DOCUMENT RESUME

ED 100 354 16 IR 001 452

AUTHOR Nickerson, Raymond S.; And Others
TITLE Some Observations on Timing in the Speech of Deaf and Hearing Speakers. Report No. 2905.
INSTITUTION Bolt, Beranek and Newman, Inc., Cambridge, Mass.
SPONS AGENCY Bureau of Education for the Handicapped (DHEW/OE), Washington, D.C. Media Services and Captioned Films Branch.
REPORT NO BBN-R-2905
PUB DATE 15 Sep 74
CONTRACT OEC-0-71-4670(615)
NOTE 62p.; For related documents see IR 001 449-455

EDRS PRICE MF-$0.75 HC-$3.15 PLUS POSTAGE
DESCRIPTORS Acoustic Phonetics; Auditory Evaluation; Aurally Handicapped; *Deaf; Deaf Education; *Deaf Research; Pacing; *Speech Evaluation; *Speech Handicaps; Speech Improvement
IDENTIFIERS Intelligibility; *Speech Rhythm

ABSTRACT

The factor of timing on intelligible speech among normally-hearing persons was studied to determine some ways in which it differs from the speech of the deaf. Additional data was gathered on the temporal aspects of the speech of deaf and hearing children and hearing adults. The data corroborated earlier studies indicating that (1) deaf speakers tend to speak at a much slower rate than hearing speakers; (2) the difference between the durations of stressed and unstressed syllables is proportionately much smaller for deaf than for hearing speakers; (3) deaf speakers tend to insert more pauses, and pauses of longer duration, within running speech. (SK)

BOLT BERANEK AND NEWMAN INC

CONSULTING · DEVELOPMENT · RESEARCH

Report No. 2905
Job No. 11580

SOME OBSERVATIONS ON TIMING IN THE SPEECH OF DEAF AND HEARING SPEAKERS

14 September 1974

Submitted to:

Media Services and Captioned Films Branch
Division of Educational Services
Bureau of Education for the Handicapped
U.S. Department of Health, Education, and Welfare
Washington, D. C. 20202

Table of Contents

Abstract

Proper timing is recognized as essential to intelligible fluent speech. Conversely, inappropriate timing has been considered by many investigators to be one of the major causes of the generally poor intelligibility of the speech of the deaf. What constitutes correct timing is not yet thoroughly understood, however, and consequently attempts to improve the temporal aspects of the speech of the deaf are necessarily somewhat ad hoc and lacking a firm theoretical basis. In this paper we review some of what is known concerning the role of timing in the speech of normally-hearing individuals, and we consider some of the ways in which the speech of the deaf tends to differ from that of hearing speakers in terms of its temporal characteristics. Additional data are presented on the temporal aspects of the speech of deaf and hearing children and hearing adults. These data corroborate the results of other studies that have found that: (1) deaf speakers tend to speak at a much slower rate than do hearing speakers, (2) the difference between the durations of stressed and unstressed syllables is proportionately much smaller for deaf than for hearing speakers, and (3) deaf speakers tend to insert more pauses, and pauses of longer duration, within running speech--particularly within phrases--than do hearing speakers.

Some Observations on Timing in the Speech of Deaf and Hearing Speakers[1]

R. S. Nickerson
K. N. Stevens
A. Boothroyd
A. Rollins

Perhaps the most fundamental property of speech is the fact that it occurs over time. One cannot say of an utterance, as one can of a visual scene, that it exists. Rather, it takes place; it happens. A sentence is a sequence of words, spoken one after the other. A spoken word is itself an unfolding event, and to describe it one must consider how it develops in time. It is not surprising, therefore, that the temporal properties of speech should play an important role in its production and perception.

Inappropriate timing has been considered by many investigators of the speech of the deaf to be a major--if not the major--cause of its generally poor intelligibility (Bell, 1916; Hood, 1966, 1967; Hudgins & Numbers, 1942; John & Howarth, 1965). The purpose of this paper is to consider some of the evidence for that claim, and, more generally, to consider how the speech of the deaf differs from the speech of individuals with normal hearing, in terms of its temporal characteristics.

TEMPORAL CHARACTERISTICS OF NORMAL SPEECH

What are the temporal characteristics of speech? To what extent do these characteristics differ from speaker to speaker, or from one speech context to another? How much deviation from statistical norms can be tolerated before the speech begins to be unintelligible or to sound unnatural? How is timing used to color speech and to convey information in addition to that carried by the words themselves? The answers to these and similar questions are not fully known; however, data are slowly accumulating that may provide the basis for a theory of the timing aspects of speech. Some of these data are considered in this paper. The intent is not to present a comprehensive review of research on this topic, however, but to provide a framework within which to view the problem of teaching speech timing and rhythm to the deaf.

Speech Rate

There are three questions that one might raise concerning speech rate: How fast can people talk? How fast do people talk? How fast should people talk?

How fast people can read aloud depends somewhat on such factors as the average number of syllables per word, and the reader's familiarity with the material read (Pierce, 1961). In general, the fewer the number of syllables per word, the higher the word emission rate; however, the relationship is not a simple tradeoff: one cannot read one-syllable words at twice the rate at which one can read two-syllable words. The effect of familiarity is seen in the fact that lists of more commonly occurring words can be read more rapidly than can lists of less commonly occurring words. For nontechnical prose, Pierce reports maximum reading rates of between four and five words per second. The limitation, he notes, appears to be a cognitive--as opposed to a mechanical--one, inasmuch as speakers in his study were able to repeat memorized phrases at much higher rates (seven to nine words per second). Data are not presented concerning how fast people can talk when generating communicative speech; however, one would guess that the limit would be somewhere between that for reading and that for emitting rehearsed material.

Most people probably do not normally talk as fast as they are able, however; and for our purposes the more important question is, how fast do they normally talk when not pushed to their limits? Several studies of the speech of radio announcers have yielded word emission rates of from 107 to 240 words per minute (Voelker, 1938). This is a very broad range, and, unfortunately, details are not given concerning how the measured rates depended on such factors as the type of material involved and whether it was read, rehearsed, or spontaneous speech. One guesses that, in general, radio announcers may be motivated to talk faster than the average individual in conversation because of the need to make effective use of limited time. Pickett (1968) gives about 3.3 syllables per second as an average speech rate. Assuming an average of between one and two syllables per word in conversational speech, this translates to between about 100 and 200 words per minute, which is within the range of the measured rates reported by Voelker.

The question of how fast people should talk is a complex one. Perhaps the question is better phrased: What types and magnitudes of deviations from statistical speech rate norms can be tolerated before the speech decreases in intelligibility or begins to sound unnatural? Voelker (1938) reports 100 to 175 words per minute as the range of speech rates recommended for radio announcers. The optimal rate for "untrained" speakers is probably closer to the lower end of this range than to the higher. The results of a study by Abrams, Goffard, Kryter, Miller, Sanford, and Sanford (1944) suggest that intelligibility falls off slightly as speech rate increases from 100 to 150 words per minute, and somewhat faster as the rate increases even more. The mean rate that these investigators obtained from 47 speakers was about 140 words per minute; the rate that was judged to be optimal by listeners in their experiment was about 120 words per minute.

Timing at the Syllabic and Phonemic Levels

Miller (1962) has suggested that in order to comprehend messages spoken at the rate of 150 words per minute, one would have, at least implicitly, to make about a dozen phonemic decisions per second. If, as the results of Abrams et al. (1944) suggest, the average speaking rate may be closer to 140 words per minute, Miller's tacit assumption of about 4.8 phonemes per word would lead to an estimate of an average phoneme production rate of about 11 per second. Assuming an average of about three phonemes per syllable (counting those phonemes that mark syllabic boundaries only once), Pickett's estimate of 3.3 per second as the average rate of syllable production again suggests roughly 10 per second as the average rate of phoneme production in continuous speech. We take 80 to 100 msec., therefore, as useful, round-figure estimates of average phoneme duration.[2] The duration of individual speech sounds may vary, however, from a few tens of milliseconds to several hundred milliseconds, depending on such factors as the type of phoneme, the phonetic environment, the speaker, linguistic stress, and the overall speech rate.

Some speech sounds are inherently shorter than others by virtue of the way they are produced. For example, the closure or constricted interval for voiceless consonants tends to be longer than that for voiced consonants, and this interval is longer for fricatives than for stops. (See, for example, Lehiste, 1970, and Klatt, 1974b.) Peterson and Lehiste (1960) and House (1961) have shown that while the durations of vowels in stressed utterance-final syllables in English may vary over a range of from less than 100 msec. to over 400 msec., a large amount of the variability can be attributed to variables that are operative across speakers. For example, lax vowels (/I, ɛ, ʌ, ʊ/) tend to be shorter on the average than their tense counterparts (/i, e, ɚ, u/).

The phonetic environment in which a vowel occurs also affects the vowel's duration. In the studies of Peterson and Lehiste and of House, for example, vowels bordered by voiced consonants were longer than those bordered by voiceless consonants, and vowels bordered by fricative consonants tended to be slightly longer than those bordered by plosives. The duration of a vowel has been shown to provide a cue for voicing of a following consonant (Denes, 1955). Likewise, duration of a consonant is also influenced by its phonetic environment, particularly in consonant clusters (Klatt, 1973).

Another factor that influences the duration of a speech sound is the stress pattern of the utterance in which it occurs. Unstressed vowels tend, for example, to be shorter than stressed vowels (Parmenter & Trevino, 1935), and this duration modification has been shown to be a cue for the perception of stress (Fry, 1958). In fact, in conversational speech an unstressed vowel in certain phonetic environments (particularly before a stressed vowel) can be as short as one or two glottal periods, and, under some circumstances, may be eliminated altogether. Consonants in unstressed syllables are also shorter than those in stressed syllables.

The durations of speech sounds are also influenced by effects that operate at the level of words and phrases. For example, the final vowel of a word is lengthened relative to its inherent duration, and the durations of the individual segments that precede the final syllable are shortened by an amount that depends on the number of syllables in the word (Lindblom & Rapp, 1973). As a result of these effects, the segment durations in longer words tend to be shorter on the average than those in shorter words, other things being equal. This shortening is greatest when the word length increases from one to two syllables, and the additional shortening becomes small when the number of syllables increases beyond two. The lengthening of a syllable in word-final position is most marked when the word occurs at the end of a phrase, particularly before a pause (Klatt, 1974b).

When a person intentionally slows down or speeds up his rate of speech, part of the rate change is accomplished by changes in pause durations and part by changes in the durations of individual speech sounds. It is clear that the latter cannot be accomplished by a linear transformation of the time scale, because the durations of some speech sounds are relatively free to vary, while those of others are not. There are other factors as well, however, that determine the nature of the changes that are made. There is some evidence, for example, that the relative durations of unstressed vowels decrease more than those of stressed vowels when the rate of speech in increased (Peterson & Lehiste, 1960).

