulus item. Usually sounds originate farther from the ear than headphone signals do. Perceiving audio signals through headphones may require special adjustment on the part of the listener (Gibson, 1966). For greater subject comfort and to increase the naturalness of the speech perception task, the second experiment stimuli were presented over a speaker.

The touchpad allowed for four choices of consonant responses, and the three confidence levels as in Experiment I, and automated data recording because of the connection to the minicomputer. As well, using a computer and a touchpad allowed subjects to control the pace of the presentation. Subjects were able to listen more critically to the stimuli. Response errors due to accidentally touching the wrong key on the pad occurred, but subjects could correct errors by touching another key before the first was recorded. When the computer received responses from two keys it ignored both responses and replayed the stimulus.

Subjects

Five native English speakers participated in Experiment II, three females and two males. The subjects were university students and faculty members. The subjects had no known hearing losses. Subject 5 had no knowledge of phonetics. Subjects 1 and 2 also participated in Experiment I, as Subjects 1 and 3.

IV. Results

A. Experiment I

Measurements

Duration

Measurements were made of the duration and intensity of the Experiment I stimuli. These measurements were made on the six original words, the original consonants, and all the altered consonant stimuli. (The alterations included variation of duration of onset, of duration, and of the amplitude of the consonant.)

Table 4.1 gives the durations of the original syllables, the original consonants, and the altered-duration consonants. The stimuli altered for intensity (amplitude) and duration of onset are not included in this table since these modifications did not change the stimulus duration. The first column in the table shows the duration of the original syllable; the second shows the duration of the vowel; the third gives the duration of the original consonant. The last column in Table 4.1 gives the duration of the stop consonant which was repeated to create the duration varying stimuli. Since the five increased duration stimuli were created by concatenating the ramped consonant, their durations are not included in the table; their durations are simply multiples of the consonant in the last column. The stop consonant measurements show that the burst of /b/ is,

Table 4.1 Durations of Experiment I stimuli.

	Duration in ms			
Syllable	Syllable	Vowel	Original C	Gradual Onset C
bee	330	324	6	10
bat	365	355	10	10
boat	300	296	4	10
dee	287	275	12	12
Dan	376	363	13	13
dough	334	322	12	12

on average, a little more than half the duration of the burst of /d/. The mean measurement of /b/ is 6.7 ms; the mean of /d/ is 12.3 ms.

Intensity

The PDP-12 minicomputer was used to take measurements of the root mean square area of the stimulus syllables and the consonant portions of the stimuli. These measurements were converted to decibels to represent the intensities of the signals. The decibel measurements are presented in Table 4.2. The measurements have been averaged over the three vowels. As well as showing the intensities of the consonants, the table gives the mean syllable intensity as an indication of the amplitude of the consonant relative to the vowel. The mean syllable intensity for both the original labial and dental stop words in Experiment I is 44 dB. The dental consonants have greater intensities than the labials. The dental stop /d/, with a mean intensity of 20 dB, is 4 dB higher in intensity than the average /b/ at 16 dB. The intensities of the stops drop 2 to 3 decibels when the burst is removed. And the intensity decreases again, by approximately 4 decibels, when the stop has a gradual 10 ms onset. The intensities increase as the durations of the consonant portions of the stimuli increase from 10 to 60 ms.

Table 4.2 Intensities of Experiment I stimuli (dB).

	Stimulus								
Consonant	Original	Burstless	10 ms	20 ms	30 ms	40 ms	50 ms	60 ms	Syllable Mean
/b/	16	14	9	11	13	14	15	16	44
/d/	20	17	14	16	18	19	20	21	44

Recognition Curves

The recognition curves of the consonants in Experiment I are generally consistent. In only a few instances did subjects reverse the trend of their responses as the stimuli varied along the continua. Mean results for all the subjects are discussed below; recognition curves for individual subjects are given in Appendix C. The individual subject curves are supplied to compare results across subjects.

Experiment I data are recorded as a proportion of stop (/b/ or /d/) responses at the various intensity, onset slopes, and duration levels. Figures 4.1 to 4.6 show the mean of all subject responses to the stimuli with various onsets. The three onsets are labelled "burst" - the original, unaltered consonant; "no burst" - the consonant was windowed by a short cosine squared wave to remove the abrupt onset; and "gradual" - the consonant was windowed by a 10 ms linear ramp. Each figure has two lines - one for each consonant intensity level, normal and double amplitude. Figures 4.1 to 4.3 depict the results for the consonant /b/; Figures 4.4 to 4.6 show the /d/ onsets. These illustrations, for the most part flat horizontal lines, show that onset slope alone does not greatly affect a consonant's recognition as a stop. Stops are still identified as stops when their bursts are removed. Even stops with linear 10 ms rises are perceived as stops more than 80 percent of the time. All subjects categorized the consonant as a stop 100

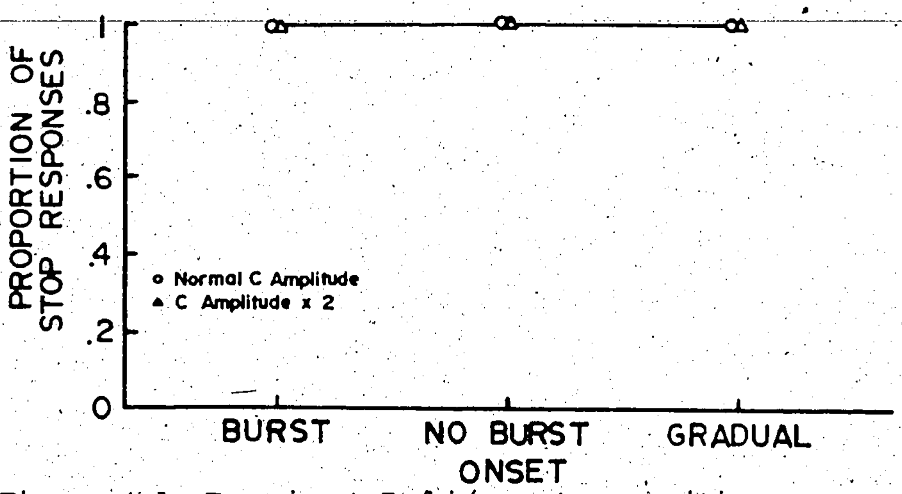


Figure 4.1 Experiment I /bi/ onset recognition curve, mean of all subjects.

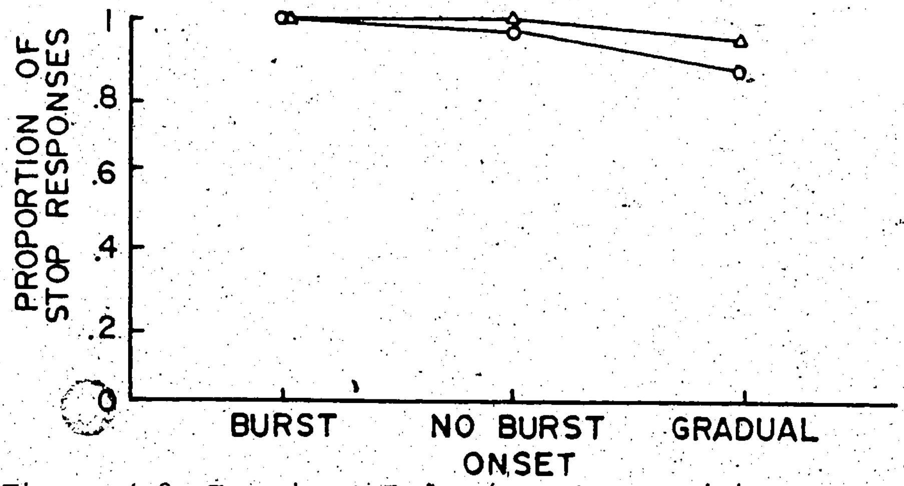


Figure 4.2 Experiment I /bae/ onset recognition curve, mean of all subjects.

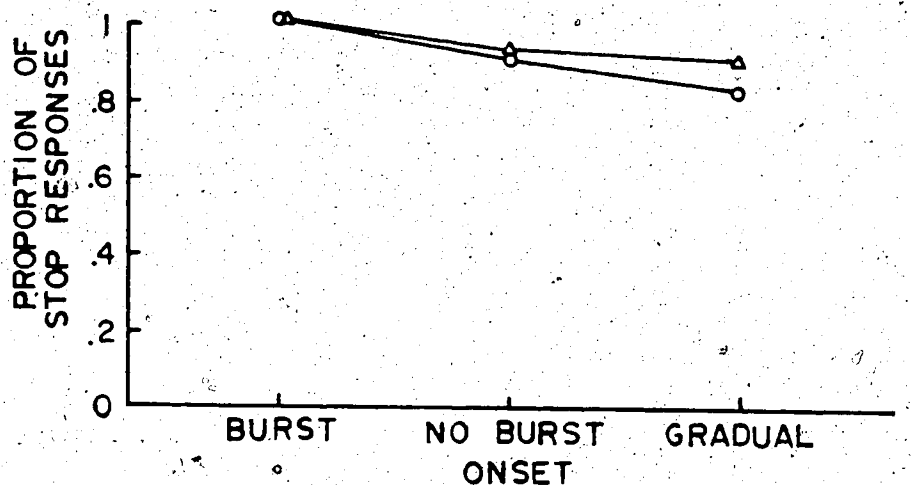


Figure 4.3 Experiment I /bɑ/ onset recognition curve, mean of all subjects.

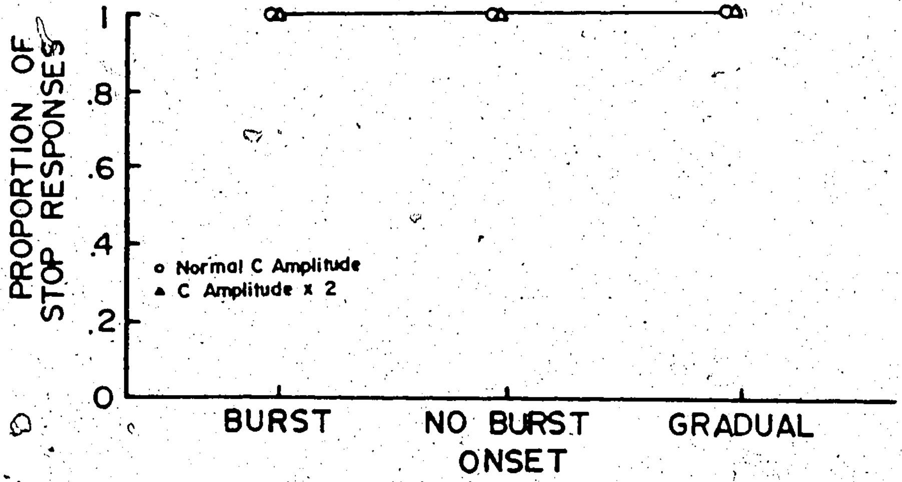


Figure 4.4. Experiment I /di/ onset recognition curve, mean of all subjects.

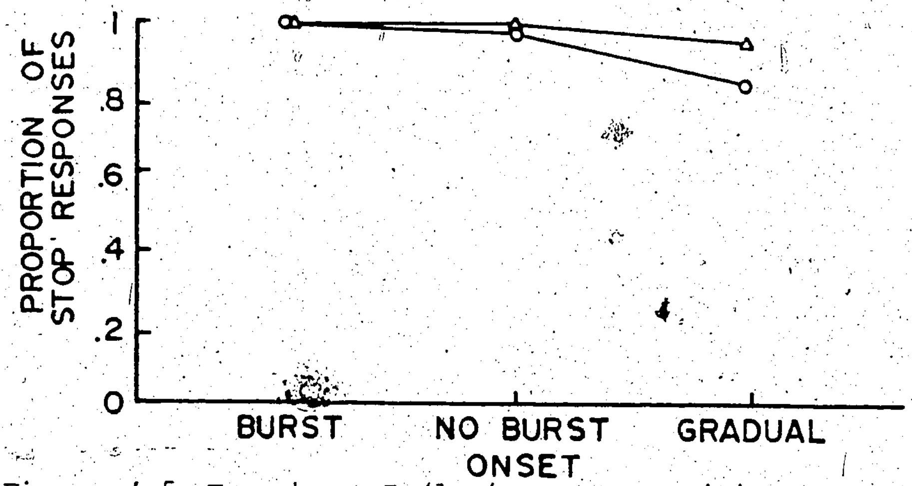


Figure 4.5 Experiment I /dae/ onset recognition curve, mean of all subjects.

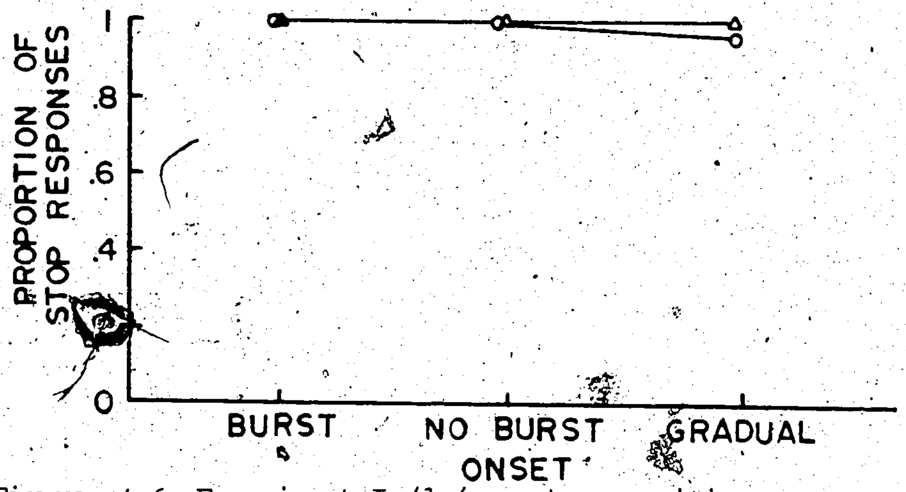


Figure 4.6 Experiment I /do/ onset recognition curve, mean of all subjects.

percent of the time under all three onset conditions when the vowel following was /i/. Only Subject 2 heard stops less than 80 percent of the time under these gradual onset conditions.

While onset slope had only a small effect on the perception of stops, duration had a large effect. Figures 4.7 to 4.12 show the mean responses over all subjects to stimuli that varied in duration in Experiment I. The first point on the abscissa, labeled "10 ms", is the same as the last point, labeled "gradual", on the abscissa of the onset figures, Figures 4.1 to 4.6. The response curves in these figures show a shift from stop to fricative response occurring between approximately 18 and 30 ms in consonant duration. The slopes of the lines are not steep, implying that the changes in perception from stop to fricative do not happen at a specific duration. Rather, the stimuli with consonant durations near the manner boundary appear somewhat ambiguous in their manner status. Since subjects are shown to be significantly different in the statistical analysis (discussed in the next section), the individual subjects' graphs in Appendix C should be scrutinized and compared. Scrutiny shows that individuals did have differing crossover points, some subjects giving fricative responses at earlier durations, some later.

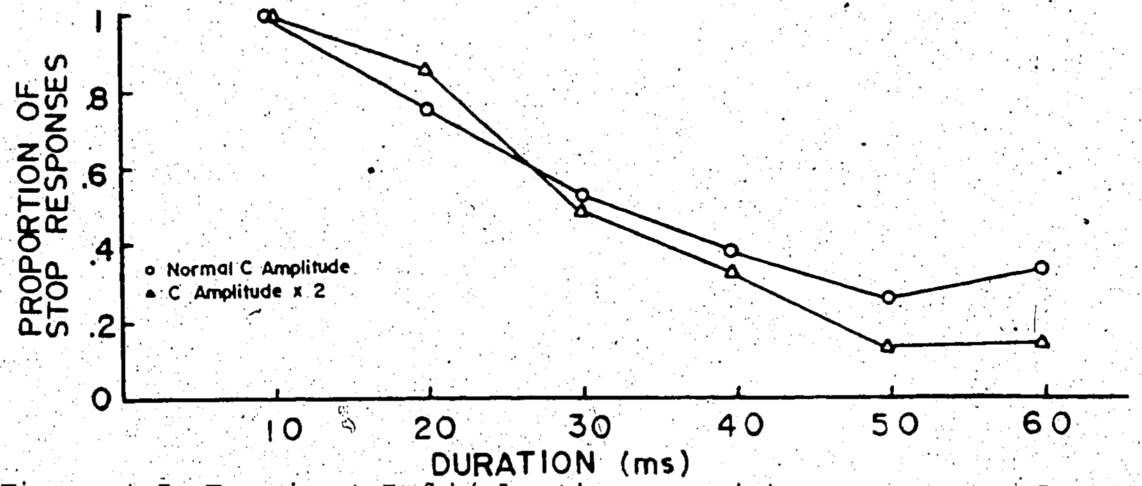


Figure 4.7 Experiment I /bi/ duration recognition curve, mean of all subjects.

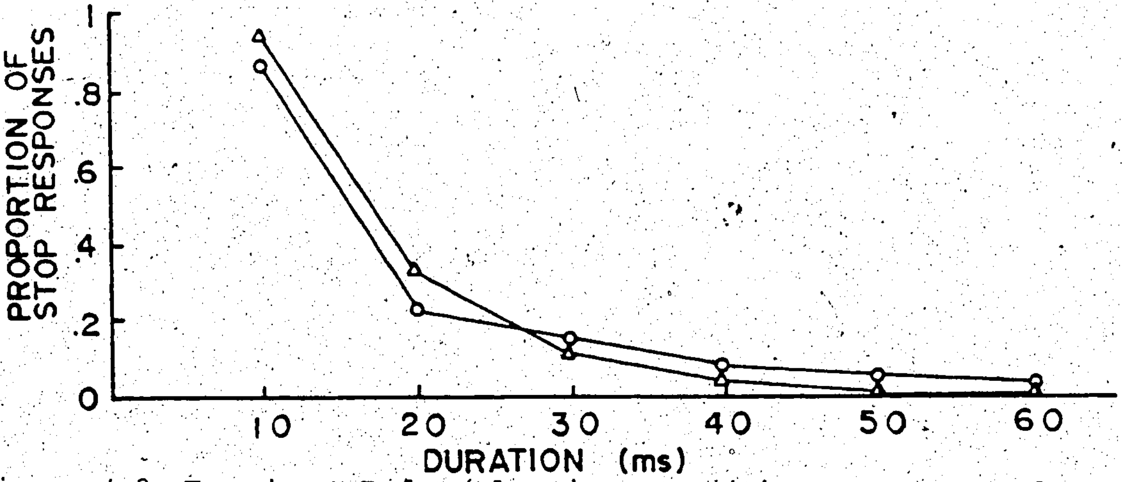


Figure 4.8 Experiment I /bae/ duration recognition curve, mean of all subjects.

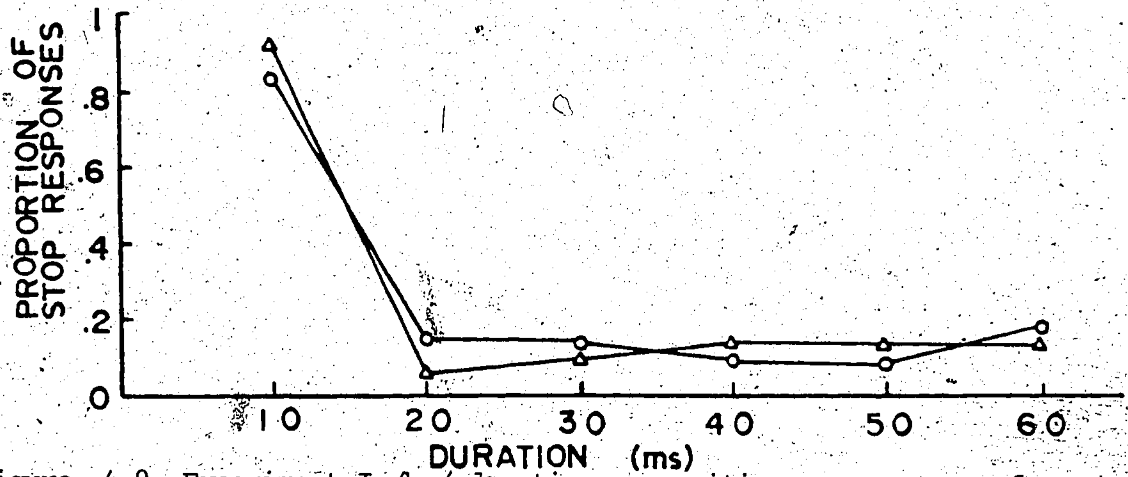


Figure 4.9 Experiment I /bo/ duration recognition curve, mean of all subjects.

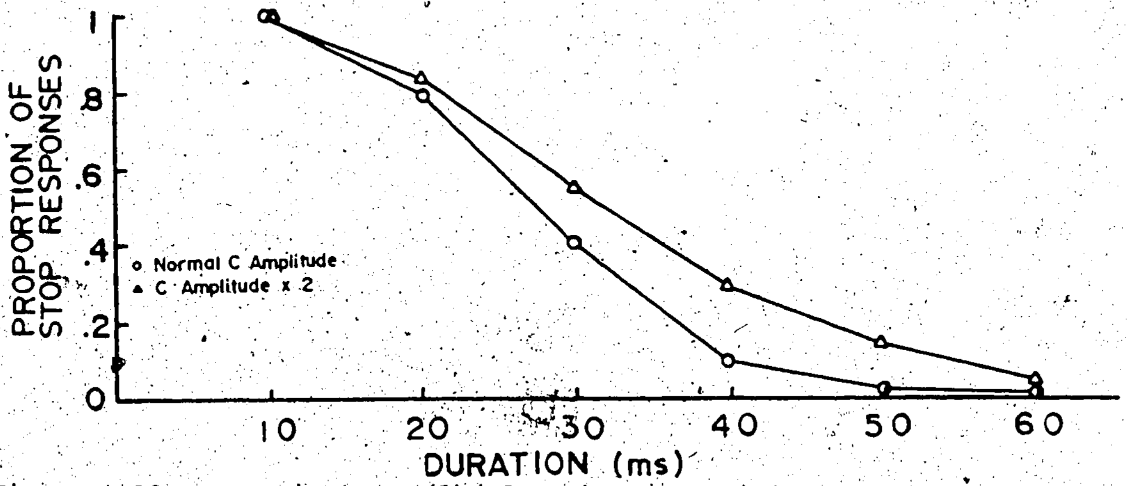


Figure 4.10 Experiment I /di/ duration recognition curve, mean of all subjects.

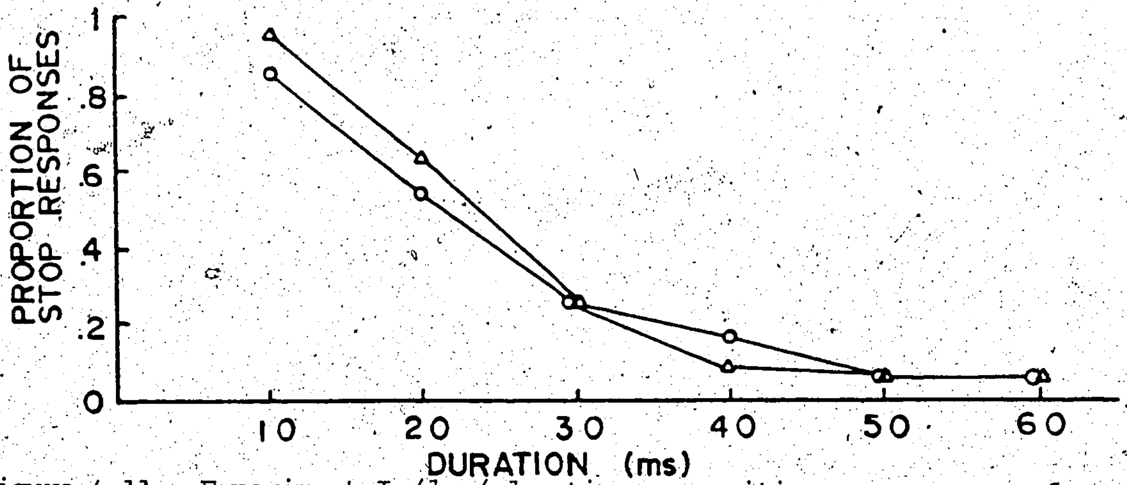


Figure 4.11 Experiment I /dae/ duration recognition curve, mean of all subjects.

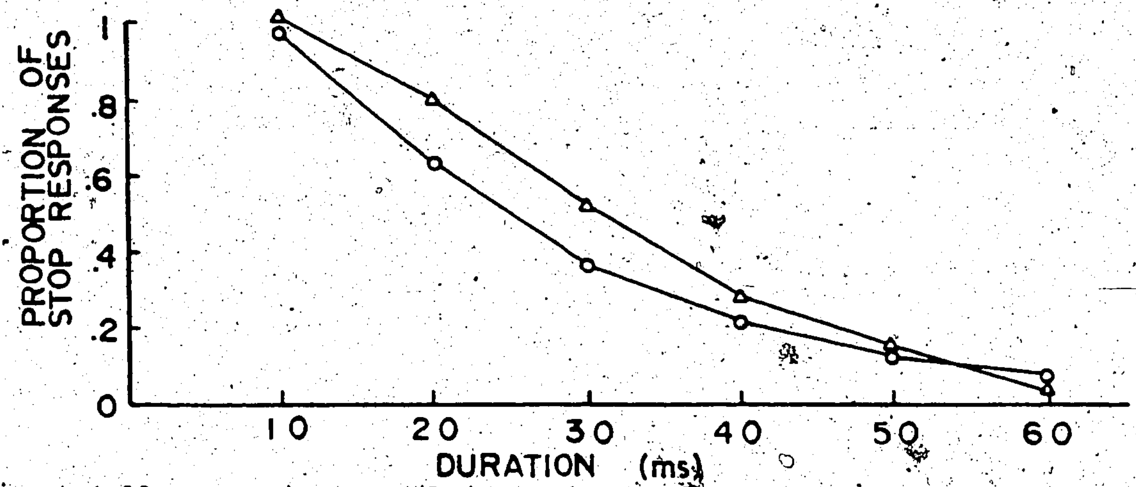


Figure 4.12 Experiment I /do/ duration recognition curve, mean of all subjects.

Confidence Ratings

As well as identifying experimental stimuli, subjects indicated their confidence in the identification using a rating of 1 (low confidence), 2 (moderate confidence), or 3 (high confidence). These confidence ratings provide evidence of the location of the subjects' stop/fricative boundaries, in addition to the evidence from the recognition curves. Table 4.3 lists the mean confidence ratings each Experiment I stimulus received. The mean confidence ratings for the original consonants are high, 2.9 out of 3. The ratings decrease as the stimulus duration increases, reaching the lowest level at 20 ms consonant stimulus duration. The mean confidence rating for the 20 ms Experiment I stimuli is 1.1. The 30 ms stimuli received the next lowest rating, 1.6. The ratings rise again as the duration of the stimulus consonant further increases toward the 60 ms durational stimuli.

The decrease in the confidence ratings indicates the location of the subjects' stop/fricative boundaries because the ratings show which stimuli subjects' felt were ambiguous. Uncertainty about the manner of articulation of the consonant is the primary reason for assigning low confidence ratings. The ratings are lowest for consonant durations of 20 and 30 ms, which is evidence that the boundary between stop and fricative durations is approximately 20 to 30 ms.

Table 4.3. Mean confidence ratings assigned to Experiment I stimuli.

Stimulus	Confidence Rating out of 3
Original Consonant	2.9
Burstless	2.8
10 ms	2.6
20 ms	1.1
30 ms	1.6
40 ms	2.0
50 ms	2.3
60 ms	2.4

The confidence ratings also reflect how natural the stimuli sounded. Some subjects told the experimenter that they [illegible] low ratings to unnatural sounding stimuli. Naturalness and confidence are closely linked since one cannot be entirely sure of the status of a consonant which does not sound like speech or which sounds like speech but has never been heard before. The confidence ratings give an indication of the quality of the stimuli as speech sounds, if they are interpreted as reflecting naturalness. The overall mean rating for Experiment I stimuli is 2.2. Subjects seemed reasonably confident of their responses and satisfied that the stimuli were speech.

Analysis of Variance

Design

An analysis of variance (ANOVA) was performed using the 50 percent stop/fricative crossover points of the duration recognition curves of the subjects as input data. The crossover is the point along the stimulus consonant duration continuum at which the subject's responses reach 50 percent stop and fricative. At durations greater than this boundary, the stimuli were heard as fricatives; at shorter durations they were perceived as stops. The trials each subject performed were totalled into two groups - two replications, so each group in the repeated measures design had a large number of responses.

The ANOVA was compiled using BMDP 8V in a four factor repeated measures design. The four factors are subjects (5), intensity (2 levels: normal and consonant amplitude times two), consonants (2: labial and dental), and vowels (3: /i/, /æ/, /o/). Subjects and replications were treated as random effects; intensity, consonants, and vowels were regarded as fixed effects. The repeated measures were the two replication groups of subject trials.

The analysis of variance sums of squares, F ratios, and probabilities of the factors in Experiment I are given in Table 4.4. A few factors and interactions were significant: subjects ($p < .0001$), vowels ($p < .01$), the subject by consonant interaction ($p < .0001$), the subject by vowel interaction ($p < .0001$), the consonant by vowel interaction ($p < .05$), and the subject by consonant by vowel interaction ($p < .0001$). The highest level interaction, subject by consonant by vowel, is diagrammed in Figure 4.19.

Proportion of Variance of Each Factor

The statistic ω^2 (omega squared) was calculated to estimate the proportion of total variance associated with each factor (Hays, 1963). ω^2 is a ratio of the estimated treatment variance to the estimated total variance. It can be calculated using the following formula:

$$\omega^2 = \sigma_\beta{}^2/\sigma_T{}^2$$

$$\text{where } \sigma_\beta{}^2 = (MS_\beta - MS_W)/n$$

$$\text{and } \sigma_T{}^2 = (MS_\beta + (n-1)MS_W)/n.$$

The ratio indicates the approximate amount of variance which

Table 4.4 Analysis of variance for Experiment I.

SOURCE	ERROR TERM	SUM OF SQUARES	D.F.	MEAN SQUARE	F	PROB.
MEAN	S	71562.782	1	71562.782	53.76	0.0018
S	R(SCIV)	5324.734	4	1331.183	69.02	0.0000
C	SC	1272.769	1	1272.769	7.12	0.0559
I	SI	148.274	1	148.274	3.81	0.1225
V	SV	2664.032	2	1332.016	9.93	0.0068
SC	R(SCIV)	714.962	4	178.740	9.27	0.0000
SI	R(SCIV)	155.488	4	38.872	2.02	0.1037
CI	SCI	63.061	1	63.061	5.97	0.0710
SV	R(SCIV)	1073.057	8	134.132	6.95	0.0000
CV	SCV	1458.989	2	729.495	5.11	0.0371
IV	SIV	20.240	2	10.120	0.28	0.7604
SCI	R(SCIV)	42.279	4	10.570	0.55	0.7011
SCV	R(SCIV)	1141.193	8	142.649	7.40	0.0000
SIV	R(SCIV)	285.551	8	35.694	1.85	0.0851
CIV	SCIV	40.872	2	20.436	1.42	0.2959
SCIV	R(SCIV)	114.846	8	14.356	0.74	0.6[illegible]3
R(SCIV)		1157.232	60	19.287		

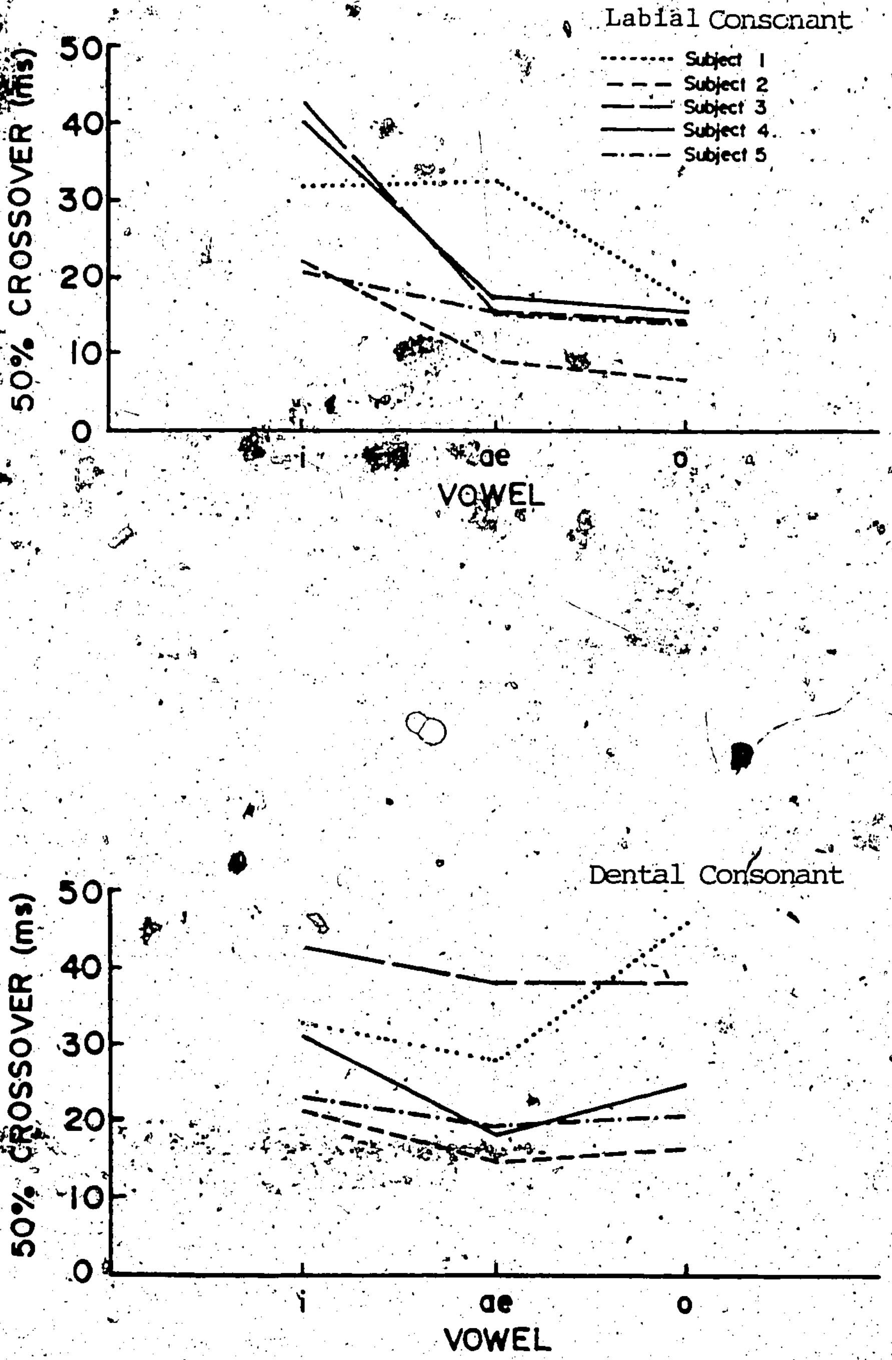


Figure 4.13 Experiment I, subject by consonant by vowel interaction.

each factor is responsible for. This statistic is presented in Table 4.5 for the subject, consonant, vowel, and intensity factors. The subject factor accounts for over half of the total variance, 74 percent. This is a very large portion of the total variance associated with a single factor. Along with the ANOVA results, the large subject variance proportion shows that subjects are a very important factor in the experimental results. The next largest portion of the variance in Experiment I is associated with the vowel factor, 18 percent. The remaining factors, consonants and intensity levels, account for 9 percent and 5 percent of the variance respectively.

