# FURTHER EXPERIMENTAL STUDIES OF FUNDAMENTAL FREQUENCY CONTOURS*

MICHAEL STUDDERT-KENNEDY AND KERSTIN HADDING

The general purpose of the present study was to increase our understanding of the mechanisms by which we perceive and interpret intonation contours formed by variations in fundamental frequency. The work replicates and extends earlier studies of synthetically manipulated fundamental frequency contours (Hadding-Koch and Studdert-Kennedy 1963, 1964, 1965).

Stimuli were prepared by means of the Haskins Laboratories Digital Spectrum Manipulator (DSM) (Cooper 1964). The utterance 'November' [no'vεmbɚ], chosen to be acceptable to both Swedish and American English listeners, was spoken by an American male voice, digitized and stored in the DSM. $F_0$ was then manipulated over a range from 85 to 220 Hz. The contours are schematized in Figure 1. The combination of three peaks, four turning-points and six end-points yielded 72 contours. Each was recorded in two forms: (1) carried on a speech wave [no'vεmbɚ], (2) as a frequency-modulated sine wave. Each set of 72 was spliced into five different random permutations and presented to 22 Swedish and 16 U.S. subjects for judgment. For the sine waves, subjects were asked to judge whether the final glide of each contour was rising or falling. For the speech waves, a psychophysical task called for judgment of the final glide as rising or falling, exactly as for the sine-waves, and, in a separate session, a linguistic task called for judgment of the entire contour as either more like a question or more like a statement.

The next four figures present results. In each figure, the columns group linguistic, speech psychophysical and sine wave results from left to right, while the rows group sets of contours according to specified parameters. In each graph, the percentage of question, or rise, responses is plotted as a function of the extent of the terminal glide in Hz from turning-point to end-point, with negative values indicating a fall, positive values a rise. The diagonal movement, upward from left to right, of each curve therefore reflects the influence of the terminal glide on subjects' responses.

Consider, first, Figures 2 and 3. Here the value of the peak is constant across each row and parameters of the curves are the turning points numbered from low to high.

* This study was supported in part by a grant to Haskins Laboratories from NICHD, Washington, D.C., U.S.A.



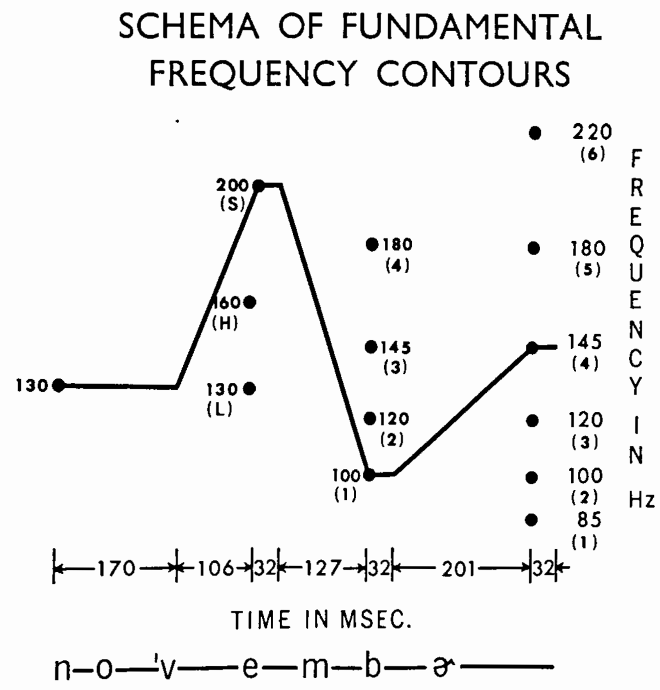


Fig. 1. Schema of fundamental frequency contours imposed on the utterance "November" [nóvεmbɚ].

Examining the speech and sine wave psycho-physical columns, we see that, in each griaph, the curves are staged across the terminal glide axis from right to left in order of turning points from low to high: the higher the turning point the smaller the termnal rise required for subjects to hear a rise. For example, a contour such as *H*1 with a strong fall from peak to turning-point requires a marked terminal rise to be judged as rising, while a contour such as *H*4 with a steady rise before the terminal glide may be heard as rising even when the glide is falling. Such effects are particularly marked in the speech psychophysical data of both groups. Evidently, listeners have difficulty in separating the terminal glide from earlier sections of the contour, if those earlier sections have a marked movement. The terminal glides of contours with a turning-point relatively close to the precontour level (e.g., *S*3, *H*3, *L*3) are more accurately perceived, perhaps because listeners are able to average across earlier sections of the contour and establish an anchor against which the terminal glide may be judged.

