ON TRANSITION IN THE LIGHT OF X-RAY FILMS

ANTTI SOVIJÄRVI

In my discourse at the IVth International Congress of Phonetic Sciences in Helsinki in 1961, I already touched to some extent the subject of transitions; namely transitions of the components of some Finnish diphthongs. For demonstrational purposes I used a special piece of equipment called ADAM (Apparatus for Demonstrating Articulatory Movements), which had just been developed at our Phonetic Institute. In the meantime some parts of this apparatus, especially the contourtape, plastic die plates, and the illumination showing through the pierced back plate have been improved to simplify the process of the diagramming of X-ray films and the photographing of the separate pictures thereof. Fig. 1.

Today I shall as examples handle some VC-, VV-, CC- and CV- transitions as they appear in some Finnish words and are shown in the successive single frames of X-ray films and in the corresponding segments of spectrograms.

When analysed transitions from that X-ray sound film for whose speed was chosen 48 frames/sec. I used the following method. Segments of speech corresponding to film frames were diagrammed with the help of the apparatus (ADAM) mentioned above. Various types of spectrograms were made of the series of words where the

Table 1. The transition of pä (in the word käytyppä)

<table>
<thead>
<tr>
<th>F</th>
<th>Print</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>c/s</td>
<td>db</td>
<td>c/s</td>
</tr>
<tr>
<td>F0</td>
<td>NB</td>
<td>321</td>
<td>30</td>
<td>318</td>
</tr>
<tr>
<td>F1</td>
<td>VP</td>
<td>675</td>
<td>30</td>
<td>859</td>
</tr>
<tr>
<td></td>
<td>WB</td>
<td>800</td>
<td>36</td>
<td>975</td>
</tr>
<tr>
<td></td>
<td>NB</td>
<td>642</td>
<td>36</td>
<td>954</td>
</tr>
<tr>
<td>F2</td>
<td>VP</td>
<td>1950</td>
<td>2925</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WB</td>
<td>1875</td>
<td>2000</td>
<td></td>
</tr>
<tr>
<td>F2</td>
<td>NB</td>
<td>1826</td>
<td>1908 + 2226</td>
<td></td>
</tr>
<tr>
<td>F3</td>
<td>VP</td>
<td>2950</td>
<td>3090 + 2226</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WB</td>
<td>2925</td>
<td>3200</td>
<td></td>
</tr>
<tr>
<td></td>
<td>NB</td>
<td>2889</td>
<td>3180</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: The transition of pä (in the word käytyppä)
transitions being examined occurred. For this purpose were used either the Kay Sonograph (Model 661-A) or the Voiceprint Laboratories Sound Spectrograph (Model 4691 A) recently obtained by our Institute in Helsinki. For analyses of the transitions the formant data were computed from the end section of the segments representing the frames. The actual speed of this film was 47.4 frames/sec.

In Fig. 2 is seen the normal spectrogram made with the Voiceprint Spectrograph using a wide band filter and flat shaping of amplitude. By vertical lines the divisions of separate segments corresponding to film frames have been made clearer. The test word is the Finnish word *tytujępa* (*tijyypa*) ‘one ought to.’

Tables 1—3 show also the results of the formant data obtained by using three spectrogram modes.

The explosions of plosives in the transitions [pee] and [tea] are included to the first segment of the vocoid [33]. The duration of each segment corresponding to the frame frequency was 21.1 msec.

In the transition [ped] formants 1, 2 and 3 rise due to the low labial loci in changing over from segment å1 to segment å2. Their frequencies are already invariable when reaching the next segment, å3, but the amplitude of F1 rises from 36 to 42 db. In Fig. 3—6 the smallest distance between the labial contours grows gradually frame by frame as indicated by the following measurements: å1 7 mm, å2 10 mm, å3 16 mm. Measuring is easy using the pointed pattern in the back plate of the apparatus; the distances of the points, which are equal in size, correspond to 4 mm of life-size of the subject’s articulatory organ. The scale of the apparatus is 5-fold. Measured from the contour spectrogram this labial transition took about 40 msec. (Fig. 7).

In the transition tä F2 is continuously falling during the course of four successive segments as shown by the frequencies 2650—2375—2250—2125 (WB). The locus of F2 in this transition [tà'] is rather high (about 2600) mostly due to the fact that the test person was very young (a 12-years old boy), but also depending on the post-dental-prealveolar position of the Finnish [t]. Examining the movements of the tip of the tongue in the corresponding frames, we find that the profile of the tongue is not completely convex and non-apical until the fifth segment, that is, around 85 msec. (4×21.1 msec. = 84.4 msec.). See Figs. 8—13. The apical distance (between the tip of the tongue and the alveolar arch) is increasing in millimeters, frame by frame, in the following way: å1 and å2 5 mm., å3 6 mm., å4 7 mm., and å5 14 mm.