Consonant Confusions

As the present experiments were designed to induce stop-fricative changes in perception, consonant confusions in manner cannot be considered errors in identification. Confusions in place of articulation, however, are true perception errors. These confusions involve labial consonants being perceived as dentals and vice versa. In Experiment I, the method of subject response did not allow for place confusions. But two subjects reported hearing the fricative /ð/ when given the choice of /b/ or /v/ as response, i.e. when hearing a labial consonant. These two subjects indicated when this perception occurred so a small amount of consonant confusion data is available from Experiment I. The response /ð/ to a labial occurred only

Table 4.5 Proportion of variance of each Experiment I factor.

Factor	ω^2	Rank
Subjects	.739	(1)
Consonant	.093	(3)
Vowel	.183	(2)
Intensity	.045	(4)

when the word "bat"/"vat" or "bee"/"vee" was presented, never when "boat"/"vote" was the stimulus, presumably because the word "thote" does not occur in English. Five percent of the responses to "bee" of these two subjects were "thee". Thirty-two percent of the responses to "bat" were "that". This limited Experiment I data seems to indicate that 19 percent of the "bat"'s and "bee"'s were perceived as beginning with dental consonants. If the responses of the other three subjects in this experiment are included, the ratio of labial/dental confusion in Experiment I drops to five percent.

B. Experiment II

Measurements

Duration

The durations of the Experiment II stimuli are given in Table 4.6. The durations of the stimuli with increased consonant amplitude are not included since the amplitude increase did not affect the duration. The first column in Table 4.6 shows the duration of the entire original syllable, the second column shows the vowel duration, and the third the duration of the original fricative. The last three columns give the exact durations of the stimuli consisting of three, two, and one period of frication. The average durations of the labial and dental fricatives are identical, 114 ms. But the duration of the /v/ stimuli

Table 4.6. Durations of Experiment II stimuli.

	Duration in ms					
Syllable	Syllable	Vowel	Original C	3-Period C	2-Period C	[illegible]-Period C
vee	360	274	86	29	20	9
vat	526	412	114	27	17	9
vote	538	396	142	27	18	9
thee	357	239	118	27	17	9
than	582	472	110	30	20	10
though	377	263	114	30	20	10

spans a much greater range. The standard deviation of the durations of the three /v/'s is 28. The standard deviation of the three /∂/ durations is 4.

Intensity

Table 4.7 shows the intensities of the Experiment II fricative-based stimuli in decibels, averaged over the three vowels. Each amplitude increase added 6 dB to the consonant intensity. The mean intensities for the original syllables are similar, 41 dB for syllables starting with /v/, 42 dB for syllables starting with /∂/. The dental fricative alone, at 23 dB, is 4 dB higher in intensity than the average /v/ at 19 dB. The original fricative consonants appear to be higher in intensity than the original stops (see Table 4.2), but the fricatives are approximately ten times as long as the stops. Ten milliseconds of frication (13 dB for /v/, 16 dB for /∂/) is lower in intensity than the original stops (16 dB for /b/, 20 for /d/). For their durations, stops are more intense than fricatives.

Recognition Curves

The mean of all subjects' responses in Experiment II are shown in the recognition curves in Figures 4.14 to 4.19. The ordinate shows the proportion of fricative responses. Although Experiment II subjects were able to respond by choosing any of the four consonants (/v/, /b/, /∂/, or /d/), the responses are given in terms of the proportion of fricative responses. The /v/ and /∂/ identifications were

Table 4.7 Intensities of Experiment II stimuli (dB).

	Stimulus				
Consonant	Original	10 ms	20 ms	30 ms	Syllable
/v/	19	13	15	16	41
/ð/	23	16	18	19	42

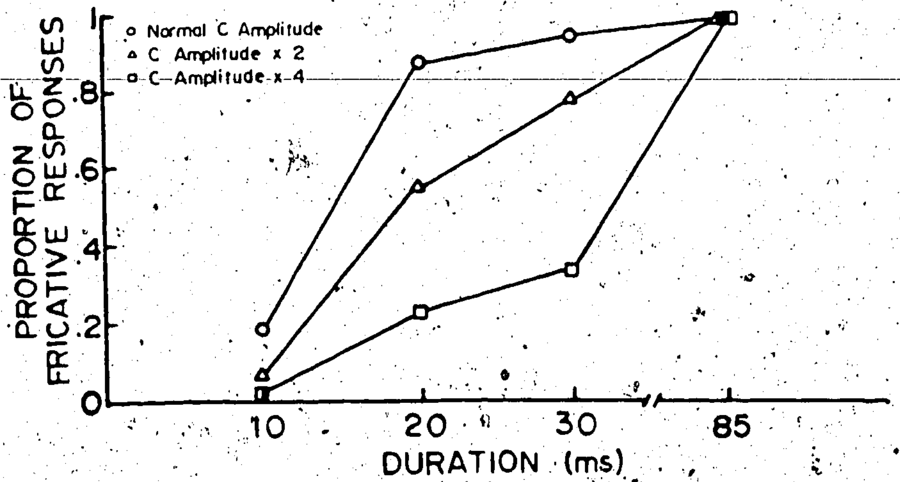


Figure 4.14 Experiment II /vi/ recognition curves, mean of all subjects.

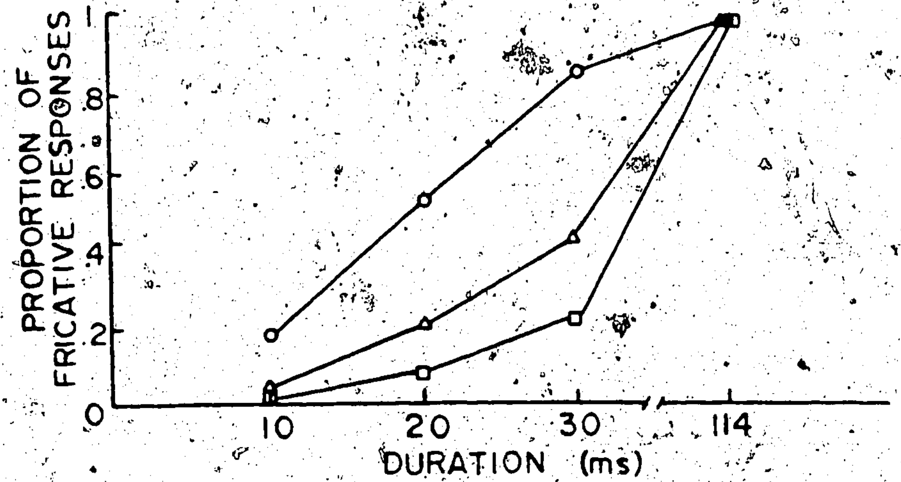


Figure 4.15 Experiment II /vae/ recognition curves, mean of all subjects.

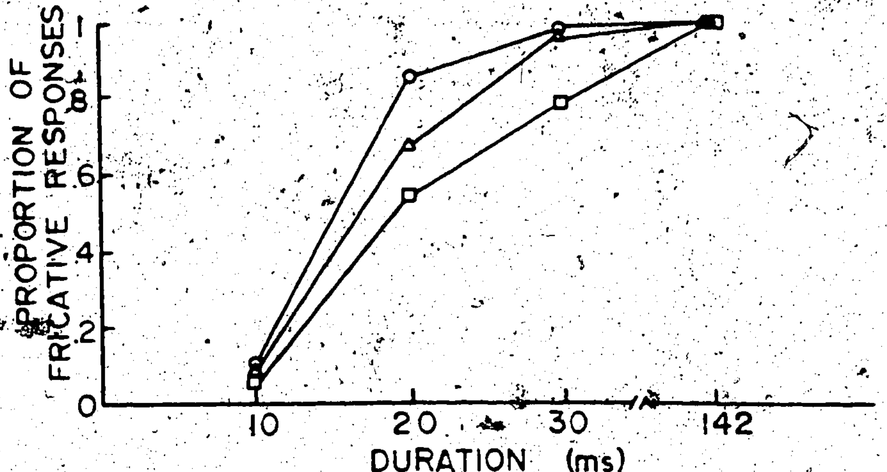


Figure 4.16 Experiment II /vo/ recognition curves, mean of all subjects.

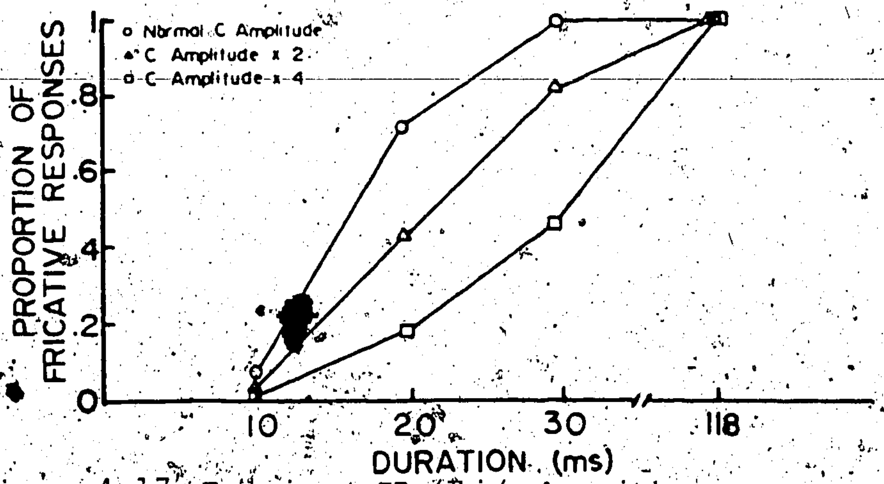


Figure 4.17 Experiment II /dhi/ recognition curves, mean of all subjects.

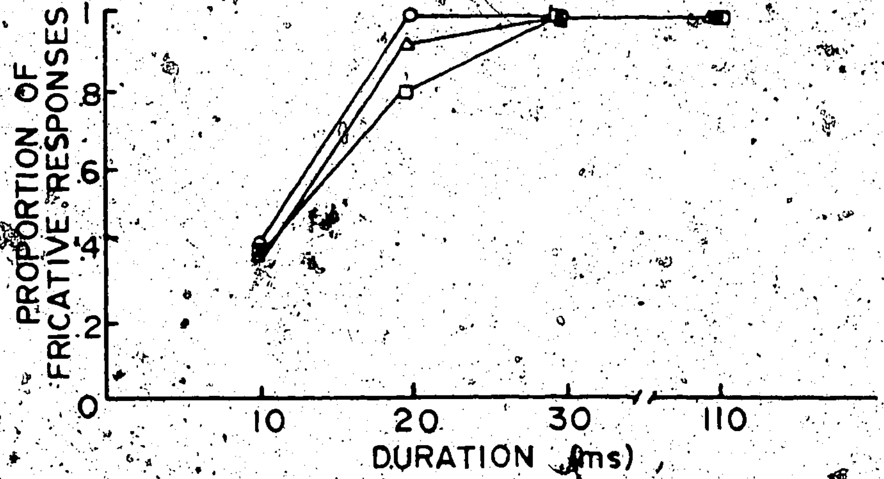


Figure 4.18 Experiment II /dhae/ recognition curves, mean of all subjects.

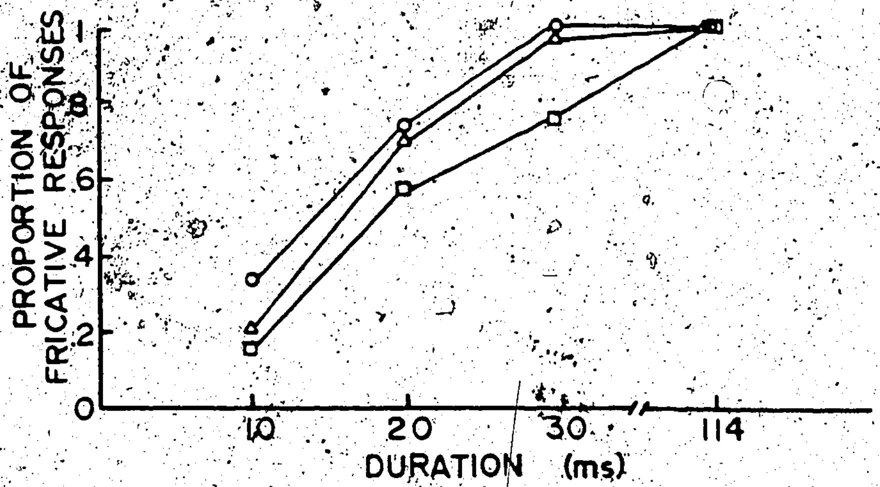


Figure 4.19 Experiment II /dho/ recognition curves, mean of all subjects.

pooled, as were the /b/ and /d/ identifications, to allow fricative versus stop response comparisons. Adding together the /v/ and /ð/ responses hides errors of incorrect place identification. Consonant confusions will be considered later in this paper with respect to place of articulation.

In these figures, 4.14 to 4.19, the original, unaltered consonant is the rightmost point on the abscissa, labeled with the duration of the original consonant. Since Experiment II involved three intensity levels, normal, double and quadruple consonant amplitude, these figures each have three lines.

The graph lines show a consistent decrease in the proportion of fricative responses as duration decreases. The change in response from 50 percent fricative to 50 percent stop tends to occur, for all consonant intensities, when the consonant is between 15 and 30 ms long. Almost all the individual subjects' curves, given in Appendix D, reflect this increase in stop responses with decrease in duration. The only exception is Subject 5, who shows no stop responses to the /ðæn/ stimuli at normal or double amplitude levels. In a few other cases, subjects do not reach 50 percent stop response: Subject 4's /ðæn/ and Subjects 5's /ðo/.

Confidence Ratings

The Experiment II stimuli received relatively high confidence ratings. The overall mean rating was 2.7 out of

3. The generally high scores for these stimuli reflect the naturalness of the stimuli since subjects gave lower scores to stimuli which they felt were not speech-like, as well as to those which were ambiguous. Subjects were confident of their manner identifications of the original fricatives, giving them a mean confidence rating of 2.9. The ratings are lower for the 30 ms stimuli, 2.6, lowest for the 20 ms stimuli, 2.5, and rise for the 10 ms stimuli to 2.8 (see Table 4.8). The lower 20 and 30 ms ratings indicate that the boundary between stop and fricative identifications occurs at approximately this duration.

Analysis of Variance

Design

In Experiment II the input to the ANOVA were the stop/fricative crossover points. Although subjects responded by choosing a particular consonant in Experiment II, the ANOVA input points are indications of where the responses changed from fricative to stop, not from /v/ to /b/ or /ð/ to /d/. The six trials each subject performed were totaled into two groups to create two replications each consisting of nine responses per stimulus.

A few difficulties with the 50 percent crossover points were encountered. When the recognition curves did not reach 50 percent stop response, if the last point on the recognition curve was close to 50 percent, the line from the second last point through the last one was extended to judge

Table 4.8 Mean confidence ratings assigned to Experiment II stimuli.

Stimulus	Confidence Ratings out of 3
Original Consonant	2.9
30 ms	2.6
20 ms	2.5
10 ms	2.8

the 50 percent crossover. This extrapolation was done three times for the Experiment II ANOVA input (out of 90 cases). In three other instances, when the curve did not near 50 percent, the mean of all the other subjects' response boundaries under that condition was inserted. This was done so that data would not be missing (especially since this data was not strictly speaking "missing"). Inserting the group mean of the condition does not increase the sum of squares of that factor.

One other complication with the 50 percent crossover points occurred in Experiment II. This problem developed when the crossover fell between 30 ms duration of the consonant and the duration of the original consonant. This affected about 30 percent of the Experiment II data. If the 50 percent crossover point were interpolated between the 30 ms and the original duration, the crossover might fall at an unreasonably long duration because the original consonants' durations ranged from 85 up to 142 ms. Thus it was possible to "find" a 50 percent crossover point at 86 ms. The Experiment I results indicated that this duration was unreasonable since by 60 ms almost 100 percent of the Experiment I responses were fricatives. More precisely, the crossover from stop to fricative response should occur before 60 ms consonant duration.

In order to resolve the dilemma of where to place the intermediate crossover point, two ANOVAs were calculated. The data for the first ANOVA assumed the crossover point occurred between 30 ms and the original fricative duration. The second data set used a more conservative estimate of the crossovers between 30 ms and original duration which involved assuming that responses would be 100 percent fricative when the consonant duration was 60 ms. So, the crossover point was interpolated between 30 and 60 ms. This resulted in crossover points at shorter durations, durations closer to the Experiment I boundaries and to the rest of the Experiment II data.

A comparison of the results of these two ANOVAs indicated that they were largely the same, although the sums of squares of the first were usually much larger than those of the second data set. Only one difference in significance of factors appeared: the vowel by intensity interaction of the first ANOVA had a probability of 0.0407 ($p < .05$); in the second its probability was 0.0063 ($p < .01$). Otherwise all factors which had probabilities less than .01 in the first ANOVA also had low probabilities in the second. Since the second represented a more conservative and more reasonable estimate of the 50 percent crossover points, it was selected as the ANOVA to represent the Experiment II data. Thus, in Experiment II, if the crossover fell above 30 ms it was calculated as lying between 30 and 60 ms.

The ANOVA was compiled using BMDP 8V in a four factor repeated measures design. The four factors are subjects (5), intensity (3 levels: normal, consonant amplitude times two, and consonant amplitude times three), consonants (2:

labial and dental), and vowels (3: /i/, /æ/, /o/). Subjects and replications were treated as random effects; intensity, consonants, and vowels were regarded as fixed effects. The repeated measures were the two replication groups of subject trials.

The ANOVA of the Experiment II 50 percent crossover points showed most factors and interactions to have very low probabilities. Table 4.9 gives these ANOVA results. The only factor significant at the .05 level and not at .01 is vowels, with a probability of .04. All other factors had probabilities of .01 or less, except for the following: consonant, the consonant by intensity interaction, and the consonant by vowel by intensity interaction. The highest level interaction, subject by consonant by vowel by intensity ($p < .01$) is graphed in Figures 4.20 and 4.21. The first three graphs show the results involving the consonant /v/ at the three intensity levels; the second three graphs show the consonant /ð/ at the three intensity levels.

Proportion of Variance of Each Factor

The ω^2 statistic, an estimate of the proportion of variance of each factor relative to the total variance, was calculated for the Experiment II factors (see Experiment I results). Table 4.10 gives the results. The subjects accounted for 52 percent of the variance, over half the total variance. Intensity also had a major impact in Experiment II; it accounts for 30 percent of the variance. Vowels and consonants are the source of much less variation

Table 4.9 Analysis of variance for Experiment II.

SOURCE	ERROR TERM	SUM OF SQUARES	D.F.	MEAN SQUARE	F	PROB.
MEAN	S	89931.45	1	89931.451	145.30	0.0003
S	R(SCVI)	2475.73	4	618.934	39.40	0.0000
C	SC	877.41	1	877.415	5.40	0.0809
V	SV	910.95	2	455.477	5.05	0.0382
I	SI	4287.84	2	2143.919	26.74	0.0003
SC	R(SCVI)	650.35	4	162.587	10.35	0.0000
SV	R(SCVI)	721.78	8	90.222	5.74	0.0000
CV	SCV	2573.53	2	1286.763	10.83	0.0053
SI	R(SCVI)	641.33	8	80.166	5.10	0.0000
CI	SCI	498.21	2	249.104	4.25	0.0554
VI	SVI	831.69	4	207.921	5.35	0.0063
SCV	R(SCVI)	950.57	8	118.821	7.56	0.0000
SCI	R(SCVI)	469.28	8	58.661	3.73	0.0008
SVI	R(SCVI)	622.24	16	38.890	2.48	0.0037
CVI	SCVI	351.29	4	87.822	2.17	0.1190
SCVI	R(SCVI)	647.54	16	40.471	2.58	0.0025
R(SCVI)		1413.81	90	15.709		

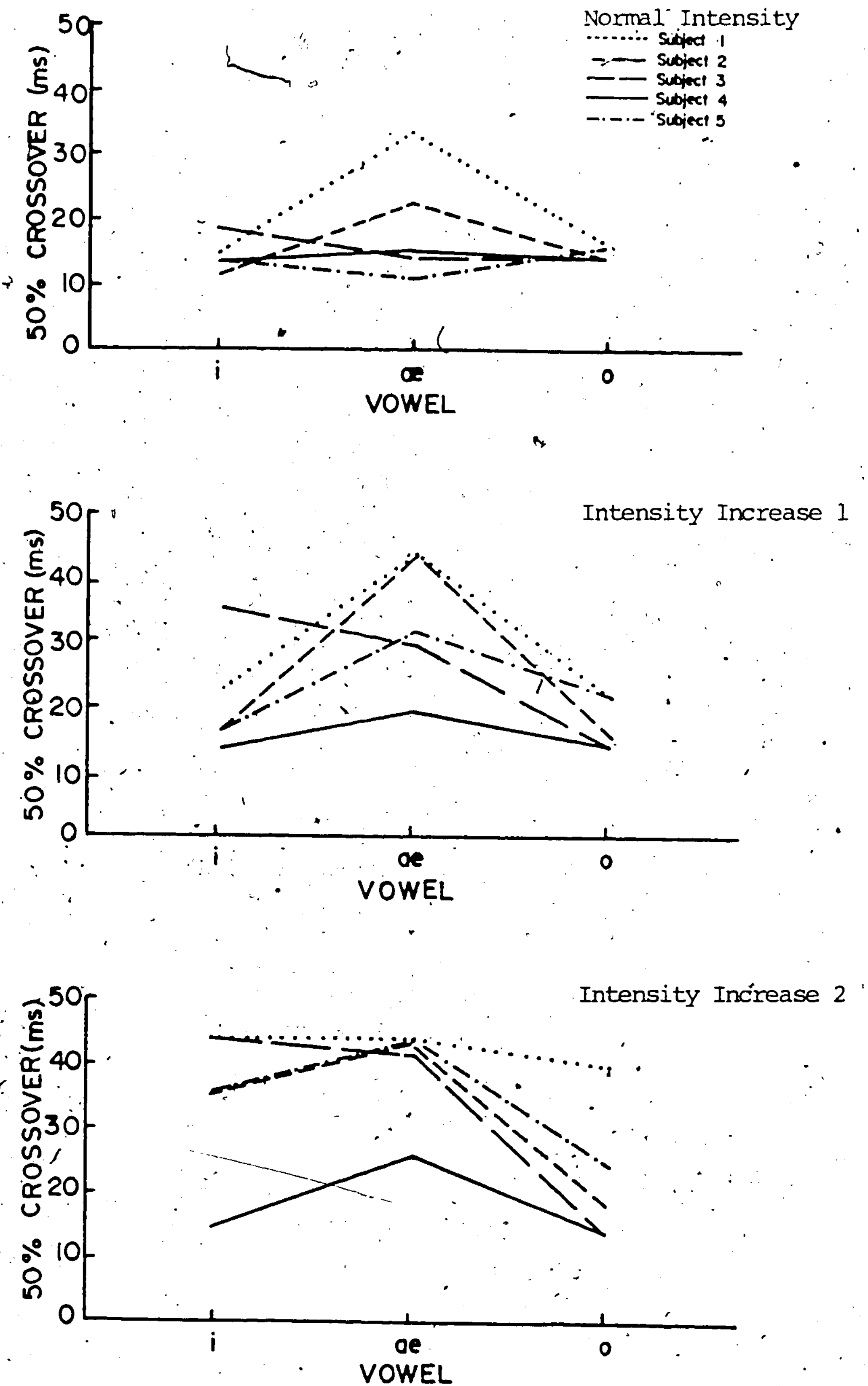


Figure 4.20 Experiment II, subject by consonant by vowel by intensity interaction - labial consonant.

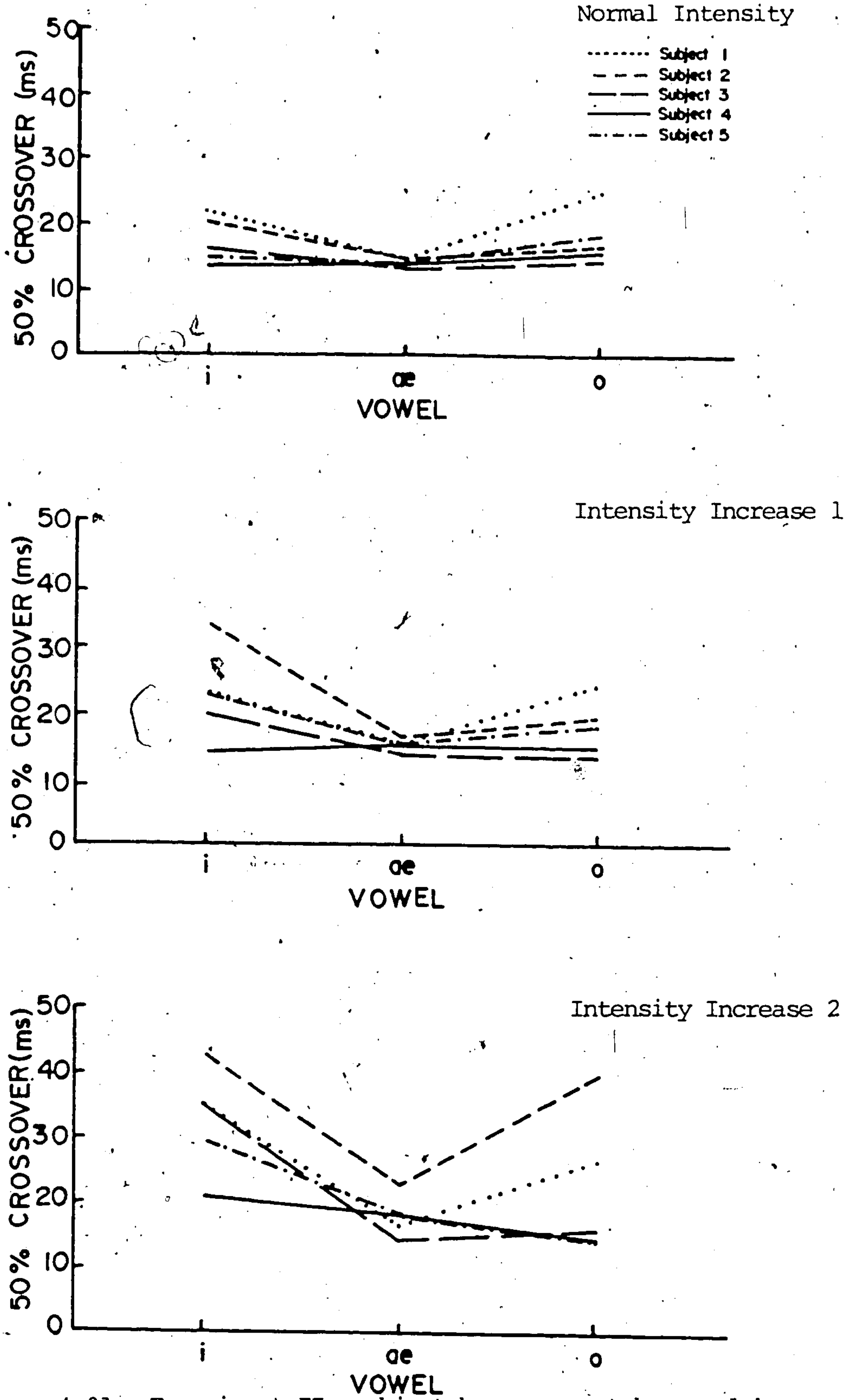


Figure 4.21 Experiment II, subject by consonant by vowel by intensity interaction - dental consonant.

Table 4.10 Proportion of variance of each Experiment II factor

Factor	ω^2	Rank
Subjects	.516	(1)
Consonant	.047	(4)
Vowel	.063	(3)
Intensity	.300	(2)

in the responses, 6 percent and 5 percent respectively.

Consonant Confusions

Experiment II provided more opportunities for consonant confusions. Subjects were allowed to respond with any of the four consonants, so it was possible to respond incorrectly with respect to place of articulation of the consonant stimulus. Six percent of the responses to labial stimuli were dental, 3 percent /d/ and 3 percent /ð/ responses. Only one percent of the dental stimuli were heard as labials. This 1 percent consisted of incorrect /b/ responses; no /v/ responses to a dental consonant occurred. The 5 percent Experiment I and 6 percent Experiment II incorrect perceptions of labials and the 1 percent Experiment II incorrect perceptions of dentals accord with the consonant confusion studies mentioned in Chapter II. The place of articulation error rates in the present experiments are actually relatively low. The average rates, however, gloss over the subject variation - one subject in Experiment II identified 17 percent of the labial stimuli as dentals. This subject's high error rate means that the place perception of the other subjects was even more accurate than 94 percent correct.

V. Discussion

A. Measurements

Duration

The measurements of the original consonants in these experiments indicate they are within a normal range of fricative and stop durations. The durations of the /v/'s, from 86 to 142 ms, shown in Table 4.6 fall between Umeda's 78 ms (1977) and Abbs and Minifie's 153 ms mean (1969). Likewise, the /ð/ durations, from 110 to 118 ms long, are greater than Umeda's 52 ms and slightly less than Abbs and Minifie's 123 ms mean. The stop consonant burst durations fall within one standard deviation of Klatt's (1975) measured durations (see Table 4.1). The /b/ mean, 7 ms, and /d/ mean, 12 ms, are typical for these consonants in English.

Intensity

The intensity measurements collected by Abbs and Minifie (1969), Malecot (1968), and Ohde and Stevens (1983) state that stops have higher intensities than fricatives and that dentals have greater intensity than labials. In this study the dentals do have greater intensities than the labials, but the intensity of the stops does not appear to be greater than that of the fricatives. The intensity of /d/ is 20 dB and /ð/ is 23 dB. The /b/ intensity is 16 dB, while /v/ is 19 dB. Considering the difference in duration between the stops and the fricatives, though, the intensity of the stops is greater, for its duration, than that of the fricatives. The fricatives are approximately ten times as long as the stops. The intensity of the stimuli (derived from fricatives) that are the same duration as the stops is less than that of the original stop consonants (see Tables 4.2 and 4.7). The measured intensities of the consonants used in the present study are as expected, the stops are more intense than the fricatives and the dentals are more intense than the labials.

B. Onset

The lack of onset slope effect in the attempt to induce fricative responses from stop consonants in Experiment I shows that the difference between a stop and a fricative is not simply plus or minus burst. Removing the burst from a stop did not cause it to be heard as a fricative. While duratation of onset is considered to be a primary factor in the fricative/affricate distinction (Gerstman, 1957), it cannot be in the stop-fricative distinction because of the duration factor. The duration of the onset slope stimuli which were originally stops is about 10 ms. This is too short for them to be identified as fricatives since fricatives typically have durations of approximately 100 ms. This duration is too short even for affricate identification since affricates typically have rise times of 50 ms. So, although the consonants did not have abrupt onsets, they were identified as stops due to their short durations. This result supports Van Heuven's (1979) claim that the primary cue to the fricative/affricate distinction is not rise time, as Gerstman states, but overall duration. Carden et al. (1980) have also claimed that the primary cue in the stop/fricative distinction is duration. The onset recognition curves for Experiment I show clearly that onset slope alone does not distinguish between stops and fricatives.

C. Duration

In both the stop to fricative and fricative to stop identification experiments, the crossover in manner identification tends to occur at approximately 25 ms consonant duration. The location of the boundary is apparent from both the subjects' recognition curves (see Figures 4.7 to 4.12 and 4.14 to 4.19) and the confidence ratings associated with the stimuli (see Tables 4.3 and 4.8). Both the recognition curves and the confidence ratings show that the stop/fricative boundary of the subjects are at approximately 20 to 30 ms consonant duration. Subjects' perceptions vary, depending on the consonant, vowel, and intensity involved, but identifications usually cross from 50 percent stop to 50 percent fricative within 15 ms of 25 ms. This duration, 25 ms plus or minus 15, seems to be a "magic" number in speech recognition. Many durational boundaries between speech sounds occur at approximately 25 ms. For example, Lieberman (1977) shows that stop VOT's of less than 25 ms are perceived as /b/ while those of more than 25 ms are heard as /p/.

