DECLINATION AND SENTENCE INTONATION IN ITALIAN

CINZIA AVESANI

Scuola Normale Superiore - 56100 Pisa - ITALY

ABSTRACT

This work on Italian intonation was designed to test Cooper and Sorensen's declination model. In their model, one half of the downdrift is accomplished in one fourth of the utterance's temporal extent; all the peaks except the first are related in a linear fashion and the first peak is correlated with sentence length. My analyses of F0 contours do not support the Cooper and Sorensen model. There is no significant correlation between the first peak height and sentence duration. In addition, a linear declination slope fitted to the whole sentence or to the sentence peaks after the first, predicts acuurately only a small percentage of sentence peaks. I suggest that local pitch accents and final pitch fall account for Italian F0 contours better than global realization rules.

INTRODUCTION

According to Bolinger (/1/), the most widely diffused intonational phenomenon is the tendency towards a low pitch at the end of declarative sentences. This "running down" pattern has been given a variety of names, of which the best known is probably declination. Declination is said to arise from the property of speech fundamental frequency to gradually decline over the course of utterances (e.g. Cohen, Collier and 't Hart /2/; Collier /3/; Thorsen /11/; Gelfer /6/), or from a rapid downmotion at the very end of an F0 contour, while non-terminal portions do not necessarily display a downwards slope (Lieberman, Katz, Jongman, Zimmerman and Miller /8/; Umeda /13/).

In the former approach, phrasal intonational properties are described in terms of global falling trends; in the latter, in terms of (final) falling features with a sentence-level parsing function.

Cooper and Sorensen's model (/4/) is an example of the first approach, since it aims to capture single F0 peak values by means of a global rule which generates falling trends within an utterance domain (Topline Rule). In this model one half of the downdrift is accomplished in one fourth of the utterance temporal extent. The first peak is correlated with sentence length and all remaining peaks are related in a linear fashion. According to Cooper and Sorensen (henceforth C&S), declination peaks are under the talker's voluntary control. The general theoretical claim is that of a statistical relation between utterance length and F0 as evidence of sentence-level global preplanning.

However, other models of English sentence intonation account for downtrends without invoking pre-planning (e.g. Fujisaki /5/; Pierrehumbert /12/). Interestingly, these are base-line models: Fujisaki's is in terms of global declination trends, where a physiological "baseline" component constitutes the phonetic framework or constraint for local phonological entities (accents) to occur (see also Ladd /7/). In this model an exponential baseline decay is posited, which may be seen as representing damping effects of the speech motor implementation. In a similar "frame of reference" approach proposed by Pierrehumbert, target F0 peaks are scaled on an abstract declining baseline.

The point I want to make here is that the weak rate of decay in F0 peaks following the first one (P1), predicted by an exponential equation, is consistent with a physiological explanation of global declination alternative to C&S's cognitive model (Fujisaki /5/; Gelfer /6/): according to Gelfer, declination is the byproduct of F0 automatic modifications in response to decreasing subglottal pressure; in Fujisaki, declination is the effect of the dynamic properties of the speech control mechanism. Also, it is not inconsistent with the "local feature" approach proposed by Liberman and Pierrehumbert /10/, in terms of exponential declination as the global statistical effect of such different local events as phonetic "final lowering" and a set of phonological "downstepping rules" characterizing the medial portion of sentences.

Moreover, as compared to C&S algorythm, what characterizes no pre-planning models such as Fujisaki's and Liberman and Pierrehumbert's (L&P), is the crucial fact that they do not posit a covariance relation between P1 and sentence length, even if for different reasons.

Therefore, as an attempt to test the C&S hypothesis, this work focuses on the statistical relation between F0 peaks and sentence length, in order to see: first, whether a correlation exists in Italian between the F0 contour first peak (P1) and sentence length; second, whether the attested downdrift in Italian declarative sentences (e.g. Magno-Caldognetto, Ferrero, Lavagnoli and Vagges /10/) can be adequately described in terms of linear declination trends fitting the peaks which follow the first one as predicted by C&S model.