A question of some interest is that of the extent to which the durations of specific phonemes can vary within a given context and still sound natural. If, as has been suggested by many investigators, speech is rhythmic, one might expect that variations in the durations of phonemes would be compensatory to some extent in order that regularities in timing above the phoneme level might be preserved. Some evidence that this is the case in Russian has been reported by Kozhevnikow and Chistovich (1965), who found a negative correlation between the durations of adjacent sound segments in running speech. Huggins (1967, 1972) investigated the possibility of compensatory durational variations in English by manipulating the durations of bordering phonemes experimentally. He was unable to demonstrate compensatory effects with phonemes that were parts of the same syllable (e.g., the stop closure of the initial p, and the following stressed vowel in "paupers"), but did find evidence of them when the phonemes involved were contained in adjacent syllables. These results were taken by Huggins as support for the view that the temporal fluency of an utterance is determined by timing relationships at the syllabic--as opposed to the segmental--level; and, in particular, that the most important factor is the maintenance of a rhythmic pattern or a syllabic "beat," in which stressed vowels are the primary elements.

Toward a Set of Rules for Timing

Although a complete theory of timing in speech has not yet evolved, the beginnings of such a theory have been proposed (Klatt, 1974a; Lindblom & Rapp, 1973). A complete theory must account both for timing effects at the level of the phonetic segment (such as the inherent difference in duration between the long vowel /e/ and the short vowel /ɛ/), and for grosser timing influences that span words, phrases and sentences and that include the influence of stress within such units. We summarize here only the major features of such a theory--features that account for the main timing effects that might be relevant to the diagnosis and speech training of deaf individuals.

As a starting point, an inherent duration is postulated for each phonetic segment, and then it is assumed that this duration is modified for particular utterances. The nature of any modification will depend on the context in which the segment occurs. A number of factors such as those discussed above determine what is the inherent duration of a segment, but three of the more important factors are: (1) lax vowels are shorter than tense vowels; (2) a vowel followed by a voiceless consonant is shorter than one followed by a voiced consonant; and (3) fricative consonants are longer than most other consonants.

When a sequence of segments is put together to form a monosyllabic word, the inherent durations are modified in accordance with several kinds of rules. If there are consonant clusters in the word, certain rules are imposed to adjust the consonant durations, usually (but not always) in the direction of shortening the individual consonant elements, i.e., in the direction of making the duration of the consonant cluster closer to the duration of a single consonant. If the word has more than one syllable, one of the syllables receives primary stress, and others may receive secondary stress or be unstressed. The durations of individual segments within the word are adjusted to make unstressed vowels shorter, and also to shorten consonants that occur in unstressed syllables. When a word consists of more than one syllable, the syllables are shortened relative to their inherent durations, and speech sounds that occur before a pause are lengthened. Speech sounds that occur in phrase-final position are lengthened relative to their inherent durations.

Spectrograms of two sentences which illustrate the operation of some of these rules are shown in Fig. 1. The number of speech sounds in these sentences is 31, and the mean duration is about 80 msec., or about 12 speech sounds per second. The duration of the longest sound (/æ/ in bad) is about 230 msec., and the shortest (/t/ in bitter) is about 15 msec. The first sentence, "My sister has a fish," shows the prepausal lengthening of the final consonant and of the vowel in fish (compare the duration of this vowel with that of the vowel in sis), and some durational effects of stress (the unstressed function word a is short, as is the s following the stressed vowel in sister). In the second sentence, "The bitter lemon was bad," several effects combine to shorten the vowel /I/-- a two-syllable word, a lax vowel that is not the final vowel in the word, and a following voiceless consonant (which in this case reduces to a flapped t)--and to lengthen the vowel /æ/--a monosyllabic word occurring before a pause, a tense vowel, and a final voiced consonant.

These comments indicate that the durations of vowels and consonants in continuous discourse differ considerably from the durations of these speech events in isolated monosyllables or words. Furthermore, the way that different speech sounds are produced in sentences may differ from one context to another in aspects other than duration. These context-conditioned variations are observable in the acoustic signal (Stevens & House, 1963; Lindblom, 1963), in the articulatory targets and movements (Daniloff & Moll, 1968; Gay, Ushijima, Hirose, & Cooper, 1973), and even in the efferent neural signals that give rise to the articulatory movements (Gay, et al., 1973). Apparently, the effects of context are not simply the results of undershoot in items in a sequence of individual invariant articulatory targets when the commands to produce these targets occur in rapid succession. There appears to be a reorganization of the motor commands for a given segment depending on the context and on the timing constraints. It is as though the speaker has a variety of ways of producing the gesture for a given speech sound, and in a given situation he selects a particular one of these. These observations indicate the difficulties that face a deaf child when he is learning to produce speech with the proper temporal characteristics.

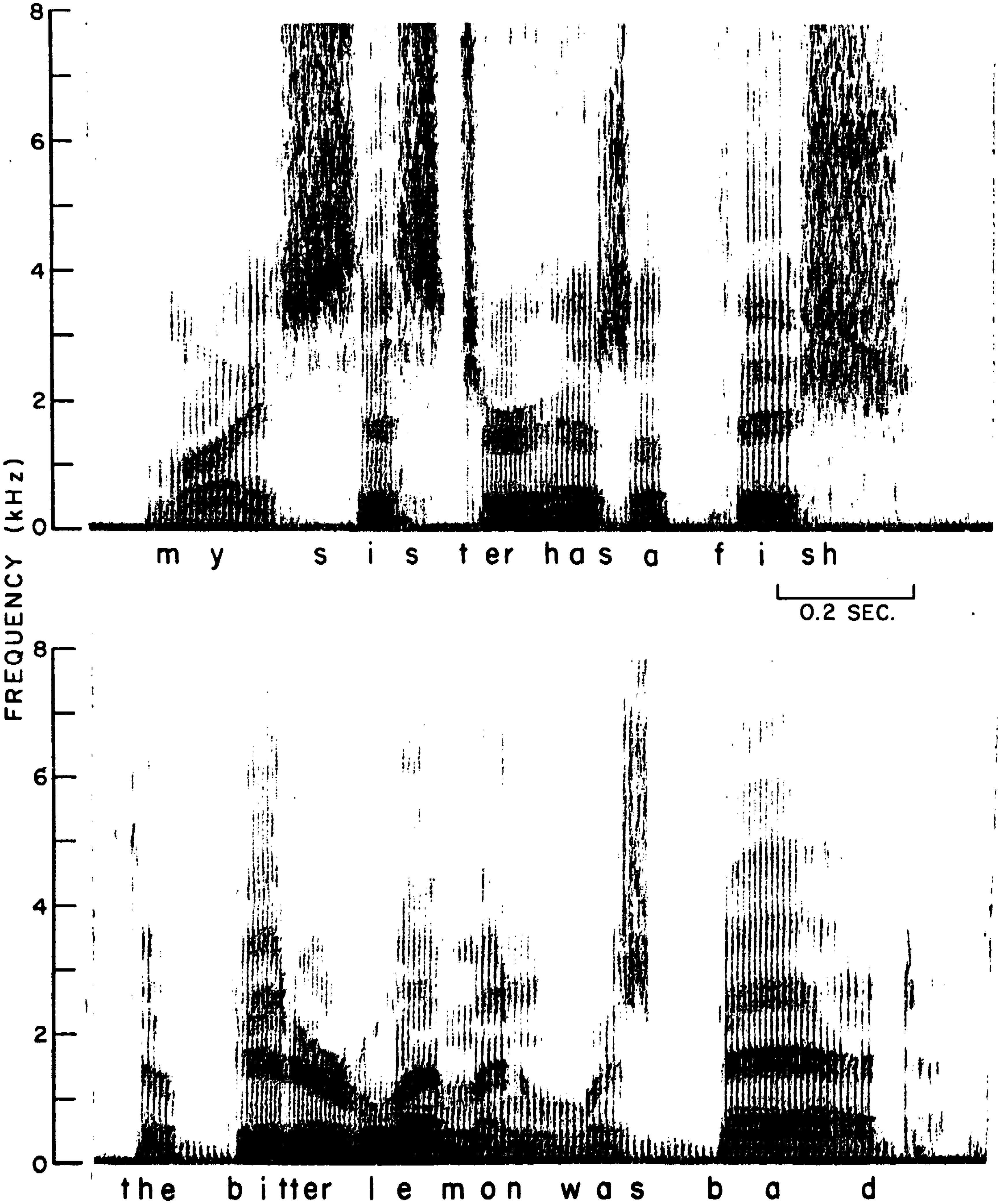

Fig. 1. Spectrograms of two sentences as shown, produced by an adult male speaker.

TIMING AND SPEECH INTELLIGIBILITY

Many investigators have called attention to the importance of timing for intelligibility. Among the first to do so was A. G. Bell. A much quoted comment of his on this topic states the case in a rather emphatic way: "Ordinary people who know nothing of phonetics or elocution have difficulty in understanding slow speech composed of perfect elementary sounds, while they have no difficulty in comprehending an imperfect gabble if only the accent and rhythm are natural" (Bell, 1916, p. 15). While other investigators might take issue with Bell's apparent deemphasis of the role of articulation, probably most would agree that timing is an important factor in determining how intelligible speech will be. A few have presented evidence on this point.

Hudgins and Numbers (1942) found that sentences spoken with rhythm that was judged by listeners to be correct were about 3.5 times more likely to be understood than were sentences whose rhythm was judged to be incorrect. Hood (1967) had listeners judge the proficiency of speech rhythm of sentences that had been recorded by deaf and hearing speakers, and also subjected the same recordings to a variety of acoustic analyses. He found that measures of duration were more highly correlated with intelligibility than were those of fundamental frequency or intensity. Cohen, Schouten, and t'Hart (1962) have shown that recognizable speech can be generated without deliberate control of formant transitions, which is to say that "spectral detail can sometimes be dispensed with provided that temporal detail is intact" (Huggins, 1972, p. 1280). Apparently, although speech can remain intelligibile in the face of many types of manipulations of the speech signal, temporal distortions, especially if added to some other type of manipulation, are likely to render it unintelligible.

Huggins (1972) has suggested that the importance of timing for intelligibility should not be surprising. He points out that such prosodic features as suprasegmental timing and rhythm are among the most resistant properties of the speech waveform to the various types of natural distortions that can occur, and argues that that fact alone should give such cues special significance in the perception of speech.

One might conclude from such findings and observations that improvement in the timing and rhythm of the speech of the deaf would invariably increase its intelligibility. In fact, studies relating to this issue have had mixed results. John and Howarth (1965) attempted to improve the timing aspects of the speech of 29 deaf children, while ignoring other aspects of their speech. The children were encouraged to use whatever residual hearing (amplified) they had for perceiving the time patterns of the speech. Phrases were used; phonemes, syllables or words were not dealt with individually. Training consisted of spending three or four minutes working with each of several sentences. The sentences that were used with a given child were originally obtained from his spontaneous speech. Untrained observers listened to recorded before- and after-training samples. Intelligibility (number of words recognized) was about 19% and 30% for the before- and after-training samples, respectively. A second method of scoring was used that was sensitive to the listener's perception of the syntactic pattern of an utterance: perception of "Put the man in the house" as "Put the (noun) in the house" was scored correct. In terms of this measure, performance was about 200% better with the after-training utterances.

