A STATISTICAL ANALYSIS OF MELODY CURVES IN THE INTONATION OF AMERICAN ENGLISH

YUKIO TAKEFUTA, ELIZABETH G. JANCOSEK AND MICHAEL BRUNT*

Pitch and duration characteristics of the melodies of American English (intonation) have been investigated by many researchers and several factors were found to influence them. In this study, in addition to pitch and duration characteristics, pitch patterns of American English were investigated statistically by introducing a method of evaluating the F_0 in the pitch shifts and that of normalizing the duration of melody curves. By the introduction of such methods, it became possible to use a simple statistical index (correlation coefficient) to compare the similarity or dissimilarity of any pair of melody curves. The authors believe that the comparative study of the variability of melody curves is important for the understanding of the functions and the structures of intonation in English.

Eight sets of ten sentences were recorded and analyzed. These were conversational sentences in American English. All were simple sentences including statements, *wh* questions and *yes-no* questions and varied from one to ten syllables in length. A sample of a sentence set is shown in Table 1. Forty-eight students of the Ohio State University served as speakers. They were assigned to eight sentence sets: six speakers in each set. All groups included the same number of male and female speakers. Speakers were requested to produce as many different intonations as they could think of in recording each of the ten sentences in the set assigned to them. They were also asked to provide a written description of the intended meaning.

In recording the sentences a combination of a dynamic microphone (Shure Bros., Model 545) and a Magnecord 1022 was used. For reproduction and extraction of the F_0, a combination of a Magnecord, Transpitchmeter and a Mingograf was used. Fundamental frequencies were measured at every 40 msec. Statistical analyses of the data were made by using the computing facilities of the Instruction and Research Computer Center of the Ohio State University.

Statistics such as the mean, standard deviation, duration, and rate of frequency change were computed from the F_0 of each utterance. They were considered to correspond with listeners' perception of pitch level, pitch range, length, and rate of

* This study was conducted under a contract between the Ohio State University and the Office of Naval Research.

TABLE 1

An example of sentence sets.

Sample Sentence Set
There
So what
Where are you
Time to get up
Aren't you ready yet
You've got a lot to learn
Where is a good restaurant
You must be a little upset
He didn't get the letter in time
How am I going to finish in time

TABLE 2

Mean of simple statistics computed for the melody curve of each utterance. The values are the average of the utterances of 48 speakers.

Number of Syllables	1	2	3	4	5	6	7	8	9	10	Units
Mean	157	162	165	162	162	159	159	156	162	152	Hz
S.D.	41	45	42	44	39	42	40	41	40	35	Hz
Duration	517	739	964	1010	1210	1512	1566	1855	1782	2219	msec
Rate up	10	9	8	9	9	9	8	9	9	8	x25 Hz/Sec
Rate down	10	12	10	10	9	9	9	9	9	9	x25 Hz/Sec

pitch change, respectively. The mean of these measures were obtained for the sentences of different lengths. Results are summarized in Table 2. The same data were then reorganized to investigate the characteristic use of pitch and duration by male and female speakers. Table 3 presents the mean values of the statistics for both speaker groups.

For Part 2 of the study, pitch patterns were compared by computing coefficients of correlation. In order to fill in the missing data in the melody curves as a result of

TABLE 3

Mean of simple statistics computed for the melody curve of each utterance. The values are the average of the utterances of 24 speakers.

Speakers	Male	Female	Units
Mean	127	186	Hz
S.D.	28	58	Hz
Duration	1192	1402	msec
Rate up	6	11	x25Hz/Sec
Rate down	6	12	x25Hz/Sec

non-vocalized portions of an utterance, the fundamental frequencies were considered to be on a straight line which connects the last value of the preceding phonation and the first value of the following phonation. With this assumption, it was possible to express a melody curve with a set of ordered linear functions. Then, by using this set of functions, F_0 values at one hundred equally distributed points were evaluated as the values of a normalized melody curve. Using these data, coefficients of correlation were calculated to compare pitch patterns of the utterances. First comparisons were between paired utterances in which the speakers tried to express 'different messages'. The median coefficients for such comparisons are summarized for each sentence length in Table 4. In this comparison, speakers and sentences of the paired utterances were the same. Only the intended messages were different.

TABLE 4

Median coefficients of correlation computed for all possible pairs of melody curves in which different messages were intended. Paired sentences were always the same.

