Durations of formant transitions

Öhman, S.

journal: STL-QPSR
volume: 6
number: 1
year: 1965
pages: 010-013

http://www.speech.kth.se/qpsr
II. SPEECH ANALYSIS

A. DURATIONS OF FORMANT TRANSITIONS

S. Ohman

The durations of formant transitions in vowel+stop consonant or stop consonant+vowel utterances are difficult to measure since the vowels are normally not absolutely stationary. The decision as to where the transition ends and the "stationary" part of the vowel starts can therefore not be made with perfect consistency. Moreover, the interpretation of this kind of data in articulatory terms is a rather risky business not only because of the relatively large uncertainty mentioned above but because the size of the articulatory transition is not immediately reflected by the size of the formant transition. A big articulatory motion may cause only a small formant transition and vice versa.

On the other hand it is of some interest to get an estimate of the order of magnitude of the formant transitions in at least a few utterances. These data are of importance for certain synthesis experiments and provide a gross picture of certain features of the articulatory events.

Speech material

The speech material studied consisted of the voiced stops /d/ and /g/ in a VCV context and was spoken by a male Swedish talker. There were a total of 50 different words obtained by forming all possible VCV combinations from the two consonants mentioned and the five vowels /y/, /ɛ/, /ɑ/, /o/, and /u/. The vowels were phonemically long and the consonants phonemically short. Each word was spoken five times. The recording was made in an anechoic chamber with an Ampex tape recorder and a Brüel & Kjær condenser microphone and microphone amplifier.

Sound spectrograms of two representative utterances (/ydɛ/ and /ɡɛɣ/) from this material are reproduced in Fig. II-A-1. The line drawings superposed on the spectrograms indicate how the measurements were made. Vertical lines were drawn at the beginning
Fig. II-A-1a

Fig. II-A-1b
and end of the stop gap. The centers of F1, F2, and F3 were then traced with a pencil. The points were marked where each formant started deviating (or finished deviating) from a horizontal line which was tangent to the center line of the formant in the most stationary part of the vowels. Finally, the point obtained in this way which was farthest away from the stop gap was taken as the end (or beginning) point of the transition. The durations of the VC and CV transitions and the stop gap were then measured and tabulated. The data are presented in Figs. II-A-2 through II-A-4.

Stop gap durations

The durations of the closed intervals of the stops are shown in Fig. II-A-2a (/d/) and Fig. II-A-2b (/g/). The data are broken up into five groups corresponding to the final vowels of the VCV words. Each of these groups is further divided into five subgroups corresponding to the initial vowels. The means of the five measurements corresponding to the individual measurements of each subgroup have been joined by straight line segments.

The durations of the d-closures are apparently quite independent of context. The grand average is 75 msec and 90% of the distribution is concentrated within ± 20 msec about the average. The mean duration of the g-closures is also 75 msec and the limits of variability are a few msec larger than those of d. The variation of the g-means along the successive positions of the horizontal axis of Fig. II-A-2b is larger than that of the d-means, however. This is because the closure boundaries of a /d/ are generally more sharply defined than those of a /g/.

Transition durations

Fig. II-A-3a shows the durations of the vowel+/d/ transitions and Fig. II-A-3b those of the /d/+vowel transitions. The variability of the data of Fig. II-A-3a is surprisingly small. The 90% limits are located at ± 25 msec about the grand mean which is 100 msec. The mean durations of the /yd/ and /ud/ transitions are slightly shorter (10 msec) than the /gd/ and /ud/ transitions. This difference is barely significant, however.
DURATIONS OF DENTAL STOP GAPS

Fig. II-A-2a

DURATIONS OF PALATAL STOP GAPS

Fig. II-A-2b
DURATIONS OF TOTAL VC TRANSITIONS

DURATION IN MSEC.
DURATIONS OF TOTAL CV TRANSITION

Fig. II-A-3b
A more striking context dependence is observable in Fig. II-A-3b. The mean /dy/, /d$, and /du/ transition durations are 115 msec, whereas the corresponding mean for the /do/ and /du/ transitions is 160 msec. If these figures reflect a corresponding difference in the durations of the articulatory transitions one must conclude that the time a transition takes is not merely a matter of the geometric distance between the two target configurations. For a normal male talker the tip of the tongue moves only a few millimeters in the transition from /y/ to /d/. In the transition from /a/ to /d/, however, it must pass a distance of several centimeters. The data of Fig. II-A-3b suggest that these two motions take the same time while the motion from /u/ to /d/ — which is not much shorter, geometrically, than that between /a/ and /d/ — is 50 msec longer in duration. These relations would not be surprising if one considered "distance" between configurations as being a matter of the difference in activity state of the articulatory muscles. The geometric effects of a certain amount of shortening activity in a given group of muscles would be expected to depend on the amount of slack which obtains in these muscles initially. The speed of shortening also depends, of course, on the strength of these muscles and on the degree of excitation.

Figs. II-A-4a and II-A-4b show the durations of the Vg and gV transitions, respectively. These data are more scattered than those of the d-transitions. The durations of the yg and gy transitions are particularly difficult to estimate consistently (cf. Fig. II-A-1b). The reason is that the Swedish /y/ has a somewhat diphthongal nature involving a progressively increasing degree of lip-rounding during the course of the vowel. The durations of the /yg/ transitions may therefore be overestimated. This error is probably not equally prominent in the case of the /yd/ and /dy/ transitions since the d-loci are much more different from the /y/ formant frequencies than the g-loci are. This fact facilitates the separation of the proper d-transition from the diphthongal component of the /y/.

With the exception of /yg/ the durations of Fig. II-A-4 seem to be monotonically related to the distance between the /g/
DURATIONS OF TOTAL VC TRANSITIONS

Fig. II-A-4a
Figure II-4b

Durations of Total CV Transition

DURATION IN MSEC.
and the adjacent vowel in the traditional vowel triangle sense. That is, the durations are shortest in /e/, /o/, /a/, /o/, and /u/ and longest in /e/, /o/, /a/, /o/, and /u/. The mean values are summarized below.

\[
\begin{array}{cccccc}
\text{ye} & \text{ge} & \text{ug} & \text{og} & \text{ug} \\
120 & 120 & 120 & 85 & 80 & \text{msec} \\
\text{ye} & \text{ge} & \text{ga} & \text{go} & \text{gu} \\
100 & 150 & 140 & 100 & 90 & \text{msec}
\end{array}
\]

In order to see if the duration of a vowel between consonants depends on the durations of the onglide and offglide transitions the following study was made. A list of symmetric CVC utterances involving the consonants /d/ and /s/ and the vowels /i:, ñ/, å:, a, o:, o, u:/ was read by the same talker as in the previous study. Each of the twenty syllables was read five times in the context "såga ... igen" (= "say ... again"). Sound spectrograms were made of all the utterances and the durations of the initial stop gap (=Cl 1), the final stop gap (=Cl 2), and the vowel (=Vow) were measured. The averages of these measurements are displayed in Fig. II-A-5.

A comparison of the durations of the vowels of Fig. II-A-5 with the sums of the corresponding onglide and offglide transition durations as estimated in the previous study did not reveal any simple relationship. Thus, the vowel of /d:rg/, for instance, is longer than that of /du:rg/ although the sum of the transition durations of /udu/ are longer than those of /u:dr/. The ratios of the durations of the CV transitions to the VC transitions of the CVC utterances were also compared with the corresponding ratios of the VCV utterances but again no simple relation was found.