Temporal organization of syllable production

Lindblom, B.

journal: STL-QPSR
volume: 9
number: 2-3
year: 1968
pages: 001-005

http://www.speech.kth.se/qpsr
I. SPEECH PRODUCTION

A. TEMPORAL ORGANIZATION OF SYLLABLE PRODUCTION *

B. E. F. Lindblom

The purpose of the present article is to demonstrate certain regularities that have been found to be characteristic of the temporal organization of syllable production in Swedish. Rules will be given that describe how the acoustic duration of vowel and consonant segments is perturbed as a function of variables such as inherent segment features, word length, position within the word and the phrase and the degree of stress. This project has been undertaken mainly to bring some new facts to the discussion of the control of timing in speech production which has been based on assumptions about mechanisms of energy expenditure (1), chain reflex (2), articulatory distance (3), equalization, stress versus syllable timing (4) and so on.

The results to be presented have been taken from a study of a nonsense vocabulary consisting of Swedish accent 1 words which were generated by systematic manipulation of the number of syllables per word and position of main stress within the word. Words were selected having 1 through 5 syllables. Each word had one main stress. For each word length this stress occurred in all possible positions. Various segmental shapes were assigned to these test words. Words such as [ˈbabb], [baˈbab], [ˈbabab] etc. exemplify the type of words examined. In the lists recorded these words occurred in a carrier phrase. Each phrase containing a test word was preceded by a reading of a normal meaningful word which was embedded in the same carrier phrase and had the stress pattern intended for the following test item. These materials were recorded by a group of Swedish subjects under high-quality recording conditions.

The Effect of Word Length

Fig. I-A-1 demonstrates a typical effect associated with word length or the number of syllables per word. The plots show how the segment durations of a syllable initial consonant and phonologically long and short

* somewhat modified version of paper presented at the 6th ICA in Tokyo, 1968.
EFFECT OF WORD LENGTH
ON SEGMENT DURATION IN
{ [CV:] SEQUENCES

![Graph showing segment duration as a function of the number of syllables per word. The measurements refer to word initial syllables with main stress.]}
Fig. I-A-2. Comparison between the data shown in Fig. I-A-1 on the duration of stressed consonant and vowel segments and measurements of the corresponding segment durations in unstressed position. The open circles pertain to word initial syllables with weak stress.
vowels vary as a function of word length. The data pertain to word initial syllables with main stress and have been taken from a single female talker. The effects shown are found for all other subjects as well. The graphs indicate that as the number of syllables per word increases consonant and vowel segments decrease in duration. The long vowel is indicated by solid circles and the short vowel by open circles. It is apparent that they behave similarly with respect to word length.

The Effect of Stress

In Fig. I-A-2 we present a comparison between the data just shown on the duration of stressed consonant and vowel segments and measurements of the corresponding segment durations in unstressed position. The open circles pertain to word initial syllables with weak stress. Only two degrees of stress will be distinguished: strong and weak stress. The data points indicating unstressed values refer to a pooled average of all the unstressed data in word initial position. From an examination of this figure and the corresponding ones for other talkers it is clear that the duration of unstressed segments is approximately constant and independent of word length whereas the stressed data exhibit longer values.

The Effect of Position

Fig. I-A-3 indicates the generally observed effect that position within the word has on segment duration. The preceding figures have presented data from word initial positions. The plots in this figure include in addition information on durations in medial and final positions. As before segment durations are plotted against the number of syllables per word. In this case the durations of unstressed [a] and stressed [aː] are shown to the left and right, respectively. For each given word length the data on initial, medial, and final positions is arranged sequentially along the x-axis. It should be pointed out that word final position is also phrase final for the data shown. Straight

* Meaningful Swedish words whose stress contours and accent patterns conform with those selected for the present nonsense vocabulary have been described (5) as having three degrees of stress, e.g. bökerna, muhammedän, egypologi etc. where 4 represents main stress and 1 and 0 two degrees of weak stress. The model words used in the present investigation seem auditorily to have been pronounced with no distinction between degree 1 and 0, and there is no clear durational basis for it in the data examined so far. In most cases, however, the perceptual reality of these levels is an unquestionable fact.
Fig. 1-A-B. The duration of unstressed [a] and stressed [a:] as a function of word length. For each given word length the data on initial, medial and final positions is arranged sequentially along the x-axis.
lines connect the points for a given word length. In the case of both the unstressed [a] and the stressed [a:] it is evident that the shortest durations are observed for the word initial positions. The medial data points tend to cluster closer to those indicating word initial durations. In no case is it possible to demonstrate for any subject that the final segment durations are shorter or equal to the initial segment durations for any given word length. Final lengthening is a feature of both unstressed and stressed syllable segments but is more marked for the vowels and syllable final consonants. As a rule, the unstressed segments exhibit little dependence on the number of syllables per word. Such a dependence, however, is clearly apparent in the right-hand plot where stressed vowel duration is seen to decay markedly as a function of word length. Final lengthening is not limited to phrase final position as additional data have shown. A tentative interpretation of this fact would be that this effect may be an attribute of the word.