Other studies mention durations close to 25 ms as being critical. Liberman (1956) reports that the stop /b/ is identified as the semivowel /w/ when the duration of its transitions reach 40 ms. Another instance of a 40 ms boundary was noted by Cutting and Rosner (1974). They found, in both speech and non-speech stimuli, a "fixed" auditory sensitivity at 40 ms. Howell and Rosen (1983) dispute Cutting and Rosner's fixed auditory sensitivity claiming that it does not correspond to the fricative/affricate rise time boundary which occurs at a duration greater than 40 ms. Howell and Rosen support their argument by showing that natural affricates usually have rise times greater than 40 ms. Howell and Rosen, and Cutting and Rosner assume that "plucked" and "bowed" (Cutting and Rosner's non-speech identification terms) correspond to affricates and fricatives. Perhaps these terms apply to stops and fricatives and perhaps the 40 ms boundary does too.

An argument against applying Cutting and Rosner's "plucked"/"bowed" boundary to stops and fricatives is that theirs is an duration of onset boundary while the stop-fricative boundary is one of duration, but duration and onset work together. In natural speech fricatives, and in stimuli created for the present study, duration covaries with rise time. Natural fricatives have roughly linear rises reaching full amplitude shortly before vowel onset. Also, Cutting and Rosner, Van Heuven (1979), and Gerstman (1957) all acknowledge that duration and rise time are much more potent together than either separately. In the present experiment, it was impossible to separate the two, using real speech, and still achieve natural sounding consonant stimuli. If stop consonants are concatenated to increase duration without being ramped to create gradual onsets, the results sound like mechanical buzzes, not like speech. Hence an increase in duration meant an increase in rise time. This also occurs in the production of fricatives. A rapid rise time followed by a long steady time would be an affricate. Since Cutting and Rosner's 40 ms boundary does not apply well to affricates and stops, it may be more aptly applied to stops and fricatives, even though the 40 ms boundary is 15 ms longer than the stop-fricative boundary found here.

The variability of the subjects' crossover points around the 25 ms mark may be related to the sensitivity of the auditory system. Lieberman (1977) notes that a time delay of 20 ms is needed between two sounds for listeners to judge which came first. Lieberman suggests there is a 20 ms auditory resolution factor. Klatt and Cooper (1975) report that the just noticeable difference (JND) for duration change of the segments /i/ and /ʃ/ is 25 ms or more. Fujisaki, Nakamura, and Imoto (1975) found the accuracy of discrimination for 100 ms of white noise is 9.1 ms. The white noise is comparable in duration and content to a voiceless fricative. People can detect differences in duration of 10 to 20 percent of the original speech segment duration. However, a just noticeable difference duration change in a consonant does not cause it to sound abnormal. Huggins (1972) found that a word-initial /ð/, originally 115 ms long, was judged "normal" when its duration was between 83 and 162 ms. Subjects, then are not aware of slight differences in duration and are tolerant of fairly large changes in duration. The crossover variation within subjects, associated with the consonant, vowel, and intensity factors in the experimental results, may be due to the inability of the auditory mechanism to discriminate differences in noise duration and tolerance of duration differences. These possible sources of crossover variation may result in chance differences in boundary placement which would cause apparent shifts of a subject's labial/dental stop/fricative boundary.

If the boundaries of individual subjects are varying randomly, for the reasons mentioned above, then statistically significant interactions between subjects will appear, just by chance. While the author does not believe that all the response variation can be attributed to chance differences in duration perception, the very low analysis of variance probabilities of the experimental factors and interactions are believed to be partly the result of a statistical system which discriminates more finely than the human auditory perceptual system does.

Production of stops and fricatives may explain boundary variation between subjects. The variation in listeners' boundaries is not important in distinguishing English voiced labial and dental stops and fricatives, because of the produced durations of these speech sounds. The mean duration of the original, unaltered stop consonants recorded for these experiments is 9.5 ms; the mean fricative duration is 114 ms. The difference between these stop and fricative durations is 104.5 ms. While it may be true that the duration differences in production of these consonants are regulated by listeners' perceptual abilities, the important point here is that differences among crossover boundaries found in these experiments are irrelevant because of the durations produced in real speech. To verify the truth of this statement, consider the range of various boundaries found for various subjects in these experiments and consider the mean durations of actual labial and dental stops and fricatives. Stop/fricative boundaries in Experiment I ranged in duration from approximately 10 ms (for Subject 2's /bæ/ at normal amplitude) up to 50 ms (for Subject 3's /di/ at increased amplitude). (The 10 ms figure is an approximation because Subject 2 heard all the duration varying stimuli based on the syllable /bæ/ as fricatives more than 50 percent of the time.) If each of the originally recorded, real speech stops and fricatives, 12 in all, were categorized based on the 150 boundaries found in the experiments (60 in Experiment I, from five subjects, two consonants, three vowels, and two amplitude levels, plus 90 in Experiment II, from five subjects, two consonants, three vowels, and three amplitude levels), 98.8 percent of the identifications would be correct with respect to manner of articulation. All the fricatives would be correctly identified; 878 of the 900 stop identifications (97.6 percent) would be correct. Despite the variation due to subjects, consonants, vowels, and consonant amplitudes, the 50 percent stop/fricative crossover points used as manner recognition criteria result in correct categorizations of the naturally occurring stops and fricatives 99 percent of the time, based on duration alone. If abruptness of onset were also used as an identification criterion, all the stops would probably be correctly identified because of their initial bursts. Intensity could also be used as a secondary characteristic on which judgements may be based.

Clearly the variation in results between subjects and within subjects, associated with consonant, vowel, and amplitude levels are inconsequential in identification of actual stops and fricatives. Actually occurring English voiced labial and dental stops are shorter than almost all the 50 percent crossover boundary points. Actual English voiced labial and dental fricatives are longer than all the crossover boundary points. The produced durations of these stops and fricatives allow listeners a large amount of discretion in establishing a stop/fricative duration boundary. Since listeners have no need to choose a certain, precise boundary location, it would be surprising if the distinguishing durational criteria of all subjects were exactly the same.

High similarity among subjects' duration boundaries might suggest a fixed, innate auditory duration sensitivity. The variation found among subjects indicates that the boundary is not fixed across all listeners. Language variations in stop and fricative durations argue against an innate, fixed boundary in humans. For instance, Danish voiced labial and dental stops are longer than the comparable English stops, approximately 20 ms long (Fischer-Jorgenson, 1954). Categorizations of Danish stops using the duration criteria found here would yield much higher error rates than the criteria would for English stops. Although Danish listeners could use other acoustic characteristics of the signal as verifying or supplementary criteria in stop vs. fricative identifications, raising the lower limit of the duration boundary would enable them to judge manner of articulation solely by duration, as English speakers can. This is not proof that Danish duration boundaries are higher, but they may well be, in which case the 25 ms duration boundary is not fixed and innate in human perceptual systems.

Now that I have argued that the statistical differences associated with the experimental treatment factors are irrelevant, I will attempt to explain the boundary variation related to the factors. Consistent variation, particularly, deserves consideration if we are to identify acoustic aspects, other than duration, which influence stop and fricative perception.

D. Factors and Interactions

The subject by consonant by vowel analysis of variance interaction of Experiment I ($p < .001$) means the stop/fricative boundaries in Experiment I cannot be discussed without considering subjects, consonants, and vowels simultaneously. Similarly, the subject by consonant by vowel by intensity interaction of the Experiment II ANOVA ($p < .01$) means Experiment II boundaries depend on all four factors. These interactions, because they involve so many factors, are complex. They can be described, as they appear in Figures 4.13 (Experiment I) and 4.20 to 4.21 (Experiment II), but their complexity makes them difficult to explain. To simplify the discussion, the factors are explained - as far as possible - one at a time in the following sections. Although the effects are considered individually, their interaction with other factors must be remembered and will be reiterated. And, although the causes for the variations in both experiments are discussed at the same time, the analyses of variance are separate.

Subjects

Subject variation accounts for more than half of the variance in these experiments, 74 percent in Experiment I and 52 percent in Experiment II (see Tables 4.5 and 4.10). As discussed earlier, the variance due to subjects, although large, is irrelevant in processing real speech - as all the factor variance is - because of the great difference between the duration of stop consonants and the duration of fricative consonants.

In the discussion of duration, the sensitivity of the auditory mechanism was proposed as part of the reason for crossover variation within and between subjects. Individual crossovers varied over a range of 40 ms, from 10 to 50 ms consonant duration. The auditory system's JND for speech sounds is 10 to 20 percent of the original duration, which may account for some of this variation. Listeners are not accurate in duration discriminations, thus variation in duration crossover points occurs.

Subject variation may also be related to the production of speech sounds. The wide difference in the duration of stop and fricative consonants allows subjects discretion in establishing a manner boundary. As well, segment duration varies according to linguistic context - whether the word is pronounced in isolation or running speech, where the word occurs within a sentence, whether or not the word is stressed, and so on. Listeners must cope with this variation by selecting an optimum duration boundary to distinguish speech sounds, one immune to variation due to linguistic context. We should be surprised, because of the variety of sources of duration variation, and because of the differences among people in such things as exposure to speech patterns, if duration boundaries for all subjects occurred at exactly the same place.

Between subject variation may simply reflect subjects adopting difference strategies for performing the experimental task. It is not possible to know exactly how subjects performed the experimental task. It is plausible that more than one method was used, even by a single subject. While subjects did not report on how they performed the task, they expressed little difficulty in making the identifications. Subjects did report variation within themselves in assigning the confidence ratings. Sometimes a subject based the rating on confidence alone; other times the ratings were based on stimulus naturalness. Hopefully, since subjects did not find the identification task unusual or difficult, the criteria they used in categorizing the consonants were ones used in everyday speech processing.

Since subjects are one factor in the highest order interactions of both experiments, subject variation is related to the consonant, vowel, and - in Experiment II - intensity of the consonant involved. An inspection of the individual subject recognition curves reveals this variation.

The between subject variation is evidence that people are different. People have been exposed to different auditory stimuli and people cope with these stimuli differently. The variation among people points to the effectiveness of the communication system. Because people are different and because individuals are not 100 percent accurate in discriminating acoustic signals, the communication system depends on gross differences in the acoustic signals of speech sounds. The gross differences are necessary so people can easily identify speech sounds and then virtually ignore the auditory level of speech perception and concentrate on higher levels of perception. The difference in duration between stops and fricatives is one of these gross differences that facilitate speech sound identification.

Intensity

Intensity accounted for 4.5 percent of the variance in Experiment I and was not involved in the highest level interaction in that experiment. The recognition curves for subjects in Experiment I show a slight tendency for greater proportion of stop identifications at higher consonant amplitudes when the stimulus duration is short, but this difference is not statistically significant. At short durations, an increase in the amplitude of the stimulus causes its intensity to increase towards that of a stop with a burst. For example, adding 6 dB (the amount added by the amplitude increase) to the 20 ms /b/ stimulus increased its intensity to 17 dB. The new intensity is very close to that of the original /b/, 16 dB. This may partly account for the increased stop responses to the increased amplitude, short duration stimuli in Experiment I.

In Experiment II, the intensity factor accounts for 30 percent of the overall variance. The intensity effect shows up well on most of the subjects' recognition curves as higher proportion of stop responses at higher consonant amplitude levels. Intensity is involved in the fourth order Experiment II interaction ($p < .01$). This means the intensity effect depends on the subject, consonant, and vowel involved. For example, Subject 4 shows variation due to amplitude for /væ/ and /ði/ but not for /vi/, /vo/, /ðæ/, or /ðo/. Subject 2 shows noticeable increase in stop responses due to intensity for all syllables. The intensity increase effect is especially apparent at consonant durations of 20 and 30 ms, durations near the fricative/stop boundary. At these durations, stimuli are somewhat ambiguous in their manner status, so their perception seems to be more susceptible to the influence of other factors. (Ohde and Stevens (1983) also found that the intensity effect in identifying stops as labial or dental was more pronounced for ambiguous stimuli.)

The reason why intensity was highly variable in Experiment II but not in Experiment I may be associated with onset slope. All the Experiment I stimuli had gradual onsets. The Experiment II stimuli onsets were not ramped to be gradual, so an amplitude increase could also cause an abrupt onset. However, the 20 and 30 ms stimuli, which showed the greatest increase in stop responses along with intensity increases, all begin with the first 10 ms of the fricative signal. This first period of a fricative is very low in amplitude and generally does not show abrupt changes in amplitude (see Figures 3.14 and 3.15). Thus the amplitude increase would not create an abrupt onset on these stimuli.

The most likely reason why an amplitude increase resulted in a stop response increase is that stops have higher intensities at short durations than fricative signals of short duration. The question of why Experiment I did not also show a significant intensity effect remains unanswered.

Vowels

The vowel factor accounts for 18 percent of the Experiment I variance and 6 percent of the Experiment II variance. Vowels were involved in the highest level interactions of both experiments. Figures 4.13, the graphs of the highest level Experiment I interaction, and Figures 4.20 and 4.21, the highest level Experiment II factor interaction graphs, show the response variation due to the vowel is irregular. When the consonant is dental, the stop/fricative crossover is lower with the vowel /æ/ than with the vowels /i/ and /o/, which have approximately equal crossovers. When the consonant is labial, the /æ/ crossover is highest in Experiment II and the /i/ crossover is highest in Experiment I. Also, there is a large amount of variation among the subjects with, especially, the labial consonant and the vowel /æ/. The irregularity in the boundaries associated with the vowels make the vowel effect difficult to explain. Three possible causes will be considered. One concerns the CV relationship, the second the frequency of occurrence of the CV pair, and the third is the influence of the particular stimulus word.

One possible explanation of the vowel effect is consonant-vowel coarticulation. Carney (1971) points out that coarticulation is greater when the tongue is not involved in the consonant production. When the tongue is not involved, it is free to assume the position of articulation of the following vowel, causing coarticulatory effects. Since the tongue is not used in /b/ production, but is in /d/ production, the vowel effect would be greater for /b/ than for /d/. Thus the greater variability found in crossovers of labial syllables than in dental syllables in both Experiment I and Experiment II may be due to consonant-vowel coarticulation. Precisely how perception is influenced by coarticulation is not clear from this study.

The vowel effect may also be related to the frequency of occurrence in real speech of the vowels /i/, /æ/, and /o/ with the consonants /b/, /d/, /v/, and /ð/. Denes (1963) counted the number of times these and other vowels occurred after these consonants in spoken, British English. (Actually Denes collected his data from phonetic readers for students of English as a second language.) Although the dialect of English of the recorded words in his study is British, not Canadian, Denes' statistics may be helpful in interpreting the vowel and consonant factors. Table 5.1 gives the percentage occurrence of each of the three vowels with the four consonants. The percentage is calculated as a proportion of the number of times a vowel occurred after a consonant divided by the total number of times any vowel occurred after that consonant. Table 5.1 also shows the rank order, based on frequency of occurrence, of each vowel after each consonant, in comparison with all 20 vowels Denes included in his study. The co-occurrences of the consonants with the vowels are variable. None of the vowels occurs consistently more often than any of the other two vowels with all the consonants, none occurs consistently least often. Perhaps the subject by consonant by vowel interactions are a result of the inconsistency of occurrence of these vowels with these consonants. Possibly some subjects considered the proportion of times /i/ occurs after /b/ (9.0 percent), for example, in comparison with /i/'s occurrence after /v/ (0.8 percent) in identifying a labial before /i/ as /b/ or /v/. It is also possible that subjects considered the CV co-occurrence frequencies as they appear in a dictionary - each word occurring only once, regardless of how often it is used in speech - in identifying

Table 5.1 Percentage frequency of occurrence, and rank order of occurrence frequency, of /i/, /æ/, /o/ after /b/, /v/, /d/, /ð/, out of twenty vowels.

	Vowel					
	/i/		/æ/		/o/	
Consonant	%	(rank)	%	(rank)	%	(rank)
/b/	9.0	(4)	6.0	(6)	2.1	(9)
/v/	0.8	(8)	1.3	(6)	1.0	(7)
/d/	1.8	(11)	0.6	(17)	13.0	(2)
/ð/	0.6	(9)	8.0	(4)	1.5	(7)
Mean	3.1		4.0		4.4	

Note. Data taken from Denes (1963).

consonants. Subjects may have considered the CV co-occurrences sometimes, but not others. Or subjects may have ignored the co-occurrence relationships completely. We do not know what subjects did do, but CV co-occurrences could have influenced their responses.

It is also possible that the vowel factor variation rests on the word involved, rather than the CV combination. Subjects may have been influenced by the particular words used as stimuli. For example, "Dan", chosen to pair with the fricative-initial word "than", might not be considered a "real" word in English since it is a proper noun. Subjects may have tended to avoid reporting "Dan" as a response because of its uncommon word status. This could be the reason why the crossovers of the /dæ/ and /ðæ/ based stimuli are lower than the stimuli with the vowels /i/ or /o/. The lower crossover duration indicates fewer stop responses, which may be attributed to avoidance of the response "Dan".

Consonant

The consonant factor alone is not statistically significant in either experiment, but in both this factor participates in the highest level ANOVA interaction. The variance associated with the consonant factor is quite low in both experiments: In Experiment I the consonant factor accounts for 9 percent of the total variance; in Experiment II it accounts for 5 percent of the total variance. So, although consonants do interact with other factors, their effect is not as large as those of most of the other factors. In both experiments the consonant effect seems to be linked to the vowel effect. The same causes suggested for the vowel variation may account for the consonant variation. These causes are the consonant-vowel coarticulation effect, the CV co-occurrence frequency bias, and the dislike-of-"Dan" hypothesis. The coarticulation effect suggests that CV coarticulation is more prevalent with consonants that are articulated without tongue involvement. Since /b/ and /v/ do not involve the tongue, while /d/ and /ð/ do, the coarticulations of the former with the following vowels will be greater. Thus there is more response variation evident with the labial consonants than with the dentals.

The variability of frequency of occurrence of vowels after the consonants in question may affect the consonant identification. If it does, it means that subjects hold off consonant identification until the following vowel is identified. Although this is possible, it seems less likely than the reverse dependency. Nevertheless it may account for some of the variation in the consonant recognitions.

The dislike-of-"Dan" hypothesis postulates that subjects may avoid some responses because the response is not a "real" word. Two subjects in Experiment II obviously avoided the "Dan" response, since their recognition curves of /d/ with this vowel do not reach 50 percent stop identification. Their avoidance of "Dan" may not have been due to their belief that it was not a real word, however. They may not have responded "Dan" because they did not hear "Dan." (The recognition curves of these subjects, 4 and 5, are presented in Appendix D.)

Although explanations have been provided for the presence of the consonant factor in the highest level ANOVA interactions, it seems that the stop-fricative continuum is independent of consonant place of articulation. Both the labials and dentals have boundaries at approximately the same durations. Both consonants respond similarly to changes in onset slope. In combination with the vowels, subjects, and amplitudes they behave differently, but this variation is relatively slight.

E. Consonant Confusions

Identifications of stops as fricatives and fricatives as stops have been induced in the present experiments. One might wonder if changes in place identification occurred along with the changes in manner identification. Carden found in his 1980 study that place identification depended, to some extent, on manner identification. In particular, Carden found the truncated voiceless dental fricative, /θ/, was often identified as a labial stop. Identification of dental fricatives as labial stops occurred only rarely in the present study.

Carden et al. suggested the reason for the labial/dental stop/fricative dependency he found is the existence of two perceptual boundaries between labials and dentals - one for stops and one for fricatives. The optimum boundary between the labial and dental stops occurs midway between the two as the boundary between the labial and dental fricatives falls midway placewise. However, the two boundaries do not coincide since the stops and fricatives have (slightly) different places of articulation. The fricative /θ/ falls on the labial side of the stop place boundary, although it is on the dental side of the fricative boundary. So when /θ/ is identified as a stop, it is categorized as a labial. This hypothesis is implausible given the difference in place of articulation of stops and fricatives. The difference is probably not sufficient to cause perceptible differences in the CV formant transitions. And the difference in place of articulation between /θ/ and /ð/, if any, would not be enough to place /θ/ on the labial side of the /b/-/d/ boundary and /ð/ on the dental side. The results of the present experiments do not support the two boundary hypothesis. No significant errors in place identification occurred, beyond what was predicted by other consonant confusion studies. Subjects did not tend to identify short, stop-like dental fricatives as labial stops. This identification occurred in only one percent of the responses. Changes in manner identification did not affect the place identification. The consonant confusions found by Carden and those found in the present experiments are more likely attributable to acoustic factors in the signals. One possible acoustic factor is lack of burst. Carden et al. propose, as an alternative to the two-boundary hypothesis, that lack of burst may be a labial cue. Thus truncated fricatives, having short noise signals with no burst, may be identified as labial stops.

The /v/-/ð/ confusions found in the present experiments are expected in light of the similarity of these two phones in intensity and duration and in light of the evidence of the prevalence of this confusion in other studies, such as Miller and Nicely (1955), Tolhurst (1954), and Wang and Bilger (1973). The labial-dental fricative confusion suggests that consonant-vowel formant transitions are not always sufficient for place identification. Labials, particularly, would be subject to misidentification on the basis of formant transitions because the short oral cavity anterior to the labial place of articulation might provide poor formant place information. On the basis of formant transitions, labials would be misperceived more often than dentals, as they are in these experiments.

"Real" English words and CV frequency of occurrence were probably other factors in the occurrence of confusions, as discussed in the sections on the vowel and the consonant effects. Kent (1979) noticed a response bias, some consonants were substituted more often then others in cases of perception errors. The four most frequently substituted consonants in Kent's study were, in descending order, /p/, /b/, /f/, and /d/. These consonants were chosen over 100 times each, constituting 57 percent of all the substitutions in Kent's study. This bias may also be in effect in these experiments, perhaps accounting for the high /d/ response to /v/ and /b/ to /ə/. It is interesting to note that the four consonants which appear as a response bias are not the most frequently appearing consonants in speech. According to Denes' (1963) data, /p/ is the fourteenth most frequently occurring consonant out of 24, /b/ is twelfth, /f/ is fifteenth, and /d/ is fourth. Apparently the response bias effect is independent of consonant occurrence frequency in speech.

Place confusions in consonant identifications were fairly low in the present studies, considering the findings of others. The amount of confusions varied, of course, according to subject. Subjects who made many errors in identification made most of these errors by responding /v/ to a /ə/ stimulus. Favouring /v/ over /ə/ may represent a bias against the consonant /ə/ or against function word responses. (A few of the identification errors can also be counted as response recording errors. Making a mistake in responding was especially easy in Experiment II where switches were pushed to indicate the choice of consonant.)

Consonant confusions in place of articulation proved to be insignificant in the present experiments. The infrequency of consonant place confusions indicates the stop-fricative continuum is independent of the place of articulation. Labial and dental consonants are both affected by changes in duration, intensity, and onset slope. Any confusions that did occur can be explained satisfactorily as functions of the acoustic properties of the signals. Dependency between the place and manner identifications need not be solved by positing a listener-internal boundary, as Carden et al. (1980) do; place and manner identifications are dependent on acoustic cues. When the acoustic signal of a labial, for example, resembles a dental, it is more likely to be identified as a dental. Likewise a dental which resembles a labial consonant will be heard as a labial. As well, fricatives which resemble stops acoustically, by being short in duration, will be recognized as stops. And stops which resemble fricatives, by being long and burstless, will be categorized as fricatives.

VI. Conclusions

A. Improvements and Further Study

Some modifications to this study are recommended in order to improve the experimental design and to increase the reliability of the results. One of these improvements is increasing the number of fricative duration conditions. As it was, there were only four durational stimuli made from fricatives: the original fricative (85 to 142 ms long), the 30 ms stimuli, the 20 ms ones, and the 10 ms ones. This range of durations did not adequately cover the stop/fricative duration boundary - 30 percent of the Experiment II boundaries fell between 30 ms and the original fricative duration. To enlarge the range to cover the entire boundary area, stimuli should also have been made with durations of 40, 50, and 60 ms. Then Experiment II would be exactly the same in terms of the number and durations of these stimuli as Experiment I, and the point at which the fricative/stop boundary falls could be judged more accurately.

Intensity in Experiment I could also have been examined more effectively. The intensity measurements of the original consonants indicated that fricatives were less intense than stops. Taking this into consideration, the way to make stops resemble fricatives would be to decrease, not increase, the intensity of the stops. This change might have caused intensity to be as statistically significant a factor in Experiment I as it was in Experiment II.

One more criticism of the design of these experiments might be made with respect to the naturalness of the presentation conditions. While investigation of specific acoustic factors cannot be performed in a completely natural, uncontrolled environment, the presentation of the stimuli can be arranged to be as representative of naturally occurring situations as possible. In this respect, presentation of an auditory signal is better via a loudspeaker than through headphones since listening through headphones is an artificial situation. As well, the stimulus items should be recorded and presented within carrier phrases. Recording this way controls the intonation and stress on the key word and helps maintain a consistent, conversational rate of production. Presentation of the items within a sentence gives the words context. Listening to a sentence is much more meaningful than hearing isolated words. For the present experiments, all the stimulus words were recorded in carrier phrases. But they were presented in sentences in Experiment I only. In Experiment II the stimuli were presented over a loudspeaker, while in I they were heard through headphones. Each experiment was natural in one of the two aspects of presentation, but neither in both.

Another consideration for experimental naturalness is the nature of the stimuli. In these experiments monosyllabic English words were used instead of nonsense syllables so that the stimuli would be meaningful. Because of the unique status of the consonant /∂/, the choice of words was limited, and it was impossible to choose all the words from the same lexical class. The voiced dental fricative occurs word initially only in function words, not in major class words. Ideally all the words involved would have been, for example, nouns. This would reduce biases in subject responses and expand the carrier phrase possibilities used in recording and presentation.

This study could also have been expanded to cover more of the acoustic aspects of stops and fricatives. For instance, onset could have been controlled more completely. In order to separate the duration changes from the onset slope changes, duration could have been increased in the stop stimuli without removing the stop's abrupt onset. At longer durations this signal might sound unnatural, or it might sound like an affricate, or, it might sound like a fricative. Another way to separate duration and onset slope is to cross the two factors so there are abrupt, burstless, and gradual onsets at every duration level. Again, some unnatural sounding stimuli might result, but allowing subjects to indicate which sounded unnatural would mark these segments. The crossing of the duration and onset slope factors would be ideal for testing the relationship between the two, seeing how each factor affects the other.

Onset abruptness also needs further study. We need to know how small an amplitude increase is perceived as an abrupt onset. A simple experiment to test this would involve increasing and decreasing the amplitude of a stop onset at the beginning of a signal. Changes in the duration of the stimuli could be varied in addition to onset amplitude, and the items might be categorized as stops or fricatives. This has been done only partially in the experiments here, only with gradual, rather than abrupt, onsets.

Finally, we must consider the validity of the 50 percent crossover point as an indication of the location of a boundary. Listeners may not have discrete boundaries they use to discriminate speech signals. Rather, they may have regions of uncertainty. For instance, consonants with durations anywhere between 20 and 30 ms may be difficult for subjects to classify since the consonants' durations are such that subjects cannot be certain about the stop/fricative status of the consonant. The 50 percent crossover point, then, indicates the approximate location of the region within which subjects are uncertain of their responses.

B. Summary and Conclusions

In summary, the primary acoustic feature which listeners use to discriminate English voiced labial and dental stops and fricatives is duration. In one experiment in this study the duration of the noise of stops was increased; in the other experiment, the duration of the noise portion of fricatives was decreased. Both the experiments indicated that listeners shift from /b/ and /d/ to

/v/ and /ð/, or vice versa, when the consonant duration is approximately 25 ms. This manner boundary varies considerably according to the subject, consonant place, and vowel involved. In the second experiment, the amplitude of the consonant also interacted in the boundary variation. The variation in both experiments, although statistically significant, is inconsequential since all the boundaries found would result in correct categorization of actually occuring fricatives and stops.

The effect of the amplitude of the consonant in the fricative-based experiment was one of increasing stop responses with increased amplitude. Since stop consonants are higher in intensity than fricatives, this effect was predictable.

In one experiment, the explosive onset of the stops was removed. Eighty percent of the resulting stimuli, burstless and gradual onset stops, were recognized as stops by subjects. Presumably, the short duration of the stops overcomes the lack of explosion. Although onset slope does not influence manner recognition of normal duration stops, we would expect that a stop burst followed by a longer duration of noise would be identified as an affricate. In English, however, there are no labial or dental affricates. Duration, therefore, is a sufficient cue to distinguish stops and fricatives at these places of articulation.

Duration is a more robust manner cue than abruptness of onset or intensity. The latter could easily be washed out by extraneous noise. Yet, while the content of a signal might not be perceived, its duration could still be estimated. Duration, then, is a reliable cue for stop/fricative identification. While onset slope and intensity may influence manner identification, the primary perceptual difference between a fricative and a stop is duration.

Bibliography

Abbs, M. S., & Minifie, F. D. Effect of acoustic cues in fricatives on perceptual confusions in preschool children. Journal of the Acoustical Society of America, 1969, 46, 1535-1542.

American Psychological Association. Publication Manual of the American Psychological Association (2nd ed.). Washington, 1974.

Baker, J. M. Quantitative characteristics of fricative and stop consonant allophones. Journal of the Acoustical Society of America, 1975, 58, S98.

Blumstein, S. E., & Stevens, K. N. Perceptual invariance and onset sprectra for stop consonants in different vowel environments. Journal of the Acoustical Society of America, 1980, 67, 648-662.

Bregman, A. S. Asking the "What for" question in auditory perception. In M. Kubovy and J. R. Pomerantz (Eds.), Perceptual organization. Hillsdale, New Jersey: Lawrence Erlbaum, 1981.

Bryans, B., & McNutt, J. A binary articulatory production classification of English consonants with derived difference measures. Journal of Speech and Hearing Disorders, 1980, 45, 346-356.

Carden, G., Levitt, A., Jusczyk, P. W., & Walley, A. Evidence for phonetic processing of cues to place of articulation: Perceived manner affects perceived place. Haskins Laboratories: Status Report on Speech Research, 1980, SR-62, 59-85.

Carney, P. J., & Moll, K. L. A cinefluorographic investigation of fricative consonant-vowel coarticulation. Phonetica, 1971, 23, 193-202.

Cutting, J. E., & Rosner, B. S. Categories and boundaries in speech and music. Perception and Psychophysics, 1974, 16, 564-570.

Delattre, P. C., Liberman, A. M., & Cooper, F. S. Acoustic loci and transitional cues for consonants. Journal of the Acoustical Society of America, 1955, 27, 769-773.

Delattre, P. C., Liberman, A. M., & Cooper, F. S. Formant transitions and loci as acoustic correlates of place of articulation in American fricatives. Studia

Linguistica, 1962, 16, 104-121.

Denes, P. Effect of duration on the perception of voicing. Journal of the Acoustical Society of America, 1955, 27, 761-764.

Denes, P. On the statistics of spoken English. Journal of the Acoustical Society of America, 1963, 35, 892-904.

Dorman, M. F., Raphael, L. J., & Isenberg, D. Acoustic cues for a fricative-affricate contrast in word-final position. Journal of Phonetics, 1980, 8, 397-405.

Dorman, M. F., Raphael, L. J., & Liberman, A. M. Further observations on the role of silence in the perception of stop consonants. Haskins Laboratories: Status Report on Speech Research, 1976, SR-48, 199-206.

Dorman, M. F., Studdert-Kennedy, M., & Raphael, L. J. Stop consonant recognition: Release bursts and formant transitions as functionally equivalent, context-dependent cues. Perception and Psychophysics, 1977, 22, 109-122.

Eimas, P. D., Tartter, V. C., Miller, J. L., & Keuthen, N. J. Asymmetric dependencies in processing phonetic features. Perception and Psychophysics, 1978, 23, 12-20.

Fischer-Jorgenson, E. Acoustic analysis of stop consonants. Miscellanea Phonetica, 1954, 2, 42-59.

French, N. R., Carter, C. W., & Koenig, W. The words and sounds of telephone conversations. Bell System Technical Journal, 1930, 9, 290-324.

Fujisaki, H., Nakamura, K., & Imoto, T. Auditory perception of duration of speech and non-speech stimuli. In G. Fant and M. A. A. Tatham (Eds.), Auditory analysis and perception of speech. London: Academic, 1975.

Gamkrelidze, T. On the correlation of stops and fricatives in a phonological system. In J. H. Greenberg (Ed.), Universals of human language (Vol. 2). Stanford, California: Stanford University Press, 1978.

Gerstman, L. J. Perceptual dimensions for the friction portions of certain speech sounds. Unpublished Ph.D. Dissertation. New York University, 1957.

Gibson, J. J. The senses considered as perceptual systems. New York: Houghton Mifflin, 1966.

Grimm, W. A. Perception of segments of English - spoken consonant-vowel syllables. Journal of the Acoustical Society of America, 1966, 40, 1454-1461.

Gupta, J. P., & Ahmad, A. Perception of (Hindi) consonants in differentiated clipped speech. Journal of the Acoustical Society of America, 1975, 58, 282-283.

Haggard, M. The devoicing of voiced fricatives. Journal of Phonetics, 1978, 6, 95-102.

Halle, M., Hughes, W. G., & Radley, J.-P. A. Acoustic properties of stop consonants. Journal of the Acoustical Society of America, 1957, 29, 107-116.

Harris, K. S. Cues for the discrimination of American English fricatives in spoken syllables. Language and Speech, 1958, 1, 1-7.

Hays, William L. Statistics. New York: Holt, Rinehart, and Winston, 1963.

Healy, A. F., & Repp, B. H. Context sensitivity and phonetic mediation in categorical perception: A comparison of four stimulus continua. Haskins Laboratories: Status Report on Speech Research, 1980, SR-63/64, 139-156.

Heinz, J. M., & Stevens, K. N. On the properties of voiceless fricative consonants. Journal of the Acoustical Society of America, 1961, 33, 589-596.

Hirsch, I. J. Auditory perception of temporal order. Journal of the Acoustical Society of America, 1959, 31, 759-767.