EXPERIMENT

Three male speakers produced four tokens of different sentences varying in length. Three sets of utterances of increasing length included: one-word sentences with five steps of length variation, i.e. ranging from one to five syllable length (e.g.: "Tom"... "indicameli"); simple sentences with four steps of length variation, ranging from one-word tofour-words length (e.g.: "'Monica" ..."'Monica compera mobili rustici"); and complex sentences with five steps of length variation. These latter consist of two coordinated clauses including /e/ ("and") and /ma/ ("but") as coordinating conjunctions. Length variations were obtained by increasing both coordinated clauses: three steps of variation derived from a symmetrical increase of both clauses, which produced short-short, medium-medium, long-long durational patterns; two other steps were obtained asymmetrically varying each clause length, in order to get a medium-long and a long-medium durational pattern. Analyses of three out of four repetitions were made using a pitch extraction algorythm in the ILS analysis system at Haskins Laboratories.

To test the first prediction of the C&S model, i.e. that P1 is strongly related to sentence length, I measured the P1 height of all the sentences across all conditions. In order to minimize segmental effects on the first peak height, both one-word and simple sentences were produced also in reiterant versions, using /da/ as the syllable for mimicking natural utterances. Pearson's correlation coefficients were calculated between the first peak height and sentence length computed in number of syllables (see C&S /4/; Fujisaki /6/) for each group of one-word, simple and complex sentences of increasing length. I expect that the P1 height will rise as sentence length increases.

To test the second prediction of the model, i.e. that all the peaks except the first in a contour are related in a linear fashion, I measured all the high values occurring on stressed words for complex sentences only. A linear regression technique for fitting theoretically linear sets of data points was used to derive "best fit" toplines, choosing peak's time of occurrence to represent the independent variable and the peak's height in Hz as the dependent variable. Correlation coefficients provide an adeguacy measure for the linear relation of the data points. Levels of significance were assigned based on $\underline{r}$ values.

RESULTS AND DISCUSSION

First peak height.

One-word sentences. Table 1 shows that no significant correlation was found betweenword length and the first peak height in both natural and reiterant versions of words of increasing length. Correlation coefficients separately calculated on the production of individual speakers reveal strong intraspeaker differences: for one speaker only (MV) $\underline{r}$ is significant in both natural and reiterant utterances; for speaker GN $\underline{r}$ is never significant; for speaker MM there is a significant and positive correlation only as far as reiterant words are concerned.

Simple sentences. Essentially the same picture emerges from the results of the correlation test on simple sentences computed on all talkers where the correlation between P1 of the first word in the utterance and global sentence length does not reach the significance level of .05 in either natural or reiterant versions.

Simple sentences were divided into three subgroups according to the stress pattern of the initial words (e.g. /'Monica; Do'nata; Mari'lu/); as P1 cooccurs with stressed sullable, I tested whether stress position in the initial word had any significant relation with sentence length. Table 2 shows that initial, medial and final peak positions in the initial word actually yield different $\underline{r}$ values, but that is statistically significant.

Intraspeaker differences again arose from correlations computed separately on each talker. Speakers MV and GN show no significant correlations in both natural and reiterant speech, while $\underline{r}$ is higly signifivant for the third speaker (MM).

An observation can be made: if a significant correlation between P1 and sentence length is an evidence of a look-ahead strategy by the subject, at the word level speaker MV alone appeared to use it consistently in natural and reiterant speech. However, in the (simple) sentence domain, the same speaker does not show any effect of sentence length on P1; on the other hand, (anticipatory) effects related to sentence length did show up significantly, in a sentence domain,for one of the other speakers (MM) who showed non-significant effects of increasing length on P1 height of natural words.

Complex sentences. These sentences are composed of two coordinated clauses; hence, a P1 raising effect if any, might be induced either by the increasing length of the whole sentence or by the increasing length of the first clause. Table 3 provides correlation scores of statistical analyses computed between P1 and first clause length.

Correlation computed on all the speakers are non significant not only for P1 and whole sentence length but also for P1 and first clause length. Intraspeakers differences are confirmed: MV and GN show non-significant correlation, consistently with results

from their simple sentences; talker MM shows a significant r value for correlation between P1 and whole sentence length, and a non significant one for correlation between P1 and first clause length.