In contrast to John and Howarth's results, however, some investigators have obtained improvement in timing accompanied either by no change, or actual decreases, in intelligibility (House, 1973; Stratton, 1973). There are several plausible explanations for the latter finding. One possibility is that focusing intensively on a single aspect of speech during training sessions may have the effect of permitting other aspects, which are not being attended to, to deteriorate. Or possibly the act of changing speech behavior with respect to certain features may naturally introduce changes, not necessarily beneficial, with respect to other features as well. For example, if a child's poor timing is due in part to difficulties he has in articulating certain phonemes, forcing him to produce more appropriate timing patterns may make it even more difficult for him to articulate those sounds properly.

If indeed training with respect to one aspect of speech has the effect of decreasing intelligibility, it does not follow that such training is ill-advised. It may be that in some instances short-term setbacks are necessary if significant long-range improvement is to be realized. In any case, these results serve as a poignant reminder of the integrity of speech and the interrelatedness of the problems associated with it. It is conceptually convenient to think of speech in terms of properties--intensity, fundamental frequency, nasality, timing, and so forth--however, speech is speech, and to modify any aspect of it is to affect it as a whole. Perhaps the important point, as far as training is concerned, is that one cannot assume that training with respect to one specific speech property will leave performance with respect to other properties unaffected. While concentration on one or a small set of properties at any given time may be a necessary training strategy, such concentration should probably be coupled with at least informal monitoring of performance with respect to the properties that are not being focused on as well.

PREVIOUS FINDINGS WITH REGARD TO TIMING PROBLEMS ASSOCIATED WITH SPEECH OF THE DEAF

We have already noted that many researchers and speech teachers have felt that timing problems are significant contributors to the lack of intelligibility of the speech of the deaf. We turn now to a consideration of some of the evidence for that claim, or at least of some data that relate to it. It will become apparent that surprisingly little of a definitive nature can be said concerning the relative importance of specific temporal features as determinants of the intelligibility or quality of the speech of the deaf, because (1) few empirical studies have been addressed to this issue, and (2) relatively little is known concerning the role of temporal properties as determinants of the intelligibility of "normal" speech. A few studies have been done, however, that have produced results that are at least suggestive of what some of the dimensions of the problem are.

Speech Rate

Many investigators have noted the relatively slow speech rate of deaf speakers (Boone, 1966; Colton & Cooker, 1968; Hood, 1966; John & Howarth, 1965; Martony, 1966; Mason & Bright, 1937; Voelker, 1937, 1938). In one study, Voelker (1938) measured the word production rate of 98 first-, second-, and third-grade students at the Ohio School for the Deaf and of a control group comprised of hearing children and teachers of the deaf. The average rates obtained were 168 and 67 words per minute for the control group and the deaf speakers, respectively. The ranges (slowest to fastest speaker) for the two groups were 134 to 210 and 29 to 145. The distributions overlapped very little; only two of the deaf speakers spoke faster than the slowest hearing speaker.

In a subsequent analysis of his speech samples, Voelker (1937; the publication of the second analysis preceded that of the first) counted individual speech sounds based on phonetic transcriptions in order to take into account the possibility that the deaf speakers were producing more sounds (because of adventitious phonetic elements) than the word count would imply. The average rates of speech-sound production were 469 and 210 sounds per minute (spm) for the hearing and deaf speakers, respectively. The ranges (slowest to fastest speaker) for the two groups were 376 to 586 spm and 80 to 406 spm. Again, the distributions overlapped little (only two of the deaf speakers were faster than the slowest hearing speaker), and the distribution of rates for the deaf had the greater spread.

Other investigators who have compared speech rates of normal-hearing and deaf speakers have also found the rates to be considerably slower in the latter case (Colton & Cooker, 1968; Hood, 1966; Mason & Bright (1937). Hood's sample of deaf speakers spoke from about two to three-and-one-half times more slowly than his normal-hearing controls. Mason and Bright (1937) found even less overlap in the speech rates of deaf and hearing speakers than did Voelker.

Speech Rhythm

Inasmuch as the rhythmic properties of the speech of the hearing are not well understood, it is not to be expected that the deficiencies of the speech of the deaf in this regard can be very precisely specified. On the point that the speech of the deaf typically is deficient in this regard, there seems to be general agreement, however. Teachers of the deaf have long stressed the importance of emphasizing proper rhythm, or phrasing, as a basic training objective (e.g., Brehm, 1922). Hudgins (1946) has noted that deaf speakers have a tendency to group syllables inappropriately. Hood (1966) had listeners rate the adequacy of the rhythm of the speech of deaf and hearing speakers, and consistently obtained lower ratings for the speech of the deaf.

DiCarlo (1964) cites some evidence that deaf subjects do more poorly than do those with normal hearing (including the blind) on tests involving the discrimination of tactile rhythm patterns. This raises the question of whether the lack of hearing inhibits not only the development of rhythmic speech, but of a sense of rhythm in general.

Timing at the Syllabic and Phonemic Levels

Given that the speech of the deaf tends to be slower on the average than that of normal-hearing speakers, a question that naturally arises is whether it differs, temporally, from the speech that results when individuals with normal hearing are asked to speak more slowly than they habitually do. There is some evidence on this point in the data reported by Hood (1966) and by John and

Howarth (1965). The average syllable duration in Hood's study was from one-and-one-half to over two times longer for the deaf speakers than for the controls, not quite as large a discrepancy as that between the word-emission rates of the two groups. Moreover, the ratio of phonation time to total speaking time was considerably higher for hearing speakers than for the deaf (.90 for hearing speakers, .76 and .66 for two deaf groups). "The abnormally slow rate of utterance of the deaf speakers, therefore, was a result of a combination of prolonged syllables and prolonged pauses between words" (p. 58). The finding of prolonged between-word pauses is borne out by the data of John and Howarth (1965), who reported that such pauses often accounted for half the time taken by a deaf child to say a sentence. One must interpret this result cautiously, however, because deaf children often have reading difficulties which could affect the durations of between-word pauses in read speech.

Hood (1966) also found that syllables produced by deaf speakers were more variable in duration than those produced by hearing speakers, although he noted large individual differences on this measure. Given that the durations of syllables produced by the deaf were longer on the average, the greater variability could have been in part a consequence of the fact that the disperson of a random variable tends to increase with its mean.

Angelocci (1962) has studied timing at the phonemic level for two-syllable nonsense words (həCVk) that were recorded by three groups of speakers: five profoundly deaf, five with normal hearing, and five with normal hearing who were attempting to imitate the speech of the deaf. Two groups of listeners heard the recorded nonsense words in random order: one group (teachers of the deaf) attempted to judge which words had been spoken by the deaf speakers; the other group (people trained in phonetic transcription) attempted to transcribe the sounds heard in each sample.

For some of the samples (apparently those for which the judges' decisions were in agreement), the durations of the following sounds were measured objectively: the unstressed vowel, the fricative or plosive that followed it, and the stressed vowel that followed the consonant. It was found that the durations of the unstressed vowels produced by the deaf speakers were typically four to five times as long as the average of those produced by hearing speakers; for stressed vowels (/æ/ and /u/) the ratio was two or three to one. Inasmuch as stressed vowels tend to be longer than unstressed vowels in normal speech (Parmenter & Trevino, 1935) the implication is that the relative difference between the duration of stressed and unstressed vowels is larger for hearing than for deaf speakers. Also for the hearing speakers in Angelocci's study, the low vowel /æ/ was typically longer than the high vowel /u/, and unstressed vowels were longer before voiced than before voiceless consonants; but neither of these relationships held for the deaf.

Fricative consonants were four to five times longer for the deaf than for the hearing speakers (see also Calvert, 1961). For the hearing, voiced fricatives were shorter than their voiceless cognates if surrounding sounds were constant, but this was not so for the deaf. The closure periods of plosive consonants were three to four times longer for the deaf than for the hearing speakers. Hearing speakers typically had longer closure durations for voiceless than for voiced plosives; for the deaf, this relationship was reversed. Angelocci noted that when the surd-sonant error occurred (production of a voiceless plosive when a voiced plosive is intended, or vice versa), the duration of the release period of the plosive was appropriate for the sound that was heard rather than for the one that was intended. "The perception of voicing for p, t, d, b seemed to be associated with the duration of the release period of the plosive" (p. 402).

Angelocci summarized his findings as follows. "Deaf speakers typically distorted the duration of phonemes in this study, first by extending their duration several times that of hearing speakers, and second by not following the relative differences in duration as a function of voicing of consonants or of the effect of one sound upon another that is commonly found among normally hearing speakers. In distorting these durations, deaf speakers destroy cues which may help us in understanding their speech" (p. 402). While the generality of these results may be questioned because of the fact that they were obtained with nonsense words spoken in isolation, the findings constitute a starting point for broader investigation of the temporal aspects of conversational speech.

THE RELATIONSHIP BETWEEN TIMING DEFICIENCIES AND OTHER SPEECH PROBLEMS

Given the fact that timing relates to the intelligibility and quality of speech in a variety of ways, it would be surprising if its relationship to particular speech problems were not a complex one. In fact, several investigators have discussed specific problems that either contribute to, or are based on, timing deficiencies to some degree.

Timing and Breathing

Apparently, some of the timing difficulties that deaf speakers have may stem from faulty breathing during speech. Scuri discussed this relationship in an Italian journal in 1935; Hudgins (1936) published a review of the article in English the following year. Scuri found that his deaf speakers ventilated a great deal more during speech than when not speaking; whereas normal-hearing speakers tend to use approximately the same amount of air volume in both cases. The normal ratio between inspiration and expiration during quiet breathing is about three to four, whereas during speech the ratio is about one to three, or one to four. Hudgins points out that the ratios are very similar to these with very young deaf children, but as the children grow older and begin to attempt consciously to acquire speech, their ratios change, particularly during speech. Some studies have found that deaf speakers use about twice as many breaths as speakers with normal hearing (Hudgins, 1934; Rawlings, 1935, 1936).

Scuri's data also suggested that the deaf tend to lack the ability to close the glottis completely, which perhaps explains in part why deaf individuals tend to have "breathy" voices and also why they lose breath before phonation starts. It is claimed that "frequently half of the breath supply is lost before the voice begins. There are two factors operating in this type of defect: (1) the air column from the chest lacks force, due to weakness and incoordination of the breathing muscles; and (2) the glottis does not close sufficiently to permit the weak air column to set the vocal cords into vibration" (p. 343). This observation again points up the interdependence of speech problems and the difficulty of treating them in isolation. If breathiness and timing aberrations that result from the need to take frequent breaths are based, to some degree, on the same glottal deficiency, it may not be possible to treat one problem effectively without also treating the other.

In a later paper, Hudgins (1946) summarizes the speech-breathing problems of deaf children with the following list: "(a) short irregular breath groups often only one or two words in length with breath pauses interrupting the speech flow at improper points; (b) excessive expenditure of breath on single syllables resulting in breathy speech; (c) false grouping of syllables resulting in the breaking up of natural groups and the misplacement of accents; (d) a slow methodical utterance resulting in a complete lack of grouping; and (e) a lack of proper coordination between breathing muscles and articulatory organs" (p. 642). Clearly, these problems all have implications for timing. Hudgins emphasizes the importance of teaching proper speech breathing, syllable and word grouping, and rhythm very early in the child's speech training. He notes that poor speech-breathing habits, once established, are difficult to modify.

