ABSTRACT

A set of fundamental frequency (F0) contours produced by one speaker can be superimposed on each other on the same time/frequency (log.) plot. If the number of contours is large enough, they together form a pragmatic approximation of the speaker's individual temporal voice range profile (TVRP). Single F0 contours or contour groups can be marked within the TVRP and indicate their acoustic characteristics in this way. Different utterance types and speech acts are illustrated with this method using data from two Finnish speakers.

1. TVRP AS A BACKGROUND OF SINGLE UTTERANCE OR UTTERANCE GROUP

Superimposition on each other of a set of F0 contours of the same time/frequency (log.) plot (Fig.1) makes possible the display of a pragmatic approximation of the speaker's temporal voice range profile (TVRP). The description is better, if the database includes variable utterances representing different speech acts and emotional states. The more the single contours included in the profile, the more representative it is of the speaker’s TVRP.

The older notion of ‘voice range profile’ with a different meaning is described in [1]. Isolated prolonged vowels (often [a]) have been used for the display of the speaker’s voice range profile. The purpose is more to promote voice analysis and to reveal the speaker’s voice resources than to study his speech behavior.

The F0 contours have been used in the prosodic analysis of languages (e.g. [2]). Empirical work with F0 contours shows that individual speakers use different voice ranges when producing linguistically identical utterances. Therefore, it seems appropriate to combine some elements from voice research with those of prosodic research. A larger F0 range may be indicative of the speaker's more variable and vivacious voice or stronger marking of accents or both. It can be seen (Fig.1) that the accumulation of data points concentrates on the lower part of the temporal distribution area. This concentration area can be understood as the default, i.e. as the indifferent level of the speaker’s intonation. It takes the form of a black declining bar. It can be assumed that those “tops” and “valleys” which indicate deviations from the default area, signal the most important information. It can be assumed that the more deviant the contours are compared to the default area, the more specific their communicative function.

In superimposing the utterances on the same plot, it is appropriate to apply the notion of intonation or tone unit, breath group or simply utterance. The length of the utterances and the speech rate can vary, but it is interesting to observe how they affect the form of the F0 contour.

Figure 1. Background (green) consisting of 20 sentences produced by a female speaker NA. Sentence Nr. (11) marked more clearly: Minne Maija vei omenan? [Where did Maija bring the apple?] High beginning and following down step effect observable.

2. PROGRAMS UTILIZED

The utterances were recorded by means of a AKG C 451E microphone and a TEAC W-850R cassette recorder. They were then digitalized by the SoundEdit 16 Version 2 (Macromedia) on the hard disk of a PowerMac 7100/66AV computer in IFF format.

Single F0 contours can be obtained by the automatic F0 analysis option of commercial programs. I have applied the F0 plot option of the SoundScope program. The measurement values with 10 ms intervals are exported (copy wave to text) to another program called PsychoF0. It is also possible to add the time points indicating the segment, syllable or word boundaries to the analysis data manually and use them in the later phases of the procedure. The analysis range has been adjusted to the speaker. The autocorrelation option was applied. Rejection of all peaks with a 25%
variation in adjacent pitch periods was applied. The reliability of the measurements was controlled by means of a zoomed (0-600 Hz) narrow band spectrogram and a FFT short time spectrum. When mistakes were observed on the computer screen, after-corrections were made.

F0 contours were processed further by a special program called PsychoF0 (created in the FutureBasic II language). The measurement values are indicated as dots on the computer plot. The silent measurement points (occurring in unvoiced segments or during pauses) indicated as "NAN15" by the SoundScope program were changed to the numerical form "0.1", which actually means their invisibility on the computer screen. The logarithmic scale according to the algorithm in [3] has been applied in order to simulate the human psychoacoustic response.

As well as showing the total variation as a background, single F0 contours, comparison of two contours or comparison of contour groups can also be presented. A contour group can be shown on a background of total variation (Fig. 2-3).

3. FUNCTIONAL USE OF F0 CONTOURS

Fundamental frequency (F0) is functionally linked to many linguistic factors [3]. The time window controlling the functional use of F0 varies from the segment to global linguistic factors [3]. The time window controlling the functional use of F0 varies from the segment to global linguistic factors [3]. The time window controlling the functional use of F0 varies from the segment to global linguistic factors [3]. The time window controlling the functional use of F0 varies from the segment to global linguistic factors [3]. The time window controlling the functional use of F0 varies from the segment to global linguistic factors [3]. The time window controlling the functional use of F0 varies from the segment to global linguistic factors [3]. The time window controlling the functional use of F0 varies from the segment to global linguistic factors [3]. The time window controlling the functional use of F0 varies from the segment to global linguistic factors [3]. The time window controlling the functional use of F0 varies from the segment to global linguistic factors [3]. The time window controlling the functional use of F0 varies from the segment to global linguistic factors [3]. The time window controlling the functional use of F0 varies from the segment to global linguistic factors [3]. The time window controlling the functional use of F0 varies from the segment to global linguistic factors [3]. The time window controlling the functional use of F0 varies from the segment to global linguistic factors [3]. The time window controlling the functional use of F0 varies from the segment to global linguistic factors [3]. The time window controlling the functional use of F0 varies from the segment to global linguistic factors [3]. The time window controlling the functional use of F0 varies from the segment to global linguistic factors [3]. The time window controlling the functional use of F0 varies from the segment to global linguistic factors [3]. The time window controlling the functional use of F0 varies from the segment to global linguistic factors [3]. The time window controlling the functional use of F0 varies from the segment to global linguistic factors [3].

4. EXAMPLES OF FINNISH

4.1. Speech material

Some examples of Finnish are illustrated in Figures 1–3. The 25 utterances were produced by a female speaker NA and a male speaker SL. The length of the utterances was 3 or 4 words. The utterances represent statements, wh-questions, answers, topicalizations, and commands. In the topicalizations, the syntactical form of the utterance implies that the first word must get a focused accent. The answers occurred in mini-dialogs. The following simple examples illustrate these classes:

(1) Statements and answers: Maija vei omenan [Maija brought the apple;] Kenelle Maija antoi omenan? Hän antoi sen Ollille. [To whom did Maija give the apple? She gave it to Olli.]

(2) Wh-questions: Minne Maija vei omenan? [Where did Maija bring the apple?]

(3) Commands: Anna Maijalle omena! [Give Maija an apple!]

(4) Topicalizations: Juoman Olli joi. [It was a drink that Olli drank.]
Figure 2. The statements (including answers) (1), wh-questions (2), commands (3), and utterances beginning with topicalization (4) illustrated on a background including all data. Female speaker NA.
Figure 3. The statements (including answers) (1), wh-questions (2), commands (3), and utterances beginning with topicalization (4) illustrated on a background including all data. Male speaker SL.