| Table 2. The transition of tä (to the word *tijyypa*) |
|---|---|---|---|---|---|---|---|
| | å1 | å2 | å3 | å4 |
| | F | Print | c/s | d/b | c/s | d/b | c/s | d/b |
| F0 | NB | 230 | | 293 | | 304 | | 304 |
| F1 | VP | 285 | 12 | 900 | 30 | 950 | 36 | 925 |
| WB | 700 | | 900 | | 925 | | 925 |
| NB | 870 | | 876 | | 922 | | 912 |
| F2 | VP | 2600 | 6 | 2375 | 24 | 2175 | 30 | 2140 |
| WB | 2650 | | 2375 | | 2250 | | 2125 |
| NB | 2610 | | 2344 | | 2128 | | 2128 |
| F3 | VP | 3700 | 18 | 3700 | 18 | 3675 | 18 | 3675 |
| WB | 3675 | | 3025 | | 3050 | | 3075 |
| NB | 3480 | | 3014 | | 3049 | | 3048 |

NB = Narrow band, linear, high shaping
WB = Wide band, linear, high
VP = contour, high voiceprint spectrogram,
In the transition of the components of the diphthong [äy] three clear stages can be distinguished. 1) F1 noticeably starts to fall (1100—950 c/s) at phase ä8—ä9. 2) F2 rises to the level of the next component [y] (2500 c/s) at phase ä10—ä11. 3) F1 falls finally to the level demanded by [y] (400 c/s) at phase ä12—y1 (Fig. 14).

In the corresponding frames we can see that the contraction of the narrowest point of the tongue channel begins at phase 28 and ends only about phase y1 (from 21 to 13 mm). At the same time we can see the narrowest point of the pharyngeal channel extending from 11—18 mm. The length of the whole transition is thus about 5 segments, or 100 msec. The labial channel in the frames ä12, y1 and y2 has the following measurements: 24 mm in the last [ä] frame, 20 mm in the first and 8 mm in the second [y] frame (Figs. 15—21).

Without explaining the corresponding acoustic phenomena, I shall show some diagrams made from a film where the rate of 80 frames per sec. was used (Fig. 23). We shall compare the last frames of the preceding sounds with the first ones of the later. In Fig. 24 the tip of the tongue has not yet quite reached the alveolar arch as the vowel [e] is just changing into the Finnish medioalveolar [s] sound in the test word Eko (a Finnish male name). The narrow opening of the front teeth and the position of the mediiodorsum of the tongue are already the same as in the next frame 12.5 msec. later (Fig. 25). The transition of [es] during two successive frames, that is 25 msec., is demonstrated in Fig. 26 where the two phases of movement are photographed simultaneously on one and the same diagram. The CO-transition [sk] of the same test word is shown in Fig. 27 as above in two successive phases of transition. One can see the tongue preparing itself to the mediopalatal occlusion of [k] in the frame of the last phase of [s]; the slight upward movement (5 mm) of the back of the tongue and at the same time the quick release movement (4 mm) of the tip very skilfully produce the transition from [s] to [k]. The jaw has hardly had time to change its position at all during this dual picture.

The following diagram (u)r + (u)r2 (Fig. 28) concerns this same [r] contoid, in which the position (u)r2 corresponds to the wider phase in the tremulation of the front part of the tongue. The narrower phase of the tremulation has been diagrammed from the preceding frame. The duration of one vibration of the tongue is about 25 msec. and the tremulation has a frequency of about 40/sec. in this contoid. The narrower phase resembles an apical medioalveolar occlusion as in the Finnish plosive [d], but in this phase the air flow is going through a very narrow groove.

In the near future we will investigate the perceptual level of these transitions of frame diagrams and spectrograms. For this purpose we will use our transistorized photocell segmentator listening to the desired cut sequences of the optical sound of the X-ray motion pictures.

LITERATURE


Fant, Gunnar (1965). Formants and Cavities. Proceed. of the Fifth International Congress of Phonetic Sciences (Münster 1964), p. 120—141.


Sovijärvi: On transition in the light of X-ray films

Fig. 1.
Šovijärvi: On transition in the light of X-ray films

Fig. 2.
Sovijärvi: On transition in the light of X-ray films

Fig. 3.

Fig. 4.

Fig. 5.

Fig. 6.
Sovijärvi: On transition in the light of X-ray films

Fig. 7.
Sovijärvi: On transition in the light of X-ray films

Fig. 8

Fig. 9
Soviär: On transition in the light of X-ray films

Fig. 10

Fig. 11
Sovijärvi: On transition in the light of X-ray films

Fig. 12

Fig. 13
Sovijärvi: On transition in the light of X-ray films

Fig. 14.

Fig. 15.

Fig. 16.

Fig. 17.

Fig. 18.

Fig. 19.
Sovijärvi: On transition in the light of X-ray films

Fig. 20.

Fig. 21.

Fig. 28.
Sovijärvi: On transition in the light of X-ray films
Sovijärvi: On transition in the light of X-ray films

Fig. 24

Fig. 25
Sovijärvi: On transition in the light of X-ray films

Fig. 26

Fig. 27