Proc. 5th int. Congr. phon. Sci., Münster 1964, pp. 392–396
(S. Karger, Basel/New York 1965).

Acoustic Correlates of Some Hungarian Emotive Intonation Patterns

By K. Magdics, Budapest

1. Acoustic analyses and various experiments were carried out on the synthesized one-word Hungarian utterance /e:p/ "It is undamaged" pronounced with 20 different emotions or attitudes. 10 of the synthesized sentences were questions, 8 were statements and 2 exclamations. One of the questions was a neutral yes-or-no question, the others were coloured by astonishment, two different degrees of curiosity, impatience, suspicion, consternation, sarcasm, indignation and threat. One of the statements was neutral, the others implied consolation, astonishment, incomprehension, warning, sarcasm, argumentation and indignation. The exclamatory sentences expressed joy and surprise.

2. The changes of fundamental frequency, duration and formant amplitudes were measured on spectograms and sections. Fundamental frequencies were plotted against time, and these graphs automatically provided measures of the durations. Formant amplitudes in dB were plotted against frequency.

a) *Fundamental frequency* of both questions and statements ranged from 80–180 cps. (this was the full fundamental frequency range of the synthesizer), that of the exclamations from 140–180 cps.

The neutral question had a medium rising pitch pattern starting at 100 cps. and rising to 140 cps.; the slightly and strongly inquisitive, impatient, indignant and threatening questions had high rising pitch patterns ranging from 100–180 cps.; the questions coloured by surprise and consternation showed low rising patterns moving between 80–120 cps.; the suspicious and sarcastic questions were characterized by rising-falling patterns, the frequency changing between 80–150 cps.

The neutral statement had a medium falling pitch pattern

moving from 110–80 cps.; the sarcastic statement showed a low falling pattern changing from 90–80 cps.; consolation, incomprehension, astonishment and joyful exclamation had high falling patterns with frequency values between 180–80 cps.; warning, argumentation, indignation and exclamation of amazement had falling-rising patterns moving from 180–120 cps. with a frequency value of 120–140 cps. between the two.

b) There was a certain relationship to be found between pitch patterns and *duration*. The sentence with a medium rising pattern was medium long (0.42 sec). The sentences with high rising and medium falling patterns were rather short (0.32–0.36 sec). Both the low rising and the low falling patterns were characterized by a fairly long duration (0.44–0.46 sec). The rising-falling patterns were extremely slow in speed (0.55–0.65 sec), whereas the fall-rise occurred with both an increased (warning = 0.36 sec, indignant statement = 0.40 sec) and a slower speed (exclamation of amazement = 0.55 sec). High falling patterns had either shorter (joyful exclamation = 0.35 sec) or longer duration (consolation = 0.49, incomprehension = 0.46, astonished statement = 0.44 sec).

In the rising pitch patterns the first, lower line was longer (0.10–0.25 sec) than the following higher one (0.07–0.15 sec), whereas in the falling patterns the first, higher line was remarkably shorter (0.08–0.15 sec) than the second, lower one (0.10–0.25 sec). In the rising-falling and falling-rising patterns the first and the last melody lines were more or less the same in length and the middle one was rather short.

c) The *amplitude* generally decreased with a medium and a high rise[1] (A_1: 0 – —2.5 dB, A_2: —1 – —4 dB), but increased with a low rise (A_1: 0 – +2 dB, A_2: —2 – 0 dB). Medium falling and high falling patterns were generally characterized by a falling amplitude (A_1: +2.5 – —1 dB, A_2: 0 – —3 dB), whereas low fall showed a rising amplitude (A_1: 0 – +2 dB, A_2: —3 – 0 dB). Both rise-fall and fall-rise had a falling-rising amplitude (A_1: +1 – —2 – +1 dB, A_2: 0 – —4 – 0 dB).

3. 10 English and 10 Hungarian subjects were asked to define the emotions or attitudes the synthesized sentences implied. Out of the 200 English answers 78 (39%), of the 200 Hungarian answers 140 (70%) turned out to be correct.

[1] The dB amounts are related to the amplitude of the A_1 of the neutral yes-or-no question.