Timing and Nasality

Colton and Cooker (1968) cite some evidence presented by Bzoch (1965) "which suggests that normal speakers tend to break the velopharyngeal seal when their rate of speech is reduced." They suggest, therefore, that the nasality that often characterizes the speech of the deaf may be a by-product of its slower-than-normal tempo. If this suggestion is valid, it corroborates Calvert's (1962) and Jones' (1967) observations concerning the importance of the role of the dynamic aspects of speech timing--the transitions from one articulatory position to another--as determinants of voice quality.

Timing and Articulation

A distinction is often made between timing problems and problems of articulation. While the distinction is a helpful one for some purposes, it should not be pressed too far. Articulation itself depends upon proper timing at the level of individual speech sounds and the transitions between them. The control of voice-onset time relative to release for a voiceless stop consonant, the timing of movements in a sequence of consonants preceding a vowel, the timing of transitions between a fricative or a nasal consonant and a vowel are all examples of articulatory timing demands that can cause problems for deaf speakers at the level of individual speech sounds and transitions.

Even the distinction between timing problems that apply to the production of individual speech sounds and those that relate to suprasegmental, or prosodic, aspects of speech cannot be maintained without qualification. The results obtained by Hood (1966), for example, suggest that deaf children who tend to make syllables of relatively long duration are likely to be judged to have poor speech rhythm.

It is clear that faulty articulation can detrimentally affect speech rhythm. As a case in point, Stewart (1969) notes the difficulty that many deaf speakers have with the articulation of the fricatives and affricates of English. More generally, the introduction of intrusive stop elements into the pronunciation of fricatives and the omission of stop elements when they should be there are both noted as problems. "When very widespread in extent, the insertion of intrusive stop elements seems to impart a slightly 'clipped' quality to the speech" (p. 42).

John and Howarth (1965) have also described intrusive sounds that can result in errors of timing: intrusive glides from one phoneme to another and intrusive sounds associated with consonants. The following of a final nasal consonant by plosions, which in turn may be followed by prolonged aspiration, is also mentioned. These investigators suggested that some of these timing errors may result from unduly slow and deliberate movements of the articulators and of unnecessary emphasis with which some phonemes are produced. "These errors in duration may be due to the children's preoccupation with the articulation of the individual phonemes in a word and with the pronunciation of words as separate items in a sentence" (p. 128).

The possibility that articulation training may interfere with the acquisition of proper timing and rhythm has considerable significance for the development of optimal strategies for teaching speech. Other writers, in addition to John and Howarth, have suggested the possibility either directly or indirectly. Boone (1966), for example, noted that drill in the production of isolated phonemes can affect the way in which deaf children synthesize phonemes in word production: Often the phonemes continue to maintain their separate identities and are not influenced by the occurrence of adjacent phonemes. Boone saw this as the reason for the tendency for deaf speakers to prolong and diphthongize vowels and some consonants. In normal speech, the duration of a particular phoneme varies considerably with the context in which it occurs, as has been noted above. Speech in which individual phonemes do not vary in duration as a function of context will sound artificial at best.

Borrild (1968) has remarked on the difficulty of teaching timing relative to that of teaching the articulation of isolated sounds. He claims that, with the exception of the voiceless consonant /s/, which often is difficult to learn, and even when learned often is not used in speech, speech teachers encounter little serious difficulty in teaching correct articulation of isolated speech sounds. Very great difficulties are encountered, however, when efforts are made to integrate articulated sounds into fluent speech. The problem, according to Borrild, seems to be with rhythm and intonation. He does not suggest that the difficulties encountered in the acquisition of correct timing and intonation patterns are direct consequences of the way in which articulation is taught, but the possibility is implicit in his observation.

SOME ADDITIONAL DATA ON TIMING FOR DEAF AND NORMALLY-HEARING SPEAKERS

In an effort to obtain some additional data on how the speech of deaf children compares with that of hearing children and adults with respect to timing, recorded speech samples obtained from deaf and hearing speakers were subjected to a variety of analyses.

Speakers

Speech samples were obtained from three groups of speakers (25 individuals per group): deaf children, normally-hearing children and normally-hearing adults. The deaf children (12 boys, 13 girls) were students at the Clarke School for the Deaf in Northampton, Massachusetts. The ages ranged from 9 to 15 years and averaged 12.1 years. All of these students were profoundly deaf, with a hearing loss (better ear, with amplification) of more than 90 dB ISO in the range 500 to 2000 Hz.

The normally-hearing children (14 boys, 11 girls) were students in Boston-Cambridge schools. Ages ranged from 8 to 13 years, with an average of 10 years. The third group consisted of 25 normally-hearing adults (14 men, 11 women).

The Speech Sample

Each speaker read the following paragraph:

> "My sister has a fish. She keeps it in a tank. The fish has five spots. I think it looks like my sister."

It is not possible, of course, to make a small sample representative of natural speech in all respects. This particular sample doubtlessly is nonrepresentative in its simplicity. Each sentence is relatively short and has a simple grammatical structure; each of the words but one (which occurs twice) has only a single syllable; and so on. The phonetic content of the paragraph was chosen in part for ease of segmentation on a sound spectrogram.

Only the first two sentences of the paragraph were analyzed in detail. These sentences contain 12 syllables, or 29 phonemes. Discounting repeating sounds, the sample contains 17 different phonemes: 11 consonants, 5 vowels, and 1 diphthong. Several classes of manner and place of articulation are represented.

Method of Recording

All speakers were recorded individually after the recording procedure had been explained to them. They were permitted to read and familiarize themselves with the paragraph before recording it so as to minimize the chances that the results would be unduly sensitive to differences in reading ability, especially among the children. If mistakes were made, words left out or substituted for one another, the speaker was allowed to rerecord the entire paragraph.

Methods of Analysis

Wide-band spectrograms were made of the first two sentences of the paragraph, and time measurements were made from them. "Sis," "fish," "keeps," and "tank" were considered stressed syllables. The sentences were separated into four phrases: "my sister," "has a fish," "she keeps it," and "in a tank." Each phrase was defined as including the pause (if any) after the last word in that phrase. Two other types of segments were defined for the purposes of duration measurements: syllables and pauses. Some of the conventions for defining syllables and pauses were selected for convenience of measurement, and because data from hearing speakers and from deaf children were to be compared. For a syllable beginning with a continuant consonant (e.g., sis, has), the onset of the syllable was taken to be the beginning of the consonant, but for a stop consonant (e.g., ter, tank), the onset of the syllable was taken to be the release of the consonant. Gaps preceding stop consonants that occurred at the beginnings of syllables were counted as pauses, whereas stop gaps within words, such as keeps and it, were not counted as pauses. This procedure was followed since, in making measurements on the speech of deaf children, for whom long pauses are likely to occur, it would often be difficult to distinguish between a pause and a stop gap. The gross differences that will be demonstrated in temporal attributes of the speech of normally-hearing individuals and deaf children are largely independent of these details of the definition of syllables and pauses.

RESULTS

Speech Rate

The results are entirely consistent with those of the several investigators who have found the rate of word emission to be considerably less for deaf than for hearing speakers. Also, the deaf speakers were somewhat the more variable in this regard. On the average, deaf children took about 1.6 times as long to say the two sentences as did the hearing speakers (Fig. 2). Hearing adults and hearing children differed very little with respect to this measure, although the children were slightly more variable. The word-emission rates for the hearing speakers averaged about 179 words per minute for both children and adults. The fact that these are relatively high (as compared, for example, with the average of 140 words per minute reported by Abrams, et al., 1944) may be due in part to the fact that all of the words but one were comprised of a single syllable. On the assumption that the speakers articulated the sentences correctly, the rate of syllable production was about 3.3 per second, and that of phoneme production about 8 per second. The first number agrees precisely with Pickett's (1968) estimate of syllable-production rate. The latter number is considerably less than the 12 phonemes per second estimated by Miller (1962). This discrepancy could be due in part to the fact that the average number of phonemes per word in our sample was not quite three, whereas Miller's estimate was based on an assumption of approximately 5 phonemes per word. It is quite probably the case that our sample has fewer phonemes per syllable (and hence per word) than a random sample of conversational speech. (There are, in fact, only two cases in which two syllabic nuclei are separated by more than one consonant.) What is particularly interesting is the fact that the syllable production-rate appears to be so little affected by this factor.

More to the main point of the paper, however, is the fact that the speech of the deaf speakers was much slower than that of the hearing speakers, independently of the speech-rate index that is used. The average word, syllable and phoneme production rates for the deaf speakers of this sample were 108 per minute, 2.0 per second, and 4.7 per second, respectively.

Phrase Duration and Sentence-Final Lengthening

When the durations of the individual phrases are compared across the three groups of speakers, the results are similar to

