

Experiments on the Stylization of British English Intonation

J. 't Hart and J.R. de Pijper
Eindhoven, the Netherlands

1. Introduction

On earlier occasions we have reported on a method applied in intonation analysis which we call the stylization method. Straight line segments are used to construct approximations to F_0 curves in such a way that these are simplified as much as possible, but nevertheless give a melodic impression similar to that of the originals. After stylization it is far less difficult to make an inventory of the melodic possibilities of the intonation system of a language, as was demonstrated for Dutch ('t Hart and Cohen, 1973; 't Hart and Collier, 1975). Since this approach boils down to analysis by synthesis, an additional advantage is that it yields reliable recipes for the control of F_0 in speech synthesis, by virtue of its continuous testing of the perceptual adequacy in the process of the analysis.

This study examines whether the stylization method is also applicable to the intonation of British English (henceforth BE).

2. Close-copy stylizations

The first question to be answered for BE intonation is whether such a straight line approach is feasible at all. Instead of aiming at fully standardized contours in one step, we introduced an intermediate step of constructing close-copy stylizations. These are defined as composed of the smallest number of straight line segments with which perceptual equality can be obtained.

Close-copy stylizations are made following a trial and error method with the experimenter as sole judge. The alleged perceptual equality is, of course, tested afterwards in an appropriate listening experiment.

In such an experiment, 64 native English subjects were presented with 60 pairs of sentences, varying in duration between one and three seconds. Each sentence was processed by means of an LPC analysis-resynthesis system in order to obtain three versions: a resynthesized original, with the original F_0 ; a close-copy stylization; an 'alternative' contour, made in such a way that the differences with the originals would be easily audible. Sentence pairs came in three categories: (A) a resynthesized original and a close-copy stylization; (B) two identical versions; (C) a resynthesized original or a close-copy stylization and an alternative contour. Subjects were asked to indicate for each pair

whether they thought its members to be exactly equal or not. In the instruction they were acquainted with the notion of close-copy stylizations, not only verbally, but with demonstrations on tape as well. Table I summarizes the results of the test.

Category A shows a very high number of responses 'equal'. One reason for this could have been that, since the pairs of category C differed so much, subjects' criteria might have been biased in favour of judging the pairs of category A to be equal. However, a post hoc analysis of the data did not substantiate this possibility. Moreover, the scores show that the number of 'equal' is lower with A than with B. This can be attributed to a number of cases in which some listeners were able to hear differences so small that they had not been observed by the experimenter. The listeners would therefore certainly have scored a higher number 'different' for category A if the objective differences would have been only slightly bigger than they actually were. The conclusion is that it is possible to stylize BE intonation by means of a restricted number of straight line segments, and nevertheless to maintain perceptual equality.

3. Standard stylizations

The ultimate aim of the stylization method is to replace the movements in close-copy stylizations by pitch movements with standard specifications of slope, duration and position in the syllable. The purpose of such a standardization is to facilitate the finding of correspondences between otherwise different contours, and thus reveal the melodic structure of the various patterns of the intonation system of the language. Such standard recipes have been developed for a limited variety of different intonation patterns. As examples of these patterns, two representative items were taken from each of the seven Primary Tones as recorded on the tape which goes with Halliday's course (1970). These are fully acceptable samples of BE intonation, as has been confirmed in the experiment to be discussed below.

Standard stylizations are generally not indistinguishable from their corresponding originals. But if we want to claim that standard stylizations are still representative of normal BE intonation, they should at least sound as acceptable as resynthesized originals.

Table I. Total number of responses, and number of cases in which members of a pair were judged equal or different. Objectively, the members in a pair were different in categories A and C, and equal in category B

Category	No. of responses	No. 'equal'	No. 'different'
A	1280	1111	169
B	1280	1222	58
C	1280	75	1205

This claim was tested in an acceptability experiment. Of each of the fourteen chosen utterances, five versions were made, only differing in intonation. The five versions were: resynthesized originals (ORIG), fully standardized stylizations (FS), partly standardized stylizations (PS), DUTCH-versions and WITTEN-versions.

In PS-versions not all parameters that characterize pitch movements are fixed, but two of them are allowed to vary freely, viz. overall position of F_0 and overall size of the excursions. DUTCH-versions were constructed following the rules developed for Dutch intonation. Care was taken to select Dutch patterns that resembled the given BE patterns as much as possible. WITTEN-versions were constructed following the recipes for the synthesis of Halliday's seven Primary Tones, as published by Witten (1977).

DUTCH-versions were incorporated with the expectation that, when recipes for another language are used to build the contours, the outcome will be less acceptable. Witten based himself on Halliday's impressionistic descriptions of the Tones. In view of the notorious unreliability of the unaided ear in intonational matters, these versions were also expected to be less acceptable than the ones based on instrumental analysis. Sixty-six native speakers of BE were asked to judge the acceptability of each of the test items on a five point scale. By means of the method of successive intervals (Edwards, 1957) a psychological continuum was derived from the raw data, and each of the stimuli assigned a scale value which refers to that continuum. See Table II.

Unlike the raw data, the scale values lend themselves to a statistical analysis. Such an analysis revealed that ORIG, FS and PS do not differ significantly in acceptability, whereas DUTCH and WITTEN do differ from each of these, and from each other. We may therefore conclude that the FS- and PS-versions sufficiently agree with the internal representations the listeners have of the intonation patterns of their language. But at the same time, the outcome for DUTCH- and WITTEN-versions shows that the recipes used for the standardized versions are necessary to agree with the listeners' internal representations: as soon as other recipes are applied, the resulting contours sound less acceptable. Thus, the attentive subjects are sensitive to any violation of the rules of their intonation system. However, their critical behaviour did not keep them from approving of the FS- (and PS-) versions. These versions can therefore be considered as fully representative of some patterns of the BE intonation system.

Table II. Scale values, averaged over the 14 stimuli, and standard deviations, for each of the five conditions used in the acceptability test

	ORIG	FS	PS	DUTCH	WITTEN
Mean	2.676	2.712	2.476	1.839	.654
s.d.	.438	.384	.541	.625	.454

Acknowledgements

This research was funded by the Foundation for Linguistic Research, funded by the Netherlands Organisation for the Advancement of Pure Research (Z.W.O.) under grant no. 17-21-07.

References

- Edward, A.L. (1957). *Techniques of attitude scale construction*. New York: Appleton-Century Crofts.
- Halliday, M.A.K. (1970). *A course in spoken English: intonation*. Oxford: Oxford University Press.
- 't Hart, J. and Cohen, A. (1973). Intonation by rule: a perceptual quest. *J. of Phonetics* 1, 309-327.
- 't Hart, J. and Collier, R. (1975). Integrating different levels of intonation analysis. *J. of Phonetics* 3, 235-255.
- Witten, I.H. (1977). A flexible scheme for assigning timing and pitch to synthetic speech. *Language and Speech* 20, 240-260.