

FALLS: VARIABILITY AND PERCEPTUAL EFFECTS

Anne Wichmann

IBM(UK) Scientific Centre
Winchester, SO23 9DR, England

ABSTRACT

This paper presents an experiment designed to test the effect of final intonation contours on the degree to which an utterance is perceived to be final. The utterances were taken from a corpus [3] of naturally occurring monologue. Each was syntactically complete and semantically unmarked for finality. Keeping the endpoint constant, the starting point of the final fall was systematically manipulated to create 5 different versions of each sentence. The results of the perception experiment suggest that the higher the starting point of the final fall, the less final that utterance is perceived to be. There is no evidence for any discrete perceptual categories.

1. INTRODUCTION

In abstract representations of intonation, the end of a declarative utterance is generally indicated by assigning a falling contour. Physically, a fall can be any pitch contour that ends at a pitch lower than its starting point. Since both starting point and endpoint are variable within the range of any one speaker, there are any number of falls which that speaker can produce. It is generally assumed, however, that these physical differences are not significant, and that the height of the fall is determined by the declining topline across the utterance. Any significant differences in the resulting overall contour have in the past been related

to the slope of the fall, residual fall, and endpoint. This study shows that the starting point also has a systematic perceptual effect.

Other experimental studies of the acoustic correlates of boundaries [1] [2] have compared the physical realisation of contours at the end of syntactically complete and incomplete utterances. In contrast, this experiment uses only syntactically complete utterances. This study also differs from others in that it uses only naturally occurring data. The availability of resynthesis techniques has allowed for at least partial control of the stimuli.

2. EXPERIMENT

The experiment described here poses two questions:

- (i) does a change in the height of the starting point influence the perception of finality?
- (ii) how does such an effect relate to f0?

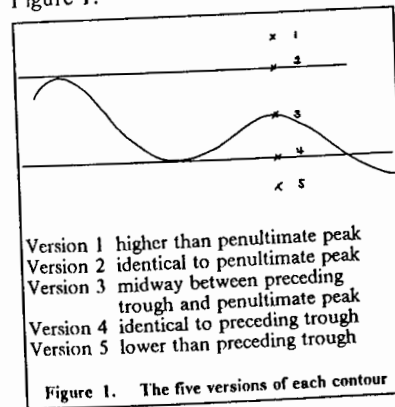
2.1. Method

Ten subjects were presented with five versions of each of 10 naturally-occurring utterances, 50 utterances in all, of which the final contours had been systematically manipulated. The utterances were all syntactically complete, and perceived (in a preliminary experiment) to be semantically unmarked for finality. Leaving the endpoint constant, the final falling contour of each one was assigned five different starting points, varying sys-

tematically in height. These five versions of each utterance were LPC-resynthesised. Listeners were asked to judge, on a four point scale, whether the speakers had finished or whether they had more to say.

2.2. Preparation of the stimuli

The ten sentences were digitised at 10 KHz, and normalised for amplitude. Pitch was extracted and the resulting f0 values were checked for octave leaps, and smoothed by hand. The peak of the last accented syllable in each sentence was manipulated to create five different versions of the f0 contour. In each case the f0 peak associated with the accent was adjusted to one of five different positions with relation to the preceding trough and penultimate peak, as illustrated in Figure 1.



In creating each version of the f0 contour, the f0 values preceding and following the manipulated peaks were adjusted to maintain as far as possible both microprosodic features and the correspondence between f0 and segments.

2.3. Procedure

The resynthesis of 10 sentences, each in five different versions, produced a set of 50 different stimuli. A stimulus sequence file was generated in which each stimulus was repeated five times, thus eliciting 250 responses from each

subject. They were preceded by a test sequence of 10 stimuli which were ignored in the analysis. The subjects were asked to judge whether the speaker of each sentence was

- definitely going on,
- probably going on,
- had probably finished, or
- had definitely finished.

For the purpose of the analysis, these responses were converted into ordered data. The response 'definitely going on' became a '1', 'probably going on' became '2', 'probably finished' became '3', and 'definitely finished' became '4'. The lower the score, the less final the utterance was perceived to be.

3. ANALYSIS

An analysis of the results must aim to investigate the significance of all effects: sentences, subjects and f0 contours. It was difficult to do this formally because the responses were ordered categories 1 to 4. This kind of response violates the usual normality assumptions for classical ANOVA. However, the package 'PLUM' [4] was used to fit the appropriate ordered responses category model. The adequacy of fit of each 'treatment' (f0 version) is done by comparing differences in deviance between models with and without the treatment effect (goodness of fit) and comparing these differences with the appropriate chi square value. The results are shown in Table 1.

Table 1. Significance of sentence, subject and f0 effects.