Number of Syllables	1	2	3	4	5	6	7	8	9	10
Coefficients of Correlation (Median)	.40	.28	.12	.42	.17	.17	.27	.20	.44	.33

Due to the design of the study, the correlation coefficients of paired utterances of the same speakers, the sentences and the messages could not be compared. However, the median coefficient of correlation for the same messages intended in different sentences by the same speaker were computed. They are summarized in Table 5.

The effect of sentence structure on pitch patterns was also investigated by classifying all utterances into three groups: statements, *wh* questions, *yes-no* questions. The median pattern for all the utterances for each sentence structure was obtained before computing coefficients of correlation among different structures of the sentence. This procedure was used to cancel the characteristic pattern of the individual messages and to extract the characteristic pattern due to the structure of the sentence.

TABLE 5

Median coefficients of correlation computed for pairs of melody curves in which the same message was intended. Paired sentences were different.

Messages	Anger	Disgust	Sarcasm	Surprise	Statement	Question
Coefficients of Correlation (Median)	.32	.36	.24	.30	.40	.27

TABLE 6

Median coefficients of correlation computed for pairs of same and different syntactic structures. Data for computing the correlation were the average of all utterances produced for each sentence.

Sentence Structure	Statement vs Statement	Wh Question vs Wh Question	Yes-No Question vs Yes-No Question	Statement vs Wh	Statement vs Yes-No	Wh vs Yes-No
Coefficient of Correlation (Median)	.28	.40	.25	.28	-.20	-.16

All possible comparisons of pitch patterns for these three selected syntactic structures are shown in Table 6.

The mean, standard deviation and rate of pitch change did not differ significantly among the utterances of different lengths. The number of syllables which constitute the sentences does not seem to affect the pitch level, pitch range nor the rate of pitch change. Sex of the speaker, however, significantly affected these statistics as they are seen in Table 3. Female speakers showed the pitch characteristics of almost twice the measures of male speakers. However, only a slight difference was found in the duration of the utterances between the male and the female speaker groups. Female speakers spent slightly more time in recording the sentences of the same length. Naturally, this measure significantly differed for the sentences of different lengths.

The coefficients of Table 5 are lower than expected. One assumption is that the speakers often failed to produce the intended messages. Another assumption is that not only the intended message but also the sentence structure (syntax as well as constituent words) affect the melody curves. The first assumption was supported in the comparison of the intended messages and the received messages. The agreement was not high. The second assumption was supported by computation of more correlation coefficients. First, only the pairs of sentences which were different both in constituent words and the intended messages were compared. The median coefficient was still lower than those in Table 5 —.11. The difference between this value and the value in Table 5 indicates that there are certain characteristic patterns used for certain messages in intonation. Further, the median of the coefficients for the pairs of melody curves obtained from the same sentence conveying the same message (heard by listeners) was found to be .79. In this case the speakers of the paired utterances were different. These two values in contrast to those in Table 5 seem to be sufficient to accept the two assumptions above. The data of Table 6 also supports the second assumption. The sentences of the same syntactic structures are pronounced by more or less similar pitch patterns. Further, it is apparent from the values of the coefficients that the statements and *wh* questions had similar patterns but *yes-no* questions had different melody patterns. *Wh* questions were found to have the least latitude in varying pitch patterns. Further, although it could not be shown by the computation of correlation coefficients, there was a definite tendency toward smaller but more peaks and valleys in the melody curves of longer sentences.

A final but significant finding in this study was the range of variability in pitch patterns. Individual variations of pitch patterns due to possible errors and the speakers' stereotype were so large that the coefficients were scattered almost from one end of the scale to the other in any category of comparison. Physical cues of intonation are very difficult to determine. They may be as complex as those of segmental phonemes.

Ohio State University

DISCUSSION

BLACK (Columbus, Ohio)
It is quite clear that there were 100 pairs of measures in each of many correlations. It is less clear exactly what each measure was. Would you clarify this for me?

JANCOSEK
These were measures of fundamental frequency at each 40 msec which were subsequently input to a computer which had been programmed to 'normalize' the data to provide 100 values for each melody curve.

ARKWRIGHT (Montréal)
Please characterize more fully your term 'normalization'.

JANCOSEK
Approximately 2400 melody curves were measured directly from enlargements of the plot provided by the Mingograf. These were rather laboriously measured 'by hand' and then those values became the input for the computer.