Howell, P., & Rosen, S. Production and perception of rise time in the voiceless affricate/fricative distinction. Journal of the Acoustical Society of America, 1983, 73, 976-984.

Huggins, A. W. F. Just-noticeable differences for segment duration in natural speech. Journal of the Acoustical Society of America, 1972, 51, 1270-1278.

Hughes, G. W., & Halle, M. Spectral properties of fricative consonants. Journal of the Acoustical Society of America, 1956, 28, 303-310.

Isenberg, D. Effect of speaking rate on the relative duration of stop closure and fricative noise. Haskins Laboratories: Status Report on Speech Research, 1978, SR-55/56, 63-79.

Jakobson, R., Fant, C. G. M., & Halle, M. Preliminaries to speech analysis. Cambridge, Massachusetts: MIT Press, 1969.

Jassem, W. The formants of fricative consonants. Language and Speech, 1965, 8, 1-15.

Jones, D. An outline of English phonetics (9th ed.). Cambridge: W. Heffer and Sons, 1972.

Kent, R. D., Wiley, T. L., & Strennen, M. L. Consonant discrimination as a function of presentation level. Audiology, 1979, 18, 212-224.

Klatt, D. H. Voice onset time, frication, and aspiration in word-initial consonant clusters. Journal of Speech & Hearing Research, 1975, 18, 686-706.

Klatt, D. H. Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. Journal of the Acoustical Society of America, 1976, 59, 1208-1221.

Klatt, D. H. Speech perception: A model of acoustic-phonetic analysis and lexical access. Journal of Phonetics, 1979, 7, 279-312.

Klatt, D. H., & Cooper, W. E. Perception of segment duration in sentence contexts. In A. Cohen and S. G. Nooteboom (Eds.), Structure and process in speech perception. New York: Springer-Verlag, 1975.

Kuhn, G. M. On the front cavity resonance and its possible role in speech perception. Journal of the Acoustical Society of America, 1975, 58, 428-433.

Lacerda, F. P. Acoustic perceptual study of the Portuguese voiceless fricatives. Journal of Phonetics, 1982, 10, 11-22.

Lariviere, C., Winitz, H., & Herriman E. The distribution of perceptual cues in English prevocalic fricatives. Journal of Speech and Hearing Research, 1975, 18, 613-622.

Liberman, A. M., Delattre, P. C., Gerstman, L. J., & Cooper, F. S. Tempo of frequency change as a cue for distinguishing classes of speech sounds. Journal of Experimental Psychology, 1956, 52, 127-137.

Liberman, A. M., & Studdert-Kennedy, M. Phonetic perception. Haskins Laboratories: Status Report on Speech Research, 1977, SR-50, 21-60.

Lieberman, P. Speech physiology and acoustic phonetics: An introduction. New York: Macmillan, 1977.

Lisker, L. On buzzing the English /b/. *Haskins Laboratories: Status Report on Speech Research*, 1978, *SR-55/56*, 181-188.

Lisker, L., Abramson, A. S., Cooper, F. S., & Schvey, M. H. Transillumination of the larynx in running speech. *Journal of the Acoustical Society of America*, 1969, *45*, 1544-1546.

Malecot, A. The force of articulation of American stops and fricatives as a function of position. *Phonetica*, 1968, *18*, 95-102.

McCasland, G. P. Noise intensity and spectrum cues of spoken fricatives. *Journal of the Acoustical Society of America*, 1979, *65*, S78-79.

Miller, G. A., & Nicely, P. E. An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America*, 1955, *27*, 338-352.

Miller, J. L., & Liberman, A. M. Some effects of later-occurring information on the perception of stop consonant and semivowel. *Haskins Laboratories: Status Report on Speech Research*, 1979, *SR-57*, 161-181.

Minifie, F. D. Speech acoustics. In Fred D. Minifie, Thomas J. Hixon, and Frederick Williams (Eds.), *Normal aspects of speech and hearing*. Englewood Cliffs, New Jersey: Prentice-Hall, 1973.

Moskowitz, B. A. The acquisition of fricatives: A study in phonetics and phonology. *Journal of Phonetics*, 1975, *3*, 141-150.

Nartey, J. On fricative phones and phonemes: Measuring the phonetic differences within and between languages. *UCLA Working Papers in Phonetics*, 1982, *55*.

Nilsen, D. L. F., & Nilsen, A. P. *Pronunciation contrasts in English*. New York: Regents, 1973.

Ohde, R. N., & Stevens, K. N. Effect of burst amplitude on the perception of stop consonant place of articulation. *Journal of the Acoustical Society of America*, 1983, *74*, 706-714.

Öhman, S. E. G. Coarticulation in VCV utterances: Spectrographic measurements. *Journal of the Acoustical Society of America*, 1966, *39*, 151-168.

Peters, G. E., & Barney, H. L. Control methods used in a study of the vowels. *Journal of the Acoustical Society* of America, 1952, 24, 175-184.

Peters, R. W. Dimensions of perception for consonants. Journal of the Acoustical Society of America, 1963, 35, 1985-1989.

Pyles, T., & Algeo, J. The origins and development of the English language (3rd ed.). New York: Harcourt Brace Jovanovich, 1982.

Raphael, L. J. Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English. Journal of the Acoustical Society of America, 1972, 51, 1296-1303.

Repp, B. H. Relative amplitude of aspiration noise as a voicing cues for syllable-initial stop consonants. Language and Speech, 1979, 22, 173-189.

Repp, B. H., Liberman, A. M., Eccardt, T., & Pesetsky, D. Perceptual integration of acoustic cues for stop, fricative, and affricate manner. Haskins Laboratories: Status Report on Speech Research, 1978, SR-52, 61-83.

Ruhlen, M. A guide to the languages of the world. Stanford University, Language Universals Project, 1976.

Sharf, D. J., & Beiter, R. C. Identification of consonants from formant transitions presented forward and backward. Language and Speech, 1974, 17, 110-118.

Sharf, D. J., & Hemeyer, T. Identification of place of consonant articulation from vowel formant transitions. Journal of the Acoustical Society of America, 1972, 51, 652-658.

Singh, S., & Black, J. W. Study of twenty-six intervocalic consonants as spoken and recognized by four language groups. Journal of the Acoustical Society of America, 1966, 39, 372-387.

Singh, S., Woods, D. R., & Becker, G. M. Perceptual structure of twenty-two prevocalic English consonants. Journal of the Acoustical Society of America, 1972, 52, 1698-1713.

Soli, S. D. Structure and duration of vowels together specify fricative voicing. Journal of the Acoustical Society of America, 1982, 72, 366-378.

Stevens, K. H. Acoustic correlates of some phonetic categories. Journal of the Acoustical Society of America, 1980, 68, 836-842.

Stevenson, D. C., & Stephens, R. C. *Alligator Reference Manual, Vol. I & II.* Edmonton, Alberta: Department of Linguistics, University of Alberta, 1978.

Strevens, P. Spectra of fricative noise in human speech. *Language and Speech*, 1960, *3*, 32-49.

Studdert-Kennedy, M. Speech perception. In Norman J. Lass (Ed.), *Contemporary issues in experimental phonetics*. New York: Academic, 1976.

Suzuki, H. Mutually complementary effect of rate and amount of formant transition in distinguishing vowel, semivowel, and stop consonant. *Massachusetts Institute of Technology Research Laboratory of Electronics Quarterly Progress Report*, 1970, *96*, 164-172.

Tobias, J. V. Relative occurrence of phonemes in American English. *Journal of the Acoustical Society of America*, 1959, *31*, 631.

Tolhurst, G. C. Audibility of the voiceless consonants as a function of intensity. *Journal of Speech and Hearing Disorders*, 1949, *14*, 210-215.

Tolhurst, G. C. Audibility-recognition sound pressure functions of the voiced cognate consonants. *Journal of Speech and Hearing Disorders*, 1954, *19*, 28-36.

Treon, M. A. Fricative and plosive perception-identification as a function of phonetic context in CVCVC utterance. *Language and Speech*, 1970, *13*, 54-64.

Trnka, B. A phonological analysis of present-day standard English. *Studies in English by Members of the English Seminar of Charles University (Prague)*, 1935, *5*.

Umeda, N. Consonant duration in American English. *Journal of the Acoustical Society of America*, 1977, *61*, 846-858.

Van Heuven, V. J. The relative contribution of rise time, steady time, and overall duration of noise bursts to the affricate-fricative distinction in English: A reanalysis of old data. *Journal of the Acoustical Society of America*, 1979, *65*, S79.

Wang, M. D., & Bilger, R. C. Consonant confusions in noise: A study of perceptual features. *Journal of the Acoustical Society of America*, 1973, *54*, 1248-1265.

Wang, W. S.-Y., & Crawford, J. Frequency studies of English consonants. *Language and Speech*, 1960, *31*, 131-139.

Weismer, G. Control of the voicing distinction for intervocalic stops and fricatives: Some data and theoretical considerations. *Journal of Phonetics*, 1980, *8*, 427-438.

A. "Randomize"

```
C THIS IS A PROGRAM TO RANDOMIZE
C A LIST OF SEGMENT NAMES.
C THE SEGMENT NAMES ARE IN A SOURCE
C FILE CALLED SEGNAMES.
C THE RANDOMIZED LIST GOES INTO
C THE SINK FILE CALLED RAND.
C
CLEAR VI
C THE ARRAY VALUE IN VARLIST IS THE
C NUMBER OF SEGMENTS TO BE RANDOMIZED
DATA VARLIST(28)*8
DATA I 1
C
SOURCE SEGNAMES
LABEL 1
READ *SOURCE &VARLIST(&I)
ADD &I 1
IF &I EQ 29 GOTO 2
GOTO 1
C
C RANDOMIZING
LABEL 2
EM D0:RAND
SINK D0:RAND
C THE FIRST PARAMETER AFTER *SINK IN THE
C RAND COMMAND IS THE SEED NUMBER
C THE SECOND IS THE NUMBER OF REPLICATIONS,
C AND THE THIRD IS THE LIST TO BE RANDOMIZED.
RAND *SINK 1000 3 &VARLIST
REL *SINK
END
```

B. "Playrand"

```
C A PROGRAM TO PLAY BACK A
C LIST OF SEGMENTS, WHETHER
C RANDOMIZED OR NOT.
DATA I 1
DATA SEG*8
DATA FILE*6
SOURCE SFNAME
LABEL 1
CL WA
READ *SOURCE &FILE &SEG
GET D1:&FILE
LQ &SEG
P
WAIT 1 SEC
IF &I EQ 10 GOTO 2
ADD &I 1
GOTO 1
LABEL 2
WAIT 3 SEC
SU &I 9
GOTO 1
END
```

C. "Measure"

```
C A PROGRAM TO MEASURE A
C SIGNAL'S DURATION, PEAK,
C AND AREA.
C THE SEGMENT NAMES ARE IN
C A FILE CALLED SEGNAMES,
C BUT THE SEGMENTS THEMSELVES ARE
C IN A FILE CALLED SEGFIL.
DATA X
DATA Y
DATA SEG*8
SOURCE D1:SEGNAMES
GET D1:SEGFIL
LABEL 1
CL WA
READ *SOURCE &SEG
L &SEG
MEAS DUR &X &Y MSEC
PRINT &SEG
PRINT DURATION= &X &Y
MEAS PEAK &X &Y
PRINT PEAK= &X &Y
MEAS AREA &X &Y
IFERROR 50 PRINT 50
PRINT AREA= &X &Y
GOTO 1
END
```

D. "Derand"

```
C A PROGRAM TO DERANDOMIZE A LIST
DATA J 0
DATA I 1
DATA CR
DATA K
DATA RESP
DATA STIMNO
DATA FILE*6
DATA FILE2*6
WRITE *TTY NAME OF FILE TO BE DERANDOMIZED?
READ *TTY &FILE2
WRITE *TTY NAME OF SINK FILE?
READ *TTY &FILE
WRITE *TTY STARTING AT LINE?
READ *TTY &K
CRE D1:&FILE SIZE=3
SINK D1:&FILE
LABEL 2
LET &J 0
SOURCE D1:&FILE2(1)(&K)
LABEL 1
ADD &J 1
C IN THIS INSTANCE, THE NUMBER OF ITEMS
C IN A REPLICATION IS 96.
IF &J EQ 97 GOTO 2
READ *SOURCE &STIMNO &RESP
IF &STIMNO NE &I GOTO 1
WRITE *SINK &RESP
ADD &I 1
IF &L LT 97 GOTO 1
END
```

E. "Switchbox"

```
C A PROGRAM TO PLAY THE SIGNALS
C FOR A CATEGORIZATION EXPERIMENT
C AND READ THE RESPONSES MADE ON A
C TOUCHPAD RESPONSE BOX.
C SOURCE FILE MUST CONTAIN:
C 1) RANDOMIZATION NUMBERS
C 2) NAME OF FILE WHERE THE SEGMENT IS STORED
C 3) NAME OF SEGMENT.
C THE SEGMENT NAMES ARE ALREADY IN RANDOM ORDER
SINK D1:DATFIL
SOURCE D2:SOUFIL
CL AL
DATA SEG*8
DATA FILE*6
DATA MASK(12) 1 2 3 4 5 6 7 8 9 10
11 12
DATA AV(12)
DATA MAXSW 12
DATA RESP
DATA SNUM
DATA I
DATA J
DATA CNT
LABEL 2
ADD &I 1
READ *SOURCE &SNUM &FILE &SEG
IFERROR 43 GOTO 6
GET D1:&FILE
LQ &SEG
LABEL 3
P
WAIT 50 MS
PULSE 5 512 5 MS
WAIT 20 MS
PULSE 5 0 5 MS
RDSW XL: 1 &MASK &AV
PULSE 5 512 50 MS
PULSE 5 0 5 MS
LET &J 1
LET &CNT 0
LABEL 4
IF &AV(&J) EQ 4095 GOTO 5
LET &RESP &J
ADD &CNT 1
LABEL 5
ADD &J 1
IF &J LE &MAXSW GOTO 4
IF &CNT NE 1 GOTO 3
IF &RESP GT &MAXSW GOTO 3
WAIT 200 MS
WRITE *SINK &SNUM &SEG &RESP
CL WA
GOTO 2
```

```
LABEL 6
REL
END
```

F. "Swscore"

```
C A PROGRAM TO DERANDOMIZE AND
C AND REORDER INTO A RESPONSE
C ARRAY THE RESPONSES GIVEN TO A
C CATEGORIZATION EXPERIMENT USING
C A TOUCHPAD RESPONSE BOX.
C (SEE "SWITCHBOX" PROGRAM.)
DATA I
DATA J
DATA SNUM
DATA RESP(12)
DATA SEG*8
DATA R
DATA K
DATA M
DATA SOUF(4)*6 SUB11 SUB12 SUB21 SUB31
DATA SINKF(4)*6 S11 S12 S21 S31
LET &M 0
LABEL 9
ADD &M 1
IF &M EQ 5 GOTO 7
CRE D2:&SINKF(&M) SIZE=11
SINK D2:&SINKF(&M)
LET &J 0
LET &K 0
LABEL 1
SOURCE D2:&SOUF(&M)
ADD &J 1
C IN THIS INSTANCE THE NUMBER OF ITEMS
C IN A REPLICATION IS 72
IF &J EQ 73 GOTO 5
LET &I 0
LABEL 2
ADD &I 1
C IN THIS INSTANCE THE NUMBER OF REPLICATIONS
C IN A RUN IS 3.
IF &I EQ 4 GOTO 4
LABEL 3
READ *SOURCE &SNUM &SEG &R
IF &SNUM NE &J GOTO 3
IF &R EQ 1 ADD &RESP(1) 1
IF &R EQ 5 ADD &RESP(2) 1
IF &R EQ 9 ADD &RESP(3) 1
IF &R EQ 2 ADD &RESP(4) 1
IF &R EQ 6 ADD &RESP(5) 1
IF &R EQ 10 ADD &RESP(6) 1
IF &R EQ 3 ADD &RESP(7) 1
IF &R EQ 7 ADD &RESP(8) 1
IF &R EQ 11 ADD &RESP(9) 1
IF &R EQ 4 ADD &RESP(10) 1
IF &R EQ 8 ADD &RESP(11) 1
IF &R EQ 12 ADD &RESP(12) 1
GOTO 2
LABEL 4
```

```
WRITE *SINK &SNUM &SEG
WRITE *SINK &RESP
C THE FOLLOWING LOOP RESETS THE
C VALUES IN THE RESPONSE ARRAY TO ZERO.
LET &K 1
LABEL 6
LET &RESP(&K) 0
ADD &K 1
IF &K LT 13 GOTO 6
GOTO 1
LABEL 5
GOTO 9
LABEL 7
REL
END
```

VIII. Appendix B: Experiment I Response Sheets

Janet Schwegel's Experiment Name: ________________ Date: ____________
83/08

Please circle the word heard at the end of the sentence "Please say the word ______." Also, indicate your confidence in your answer, on a scale of one to three, by circling the appropriate number.

The first ten are practice sentences. There is a pause after every tenth.

Scale: 1 (not very sure) 2 (fairly sure) 3 (very sure)

a	bat	vat	1	2	3
b	though	dough	1	2	3
c	dough	though	1	2	3
d	thee	dee ("D")	1	2	3
e	bee	vee ("V")	1	2	3
f	vat	bat	1	2	3
g	dee	thee	1	2	3
h	than	Dan	1	2	3
i	boat	vote	1	2	3
j	vee	bee	1	2	3
1	bat	vat	1	2	3
2	than	Dan	1	2	3
3	dough	though	1	2	3
4	though	dough	1	2	3
5	bat	vat	1	2	3
6	though	dough	1	2	3
7	dough	though	1	2	3
8	though	dough	1	2	3
9	dee	thee	1	2	3
10	thee	dee	1	2	3
11	dee	thee	1	2	3
12	vee	bee	1	2	3
13	bat	vat	1	2	3
14	thee	dee	1	2	3
15	boat	vote	1	2	3
16	vote	boat	1	2	3
17	Dan	than	1	2	3
18	though	dough	1	2	3
19	Dan	than	1	2	3
20	though	dough	1	2	3
21	bee	vee	1	2	3
22	thee	dee	1	2	3
23	bee	vee	1	2	3
24	vat	bat	1	2	3
25	dee	thee	1	2	3
26	vote	boat	1	2	3
27	bee	vee	1	2	3
28	vote	boat	1	2	3
29	dough	though	1	2	3
30	vat	bat	1	2	3
31	bat	vat	1	2	3
32	though	dough	1	2	3
33	dough	though	1	2	3
34	thee	dee	1	2	3
35	bee	vee	1	2	3
36	vat	bat	1	2	3
37	dee	thee	1	2	3
38	than	Dan	1	2	3
39	boat	vote	1	2	3
40	vee	bee	1	2	3
41	bat	vat	1	2	3
42	thee	dee	1	2	3
43	boat	vote	1	2	3
44	thee	dee	1	2	3
45	bee	vee	1	2	3
46	vat	bat	1	2	3
47	dee	thee	1	2	3
48	vee	bee	1	2	3
49	dee	thee	1	2	3
50	vee	bee	1	2	3

Scale: 1 (not very sure) 2 (fairly sure) 3 (very sure)

51	bee	vee	1 2 3	81	bat	vat	1 2 3
52	vote	boat	1 2 3	82	though	dough	1 2 3
53	Dan	than	1 2 3	83	boat	vote	1 2 3
54	thee	dee	1 2 3	84	vat	bat	1 2 3
55	dough	though	1 2 3	85	boat	vote	1 2 3
56	though	dough	1 2 3	86	than	Dan	1 2 3
57	Dan	than	1 2 3	87	dough	though	1 2 3
58	vote	boat	1 2 3	88	vote	boat	1 2 3
59	Dan	than	1 2 3	89	Dan	than	1 2 3
60	thee	dee	1 2 3	90	vat	bat	1 2 3
61	Dan	than	1 2 3	91	boat	vote	1 2 3
62	than	Dan	1 2 3	92	vat	bat	1 2 3
63	bat	vat	1 2 3	93	boat	vote	1 2 3
64	thee	dee	1 2 3	94	vee	bee	1 2 3
65	boat	vote	1 2 3	95	bee	vee	1 2 3
66	vote	boat	1 2 3	96	than	Dan	1 2 3
67	dee	thee	1 2 3	97	bee	vee	1 2 3
68	than	Dan	1 2 3	98	vat	bat	1 2 3
69	bee	vee	1 2 3	99	bat	vat	1 2 3
70	vee	bee	1 2 3	100	though	dough	1 2 3
71	bat	vat	1 2 3	101	bee	vee	1 2 3
72	vee	bee	1 2 3	102	thee	dee	1 2 3
73	Dan	than	1 2 3	103	bat	vat	1 2 3
74	though	dough	1 2 3	104	vat	bat	1 2 3
75	boat	vote	1 2 3	105	dough	though	1 2 3
76	than	Dan	1 2 3	106	thee	dee	1 2 3
77	Dan	than	1 2 3	107	dee	thee	1 2 3
78	vat	bat	1 2 3	108	vee	bee	1 2 3
79	dough	though	1 2 3	109	bat	vat	1 2 3
80	vee	bee	1 2 3	110	vee	bee	1 2 3

. . . . 3

Scale: 1 (not very sure) 2 (fairly sure) 3 (very sure)

111	dough	though	1	2	3
112	thee	dee	1	2	3
113	Dan	than	1	2	3
114	than	Dan	1	2	3
115	boat	vote	1	2	3
116	vee	bee	1	2	3
117	bat	vat	1	2	3
118	vat	bat	1	2	3
119	bee	vee	1	2	3
120	though	dough	1	2	3

121	Dan	than	1	2	3
122	vat	bat	1	2	3
123	Dan	than	1	2	3
124	though	dough	1	2	3
125	bat	vat	1	2	3
126	vee	bee	1	2	3
127	bee	vee	1	2	3
128	vee	bee	1	2	3
129	dee	thee	1	2	3
130	vote	boat	1	2	3

131	boat	vote	1	2	3
132	than	Dan	1	2	3
133	dee	thee	1	2	3
134	vote	boat	1	2	3
135	dee	thee	1	2	3
136	vee	bee	1	2	3
137	Dan	than	1	2	3
138	vote	boat	1	2	3
139	boat	vote	1	2	3
140	vat	bat	1	2	3

141	dee	thee	1	2	3
142	vote	boat	1	2	3
143	dee	thee	1	2	3
144	though	dough	1	2	3
145	bat	vat	1	2	3
146	than	Dan	1	2	3
147	boat	vote	1	2	3
148	vote	boat	1	2	3
149	dough	though	1	2	3
150	vote	boat	1	2	3

151	bee	vee	1	2	3
152	though	dough	1	2	3
153	dee	thee	1	2	3
154	vee	bee	1	2	3
155	bee	vee	1	2	3
156	though	dough	1	2	3
157	Dan	than	1	2	3
158	than	Dan	1	2	3
159	dee	thee	1	2	3
160	thee	dee	1	2	3

161	bat	vat	1	2	3
162	vat	bat	1	2	3
163	dee	thee	1	2	3
164	though	dough	1	2	3
165	boat	vote	1	2	3
166	though	dough	1	2	3
167	boat	vote	1	2	3
168	thee	dee	1	2	3
169	dough	though	1	2	3
170	than	Dan	1	2	3

. . . 4

Scale: 1 (not very sure) 2 (fairly sure) 3 (very sure)

171	bat	vat	1	2	3	201	bat	vat	1	2	3
172	than	Dan	1	2	3	202	than	Dan	1	2	3
173	boat	vote	1	2	3	203	dough	though	1	2	3
174	though	dough	1	2	3	204	vee	bee	1	2	3
175	bee	vee	1	2	3	205	dee	thee	1	2	3
176	vat	bat	1	2	3	206	vat	bat	1	2	3
177	boat	vote	1	2	3	207	Dan	than	1	2	3
178	vee	bee	1	2	3	208	thee	dee	1	2	3
179	Dan	than	1	2	3	209	bee	vee	1	2	3
180	though	dough	1	2	3	210	vee	bee	1	2	3

181	boat	vote	1	2	3	211	bee	vee	1	2	3
182	thee	dee	1	2	3	212	vat	bat	1	2	3
183	Dan	than	1	2	3	213	boat	vote	1	2	3
184	vat	bat	1	2	3	214	than	Dan	1	2	3
185	bee	vee	1	2	3	215	Dan	than	1	2	3
186	though	dough	1	2	3	216	thee	dee	1	2	3
187	boat	vote	1	2	3	217	bee	vee	1	2	3
188	thee	dee	1	2	3	218	thee	dee	1	2	3
189	Dan	than	1	2	3	219	boat	vote	1	2	3
190	though	dough	1	2	3	220	though	dough	1	2	3

191	Dan	than	1	2	3	221	boat	vote	1	2	3
192	than	Dan	1	2	3	222	thee	dee	1	2	3
193	dee	thee	1	2	3	223	bat	vat	1	2	3
194	thee	dee	1	2	3	224	vee	bee	1	2	3
195	boat	vote	1	2	3	225	boat	vote	1	2	3
196	vee	bee	1	2	3	226	vee	bee	1	2	3
197	Dan	than	1	2	3	227	bee	vee	1	2	3
198	though	dough	1	2	3	228	vee	bee	1	2	3
199	boat	vote	1	2	3	229	boat	vote	1	2	3
200	vat	bat	1	2	3	230	though	dough	1	2	3

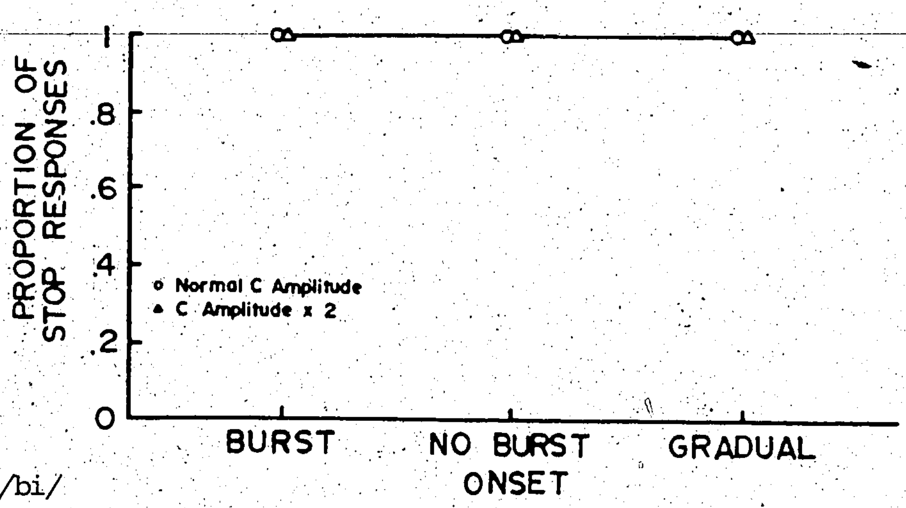
. . . 5

Scale: 1 (not very sure) 2 (fairly sure) 3 (very sure)

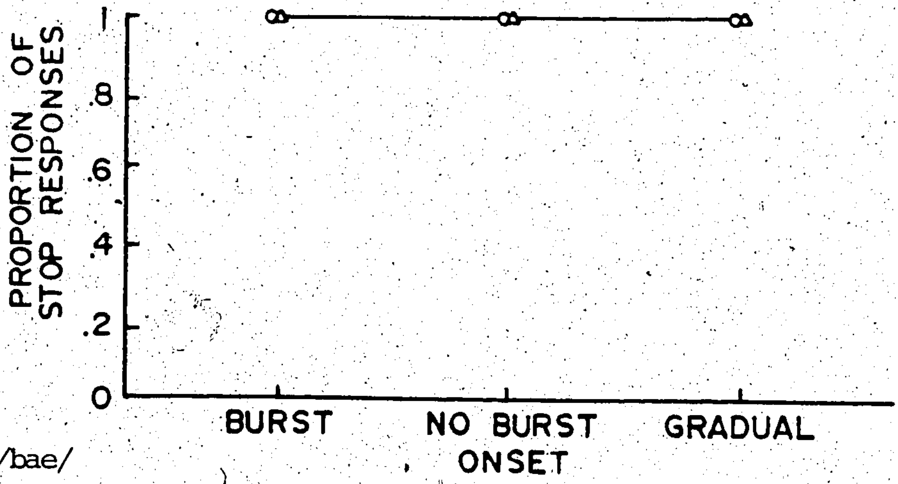
231	Dan	than	1	2	3
232	than	Dan	1	2	3
233	Dan	than	1	2	3
234	vote	boat	1	2	3
235	dee	thee	1	2	3
236	vat	bat	1	2	3
237	bat	vat	1	2	3
238	vee	bee	1	2	3
239	boat	vote	1	2	3
240	vat	bat	1	2	3
241	bat	vat	1	2	3
242	though	dough	1	2	3
243	dough	though	1	2	3
244	than	Dan	1	2	3
245	dough	though	1	2	3
246	thee	dee	1	2	3
247	bat	vat	1	2	3
248	vat	bat	1	2	3
249	bee	vee	1	2	3
250	vee	bee	1	2	3
251	bee	vee	1	2	3
252	than	Dan	1	2	3
253	dough	though	1	2	3
254	than	Dan	1	2	3
255	bee	vee	1	2	3
256	vote	boat	1	2	3
257	dough	though	1	2	3
258	than	Dan	1	2	3
259	dee	thee	1	2	3
260	vee	bee	1	2	3
261	Dan	than	1	2	3
262	though	dough	1	2	3
263	dough	though	1	2	3
264	than	Dan	1	2	3
265	dough	though	1	2	3
266	vat	bat	1	2	3
267	dough	though	1	2	3
268	than	Dan	1	2	3
269	bat	vat	1	2	3
270	thee	dee	1	2	3
271	bat	vat	1	2	3
272	thee	dee	1	2	3
273	Dan	than	1	2	3
274	thee	dee	1	2	3
275	boat	vote	1	2	3
276	vat	bat	1	2	3
277	boat	vote	1	2	3
278	though	dough	1	2	3
279	boat	vote	1	2	3
280	thee	dee	1	2	3
281	dee	thee	1	2	3
282	though	dough	1	2	3
283	dee	thee	1	2	3
284	vote	boat	1	2	3
285	dough	though	1	2	3
286	vote	boat	1	2	3
287	boat	vote	1	2	3
288	vat	bat	1	2	3
289	bee	vee	1	2	3
290	than	Dan	1	2	3

END

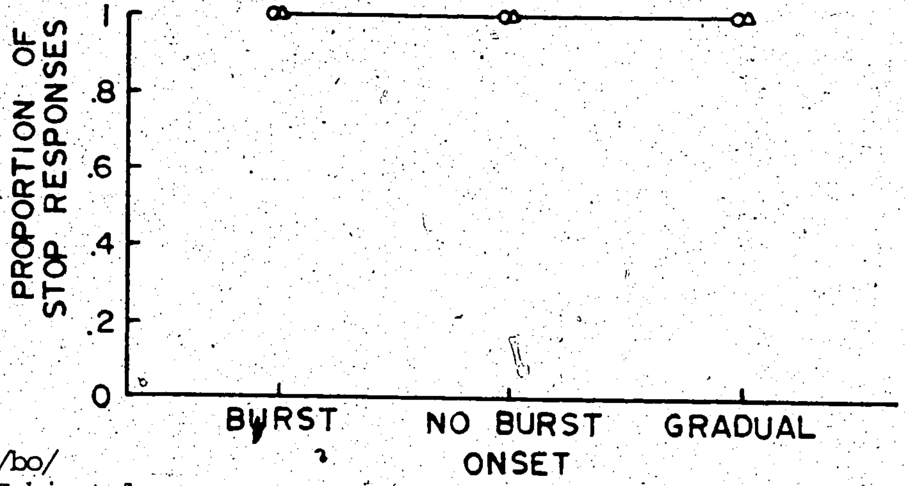
IX. Appendix C: Experiment I Individual Subject Recognition Curves