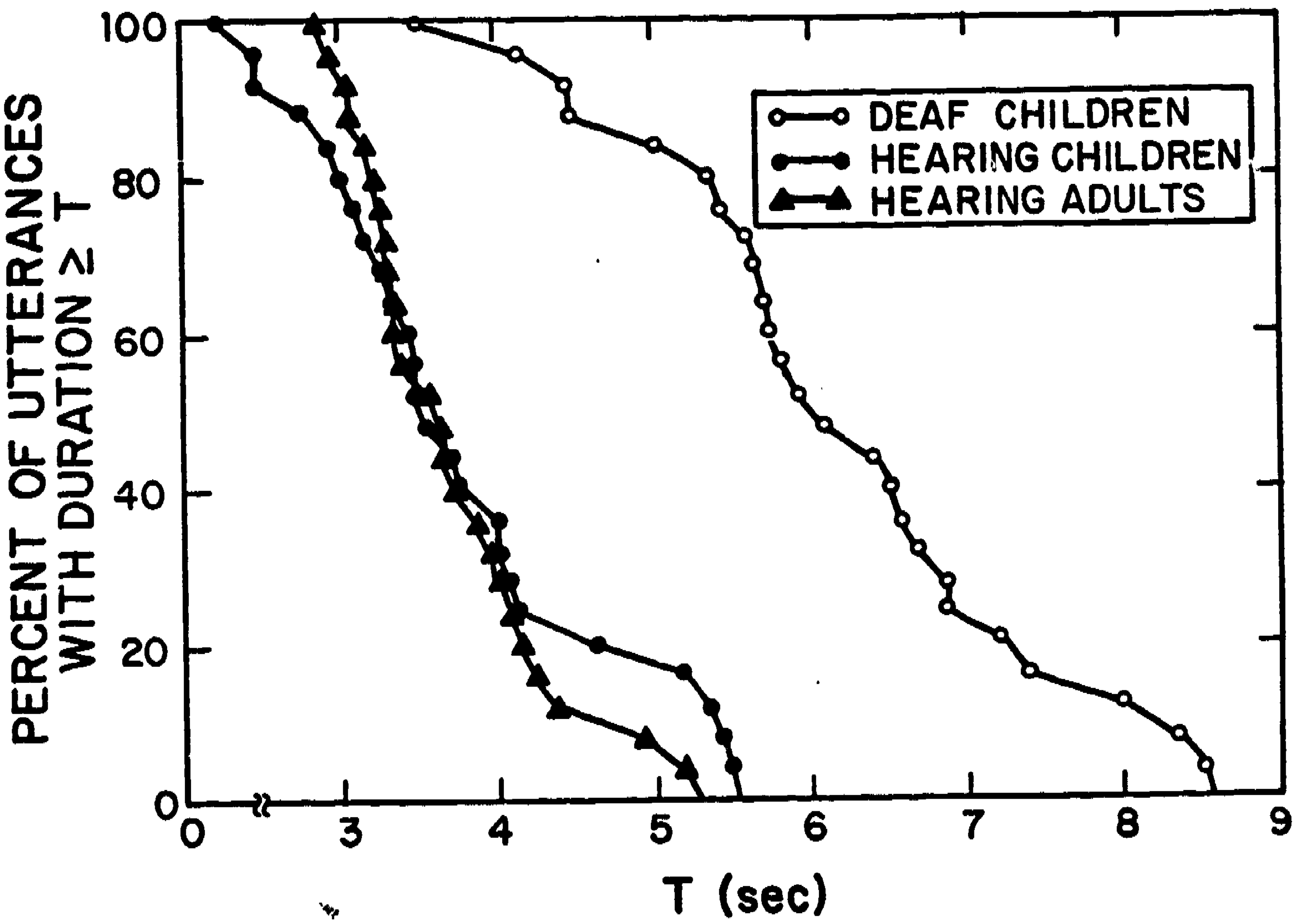

Fig. 2. Cumulative distributions of duration, of utterance ("My sister has a fish. She keeps it in a tank."). The sentences were the first two sentences of a four-sentence paragraph that the speakers read. The durations represent the time from the beginning of the first word in the first sentences to the beginning of the first word in the third sentence; thus, they include not only the pause between the first and second sentences, but that between the second and third as well.

those observed with respect to the durations of the whole utterance: The deaf speakers took longer to produce each phrase, on the average, than did either of the hearing groups, and their range of durations tended to be greater (Fig. 3). For the hearing adults and the hearing children, the average phrase duration was greater for the second (and in these cases, last) phrase of each sentence than for the first phrase, although all phrases have the same number of syllables. For the hearing children, the median percentage increases in duration of the second phrase relative to the first were 51 and 15 for the first and second sentences, respectively. For the hearing adults, the comparable figures were 46 and 15. These increases can, perhaps, be ascribed in part to the phonetic content and contrasting stress patterns of the first and second phrases, but probably the principal source of the differences is the prepausal lengthening of the final syllables, which adds significantly to the length of each terminal phrase. On the average, the final syllables (fish and tank) account for well over one-half of the total duration of the second phrase in each sentence for the normally-hearing speakers. These were the longest syllables in their respective sentences for 96 percent of the sentences produced by these speakers.

For the deaf children, the situation is quite different. Although there is some evidence of prepausal lengthening in the

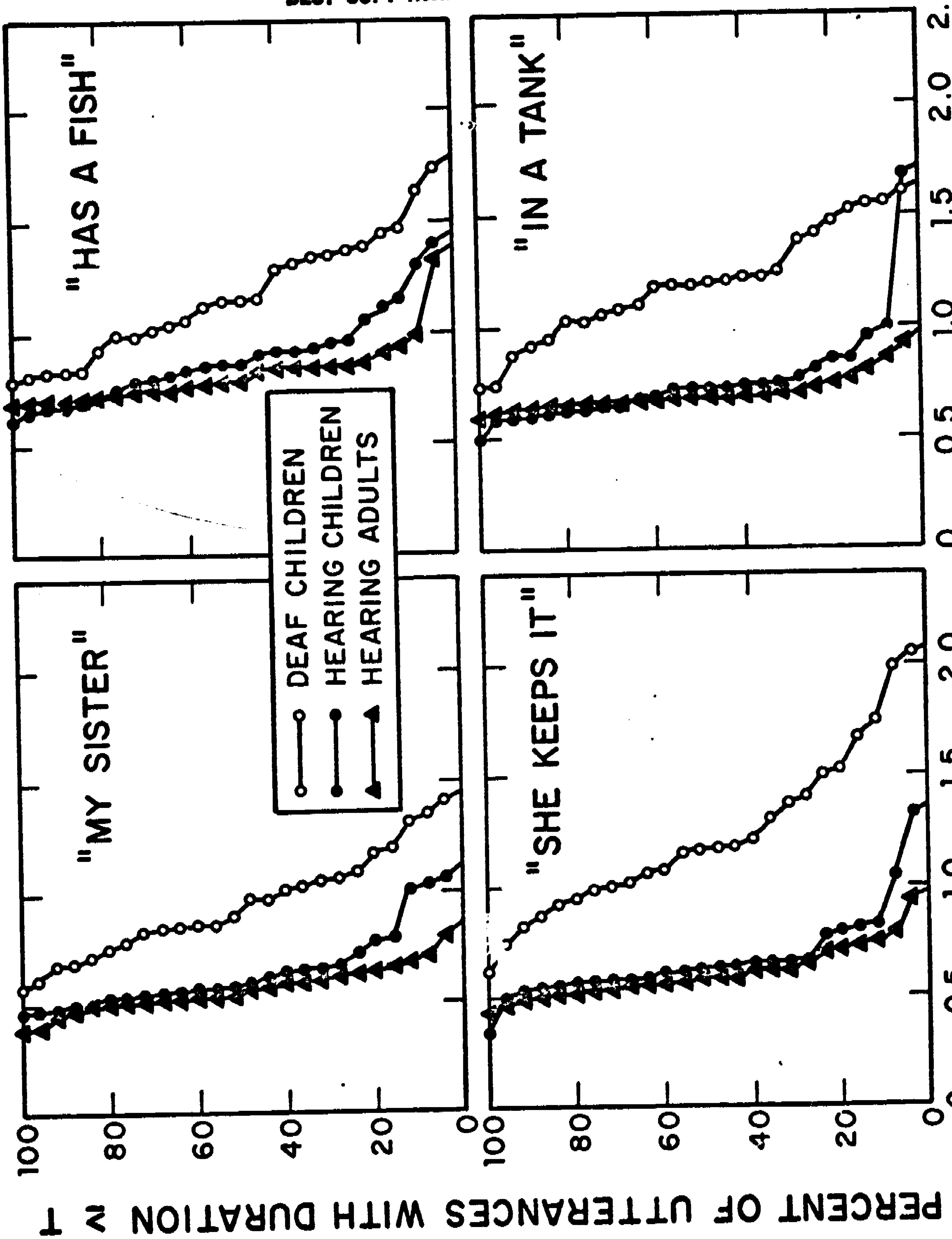

Fig. 3. Cumulative distributions of durations of individual phrases. These durations do not include the pauses (if any) following the phrases or sentences.

fact that the syllables fish and tank were the longest syllables in 68 percent of the sentences produced by these speakers, the increase in the duration of the second phrase relative to that of the first was much smaller, on the average (the median percentage increases were 29 for the first sentence and 2 for the second) for the deaf speakers than for those with normal hearing. A possible explanation of this finding is that many of the deaf children tend not to signal a sentence-final syllable by adjusting the duration of that syllable relative to the durations of the other syllables; or if they do make an adjustment, it tends not to be as large as that made by normally-hearing speakers.

Relative Durations of Stressed and Unstressed Syllables

One of the ways in which a speaker reduces the stress on a syllable is by shortening it relative to its intrinsic duration. The syllables that normally would receive primary stress in the sample sentence are "sis," "fish," "keeps," and "tank." In order to compare the performance of the deaf and hearing speakers with respect to their use of duration as a stress cue, the ratio of the duration of each of these stressed syllables to the duration of an adjacent syllable that should not have been stressed was determined. The ratios obtained apply, of course, only to this particular context, since the durations of individual syllables are influenced by many factors other than stress, as noted earlier.

Figure 4 shows cumulative distributions of ratios of the sum of the durations of the four stressed syllables to the sum of the durations of four bordering unstressed syllables. (The bordering syllables were "ter," "a," "it," and "a.") It is apparent that while both deaf and hearing speakers made the average duration of the unstressed syllables shorter than that of the stressed syllables, the proportional shortening was smaller on the average for the speech produced by deaf children than for that produced by either the hearing children or the hearing adults.

Figure 5 shows the ratios of the durations of specific stressed and unstressed syllables. On the average, both hearing and deaf speakers made the unstressed syllables shorter than the stressed syllables; that is, the ratios tended to be greater than one for each of the syllable pairs. (Fifteen of the 17 exceptions to this rule are found in the speech of the deaf children). In all cases, however, the ratios tended to be greater for both groups of normally-hearing speakers than for the deaf speakers. Median ratios were obtained for each speaker group and each syllable pair. The means of these medians were 4.1, 3.7, and 2.4 for the hearing adults, hearing children, and deaf children, respectively.