These data skow that, when it occurs, the raising of P1 height is due rather to a global sentence length increase than to a first clause length increase.

As a first conclusion, it may be said that my results do not support C&S's predictions of a systematic relation between first peak values and sentence length. Intraspeakers differences were such that only one out of three speakers consistently showed higher P1 values as simple and complex sentence length increased. Moreover, interspeaker variations point out the unstable character of such an effect (see L&P /10/

Topline declination.

To test the hypothesis of a linear declination over the peaks which follow the first one, a simple linear regression was computed on all complex sentences varying in length. In this analysis I took into account all the peaks except the first one, according to C&S's formulation. The relative correlation coefficient was significant ($p < .001$), showing that complex sentences actually display peak downtrend, but the coefficient's low value ($r = -.294$) suggests that only a small percentage of the data peaks (8%) can be fitted by a straight line .

As complex sentence length was substantially varied systematically modifying each clause's length, I checked whether correlation coefficients varied because of the different durational patterns of sentences: regressionlines computed separately on sentences of differnt lengths show that differences in $\underline{r}$ values ranged from $\underline{r} = -.230$ (for sentences with a long first clause and a medium second one) to $\underline{r} = -.691$ (for sentences where both clauses were short).

In presence of complex sentences, a possible interpretation of these outcomes (e.g. C&S) is that topline declination is "reset" at the clause boundary, generating individual toplines for each clause. If this is the case, single linear regression lines separately computed for each clause on all their peaks except the first should give better fits than the global one. My results, however, do not support this hypothesis . First clause: computed on the whole set of sentences, the analysis gives non-significant $\underline{r}$ values; separately computed on all first clause tokens belonging to different durational patterns, the analysis provides quite different $\underline{r}$ values, which in itself is not a result that confirms C&S's predictions. Second clause: a linear regression line has a negative value and a significant correlation score ($\underline{r} = -.238$; $p < .001$). However, it shows a lower $\underline{r}$ value than that obtained for the entire complex sentence.

These outcomes seem to suggest that the regular falling intonational patterns of the second clause mainly contribute to sentence level topline declination.

P1. based on regression lines computed on all peaks except the initial one on complex sentences, the P1 observed values are not systematically higher than the predicted ones. In one durational pattern only (long-long) the average of observed P1's is 11 Hz higher than the predicted first peak value.

As a consequence, linear regressions computed on all the peaks in complex sentence contours give better fits than regression computed excluding P1: the coefficients for all talkers on all sentences being respectively $\underline{r} = -.385$ vs. $\underline{r} = -.294$.

So, the outcome of no substantial drop between the first and the second peak couples with the L&P hypothesis of local final lowering as an effect contributing to declination.

Final lowering effects.

To test final lowering, new regression lines computed on all peaks except the last were calculated. Results give worse fits and flatter slopes compared with all peaks regression, as I would expect if final lowering substantially contributed to declination: on all sentences scores are respectively: $\underline{r} = -.256$ vs. $\underline{r} = -.385$; slope = -4.46 vs. -6.81 Hz per sec. Looking at the overall peak regression, it is final peak - as opposed to the first one - that is crucial in determing declination: this might suggest that sentence declination trends are mainly generated by local (initial and final) F0 values.

Medial downdrift.

Moreover, excluding both initial and final peaks from linear regression computation, an almost negligible relationship emerges between the peaks' height and their position in an utterance; and the slope of the line reduces drastically to 2.7 Hz per sec. ($\underline{r} = -.146$; $\underline{p} = .002$).

CONCLUSION

Results suggest that the height of the first peak in a F0 contour is not systematically related to sentence length for all speakers. Moreover, linear regression lines fitted onto all peaks in a contour except the first one give an account of topline declination that captures only a small set of the data.

Despite the fact that P1 is higher than the subsequent peaks, it does not lie well above the line connecting all the others, as predicted by the C&S model.

On the other hand, a weak decay in the medial portion of F0 contours and a substantial effect of final peak lowering seem to point to an explanation of topline declination as statistically arising from local (initial and final) F0 values, while the attested weak medial downdrift may be consistent with a physiological explanation of declination as the by-product of declining subglottal pressure (Gelfer et al. /6/).