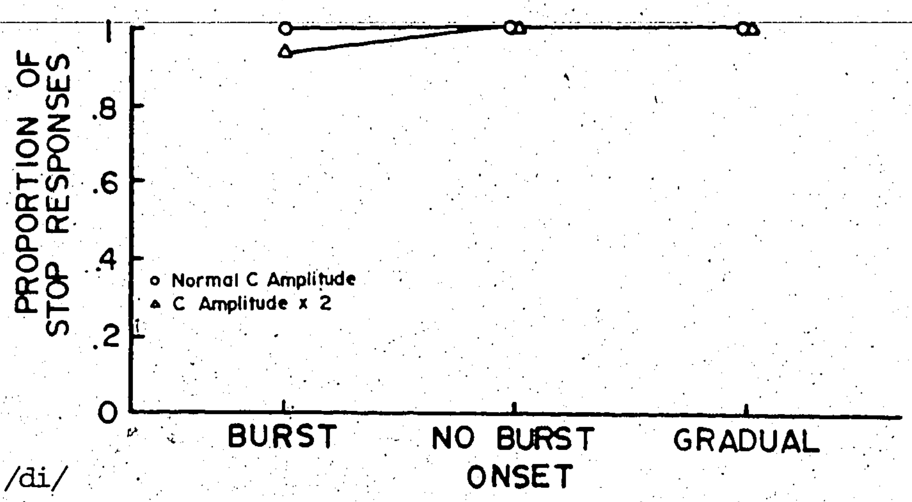
/bi/



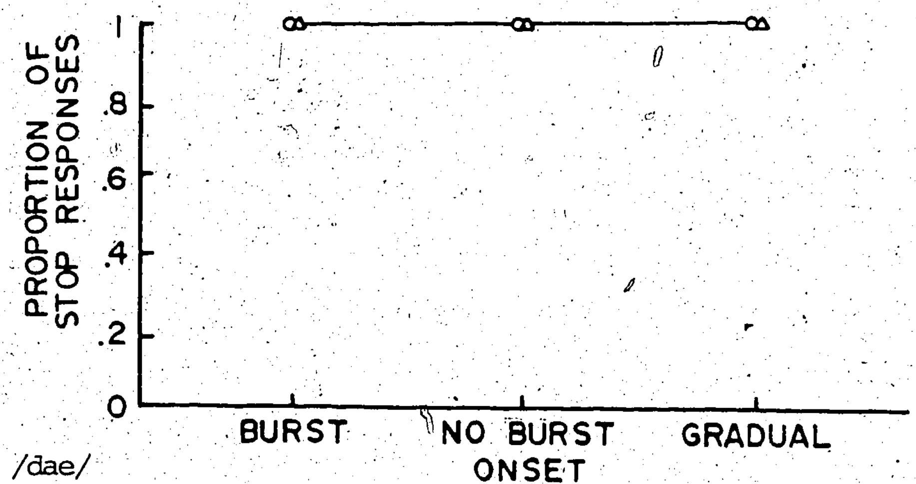
/bae/



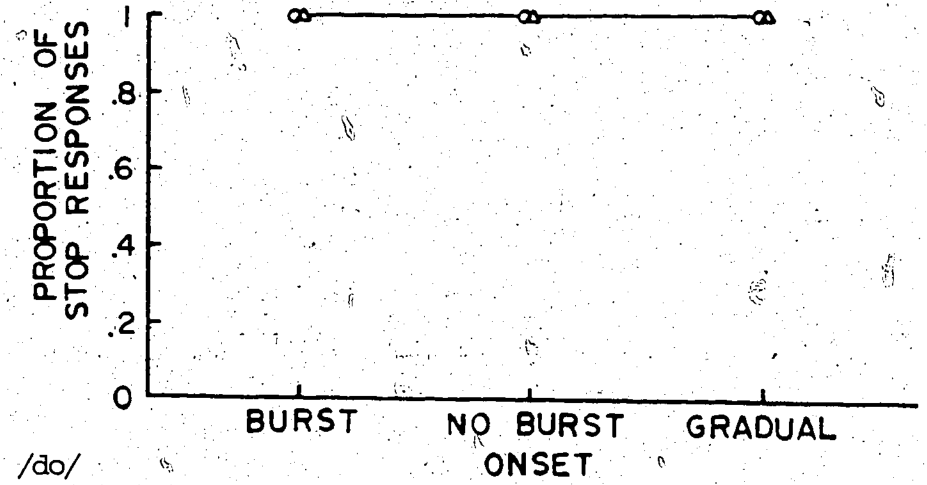
/bo/
Subject 1



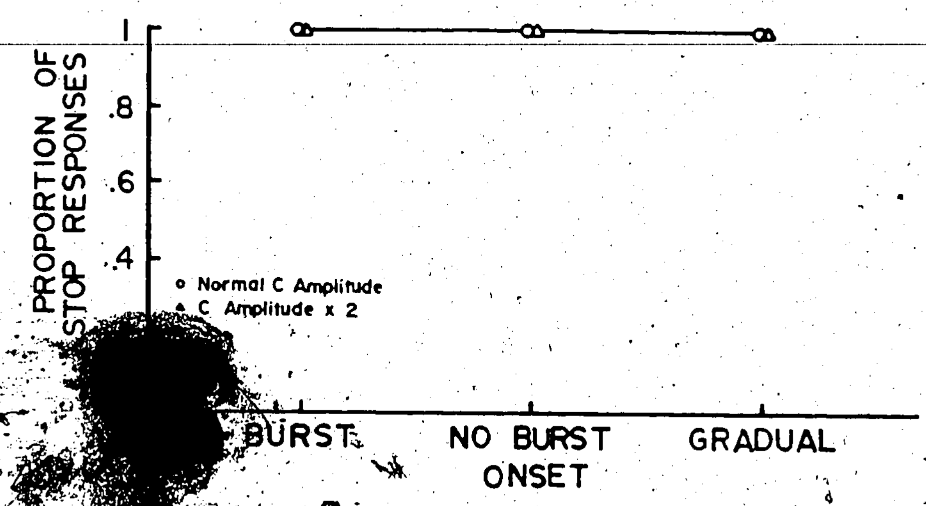
/di/

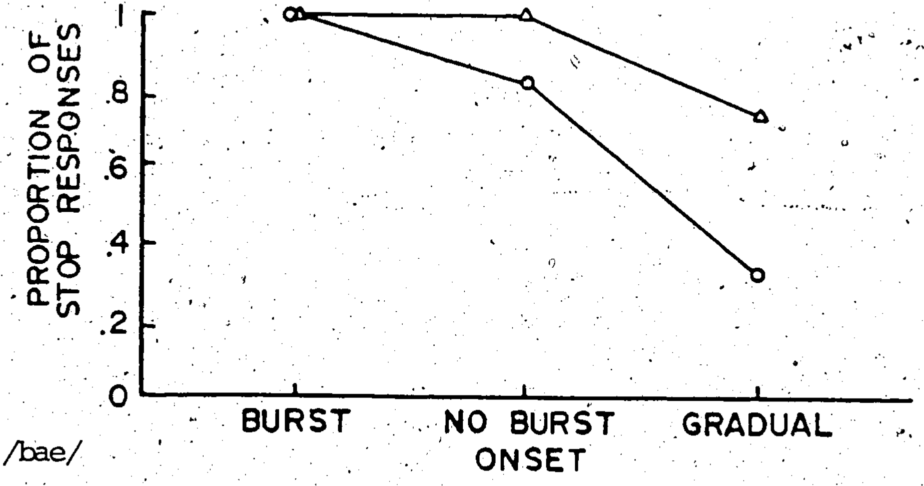


/dae/

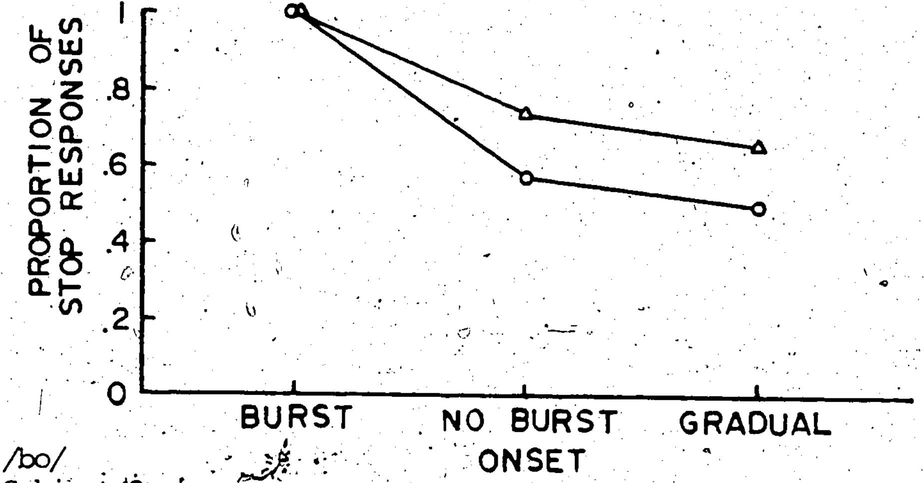


/do/
Subject 1

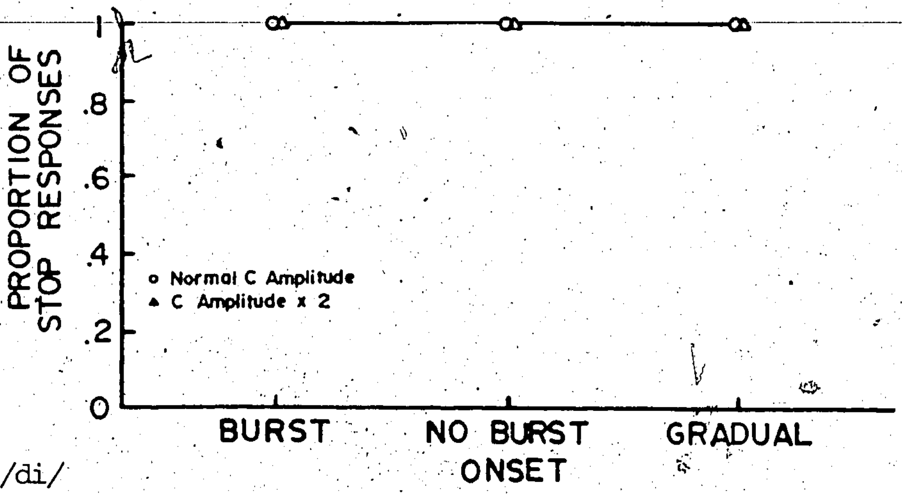




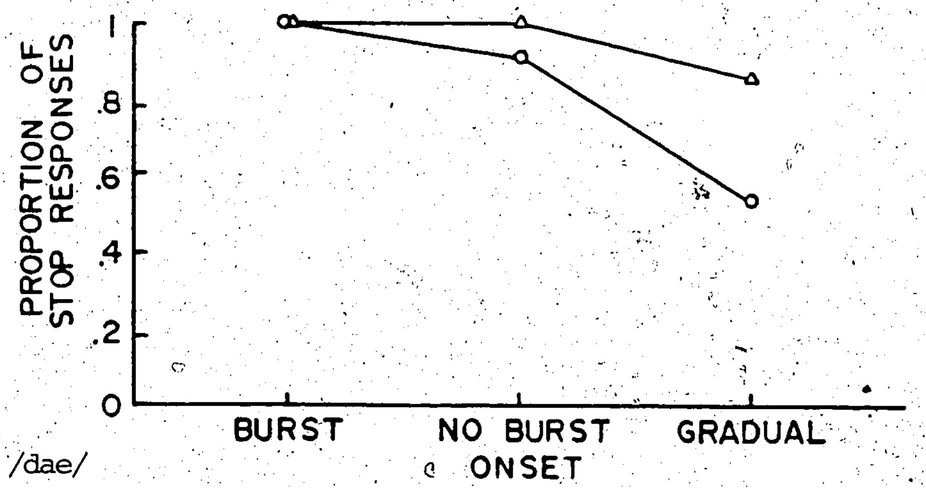
/bae/



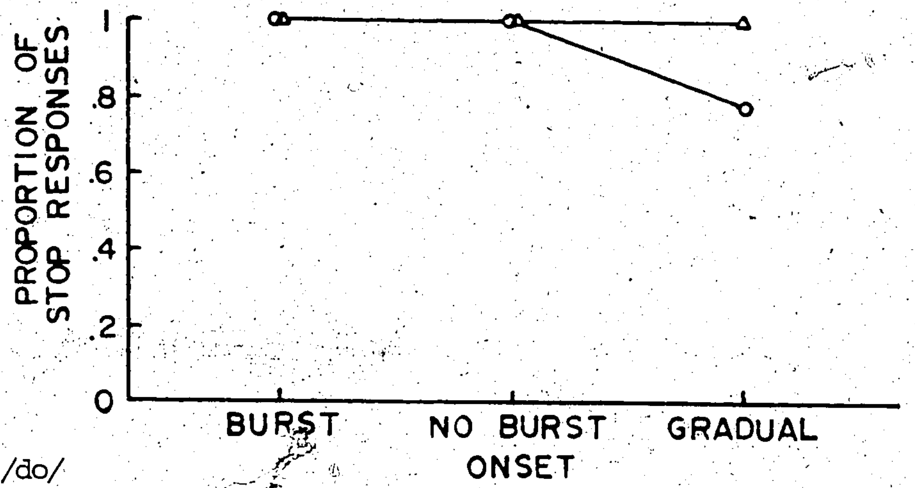
/bo/
Subject 2



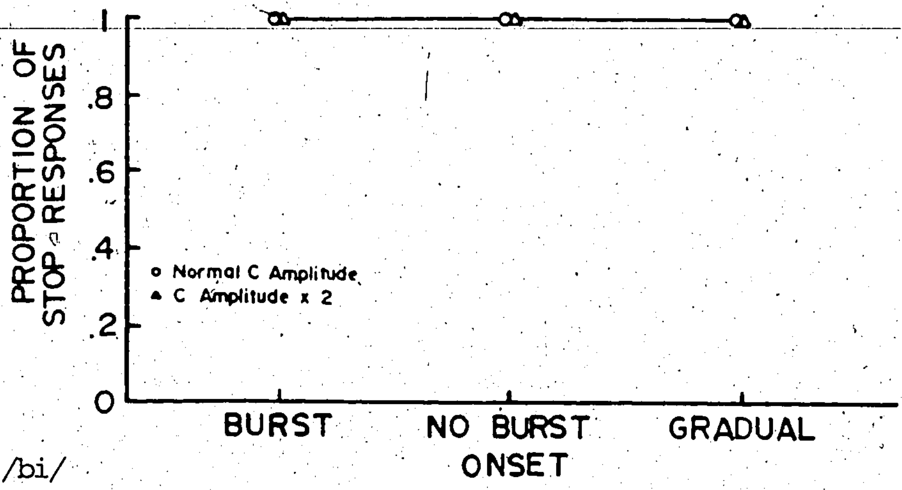
/di/



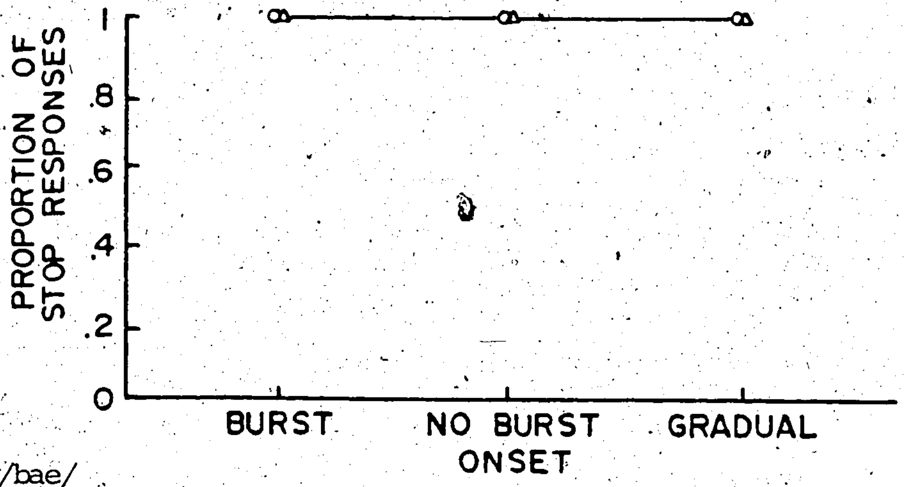
/dae/



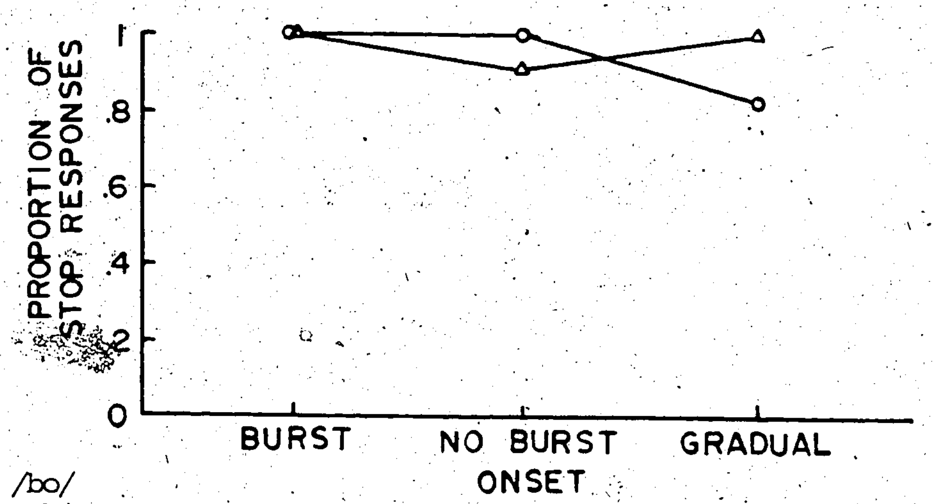
/do/
Subject 2



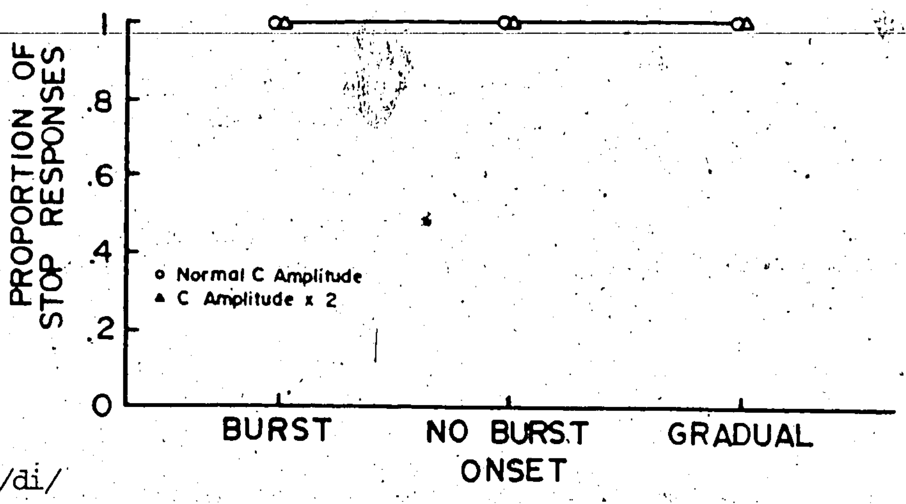
/bi/



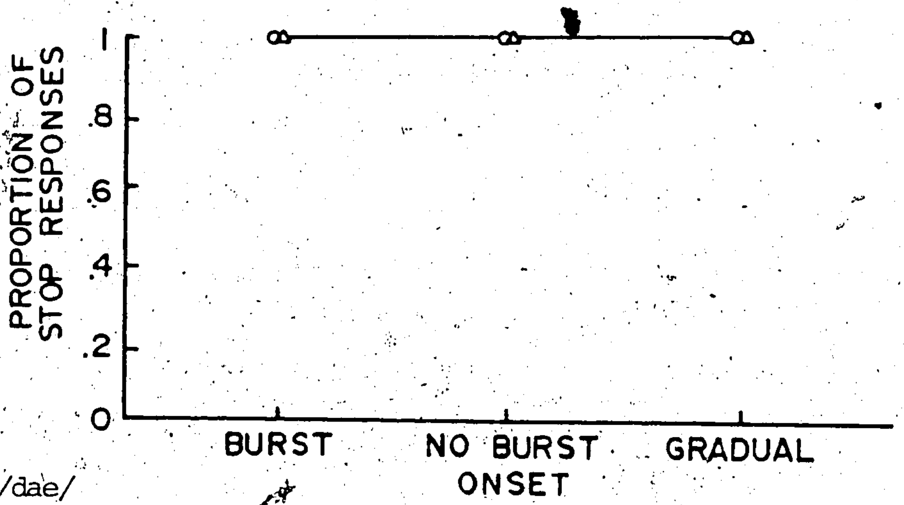
/bae/



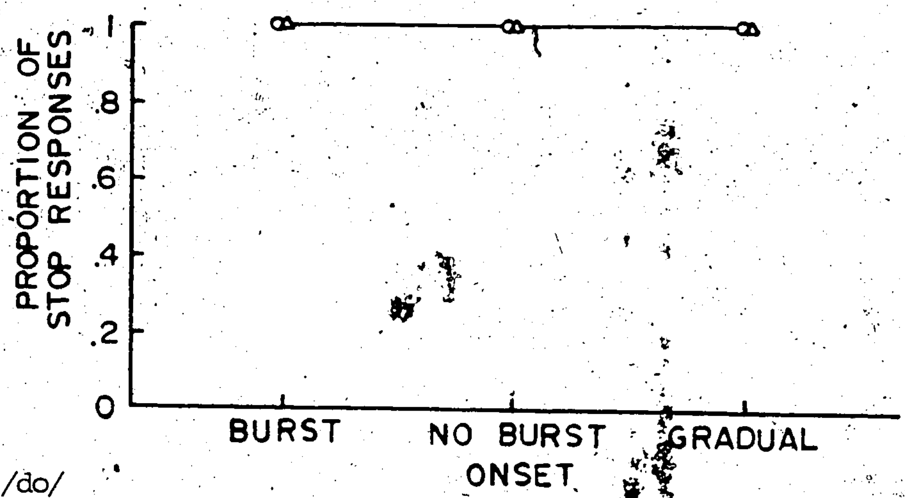
/bo/
Subject 3



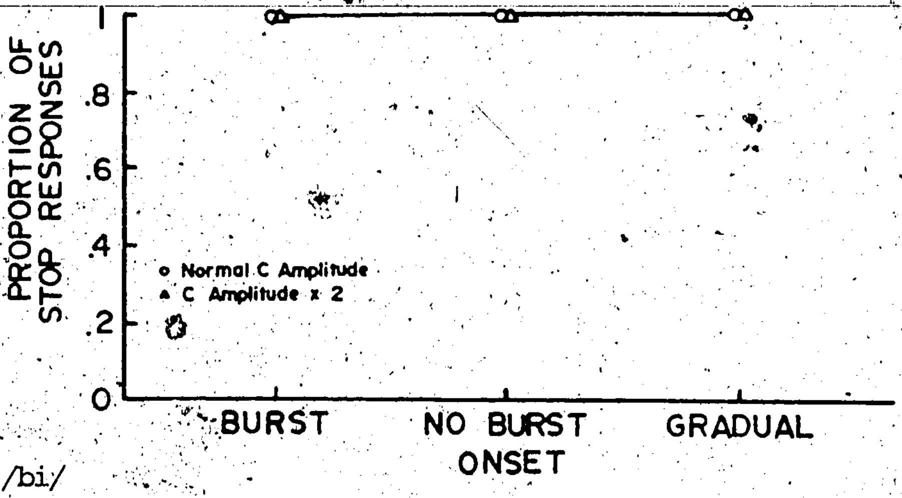
/di/



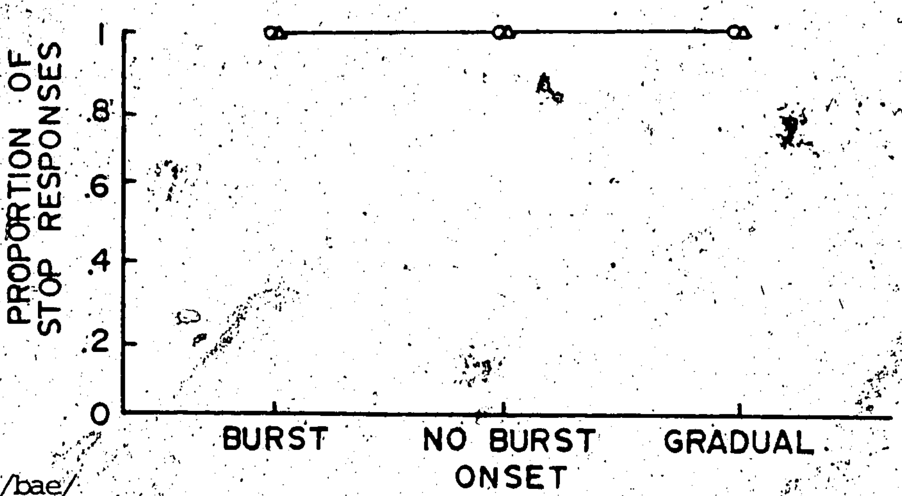
/dae/



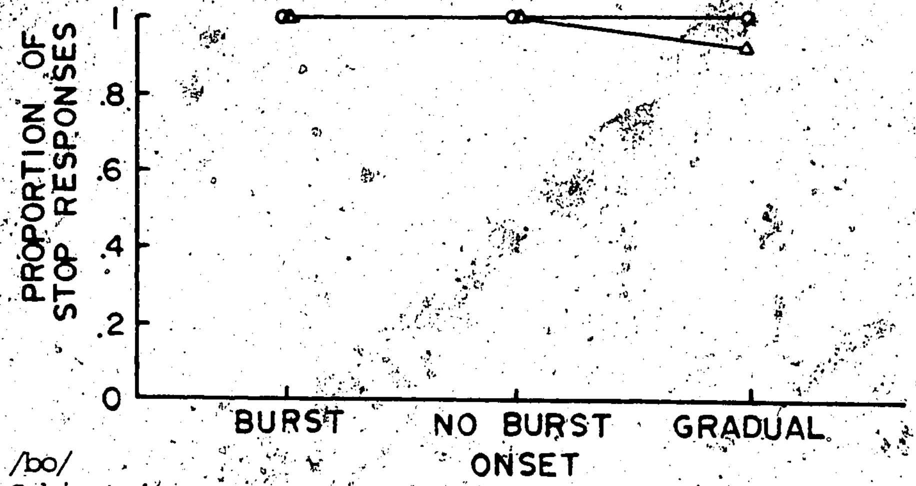
/do/
Subject 3



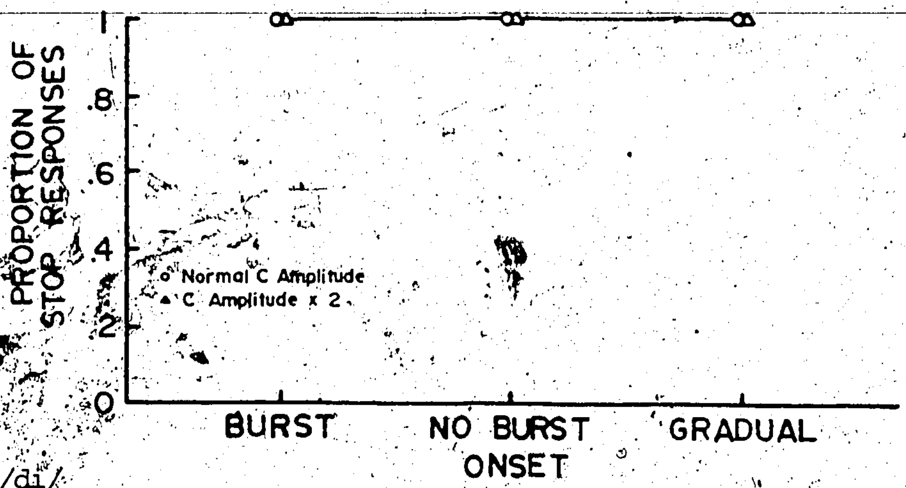
/bi/



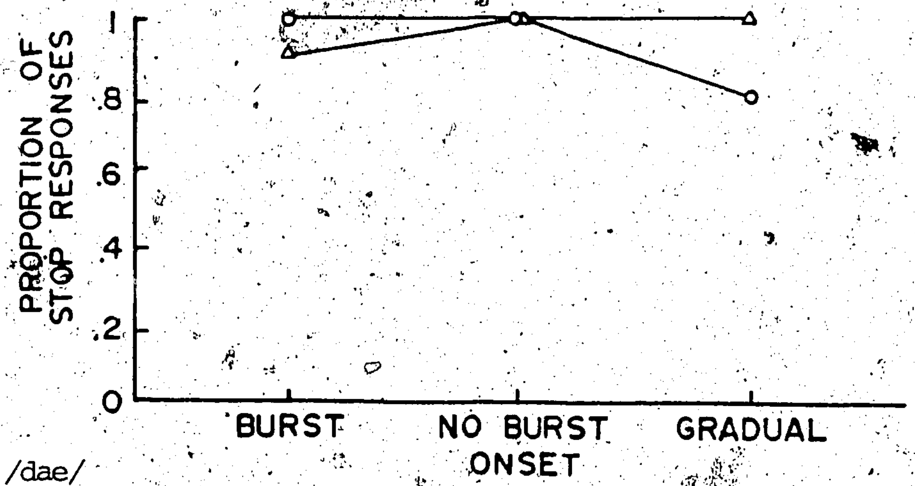
/bae/



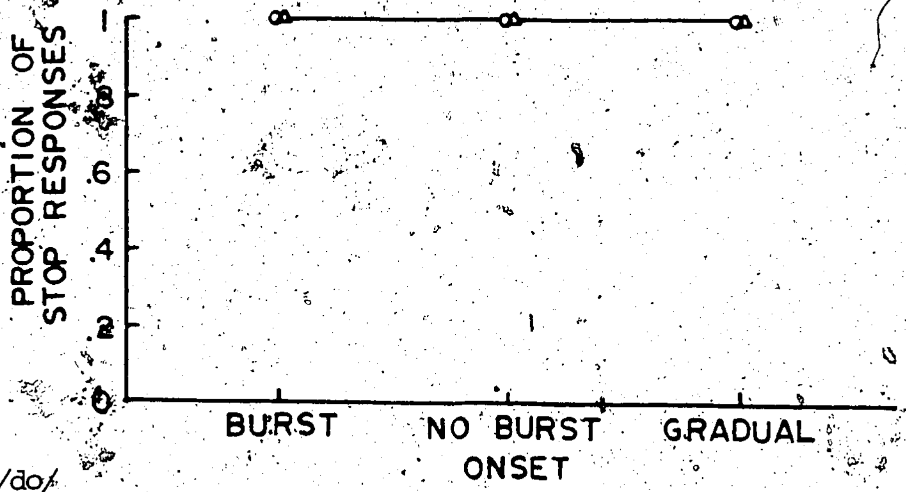