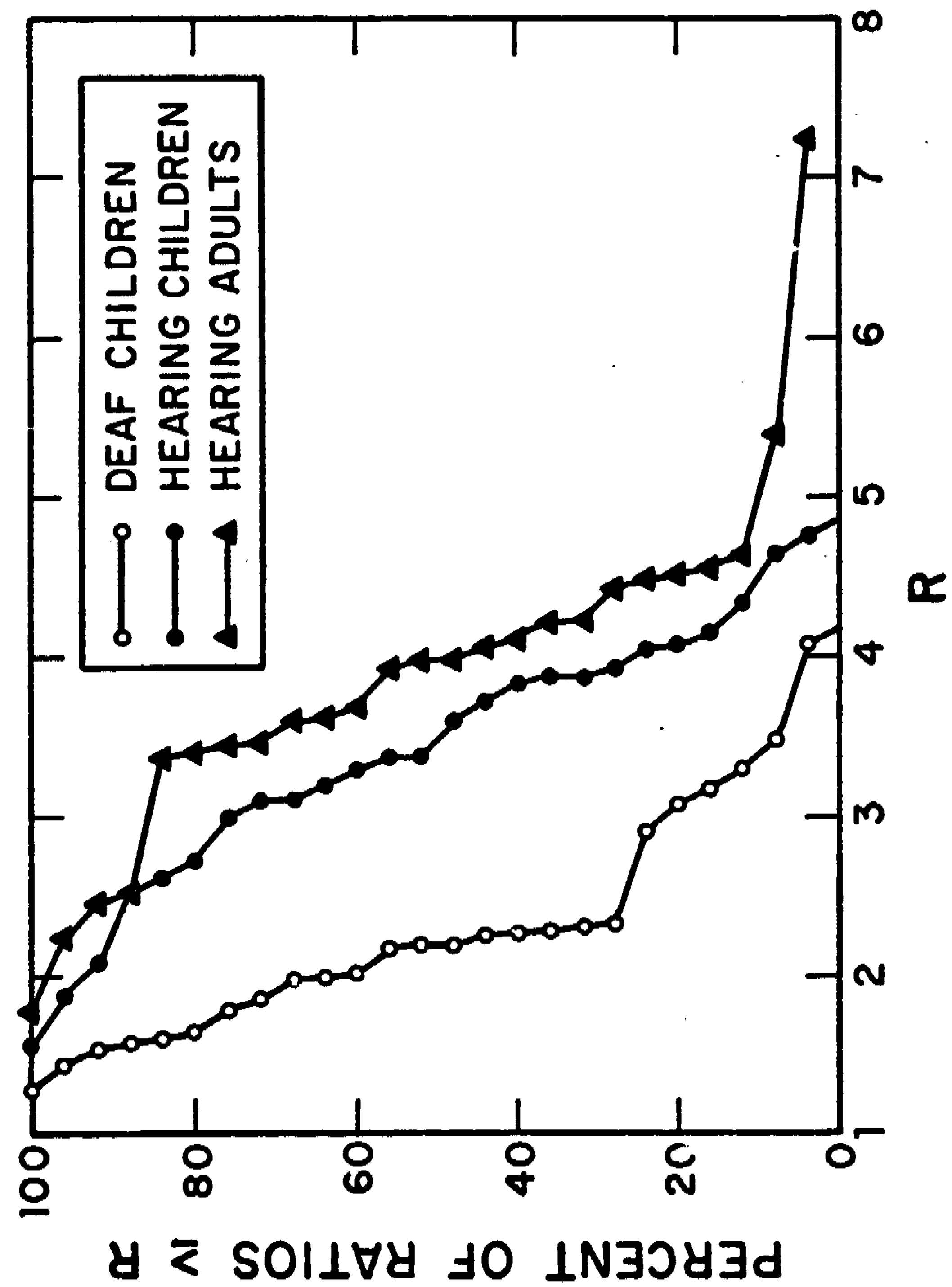

Fig. 4. Cumulative distributions of the ratios of the sum of the durations of four stressed syllables ("sis," "fish," "keeps," and "tank") to the sum of the durations of four bordering unstressed syllables ("ter," "a," "it," and "a").

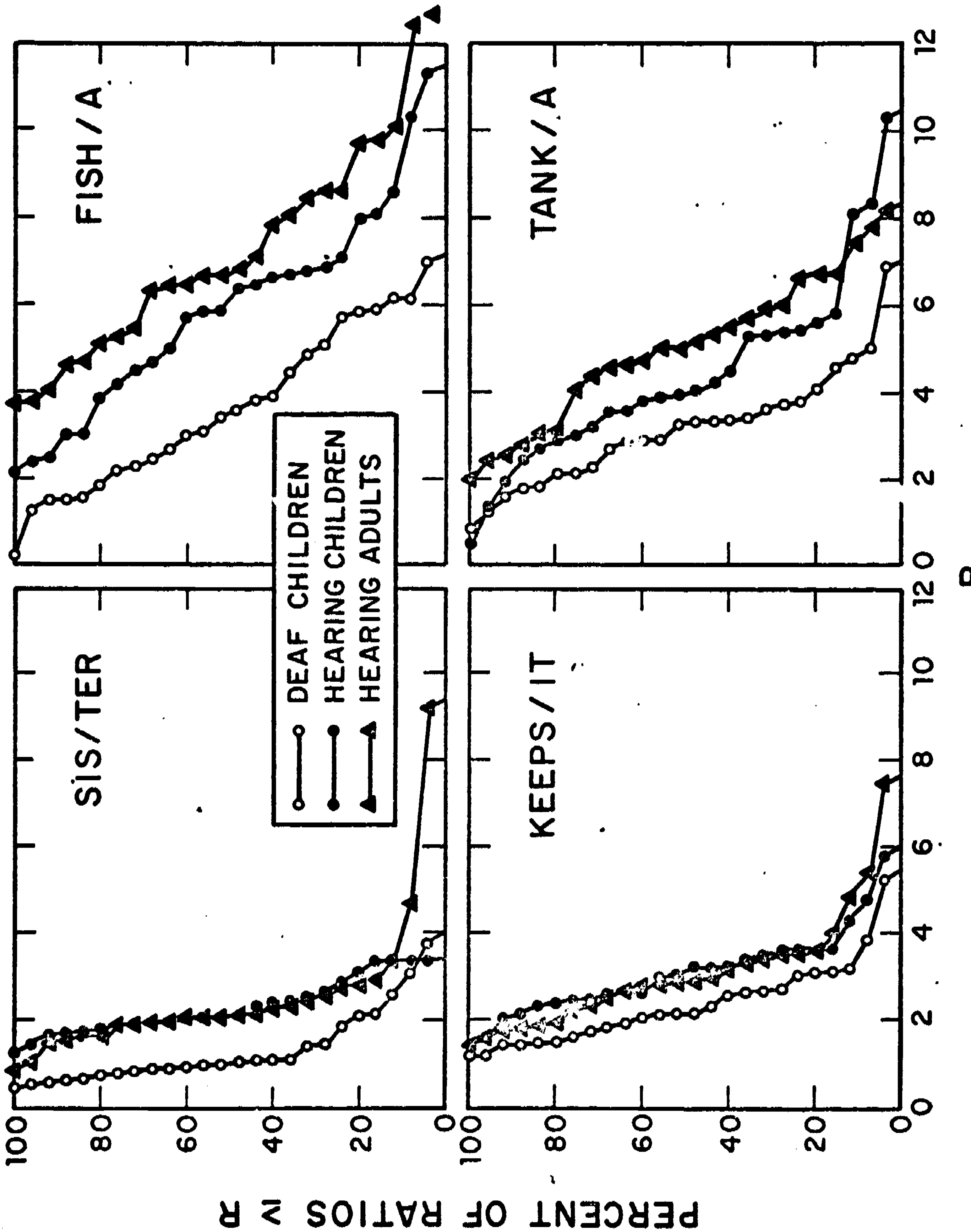

Fig. 5. Cumulative distributions of ratios of the durations of specific stressed and unstressed syllables.

For all groups, the ratio of the durations of stressed and unstressed syllables was greater for the syllable pairs fish/a and tank/a than for sis/ter and keeps/it. The means of the medians for fish/a and tank/a and for sis/ter and keeps/it were 5.9 and 2.4 for the hearing adults, 4.9 and 2.5 for the hearing children, and 3.3 and 1.6 for the deaf children. The first two syllable pairs differ from the latter two in three respects that have implications for timing: (1) Each of the stressed syllables of the first pair is the last syllable of the sentence in which it occurs, and hence undergoes prepausal lengthening; (2) the unstressed syllables of the second pair occur in phrase-final position and would therefore tend to be lengthened (although to a lesser degree than would be stressed syllables in the same positions); and (3) the unstressed syllables of the first pair have fewer phonemes than do those of the second pair. It is apparent from inspection of Figs. 4 and 5 that, while the deaf children did make adjustments in the durations of syllables of the sort that are to be expected as a result of stress and syllable positioning within phrases and sentences, these adjustments were, in most cases, not sufficiently large to give the speech a normal temporal pattern.

A question that is prompted by these results is whether the deaf children failed to produce as large differences between the durations of stressed and unstressed syllables as did hearing speakers because they made the stressed syllables too short or the unstressed syllables too long. Inspection of the data suggests that the latter is the case--which raises again the question of the effect of articulation training on timing (John & Howarth, 1965; Boone, 1966). If a child is trained to articulate each phoneme distinctly, it may be that this has the effect of teaching him to produce unstressed syllables that have longer-than-normal durations.

Pauses

Speech is punctuated with silence. Some pauses must occur, of course, in order to permit the speaker to breathe; however, not all pauses have that function. Moreover, the placement and duration even of those that do are determined by factors other than the speaker's need for additional breath. Pauses typically occur following the ends of sentences and major phrases. The durations of these pauses depend in part on syntax and in part on nonsyntactic factors. In general, it seems to be the case that the larger the syntactic unit that is being delimited, the longer will be the pause that is used to delimit it: Pauses between sentences tend to be longer than pauses between major clauses of a compound sentence, which in turn will be longer than pauses between simple phrases. Notwithstanding these general rules, however, the individual speaker has considerable latitude in varying pause durations as a means of emphasis. A lengthened pause tends to call attention to the sentence or phrase immediately preceding or following it.

Figure 6 compares the speech of the deaf and hearing speakers in our sample with respect to the cumulative duration of the pauses in the sample. It is clear that the speech of the deaf speakers had a greater total amount of silent time than did that of the hearing speakers.

Figure 7 shows the durations of pauses occurring in different syntactic contexts. In interpreting these data, it should be remembered that our convention is to count a stop gap at a syllable onset as a pause. As a consequence, there is a net within-phrase pause for normally-hearing speakers. The figure suggests that the hearing and deaf groups differed more with respect to the durations of inter- and intra-phrase pauses than with respect to the pauses between sentences. Between-phrase pauses were nearly nonexistent for the hearing speakers in this sample, a result that may be attributed to the fact that the sentences were very short, certainly too short to require more than a single exhalation to produce. It should be noted, also, that in fluent speech the phrasing may be adjusted to conform to some as yet ill-defined principle of "ease of production." For example, a normal speaker might produce "She keeps it in a tank" by grouping "keeps it in" together, with the same stress pattern as a word like "Canada." Part of the speech training of the deaf children, on the other hand, is to phrase a sentence by inserting pauses at certain places in a sentence, such as before a prepositional phrase.