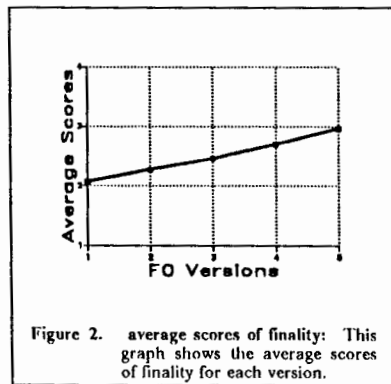
	deviance difference	d.f.	sig. level
treatments (f0 versions)	(2640 - 2199) = 441	4	.01
subjects	(2438 - 2199) = 239	9	.01
sentences	(3113 - 2199) = 914	9	.01

As expected, all effects are highly sig-

nificant at the 1% level.

3.1. F0 effects

If we average the finality scores for all sentences and all subjects we see that there is a systematic gradient difference between the versions. It is clear that the lower the starting point of the final fall, the greater the degree of perceived finality. See Figure 2.

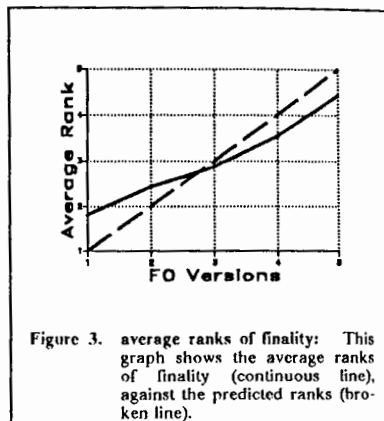


We know, however, that both the sentences and the subjects had a significant effect on the results. In order to see the effects of f0 manipulation without the influence of the significant between-sentence differences, the average scores across replicates for each version were ranked from 1 to 5. The least degree of perceived finality was ranked 1 and the highest score ranked 5. Tied scores were ranked equally.

For each of the versions 1 to 5, the average rank was calculated across all sentences and all subjects.

The results of this ranking are shown in Figure 3.

It can be seen that the different f0 versions have a marked and consistent effect on the perception of finality. A divergence from the predicted ranks is only possible in one direction, since no utterance could be perceived as less final than 'definitely going on' or as more final than 'definitely finished'. The deviations of the



endpoints of the line are therefore to be expected. The greater deviance at the non-final end of the scale is also predictable, since we can assume that the presence of any falling contour which falls to the speaker's base line will indicate at least some degree of finality. There are minor differences in the degree of slope between any two consecutive points, but the overall trend is a straight line. The f0 versions therefore have a significant effect on the way subjects ranked utterances in terms of degrees of finality. The closest fit would be a straight line, and the results must therefore be interpreted as gradient rather than categorical.

4. DISCUSSION

4.1. The perception of finality

The first question posed by this experiment must be answered with yes: there is clear evidence that the starting point of a final fall influences the listener's perception of finality. There is no evidence to suggest that this perception is a categorical one. Whether we take the average scores of finality, or the average ranking of the different versions relative to one another, the result is a gradient. This suggests that finality is perceived in terms of degrees rather than in terms of categories, binary or otherwise.

4.2. Finality and f0

The answer to the second question - how does this relate to f0 - is not so clear. Is there such a distinction to be made as 'high' and 'low' fall, and if so, how are these to be defined?

It may be that the height of final falls is not perceived relative to any preceding syllables but to the speaker's overall range. All the versions of sentence 10, for example, were fairly low in the speaker's range. This might explain the tendency to judge this sentence as inherently more final than others.

Menn and Boyce [5] claim that the endpoint of a sentence-final falling contour can be regarded as constant in relation to the speaker's norm. This was assumed to be the case in the experiment described above. Nonetheless, a similar experiment in which the endpoint of each final contour was systematically changed would complement the present study and perhaps throw light on results which are not accounted for here.

5. CONCLUSION

The tentative conclusions to be drawn from this study are as follows.

- The height of a fall given a constant low endpoint is perceived as a gradient. There is no evidence to support a categorisation of falls into high and low.
- The starting point of the final fall influences the degree of perceived finality. The lower the starting point of such a fall the more final the utterance is perceived to be. There are however other prosodic influences on the degree of perceived finality which cannot be accounted for here.

- Height seems to be perceived in relation to a speaker's norm and not in relation to the pitch of preceding syllables, accented or not.

REFERENCES

- [1] BERKOVITS, R. (1984), "Duration and fundamental frequency in sentence-final intonation." *Journal of Phonetics* 12, 255-265
- [2] FARNETANI, E. (1989), "Acoustic correlates of linguistic boundaries in Italian: A study on duration and fundamental frequency." *European Conference on Speech Communication and Technology. Proceedings Eurospeech '89 Paris Vol 2*, 332-335
- [3] Lancaster/IBM Spoken English Corpus
- [4] McCULLAGH, P. (1979), *PLUM An Interactive Computer Package for Analysing Ordinal Data Working Paper*, Department of Statistics, University of Chicago
- [5] MENN, L., BOYCE, S. (1982), 'Fundamental frequency and discourse structure.' *Language and Speech* Vol 25 Part 4