/bo/
Subject 4



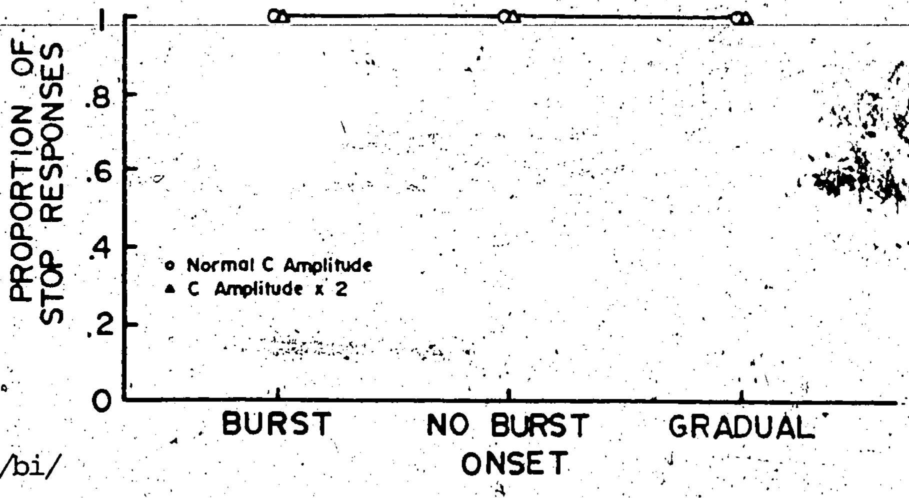
/di/



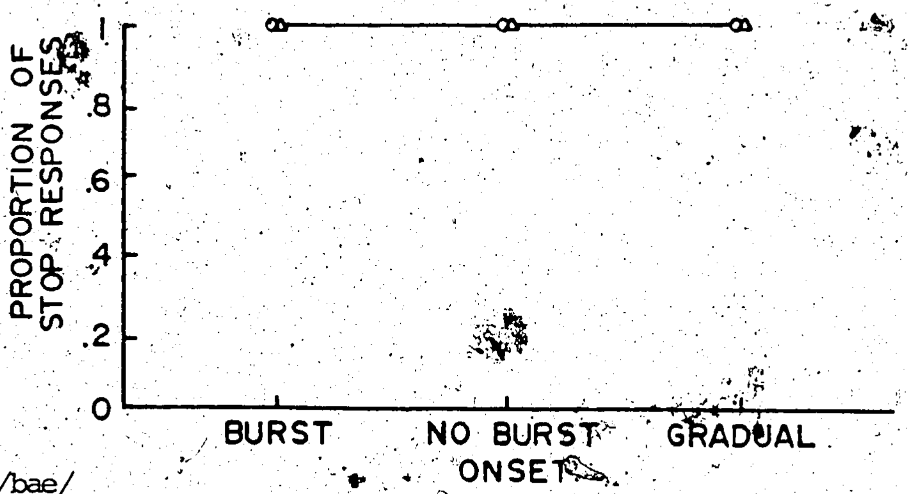
/dae/



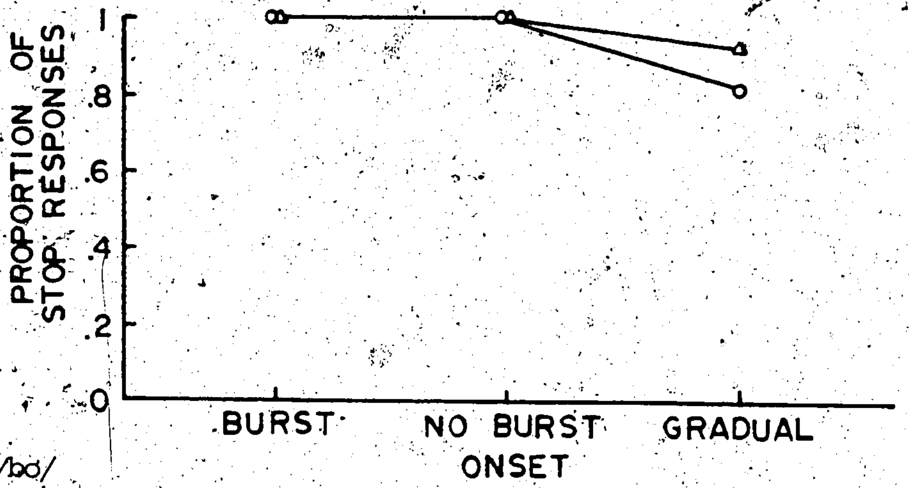
/do/
Subject 4



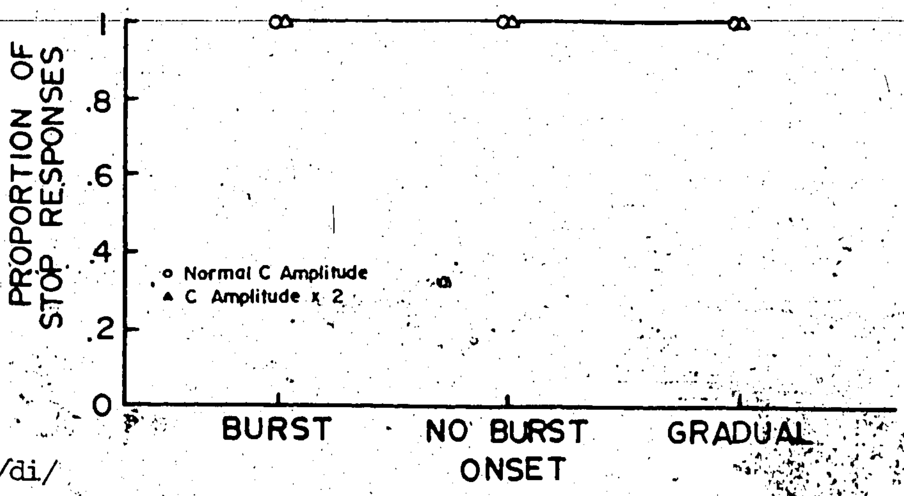
/bi/



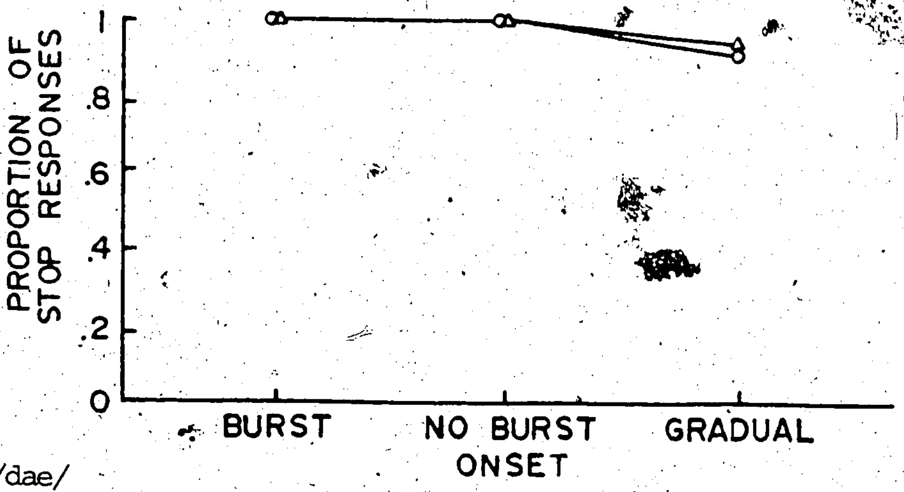
/bae/



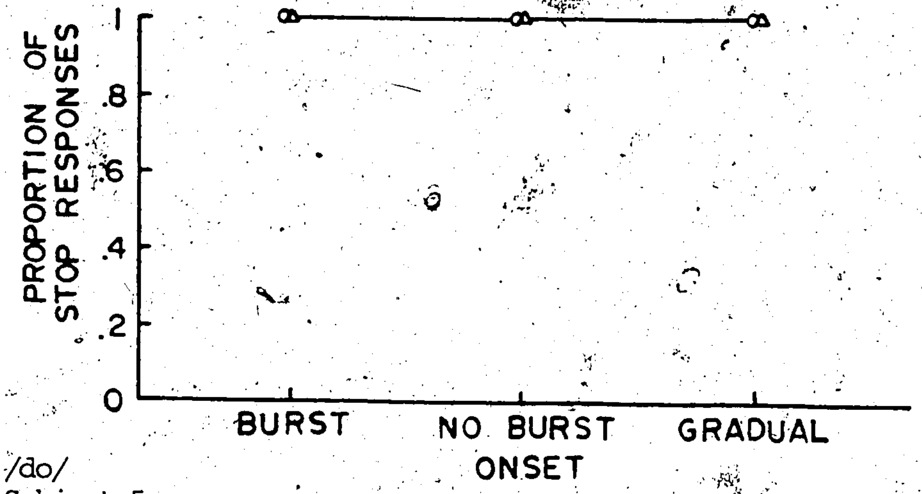
/bɔ/
Subject 5



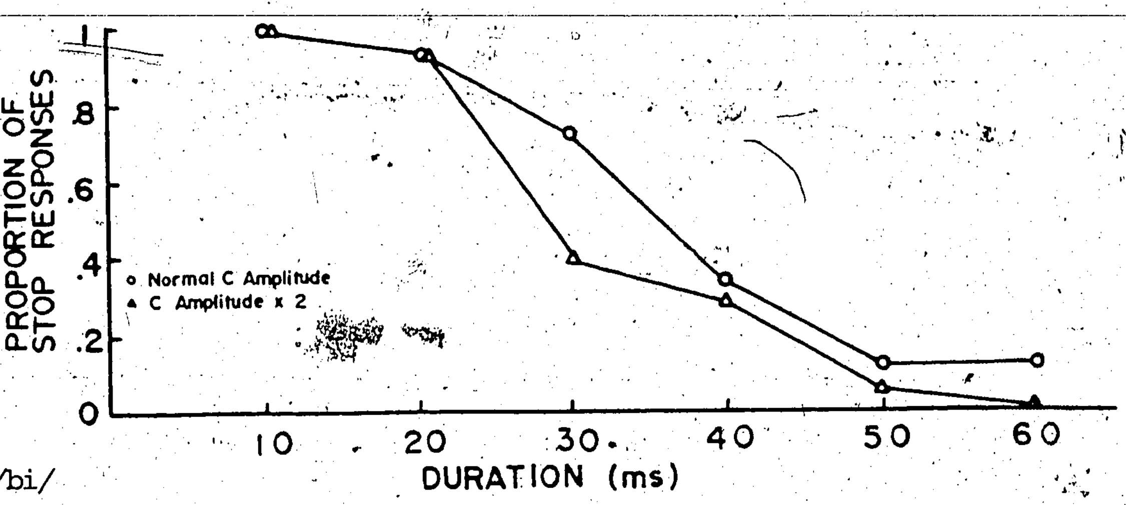
/di/



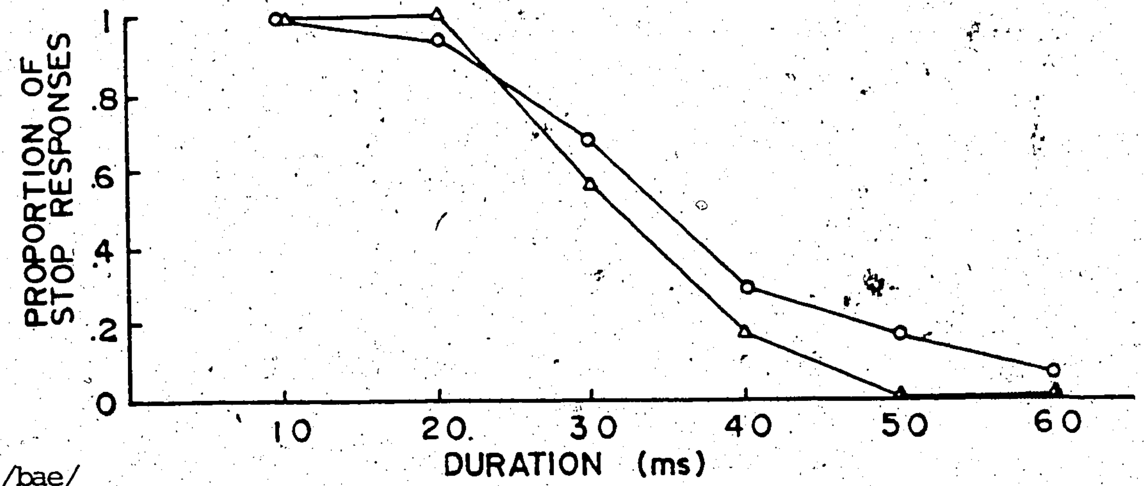
/dae/



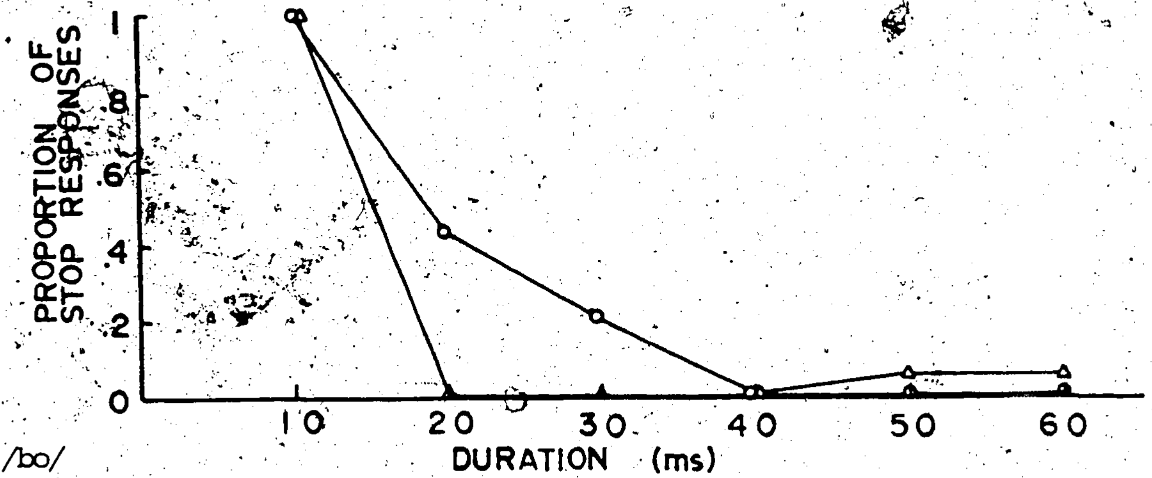
/do/
Subject 5



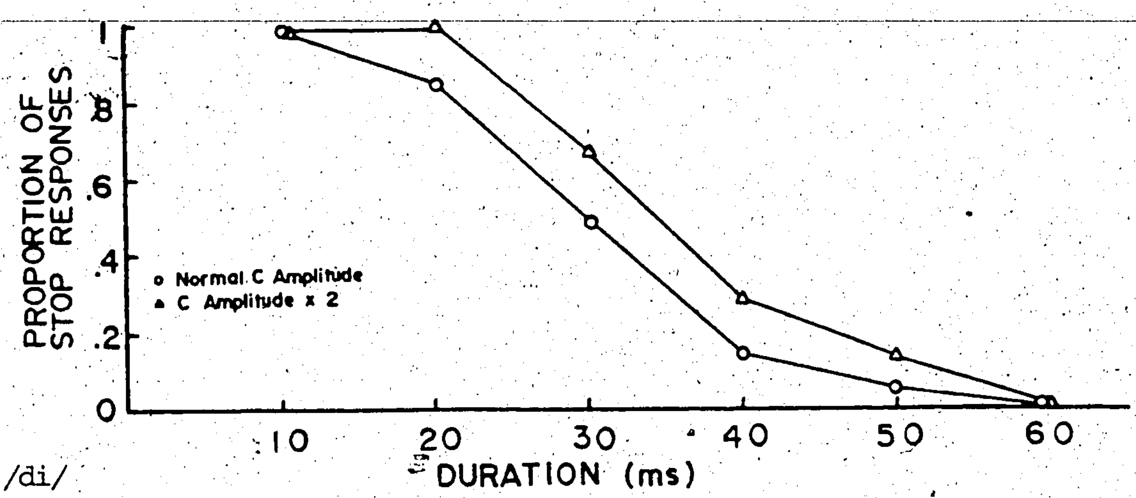
/bi/



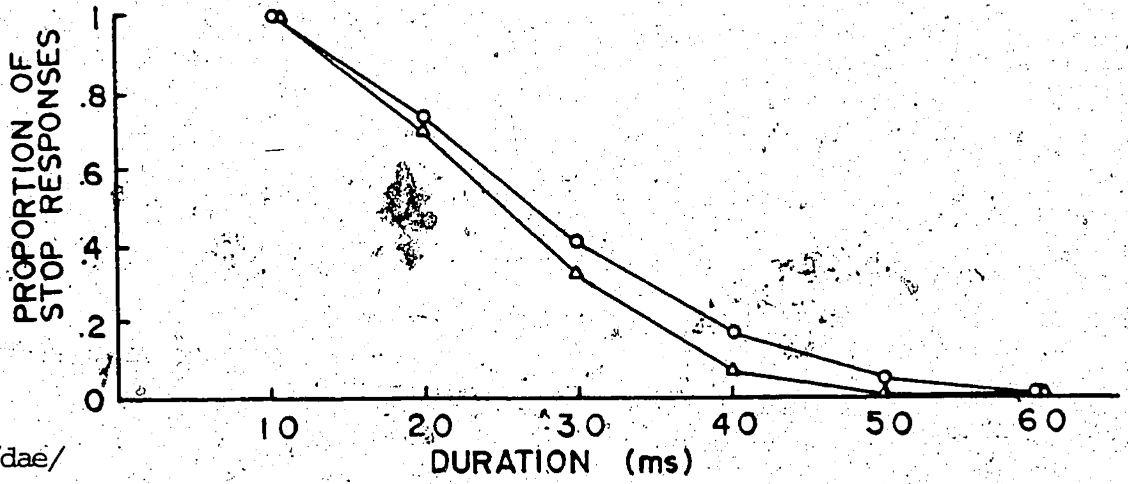
/bae/



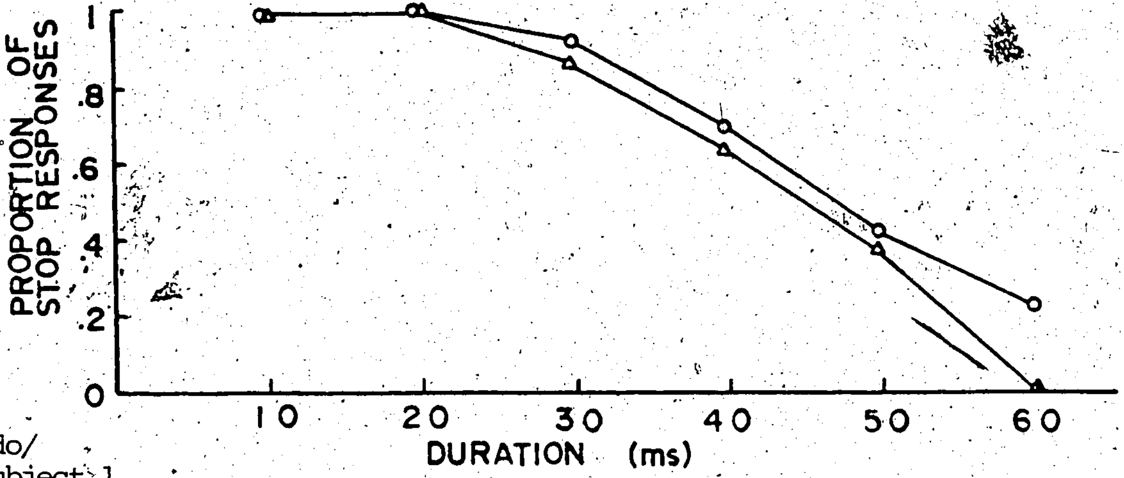
/bo/
Subject 1



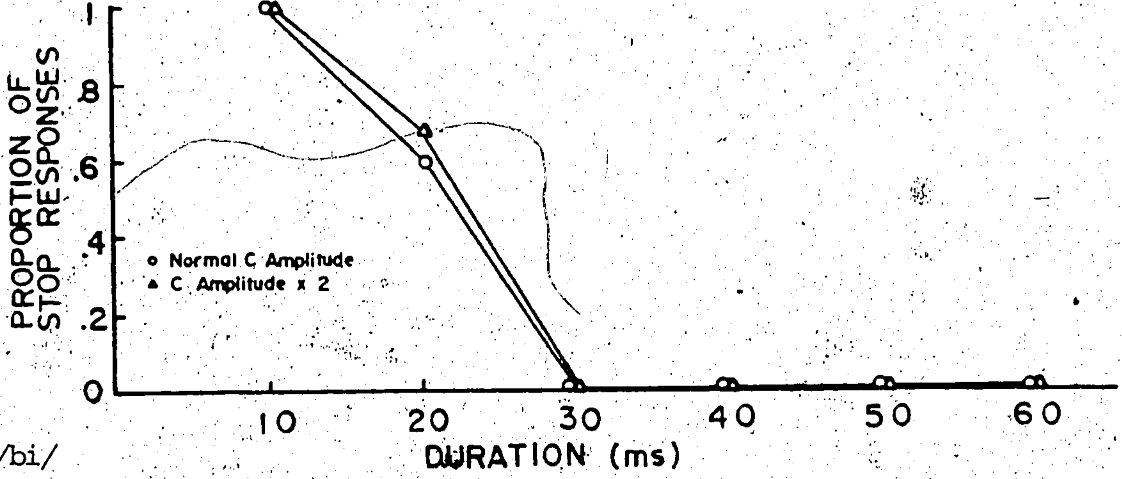
/di/



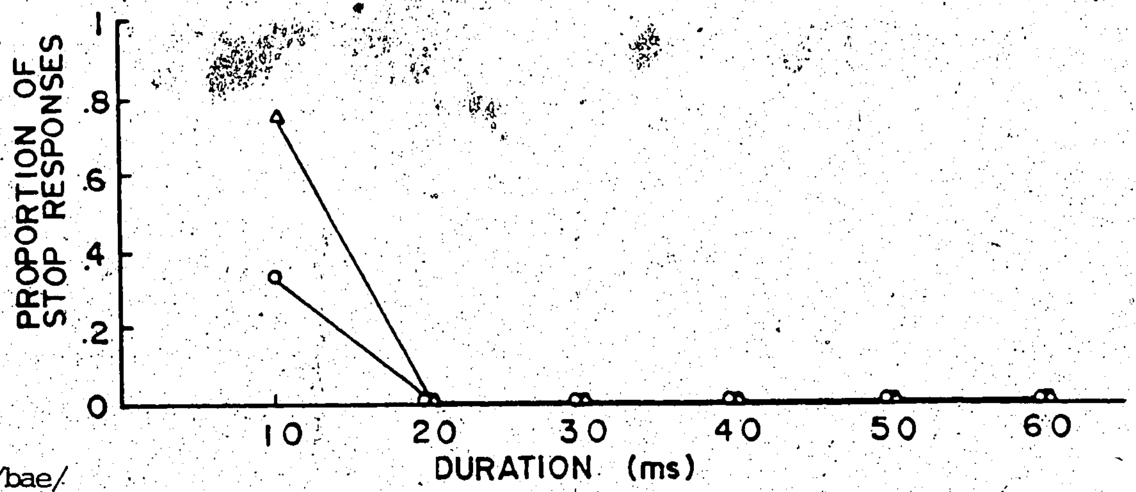
/dae/



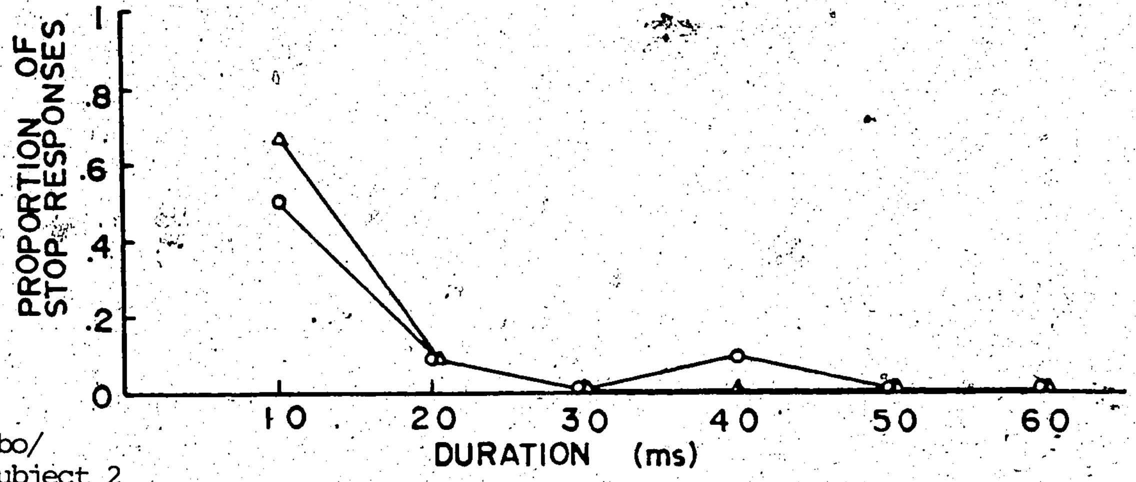
/do/
Subject 1



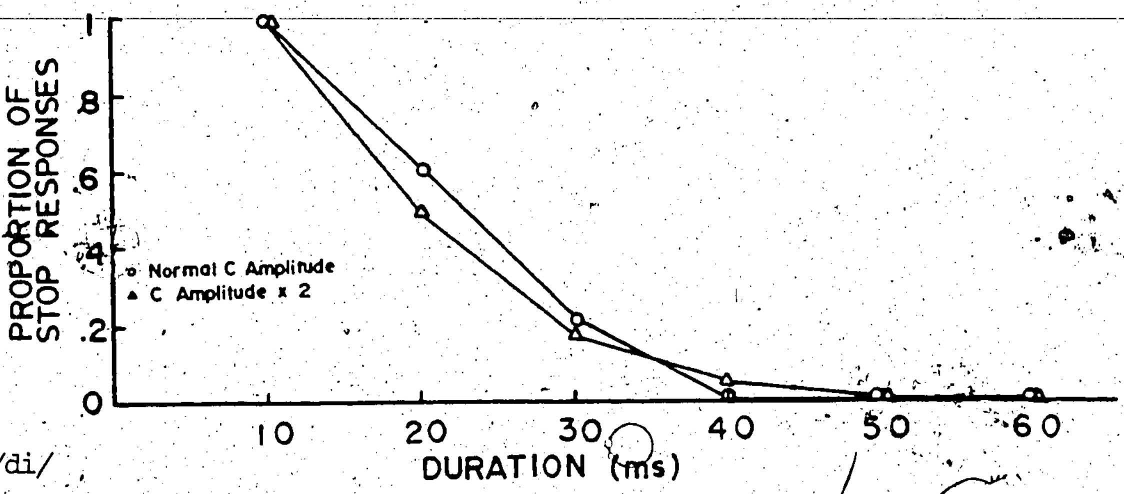
/bi/



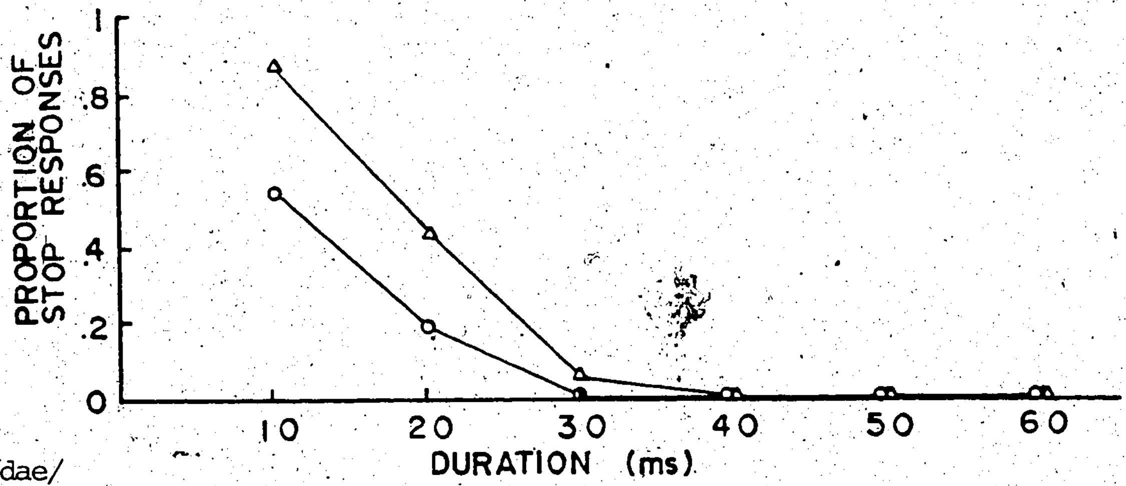
/bae/



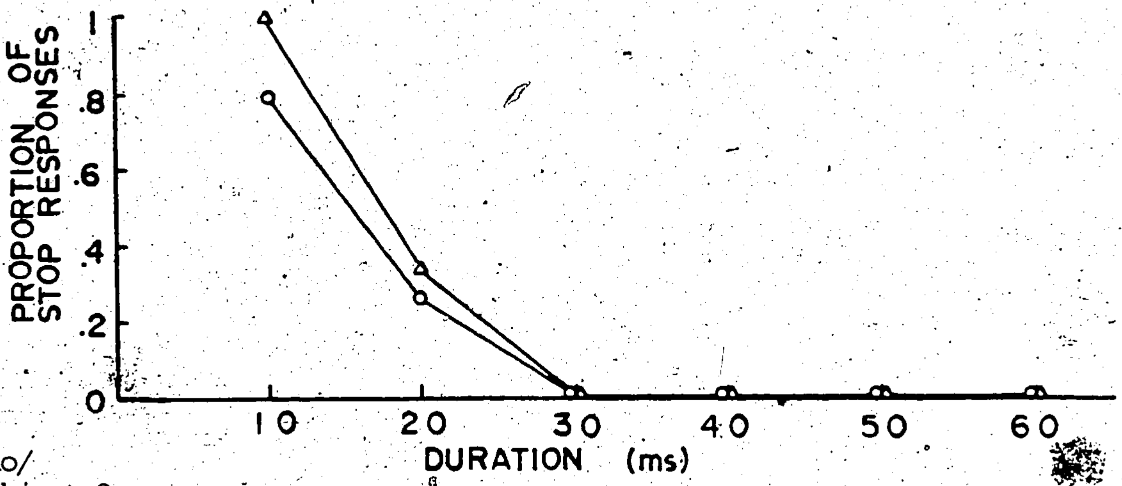
/bo/
Subject 2



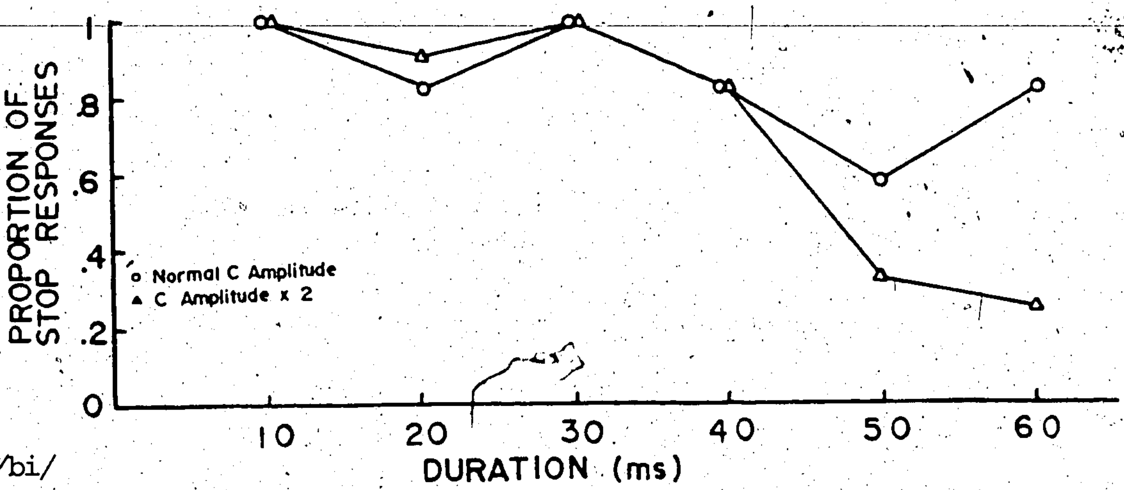
/di/



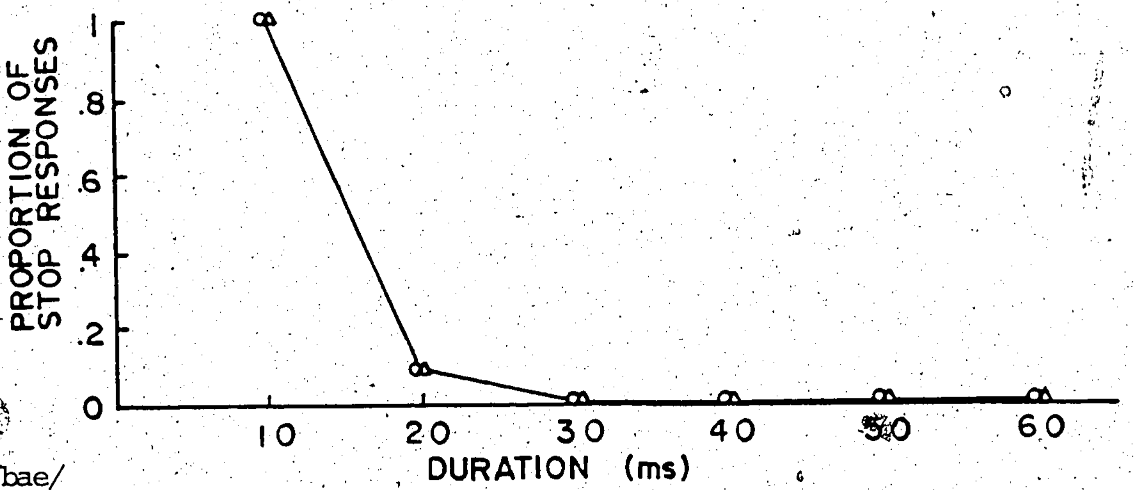
/dae/