Final lowering may be the result of the physiological relaxation of the articulatory system. As for the shape of the medial downdrift, the experiment was not designed to answer the question whether local phonological rules (e.g. "downstepping") determine the shape of intermediate F0 contours (L&P /9/). This will certainly be a line for future research, as Ladd indicates (/7/). The experiment aimed to collect evidence regarding Italian intonation which might allow to rule out the strong assumption of global preplanning in explaining falling intonational trends.

TABLE 1

One-word sentences

	Reiterant		Natural	
	r	p	r	p
gen.	.341	>.10	.214	>.25
MV	.939	<.025	.962	<.05
GN	.73	> .10	.465	>.25
MM	.8	<.05	.632	>.25

Table 1. Correlations between first peak height and sentence length.Gen.: on all speakers; MM, GN, MV: on single speakers.

TABLE 2

Simple sentences

	Reiterant		Natural	
	r	p	r	p
gen.	.304	<.10	.21	>.10
MV	-.136	>.25	-.03	>.25
GN	.309	>.25	.268	>.25
MM	.905	<.001	.866	<.005
/'---/	.264	>.25	.149	>.25
/-'--/	.16	>.25	.163	>.25
/--'-/	.471	>.10	.318	>.25

Table 2. Correlations between first peak height and sentence length. /---/ schematically represents the different stress patterns of the sentence initial words.

TABLE 3

Complex sentences

	P1/sentence		P1/first clause	
	r	p	r	p
gen.	.226	<.10	.093	>.10
MV	-.341	>.10	-.412	<.10
GN	.216	>.10	.169	>.10
MM	.504	<.05	.410	<.10

Table 3. Correlations between first peak height and sentence length, and between first peak height and the length of the sentence's first clause.

REFERENCES

/1/ D. Bolinger, "Intonation across languages". J. Greenberg, C. Ferguson & E. Mravcsik, eds., Universalsof human language, Standford University Press, 1978.

/2/ A. Cohen, R. Collier & J. 't Hart, "Declination: construct or intrinsic feature of speech pitch?", Phonetica, 39, 1982.

/3/ R. Collier, "Physiological correlates of intonation patterns", J.A.S.A., 58, 1975.

/4/ W. Cooper & J. Sorensen, Fundamental frequency in sentence production, Springer, 1981.

/5/ H. Fujisaki, "Dynamis characteristics of voice fundamental frequency in speech and singing". F.Ferrero, ed., Proceedings of the Fourth F.A.S.E. Symposium on Acoustics and Speech, Venice, 1981.

/6/ C. Gelfer, K. Harris, R. Collier & T. Baer, "Speculations on the control of fundamental frequency declination", Haskins Laboratories Status Report, 76, 1983.

/7/ R. Ladd, "Declination: a review and some hypotheses". C. Ewen & J. Anderson, eds., Phonology Yearbook, 1, 1984.

/8/ P. Lieberman, W. Katz, A. Jongman, R. Zimmerman & A. Miller, "Measures of the sentence intonation of read and spontaneous speech in American English", J.A.S.A., 77, 1985.

/9/ M. Liberman & J. Pierrehumbert, "Intonational invariance under changes in pitch range and length". M. Aronoff & R. Oehrle, eds., Language sound structure, MIT Press, 1984.

/10/ E. Magno-Caldognetto, F. Ferrero, C. Lavagnoli, & K. Vagges, "F0 contours of statements, yes-no questions, and wh-questions of two regional varieties of Italian", J. of Italian Linguistics, 3, 1978.

/11/ N. Thorsen, "Sentence intonation in textual context - Supplmentary data", J.A.S.A., 80, 1986.

/12/ J. Pierrehumbert, "The phonology and phonetics of English intonation", PhD dissertation, MIT, 1980.

/12/ N. Umeda, "F0 declination is situation dependent", J. of Phonetics, 10, 1982.

ACKNOWLEDGEMENT

I thank Haskins Laboratories for their friendly hospitality and continuous support during this research. My particular thanks go to Carol Fowler and to Mario Vayra.