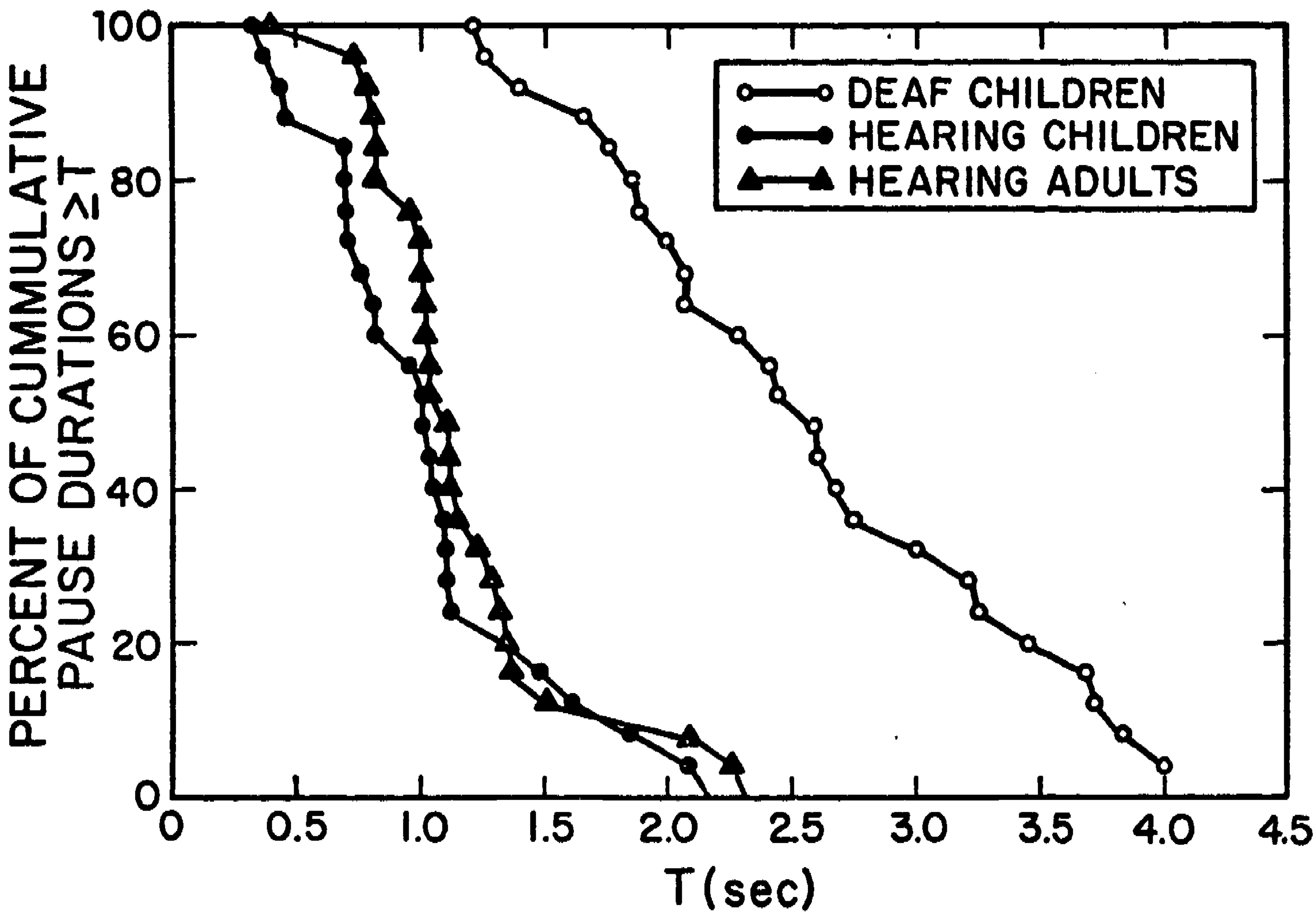

Fig. 6. Cumulative distributions of durations of pauses. (See "Methods of Analysis" section for procedure for identifying pauses.)

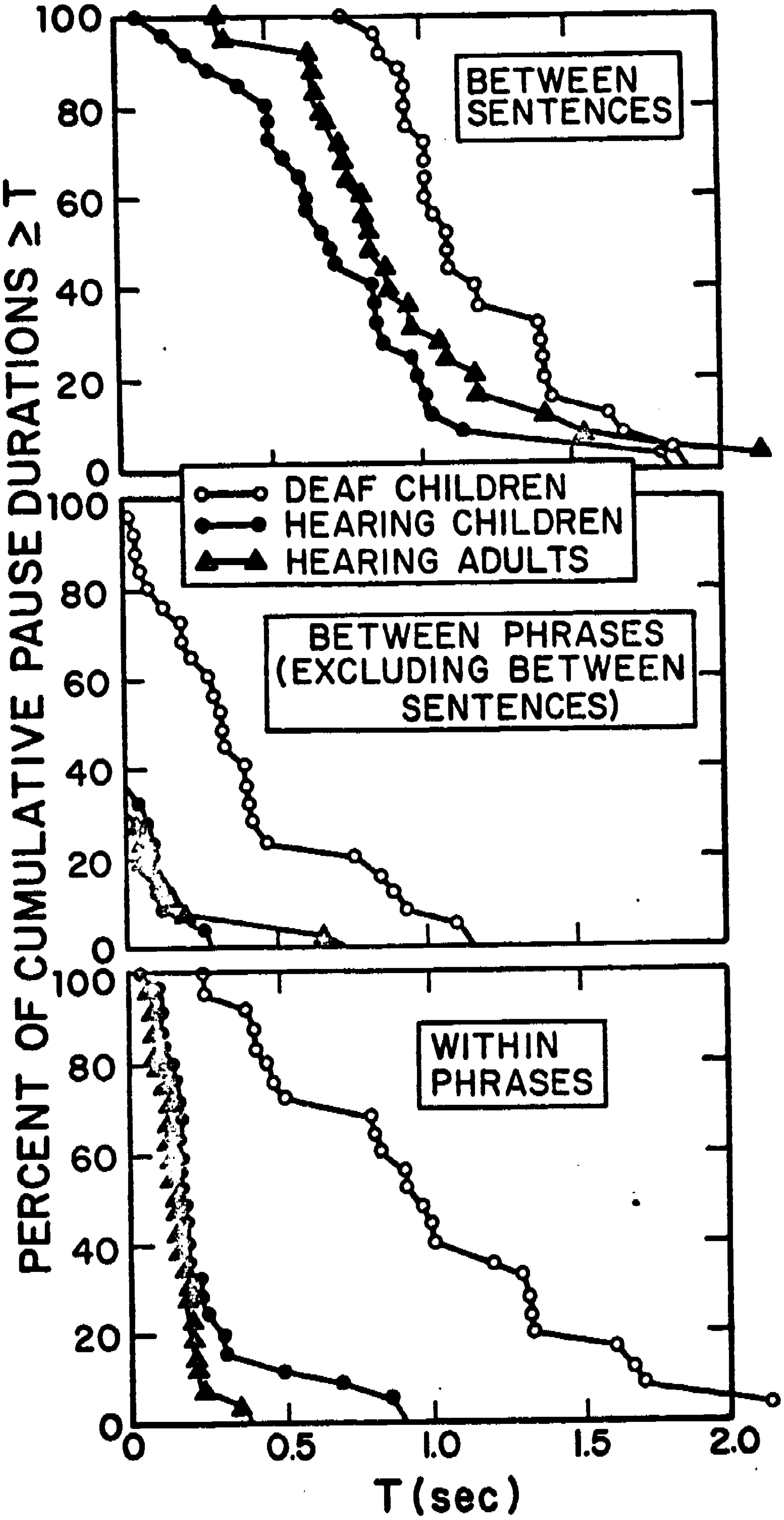

Fig. 7. Cumulative distributions of durations of pauses occurring in different contexts. (See "Methods of Analysis" section for procedure for identifying pauses.)

Speaking Time Versus Pause Durations

Perhaps as important as the total pause times of various types is the amount of pause time relative to the amount of time taken to produce a given utterance. Figure 8 shows data on this relationship for each of the speaker groups. On the average, the ratio of pause time to total time was greater for the deaf speakers than for either of the hearing groups. For hearing adults, pauses occupied about 31% of the total time required to produce the two sentences (including the pauses following both sentences). For hearing children the comparable figure was about 25%, and for the deaf children it was about 40%.

CONCLUSIONS

The results from the studies reviewed in this paper, and the data presented herein, permit the following conclusions concerning timing deficiencies in the speech of the deaf:

1. Deaf speakers tend to speak at a much slower rate than do hearing speakers (Boone, 1966; Colton & Cooker, 1968; Hood, 1966; John & Howarth, 1965; Martony, 1966; Voelker, 1938; present study).

2. When deaf speakers produce a phrase or sentence, they frequently fail to modify sufficiently the durations of

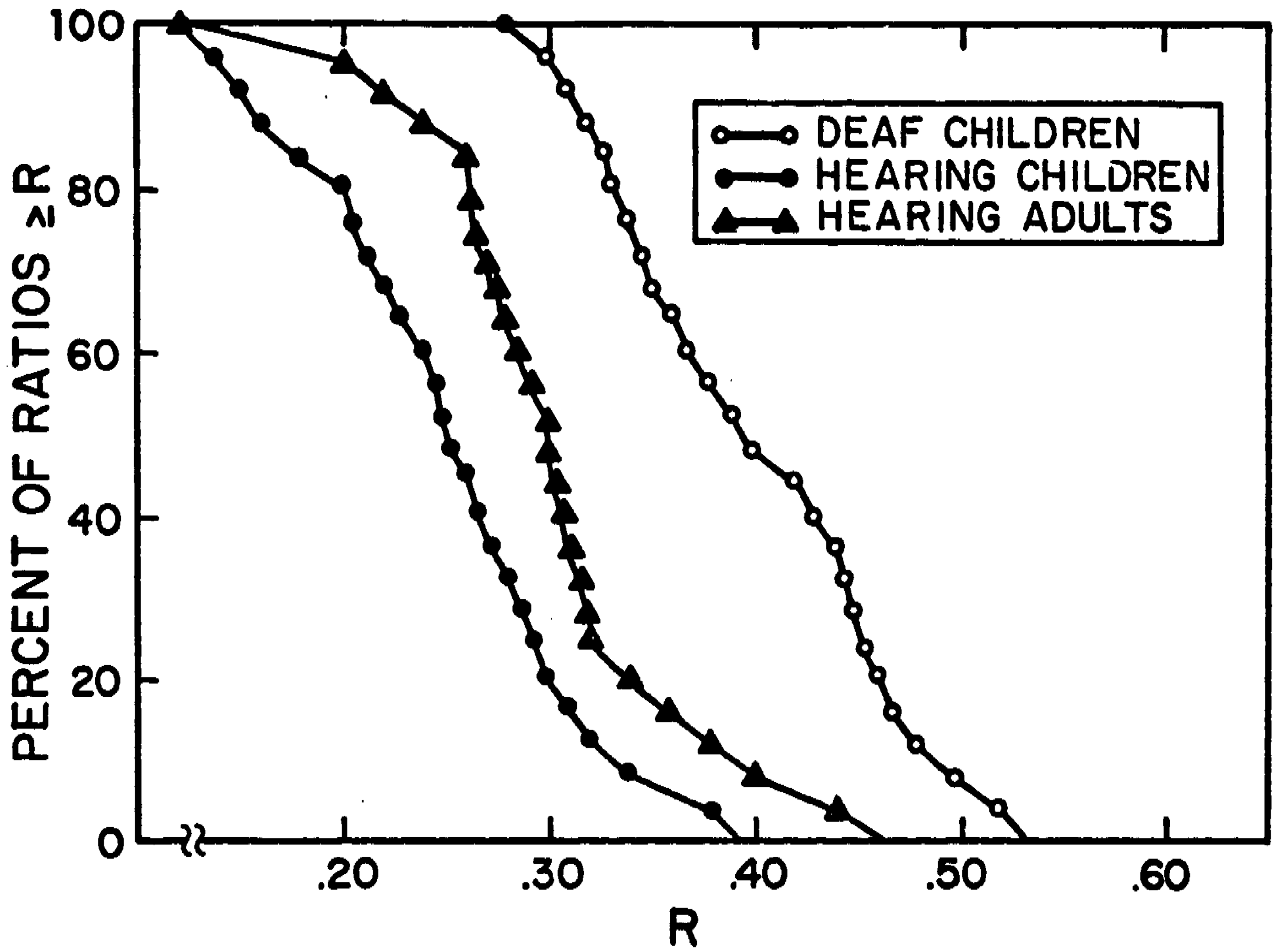

Fig. 8. Cumulative distributions of ratios of total pause time to total time taken to produce utterance. (See "Methods of Analysis" section for procedure for identifying pauses.)

the syllables relative to the durations of the same syllables produced in·isolation. In particular, deaf speakers often do not increase the duration of a phrase or sentence-final syllable relative to other syllables in the utterance, and do not sufficiently decrease the durations of unstressed syllables relative to those of stressed syllables (Angelocci, 1962; present study).