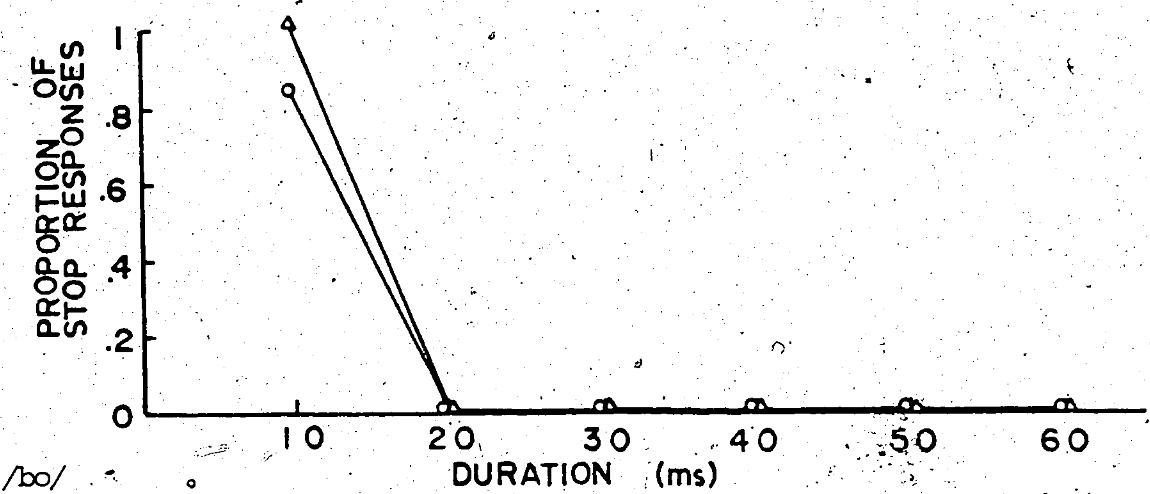
/do/
Subject 2



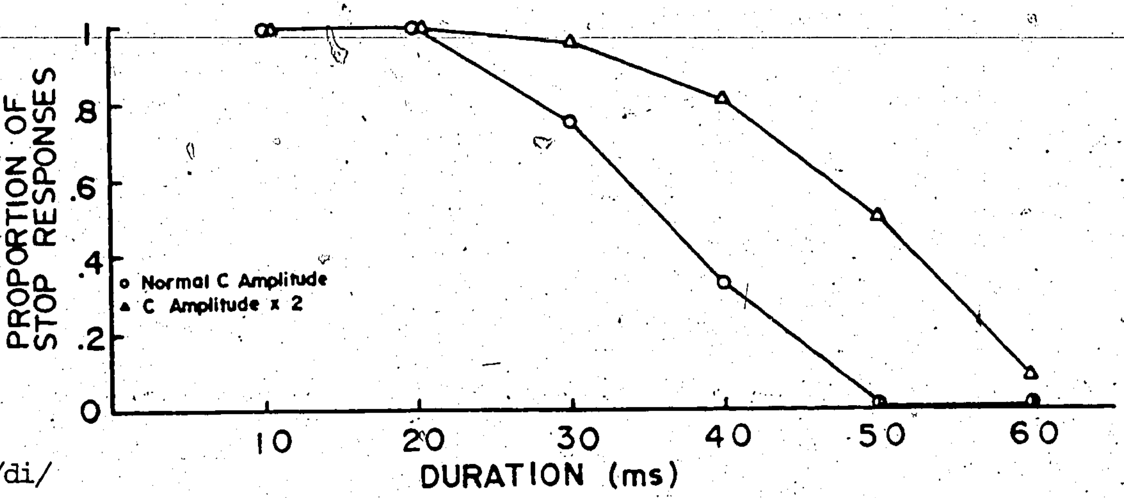
/bi/



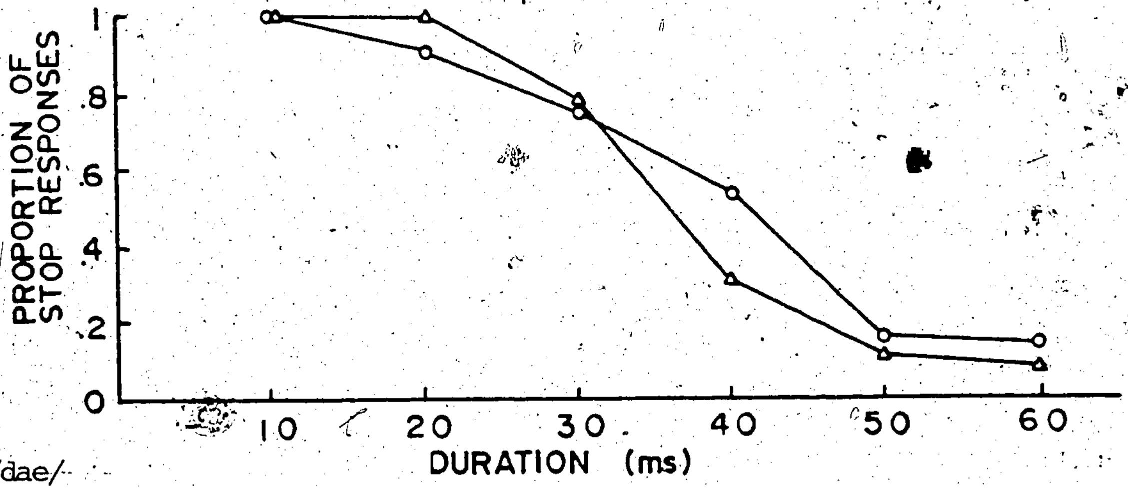
/bae/



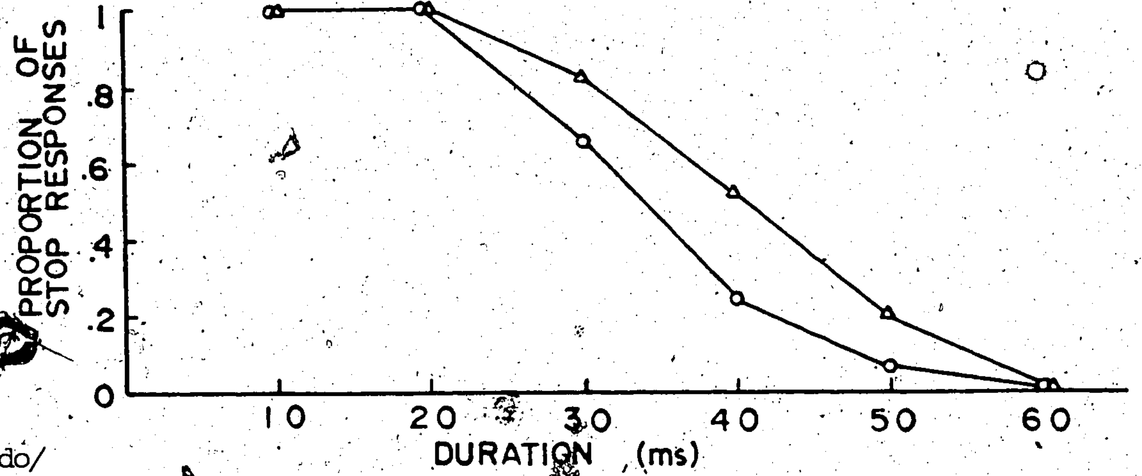
/bo/
Subject 3



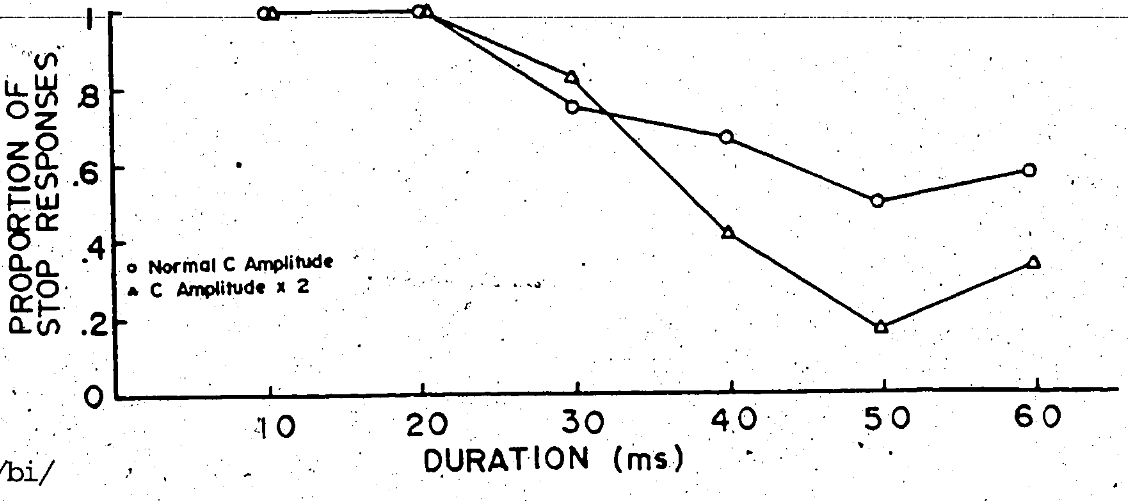
/di/



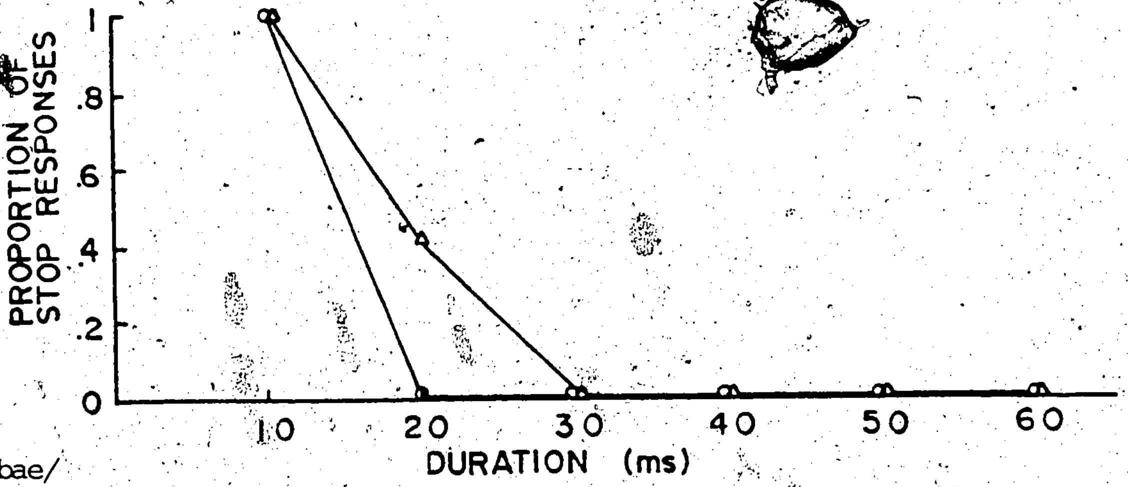
/dae/



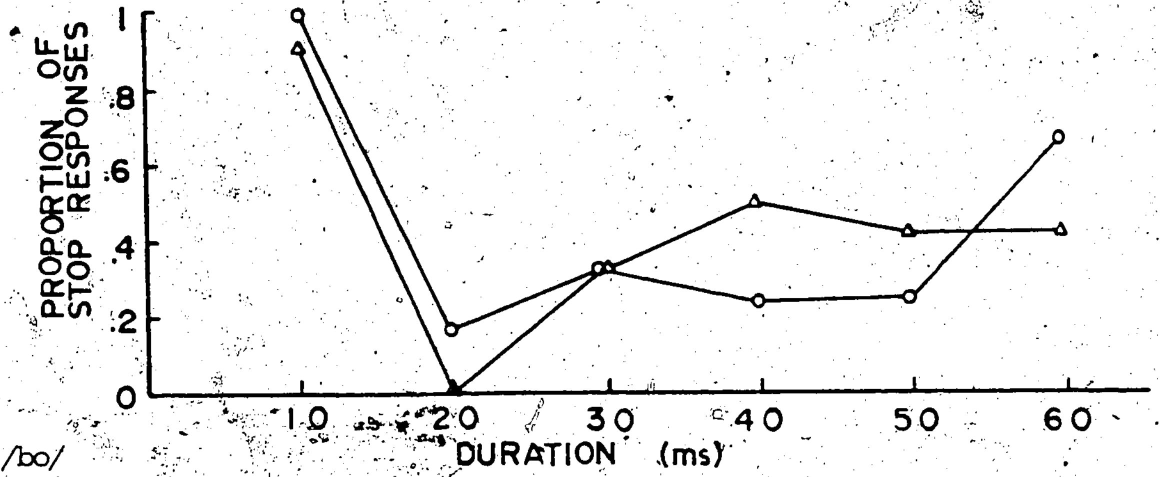
/do/
Subject 3



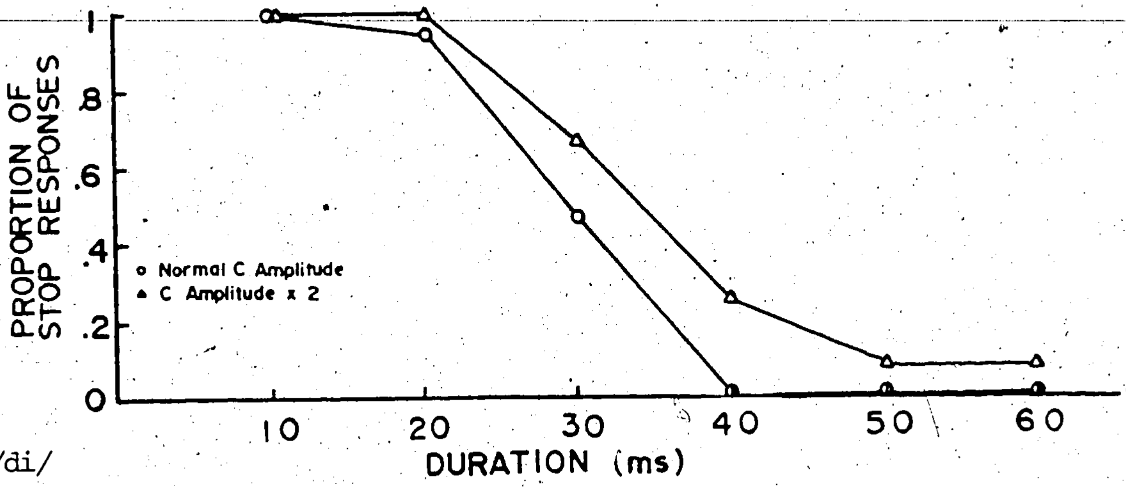
/bi/



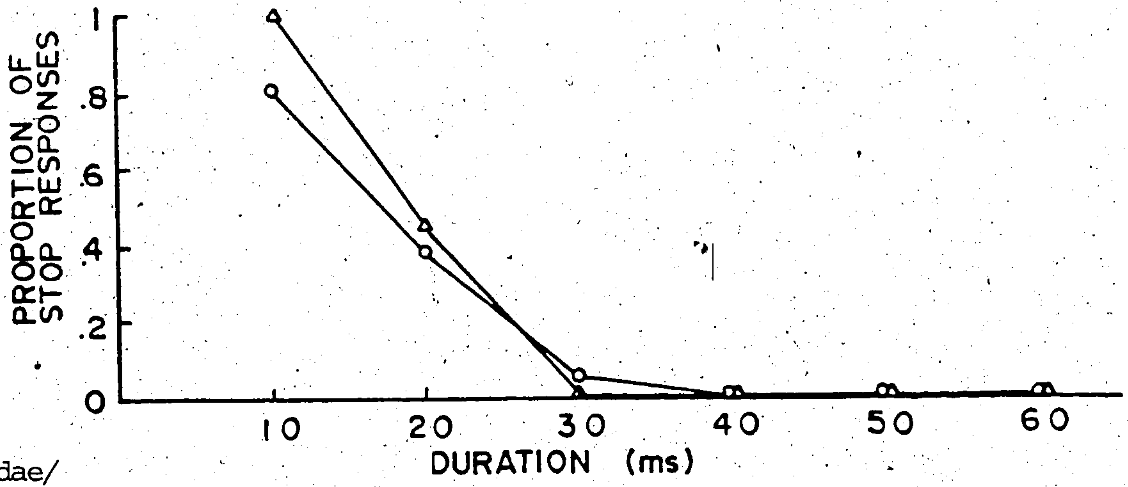
/bae/



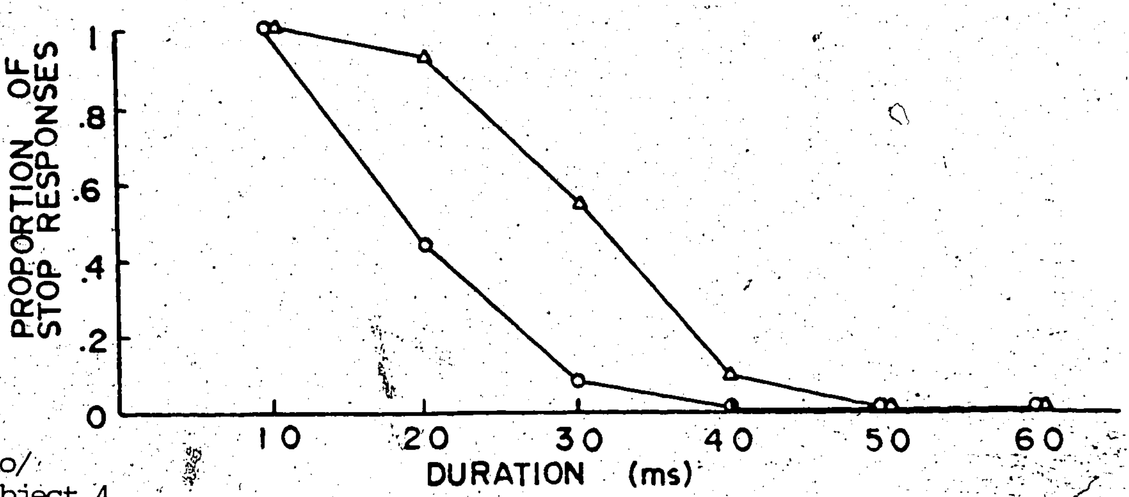
/bo/
Subject 4



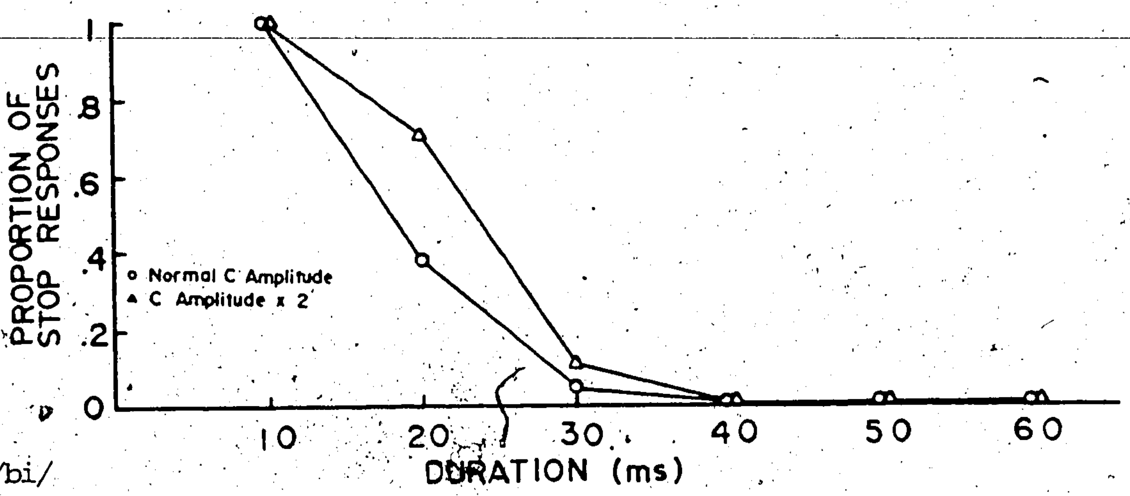
/di/



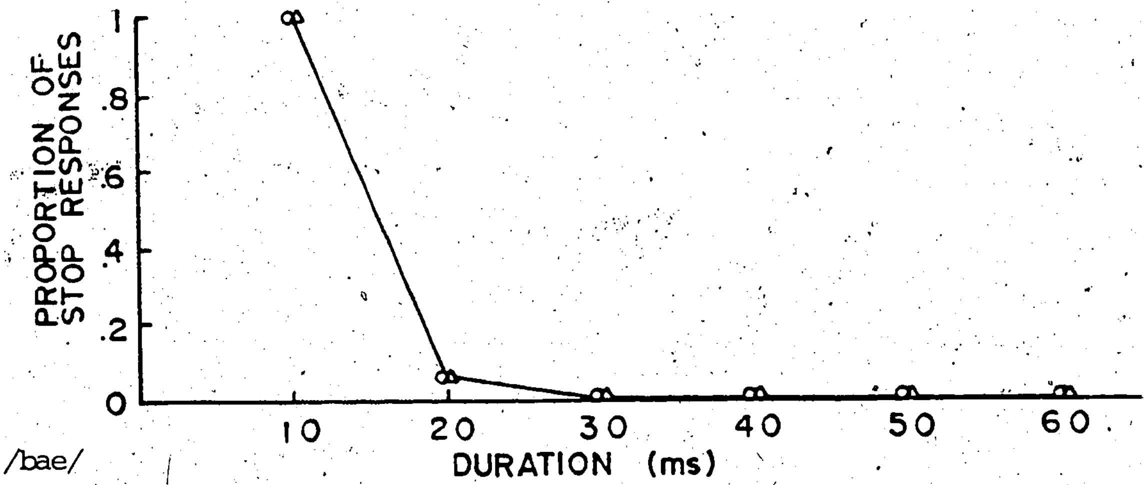
/dae/



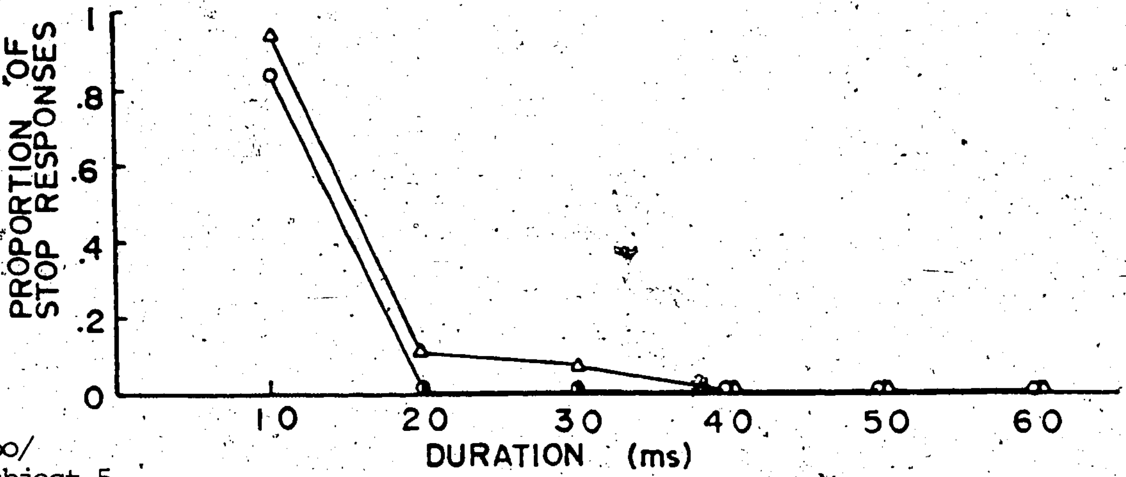
/do/
Subject 4



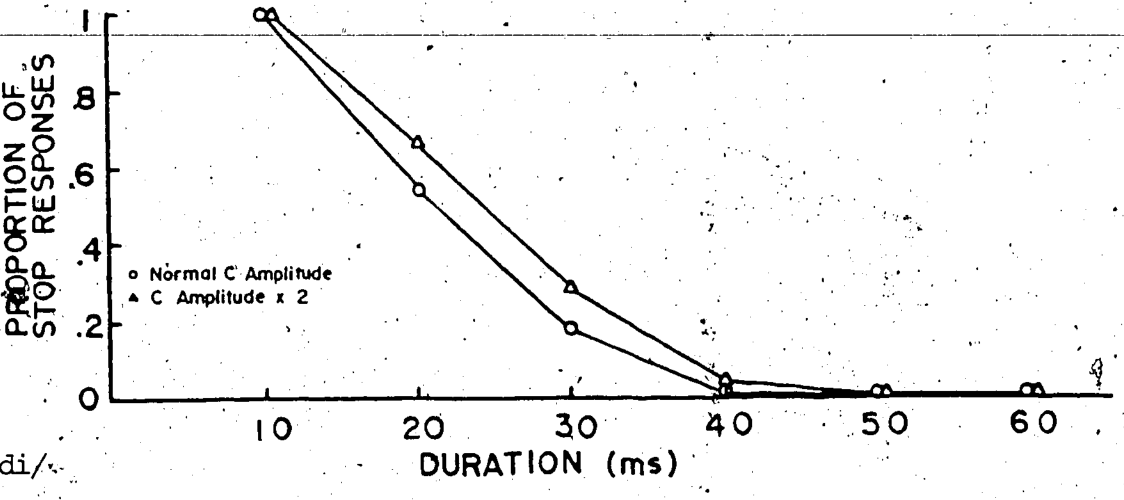
/bi/



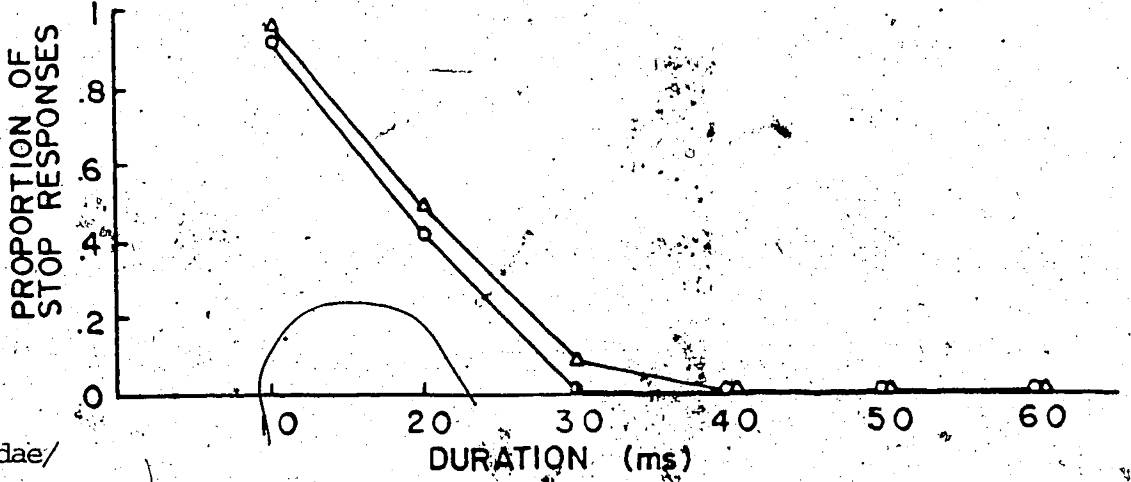
/bae/



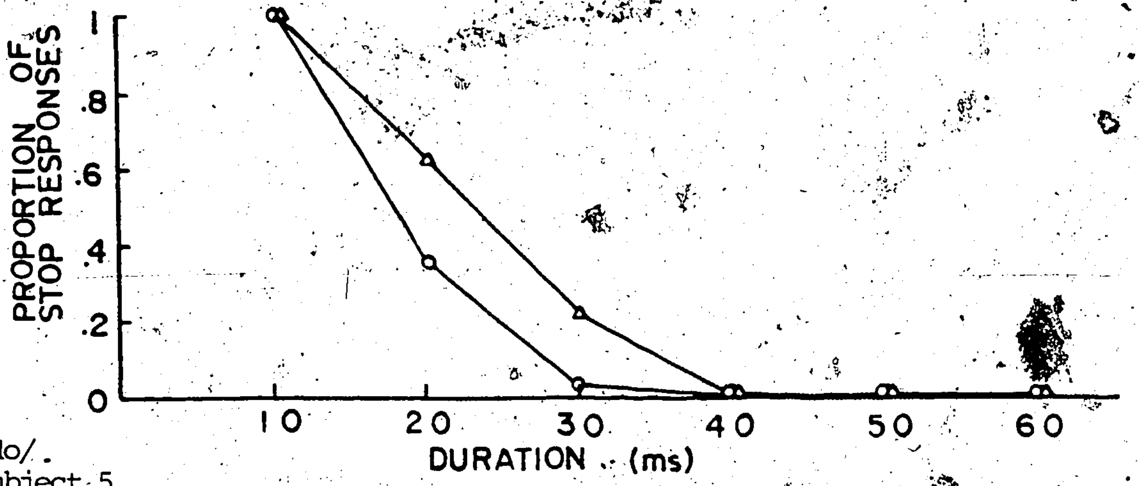
/bo/
Subject 5



/di/



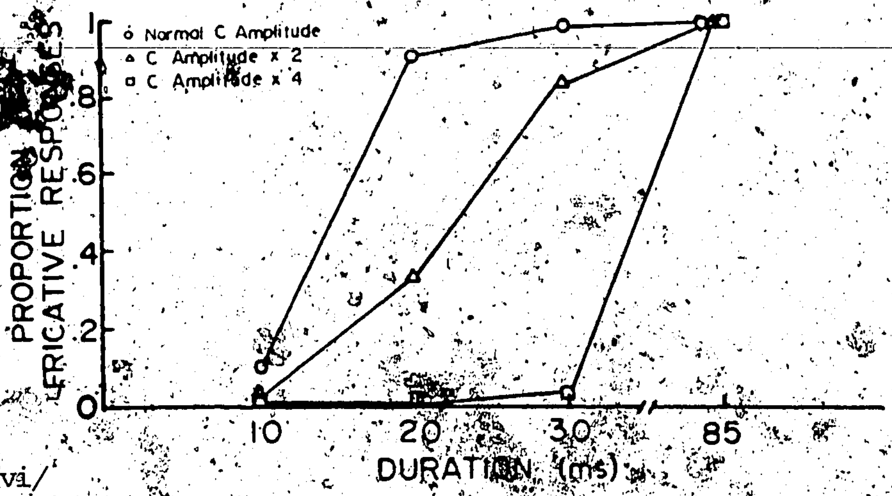
/dae/



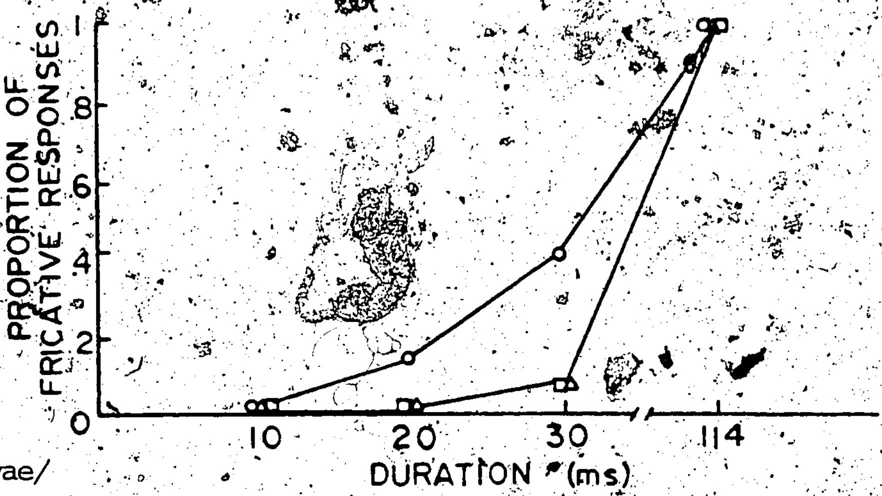
/do/

Subject 5

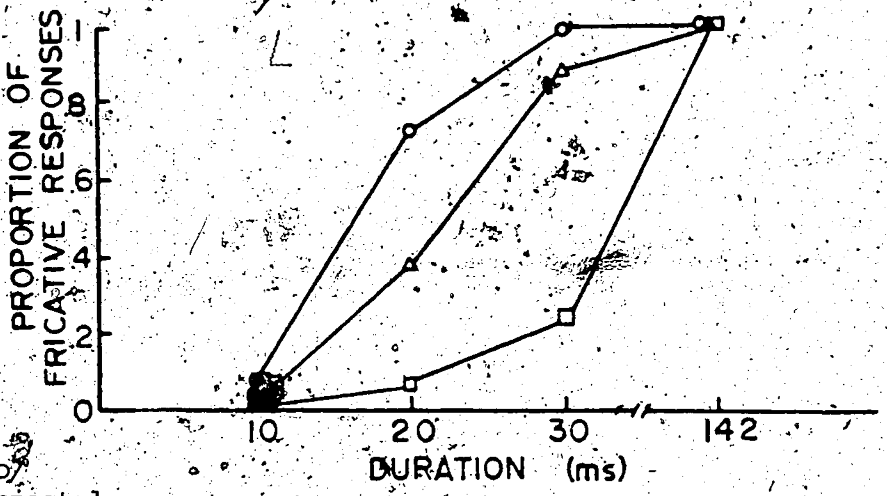
X. Appendix D: Experiment II Individual Subject Recognition Curves