In the linguistic judgments of Figures 2 and 3, essentially the same effect of turning-point is present. But we take its origin to be auditory rather than linguistic since it is also present, though reduced, in the sine wave data: the turning point evidently affects linguistic judgments indirectly by influencing auditory perception of the terminal glide. We may therefore reduce the two acoustic variables of turning-point

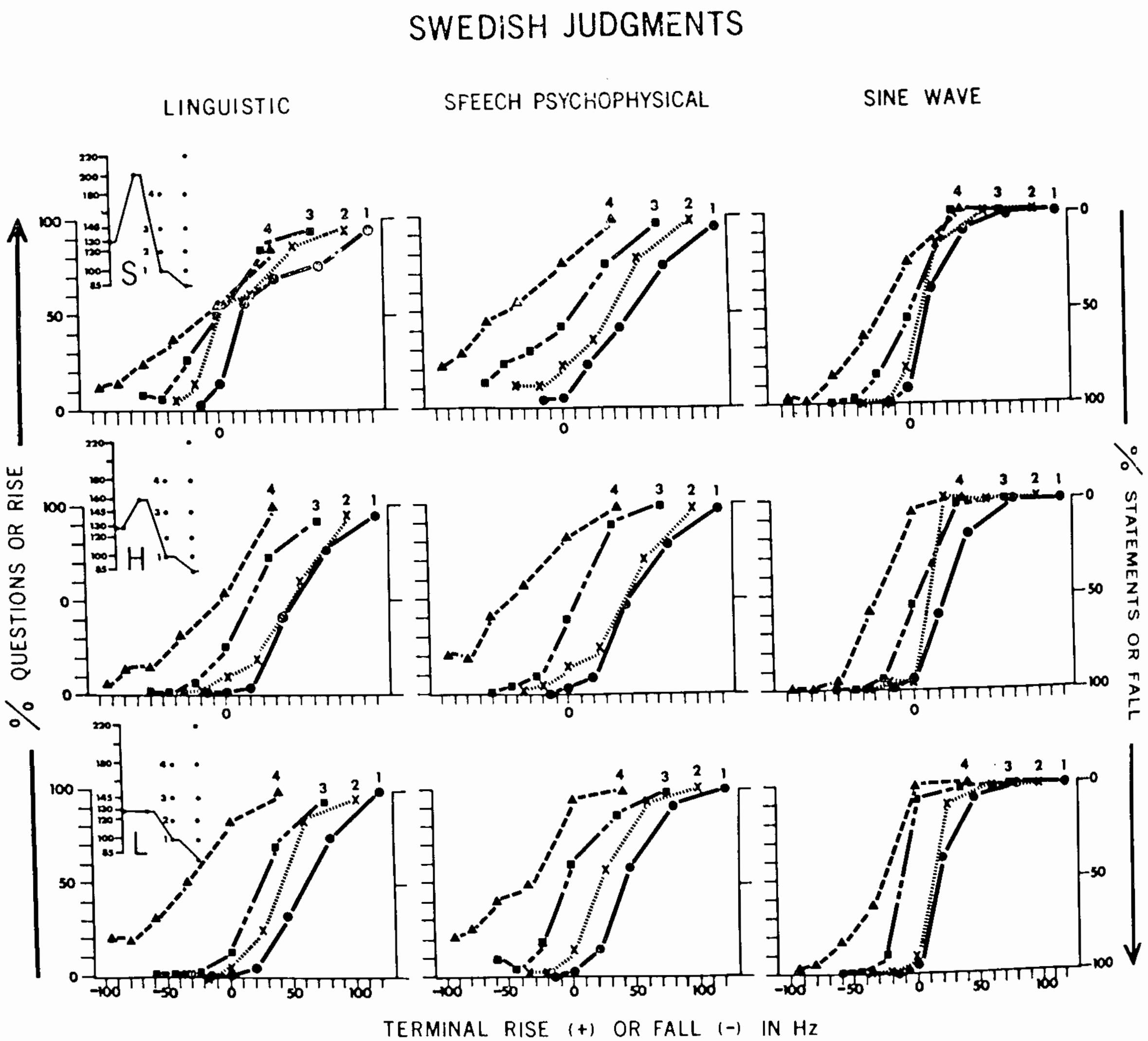


Fig. 2. Percentage of question or rise responses (left-axis) and statement or fall responses (right-axis) plotted as a function of terminal glide in Hz. Peak values are constant across rows and turning-points are parameters of the curves. For Swedish subjects.

and terminal glide to a single perceptual variable: perceived terminal glide. This is the principal determinant of the linguistic judgments.