3. Deaf speakers tend to insert more pauses, and pauses of longer duration, in running speech--particularly within phrases--than do hearing speakers (Hood, 1966; Hudgins, 1946; John & Howarth, 1965; present study).

4. The durations of certain sounds appear not to show the same context dependencies when spoken by the deaf as when spoken by the hearing (Angelocci, 1962).

5. Individual speech sounds are often produced with inappropriate durations by deaf speakers, whether they occur in one-syllable utterances or in running speech. In particular, fricative consonants may have an inordinately long duration for deaf speakers (Angelocci, 1962; Calvert, 1961), as may the closure periods of plosive consonants (Angelocci, 1962).

6. The speech of the deaf tends to be judged inferior to that of the hearing, when compared by listeners with respect to rhythm or syllable grouping (Hood, 1966; Hudgins, 1946).

There appears to be general agreement among researchers that such timing deficiencies contribute significantly to the lack of intelligibility of the speech of the deaf. While the results of some studies (Houde, 1973; Stratton, 1973) cast doubt on the validity of the assumption that improvements in timing alone will invariably lead to an increase in intelligibility, the weight of evidence suggests that timing is at least as important a determinant of intelligibility as any other aspect of speech.

These observations indicate that speech training should direct considerable attention to timing at the level of the phrase and sentence. The deaf student should have a grasp of the properties of such units independent of the sequence of articulations that form the fine structure of the units. He must learn the proper way to initiate and terminate the unit, and must learn to produce within the unit syllables of greater or lesser prominence.[3]

As has been noted earlier, the importance of timing in speech training has been recognized by teachers of the deaf for many years, and attention is devoted to this aspect of speech in many speech-training programs. This training has, however, been hampered by the fact that the general principles governing the timing of running speech have not been adequately formulated, and by the difficulties in providing the deaf speaker with a means for perceiving the timing of his own articulations as well as those of others. This situation is now changing, however.

The rules that underlie the timing of sentences are being quantified (although much has yet to be learned), and procedures are becoming available for displaying to the deaf speaker an objective representation of the temporal pattern of his utterances or those of a teacher (Houde, 1973; Nickerson & Stevens, 1973; Stratton, 1973). These developments should help provide a more solid basis for the training of the temporal aspects of speech to deaf speakers.

NOTES

1. This work was supported by the U.S. Office of Education Media Services and Captioned Films Branch of the Bureau of Education for the Handicapped, under Contract No. OEC-0-71-4670(615). Several people contributed to the programming and data gathering aspects of this study. The assistance of the following individuals is gratefully acknowledged: Rob Adams, Patricia Archambault, Douglas Dodds, Barbara Freeman, Daniel Kalikow, and Robert Storm. The initial planning and performance of the work benefitted from the advice and guidance of Lois Elliott.

2. We recognize that a phoneme duration cannot be unambiguously defined, inasmuch as the phoneme or phonetic segment is an abstraction that does not have a representation in terms of a fixed length of the speech signal. Nevertheless, it may not be unreasonable to talk about number of phonemes per second, and hence average phoneme duration, or about the duration of a speech event such as a vowel or a consonant with well-defined acoustic boundaries.

3. While the focus of this paper is the temporal aspects of speech, these comments on the role of phrases and sentences in speech training apply to intonation as well as to timing.

REFERENCES

Abrams, M. H., Goffard, S. J., Kryter, K. D., Miller, G. A., Sanford, J., & Sanford, F. H. Speech in noise: A study of factors determining its intelligibility. OSRD No. 4023, Psychoacoustics Laboratory, Harvard University, 1944. PB 19805

Angelocci, A. A. Some observations on the speech of the deaf. Volta Review, 1962, 64, 403-405.

Bell, A. G. The mechanism of speech. 8th ed., New York: Funk and Wagnalls, 1916.

Boone, D. R. Modification of the voices of deaf children. Volta Review, 1966, 68, 686-694.

Borrild, K. Experience with the design and use of technical aids for the training of deaf and hard of hearing children. American Annals of the Deaf, 1968, 113, 168-177.

Brehm, F. E. Speech correction. American Annals of the Deaf, 1922, 67, 361-370.

Bzoch, K. R. Variations in velopharyngeal valving as a function of syllabic change in repeated CV syllables. Paper presented at the National Convention, American Speech and Hearing Association, Chicago, November, 1965.

Calvert, D. R. Some acoustic characteristics of the speech of profoundly deaf individuals. Ph.D. dissertation, Stanford University, Stanford, California, 1961.

Cohen, A., Schouten, J. F., & t'Hart, J. Contribution of the time parameter to the perception of speech. In Proceedings of the 4th International Congress of Phonetic Sciences (Mouton, The Hague), 1962, 555-560.

Colton, R. H. & Cooker, H. S. Perceived nasality in the speech of the deaf. Journal of Speech and Hearing Research, 1968, 11, 553-559.

Daniloff, R. & Moll, K. Coarticulation of lip rounding. Journal of Speech and Hearing Research, 1968, 11, 707-721.

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

DiCarlo, L. M. The deaf. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1964.

Fry, D. B. Experiments on the perception of stress. Language and Speech, 19[illegible] 1, 126-152.

Gay, T., Ush[illegible]ima, T., Hirose, H., & Cooper, F. S. Effect of speaking rate on labial consonant-vowel articulation. Haskins Laboratories Status Report on Speech Research, 1973, SR-35/36, 7-70.

Hood, R. B. Some physical concomitants of the perception of speech rhythm of the deaf. Ph.D. thesis, Stanford University 1966.

Hood, R. B. Some physical concomitants of speech rhythm of the deaf. Proceedings of the International Conference on Oral Education of the Deaf, 1967, 1, 921-925.

Houde, R. A. Instantaneous visual feedback in speech training for the deaf. Paper presented at the 1973 Annual Convention of the American Speech and Hearing Association, Detroit, Michigan, October, 1973.

House, A. S. On vowel durations in English. Journal of the Acoustical Society of America, 1961, 33, 1174-1178.

Hudgins, C. V. A comparative study of the speech coordination of deaf and normal subjects. Journal of Genetic Psychology, 1934, 54, 3.

Hudgins, C. V. A study of respiration and speech. The Volta Review, 1936, 38, 341-343, 347.

Hudgins, C. V. Speech breathing and speech intelligibility. Volta Review, 1946, 48, 642-644.

Hudgins, C. V. & Numbers, F. C. An investigation of the intelligibility of the speech of the deaf. Genetic Psychology Monographs, 1942, 25, 289-392.

Huggins, A. W. F. How accurately must a speaker time his articulations? IEEE Transactions on Audio and Electroacoustics. 1968, AU-16, 112-117.

John, J. E. J. & Howarth, J. N. The effect of time distortions on the intelligibility of deaf children's speech. Language and Speech, 1965, 8, 127-134.

Jones, C. "Deaf voice"--A description derived from a survey of the literature. The Volta Review, 1967, 69, 507, 508, 539, 540.

Klatt, D. H. Durational characteristics of prestressed word-initial consonant clusters in English. Quarterly Progress Report No. 108, Research Laboratory of Electronics, M.I.T., 1973, 253-260.

Klatt, D. H. Word verification in a speech understanding system. Invited paper, IEEE Symposium on Speech Recognition, Carnegie-Mellon University, April 15-19, 1974; to be published by Academic Press, 1974a.

Klatt, D. H. Vowel lengthening is syntactically determined in a connected discourse. To be submitted for publication, 1974b.

Kozhevnikov, V. A. & Chistovich, L. A. Speech: Articulation and perception (Moscow-Leningrad), 1965. English Translation: J.P.R.S., Washington, D.C., No. JPRS 30543.

Lehman, W. P. & Heffner, R. M. S. Re effect of context on vowel durations. American Speech, 1943, 18, 208.

Lehiste, I. Suprasegmentals. Cambridge; M.I.T. Press, 1970.

Lindblom, B. Spectrographic study of vowel reduction. Journal of the Acoustical Society of America, 1963, 34, 1773-1781.

Lindblom, B. & Rapp, K. Spectrographic study of vowel reduction. Journal of the Acoustical Society of America, 1967, 34, 1773-1781.

Lindblom, B. & Rapp, K. Some temporal regularities of spoken Swedish. Papers from the Institute of Linguistics, University of Stockholm, Publication 21, September 1973.

Martony, J. Studies on the speech of the deaf. II. Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, Sweden, Quarterly Progress and Status Report 31, Jan. 15, 1966.

Mason, M. K. & Bright, M. G. Tempo in rhythmic speech education. American Annals of the Deaf, 1937, 82, 385-401.

Miller, G. A. Decision units in the perception of speech. IRE Transactions on Information Theory, 1962, IT-8, 81-83.

Millin, J. P. Therapy for reduction of continuous phonation in the hard-of-hearing population. Journal of Speech and Hearing Disorders, 1971, 36, 496-498.

Nickerson, R. S. & Stevens, K. N. Teaching speech to the deaf: Can a computer help? IEEE Transactions on Audio and Electro-acoustics, 1973, AU-21, 445-455.

Parmenter, C. E. & Trevino, S. N. The length of the sounds of a Middle Westernern. American Speech, 1935, 10, 129-133.

Peterson, G. E. & Lehiste, I. Duration of syllabic nuclei in English. Journal of the Acoustical Society of America, 1960, 32, 693-703.

Pickett, J. M. Sound patterns of speech: An introductory sketch. American Annals of the Deaf, 1968, 113, 120-126.

Pierce, J. R. Symbols, signals and noise. Harper & Brothers, New York, 1961

Rawlings, C. G. A comparative study of the movements of the breathing muscles in speech and in quiet breathing of deaf and normal subjects. Part I. American Annals of the Deaf, 1935, 80, 147-156. Part II. American Annals of the Deaf, 1936, 81, 136-150.

Scuri, D. Restirazione e Fonazione nei Sordomuti. Rassegna di Educazione e Fonetica Biologica, 1935, 14 , 82-113.

Snidecor, J. C. The pitch and duration characteristics of superior female speakers during oral reading. *Journal of Speech and Hearing Disorders*, 1951, *16*, 44-52.

Stevens, K. N. & House, A. S. Perturbation of vowel articulations by consonantal context: An acoustical study. *Journal of Speech and Hearing Research*, 1963, *6*, 111-128.

Stewart, R. B. The speech of children with high-frequency losses of hearing--A small phonetic study. *Sound*, 1969, *3*, 40-43.

Stratton, W. D. Intonation feedback for the deaf through the tactile sense, Massachusetts Institute of Technology Thesis, 1973.

Voelkner, C. H. An objective study of the comparative number of speech sounds spoken per minute by the deaf and the normal. *Annals of Otology, Rhinology and Laryngology*, 1937, *46*, 471-476.

Voelkner, C. H. An experimental study of the comparative rate of utterance of deaf and normal hearing speakers. *American Annals of the Deaf*, 1938, *83*, 274-284.