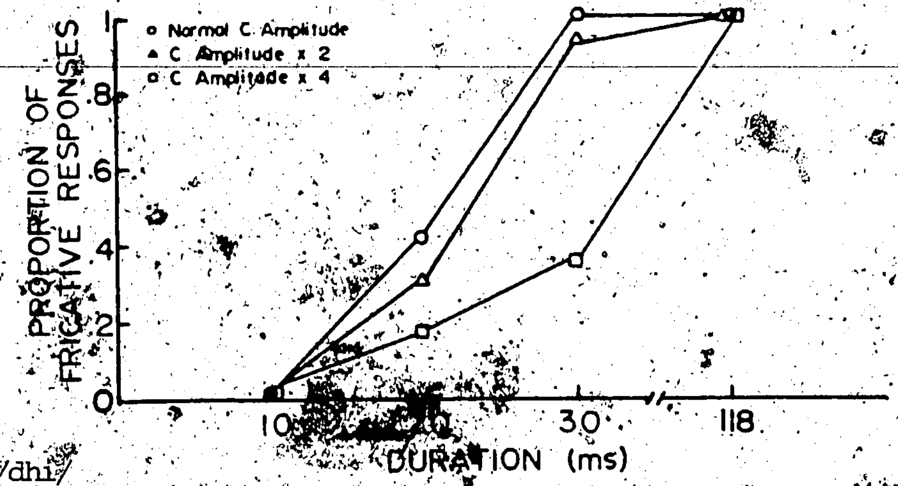
/vi/



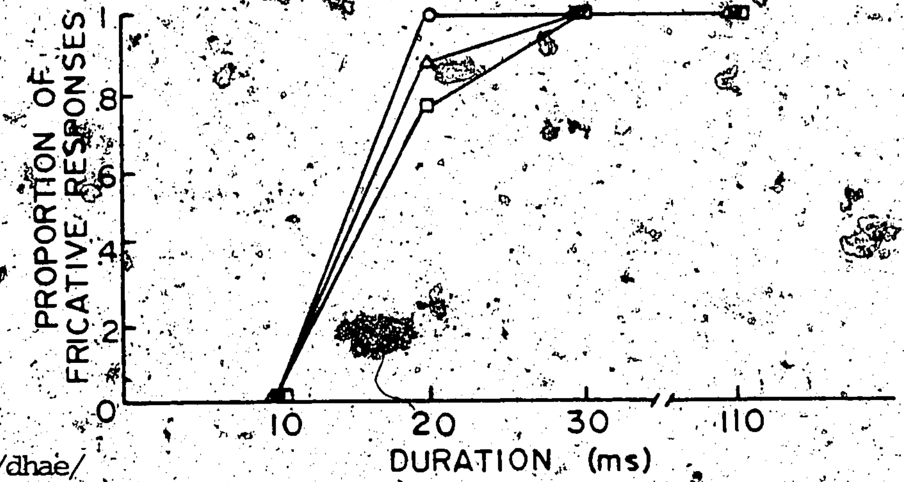
/vae/



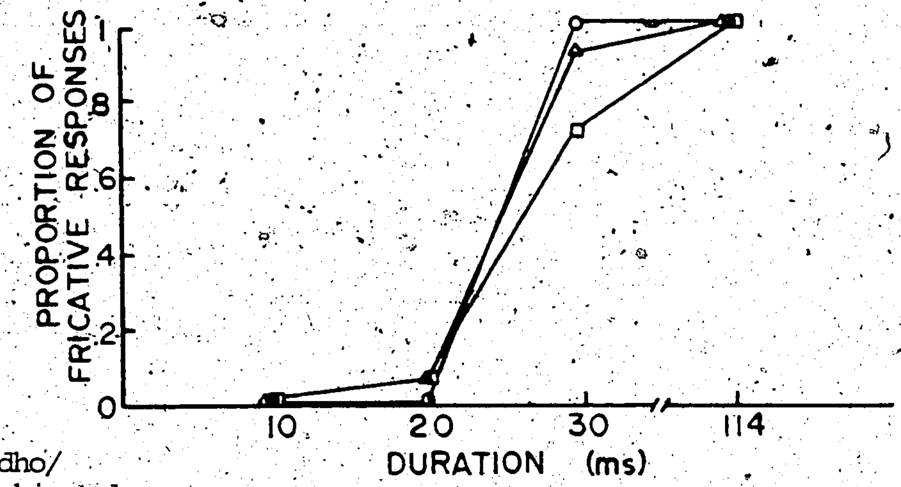
/vo/
Subject 1



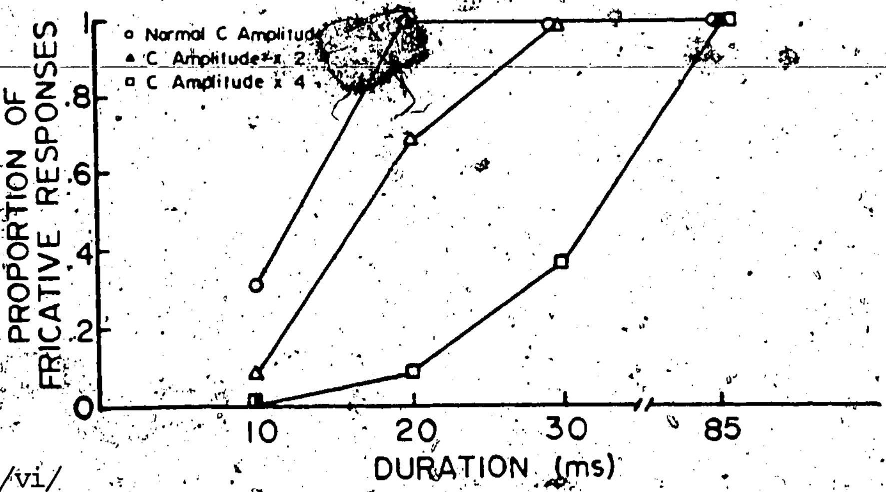
/dhi/



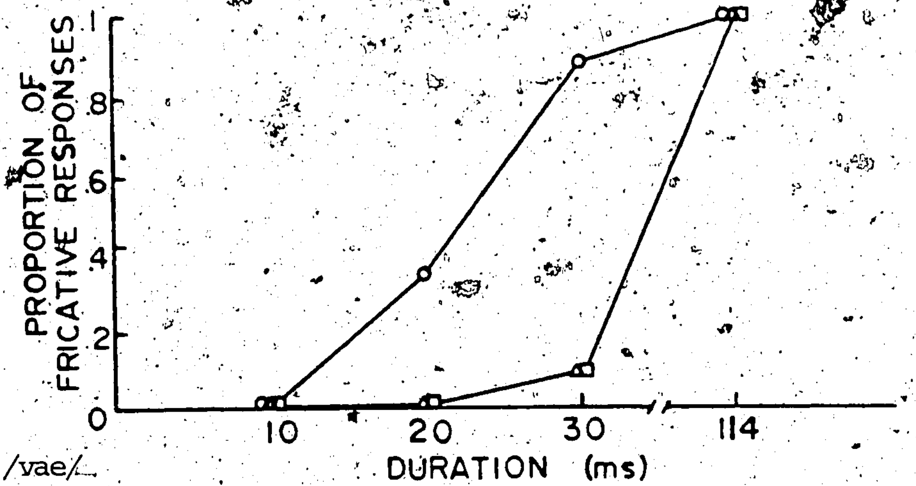
/dhae/



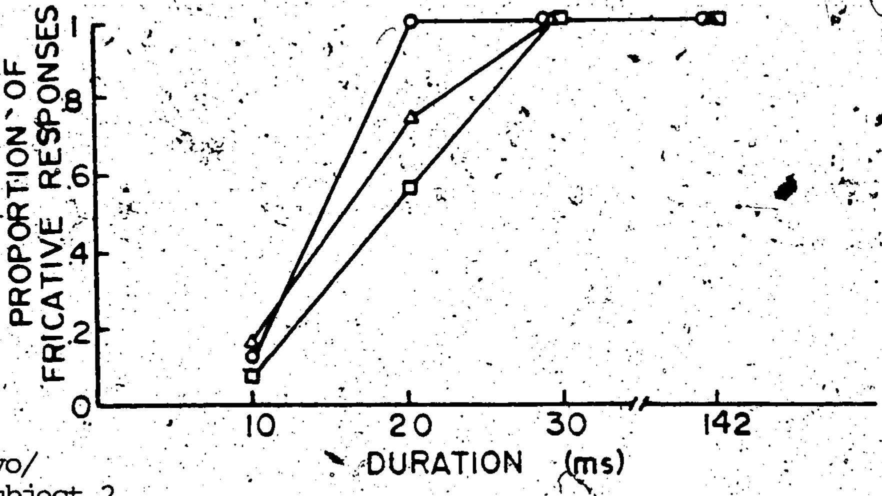
/dho/
Subject 1



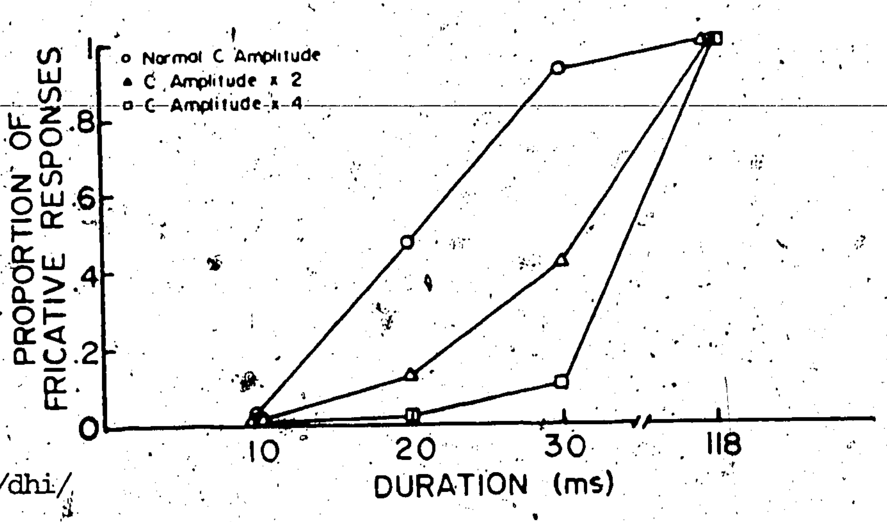
/vi/



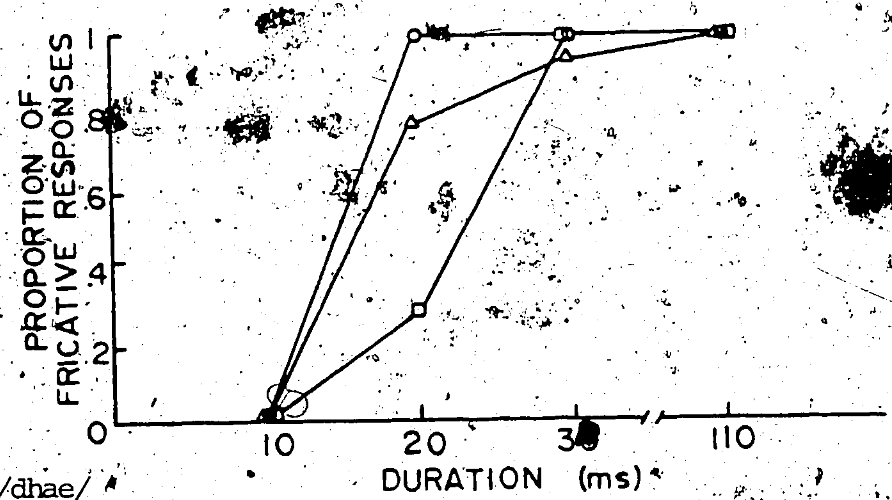
/vae/



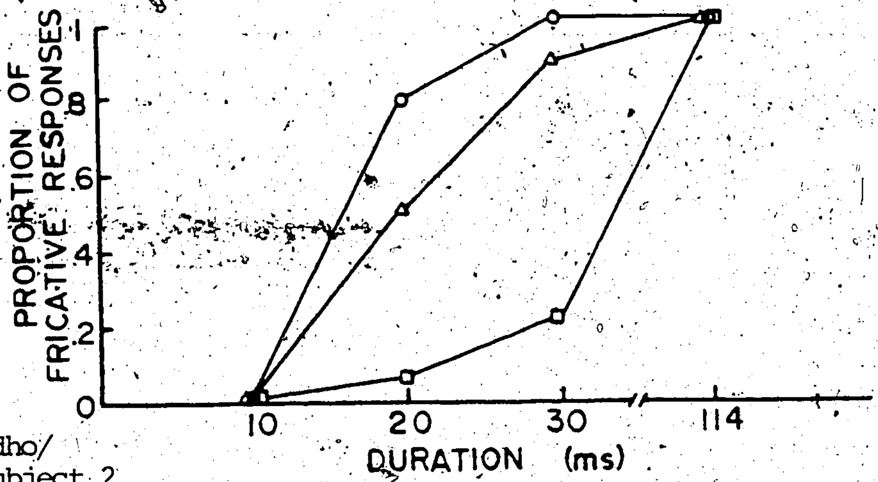
/vo/
Subject 2



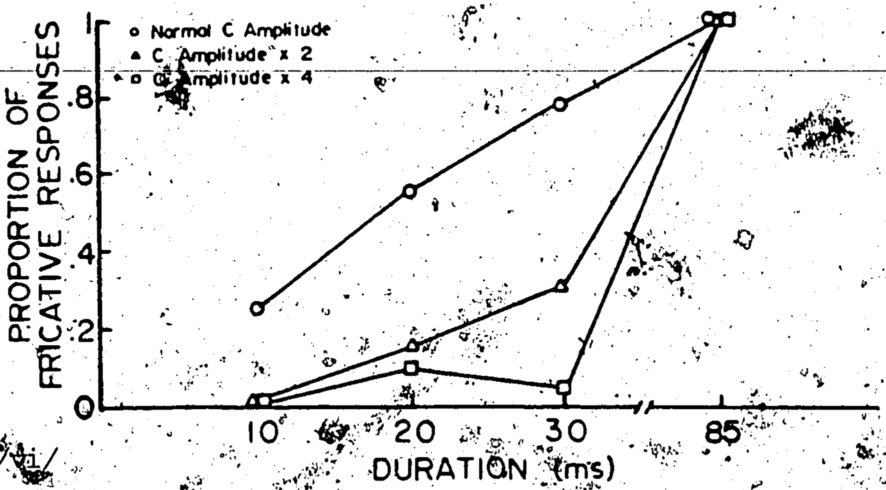
/dhi/



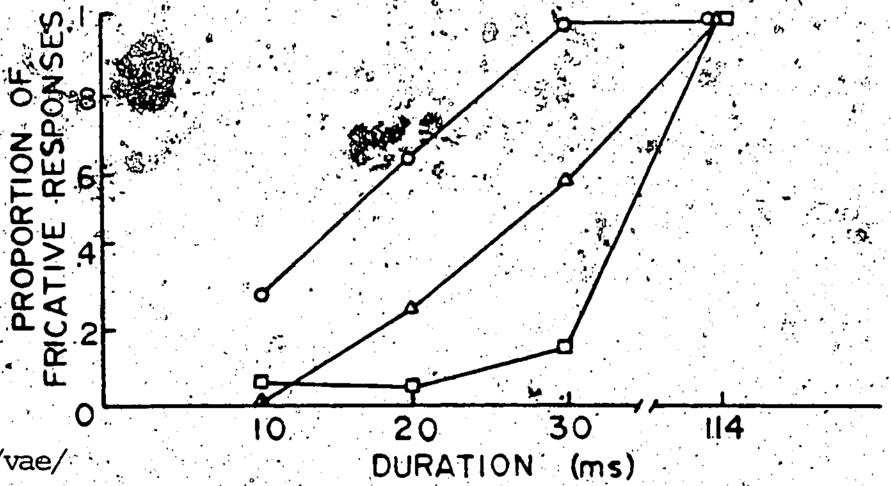
/dhae/



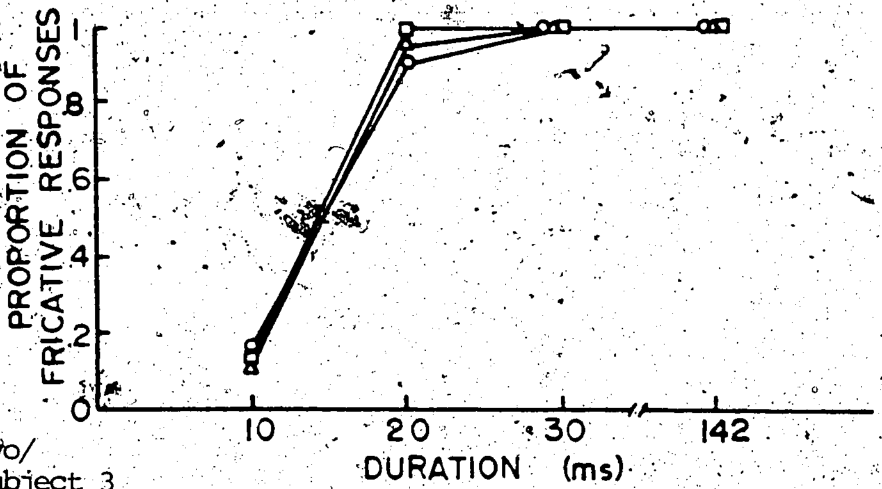
/dho/
Subject 2



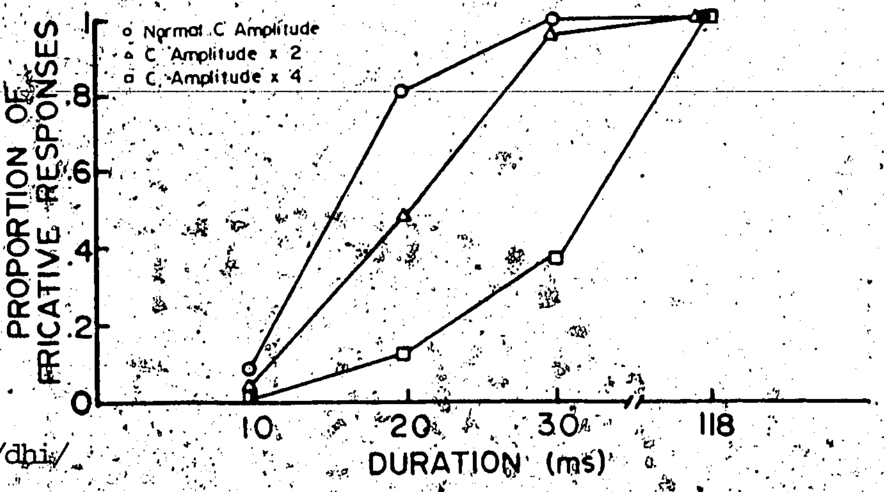
/vi/



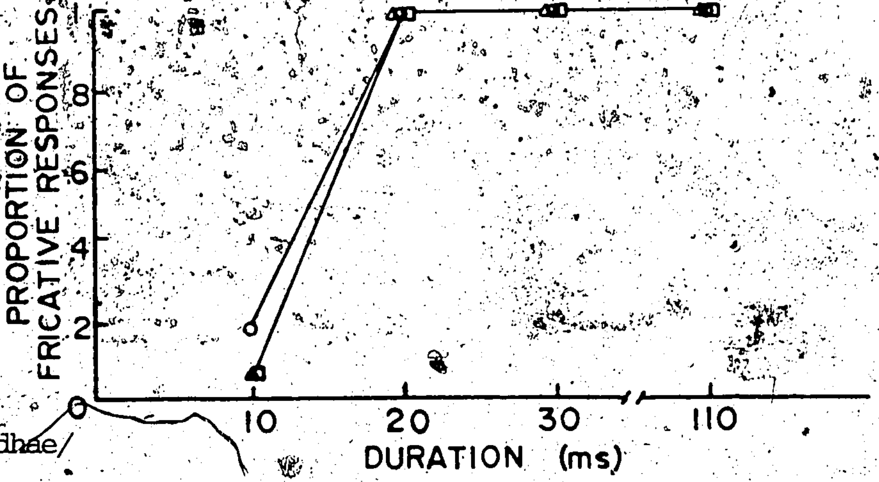
/vae/



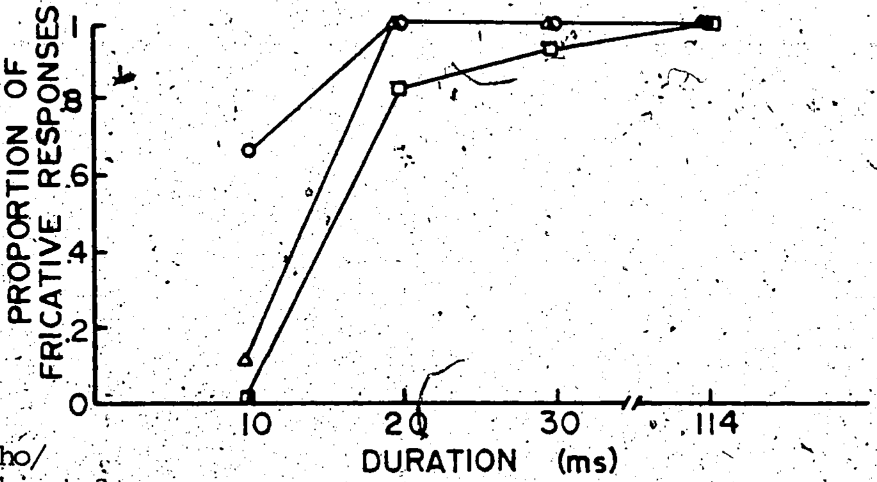
/vo/
Subject 3



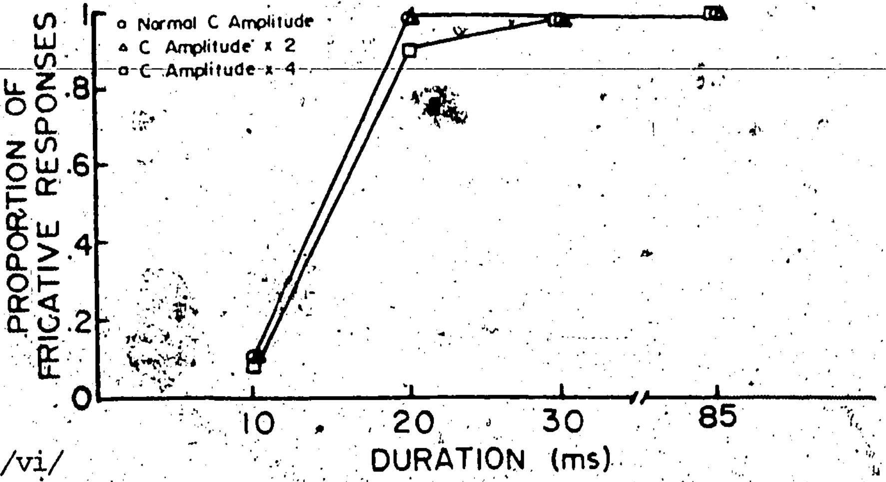
/dhi/



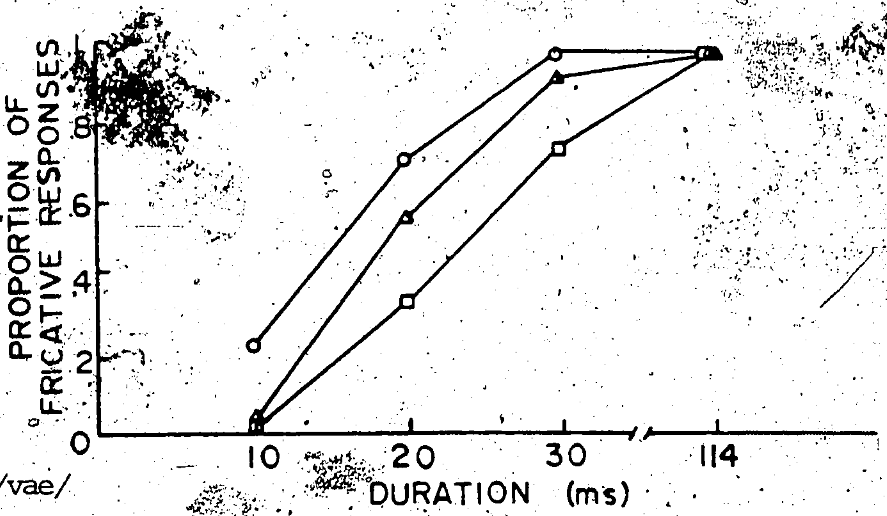
/dhae/



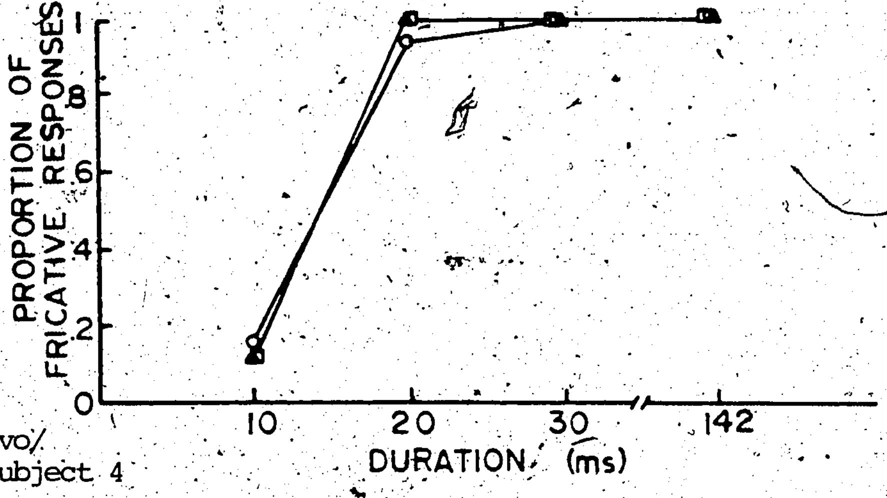
/dho/
Subject 3



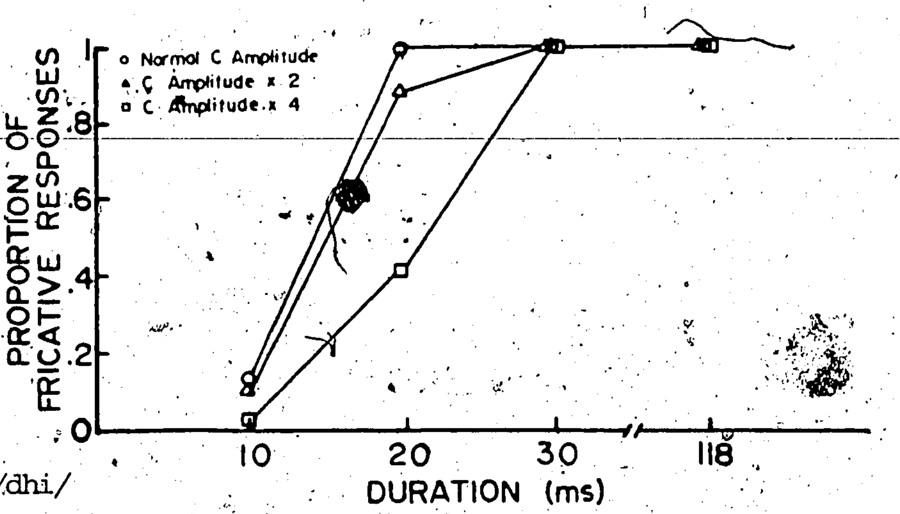
/vi/



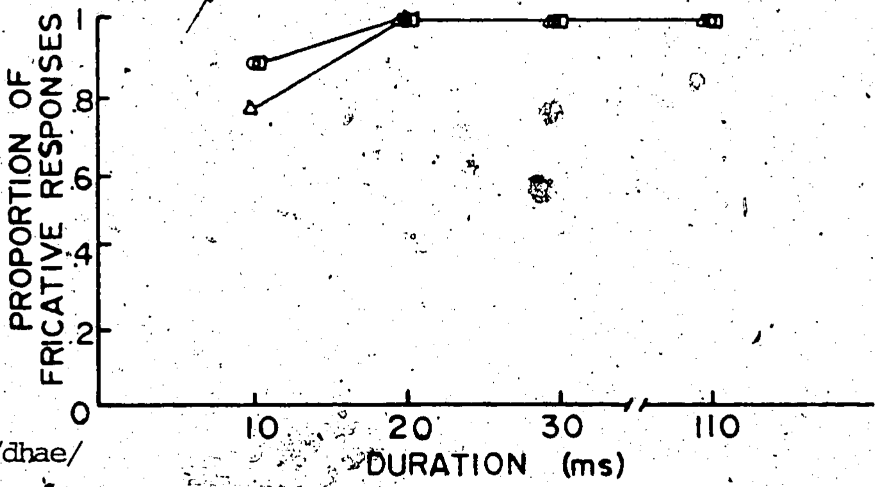
/vae/



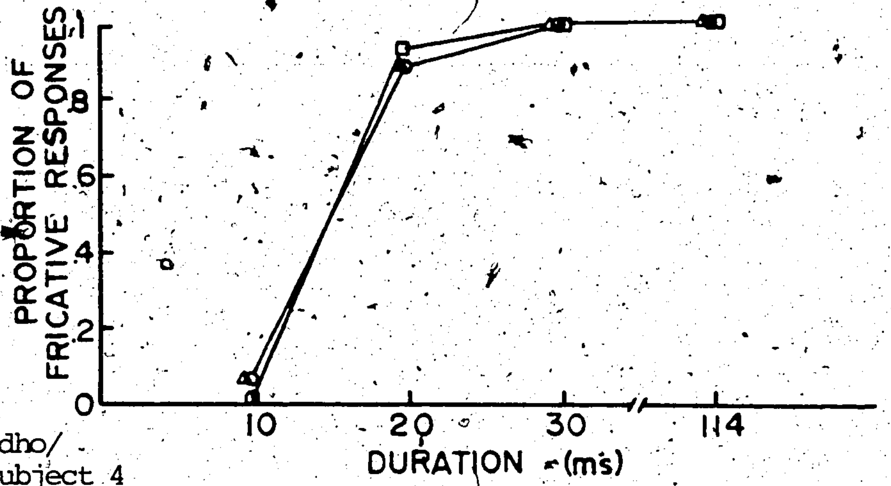
/vo/
Subject 4



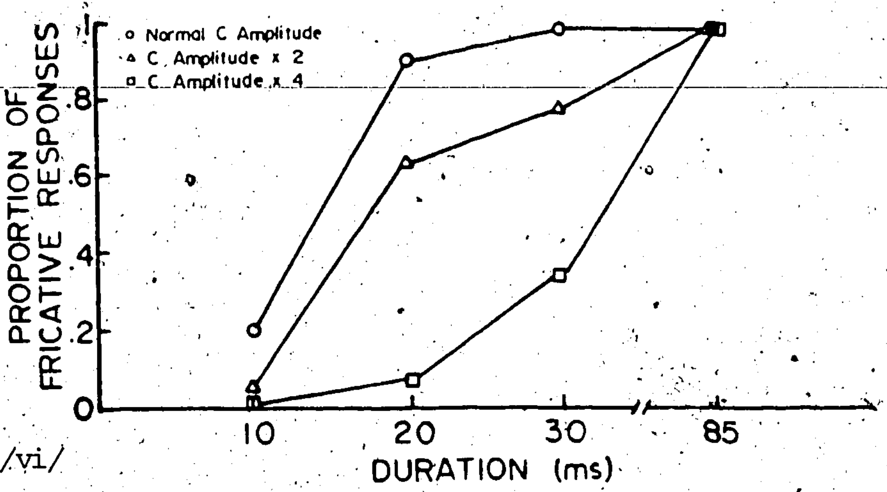
/dhi/



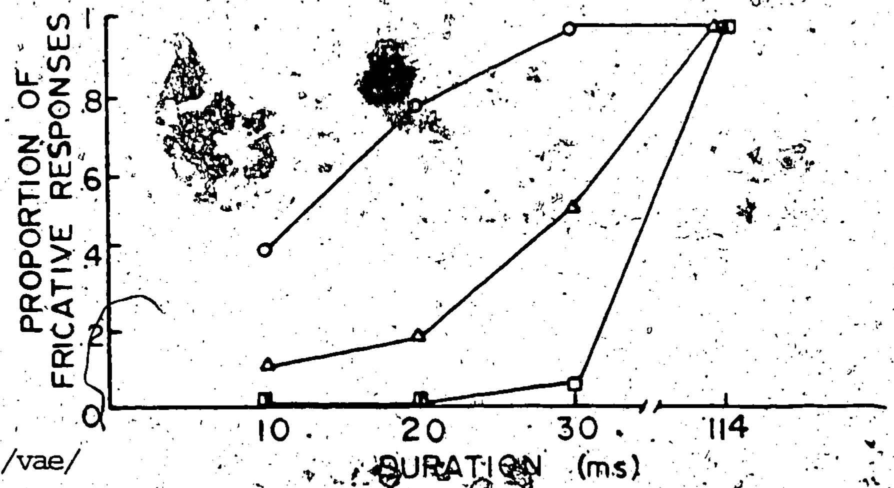
/dhae/



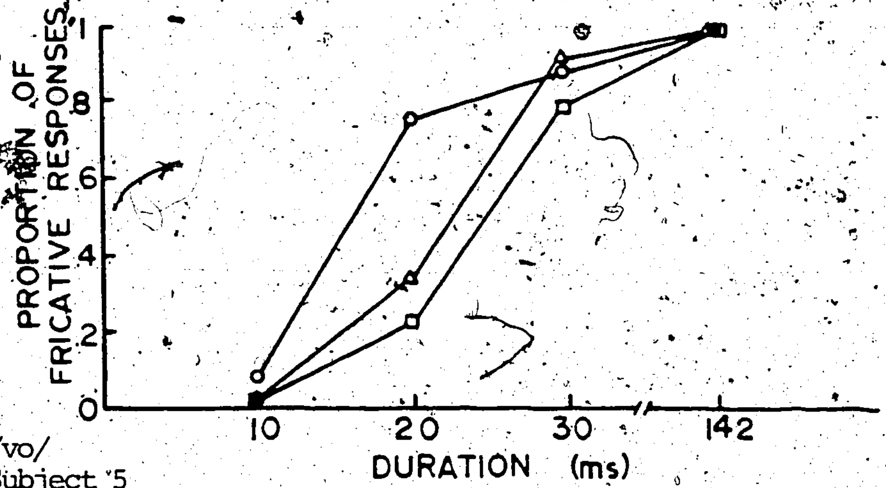
/dho/
Subject 4



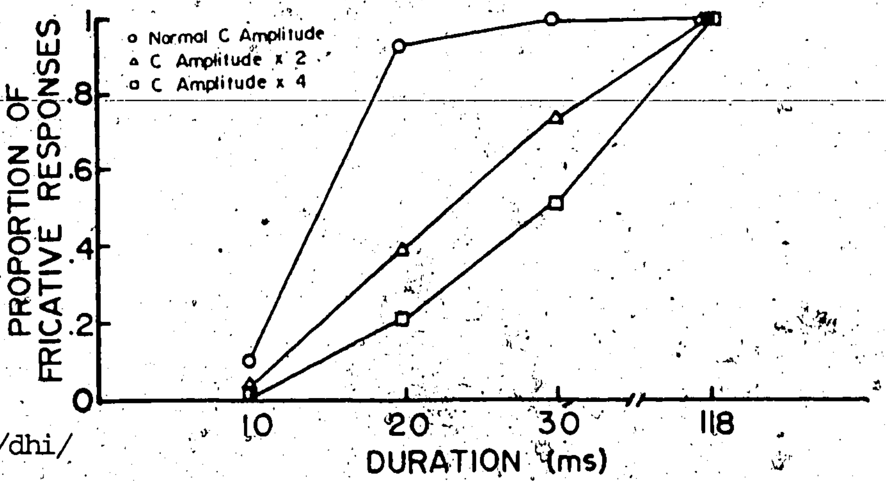
/vi/



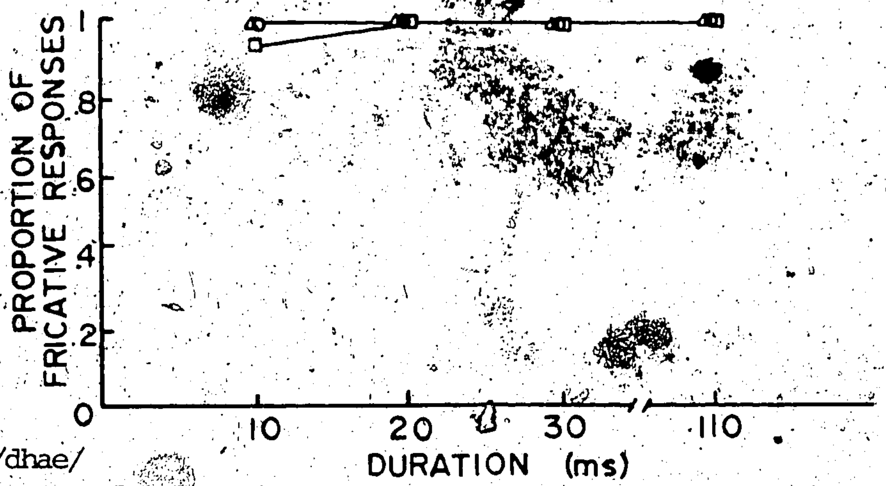
/vae/



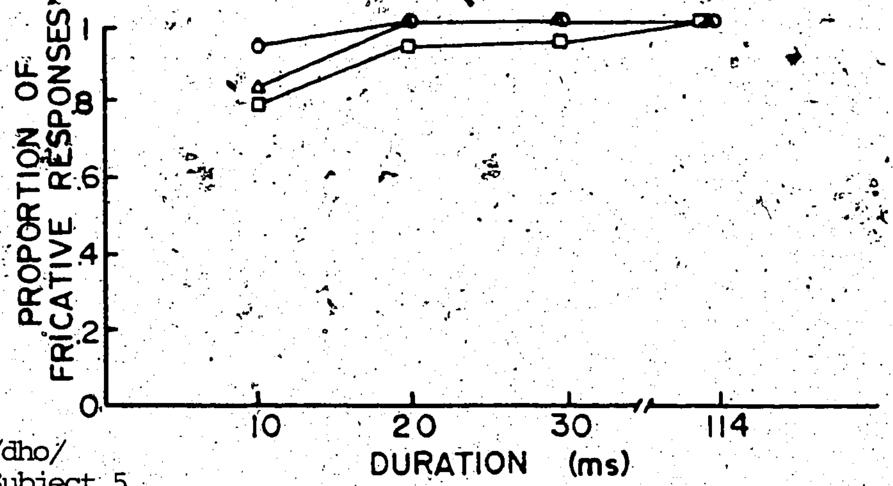
/vo/
Subject 5



/dhi/



/dhae/



/dho/
Subject 5