4. Modifications of the position, the pitch intervals, the duration and the amplitude were carried out one after the other in each of the sentences. Spectograms and tape recordings were made of each modification. Then 10 English and 10 Hungarian subjects were asked to define the emotions or attitudes the modified utterances implied.

a) *Position was raised or lowered* with 40 cps. This time the English subjects produced 50 (25%) and the Hungarian 104 (52%) correct answers. The following sentences were generally misjudged: neutral question and statement, surprised and dismayed question, statements coloured by sarcasm, incomprehension and surprise, joyful exclamation.

b) *Pitch intervals were increased or decreased* with 20 cps. In this case the percentage of the correct answers decreased to 10 (20 in number) with English subjects and to 32 (64 in number) with Hungarians. Correct judgements were given only in the case of consolation, sarcastic statement and questions coloured by astonishment, impatience, sarcasm and threat.

c) The *duration of the whole utterances were increased or decreased* by 0.1 sec. The number of correct answers was 30 (15%) with English subjects and 82 (41%) with Hungarians. Neutral question and statement, impatient and surprised questions as well as exclamation of surprise were generally not recognized.

The *durations of the first and the last pitch level* in the rising and the falling patterns *were inverted:* only 50 (38%) out of 130 answers given by Hungarian subjects proved to be correct in comparison with the 82 (63%) given by the same subjects listening to the original forms of the sentences.

d) The *amplitude* of F_1 or F_2 *was increased or reduced* with 4 dB in the beginning or at the and of the utterance. The percentage of the correct answers given by English subjects was 21 (42 in number), those given by Hungarians was 35 (70 in number). Nobody recognized the neutral question, the slightly and strongly inquisitive, surprised, dismayed, sarcastic and threatening questions nor the statements coloured by consolation, incomprehension astonishment and sarcasm.

5. *Conclusions:* Four acoustic cues: position, pitch intervals, duration and amplitude were determined in a synthesized one-word Hungarian utterance pronounced with 20 different emotions or attitudes. The synthesized sentences were presented as tests material to 10 English and 10 Hungarian subjects. The experiment resulted in 39% correct answers by the English and 70% by the Hungarian subjects. The successive modifications of the acoustic components reduced the number of the correct answers and revealed the acoustic cues relevant in the recognition of the emotive patterns in question. It turned out that *each of the four cues* were *equally relevant* in the recognition of the neutral question. *Three of the cues:* position, duration and amplitude determined the astonished question; position, pitch intervals and amplitude made the dismayed question, the astonished and uncomprehended statement recognizable; position, pitch intervals and duration were important in the neutral statement. *Two of the cues*: pitch intervals and amplitude were necessary for the recognition of both kinds of inquisitive question; position and amplitude were relevant for the sarcastic statement, position and pitch intervals for the joyful exclamation, pitch intervals and duration for the exclamation of surprise. *One of the cues:* duration was enough for the recognition of the impatient question, the pitch intervals for the suspicious and indignant question, as well as the warning, the arguing and the indignant statement, the amplitude for the consolation, the sarcastic and threatening question.

The synthesis was produced by using the speech synthesizer constructed by *P. Denes* and *J. E. West*, and a Kay Sona-Graph was used for the analyses at the Phonetics Department of University College, London. The synthesis followed the author's pronunciation, who had previously examined Hungarian intonation including emotive patterns; see *I. Fónagy* and *K. Magdics:* Das Paradoxon der Sprechmelodie. Ural-Altaische Jb. *35:* 1–55 (1963). – Emotive Patterns in Intonation and Music. Z. Phon. *16:* 293–326 (1963).

Author's address: Klara Magdics, MTA Nyelvtudományi Intézete, Szalay u. 10–14, *Budapest V* (Hungary).

Discussion

Von Kürthy (Aachen): You can distinguish Hungarian from other languages without understanding because

a) the syllables are very short
b) the words are either intonated in a high or in a low cue

For example:
High intonation: Egynegyed egy
Low intonation: Bútorraktár

Gårding (Lund): I am surprised to hear that it is possible to make listeners identify as much as twenty different intonation contours differing only in the fundamental frequency of the carrier phrase. Together with *Arthur Abramson* at Haskins Laboratories I worked on a project similar to the one reported on by Miss *Magdics*. We then started with a great number of intonation contours that we had elicited from American English informants. These were put in a listening test. I cannot describe the test procedure here

but it may be of interest to mention that the recorded sentences chosen for the test were put on cards and the card reader that Mr. *Cooper* described in his talk was used. Out of the original contours, twelve could be identified by listeners. Later, when we applied the pitch curves from these twelve contours to a common carrierphrase only five were identified by listeners.