Summary of Results

For **unstressed** syllables it was found that

(a) the acoustic duration of a segment is approximately constant;

(b) the acoustic duration of a segment is longer in final position than in initial and medial positions.

For **stressed** syllables it was found that

(a) the acoustic segment duration is approximately equal to its unstressed duration plus an increment due to the superimposed stress;

(b) in initial position the value of this stress-dependent increment is inversely proportional to the number of syllables of the word;

(c) in final position the duration of a stressed segment is likewise lengthened.

In Fig. I-A-4 an attempt is made to quantify the effects observed in terms of a simple mathematical model. As can be seen vowel segment duration \( V \) is given as

\[
V = (d + D/n)k \tag{1}
\]

where \( d \) represents a constant duration characteristic of unstressed vowels in initial and medial positions within the word and \( D \) represents...
The previously presented data are summarized by means of a simple mathematical description according to which vowel segment duration, $V$, is given by $V = (d + D/n)k$, where $d$ represents a constant duration characteristic of vowels in unstressed syllables and $D$ is an increment due to the superposition of main stress and $n$ the number of syllables per word. $k$ is the position-dependent factor which is $>1$ (final position) and $1$ (otherwise).
an increment due to superimposed stress and \( n \) the number of syllables per word. \( k \) is the position dependent factor which is > 1 (final position) or 1 (otherwise). It is possible to regard the term \( D/n \) as reflecting the well-known principle of isochrony or equalization (4), that is, the tendency for an utterance unit to assume constant total duration irrespective the number of segments that the unit is made up of.

**Discussion**

It was suggested above that the lengthening of segments in final syllables might be a property of the Swedish word. Other explanations can also be suggested to account for lengthening in phrase-final position. Investigators have noted for many languages that the subglottal pressure and the fundamental frequency of phonation and acoustic amplitude remain relatively stable in the beginning of an utterance but tend to fall at the end of a normal declarative sentence. The underlying phonatory and articulatory activity is assumed to be characterized by a general relaxation of speech gestures towards the end of an utterance. This pattern has been termed normal breath-group (6) or basic phrase contour (7). In his model of word and sentence intonation presented at the 6th ICA in Tokyo Öhman postulates the existence of an abstract "physiological intensity contour" that underlies the observed variations of subglottal pressure and fundamental frequency.

It is constant during the beginning of an utterance but then falls towards the end. It is further assumed that

(a) the physiological energy expended per syllable is constant (apart from variations with stress and, in stressed syllables, with word length), and

(b) that this energy is equal to the time interval of the so-called "physiological intensity" over the duration of the syllable.

Granted the assumption of the energy per syllable being constant final lengthening of segments becomes a consequence of the intensity being lower in the final part of the basic phrase contour.

**Effect of Inherent Features**

In Fig. 1-A-5 we note also that vowels like [a:] and [i:] although characterized by differences in so-called inherent duration behave similarly under these prosodic perturbations and fit the description just given. To account for the dependence of vowel duration on vowel identity some investigators have suggested principles of energy expenditure
EFFECT OF VOWEL IDENTITY

NUMBER OF SYLLABLES PER WORD

Fig. 1-A-5. The duration of [aɪ] and [iː] in initial stressed syllables as a function of word length. The curves fitted to the data have been derived from Eq. [1].
ARTICULATORY INTERPRETATION OF VOWEL DURATION

DRIVING FORCE    DISPLACEMENT    MIDSAGITTAL SEPARATION

OF LIPS

HALF-OPEN
CLOSE

Fig. I-A-6. Idealization of the dynamics of lip and jaw movement in a [bVb] type of syllable. The model predicts that the acoustic duration of a vowel is a function of its degree of mandibular opening.
similar to those given above (3). Our own approach to this problem has consisted in constructing a model of lip and mandible coordination that gives a mechanical interpretation of the durational variance of vowels (see Fig. I-A-6).

In this model the lips are represented by damped spring-mass systems. The rectangular input forces corresponding to a hypothetical [bVb] utterance are shown to the left of this figure and the responses in the middle. The displacement of the lips and jaw determine the course of the midsagittal separation of the lips. Vowel duration is equal to the interval during which this parameter is larger than zero. The right part of the figure shows the course of the separation as a function of jaw opening. It is seen that vowel duration increases with larger jaw opening. The point we want to make with this graph is that the variations of acoustic vowel duration may have a mechanical cause, namely the inertia of the jaw.

References


(3) Jespersen, O.: Lehrbuch der Phonetik (Leipzig/Berlin 1926).