However, the frequency level of the peak is also important, particularly when later sections of the contour are low. In Figures 4 and 5 the data have been re-arranged: turning-point value is fixed across each row, and parameters of the curves within a row are peak values. Consider, first, the top two rows of the linguistic columns. Here we see that the curves for the superhigh *(S)* contours are leftward of those for the other two peak values: a smaller terminal rise is required for contours to be judged



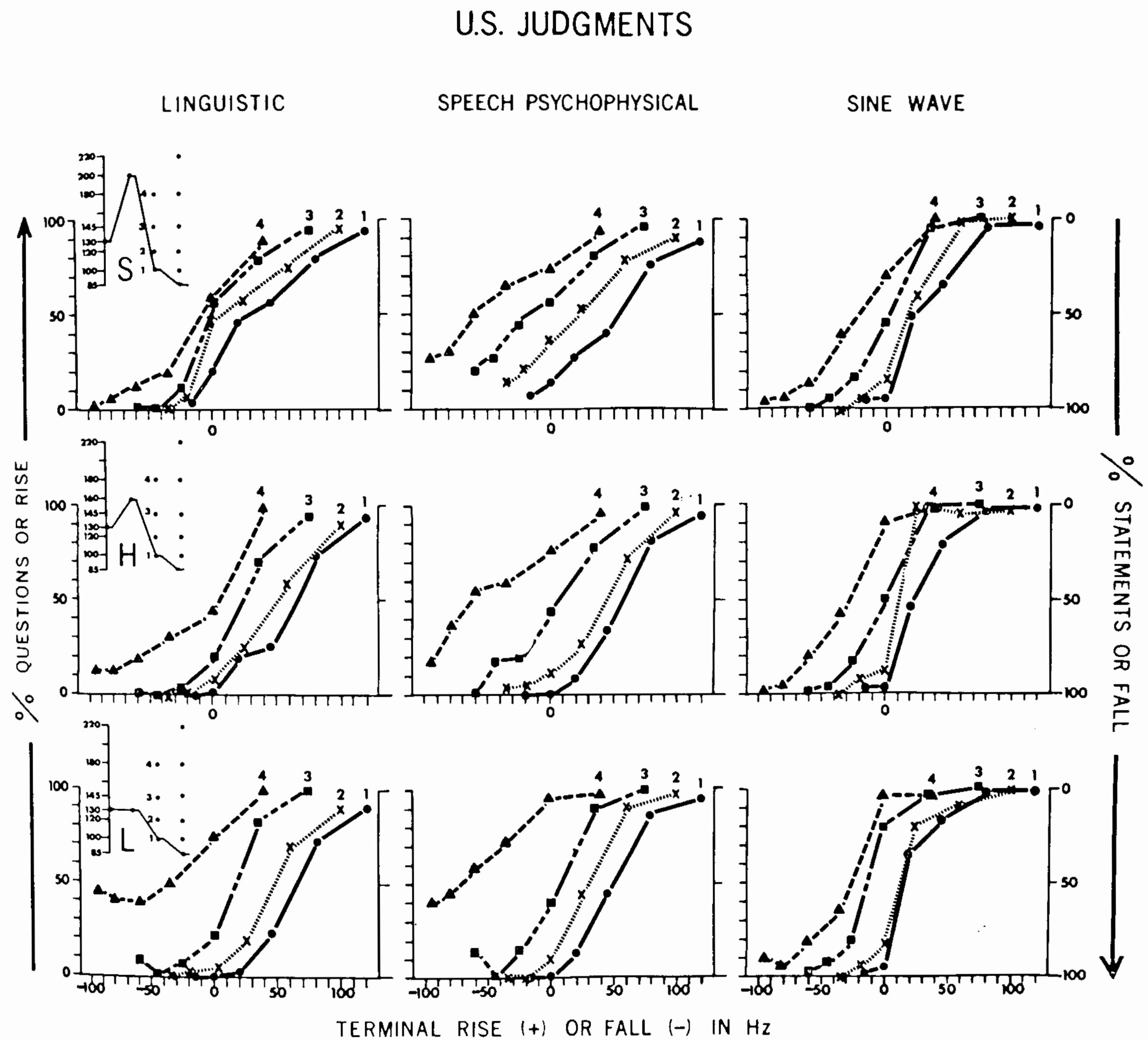


Fig. 3. As for Figure 2, for the American subjects.

as questions 50% of the time, if the peak is unusually high (200 Hz). The effect is reduced as the turning-point rises, and with a turning point at 180 Hz (fourth row), it disappears. Here, subjects (particularly Americans) find the continuously rising contours of the L4 series difficult to accept as statements, even if terminated by a sharp fall. Comparing these results with those for the speech and sine wave psychophysical judgments of Figures 4 and 5, we see that traces of the peak effect are present in the speech wave psychophysical judgments, particularly of American subjects, but are totally absent from the sine wave data. We conclude that the peak effect is linguistic and unrelated to auditory judgments of the terminal glide.

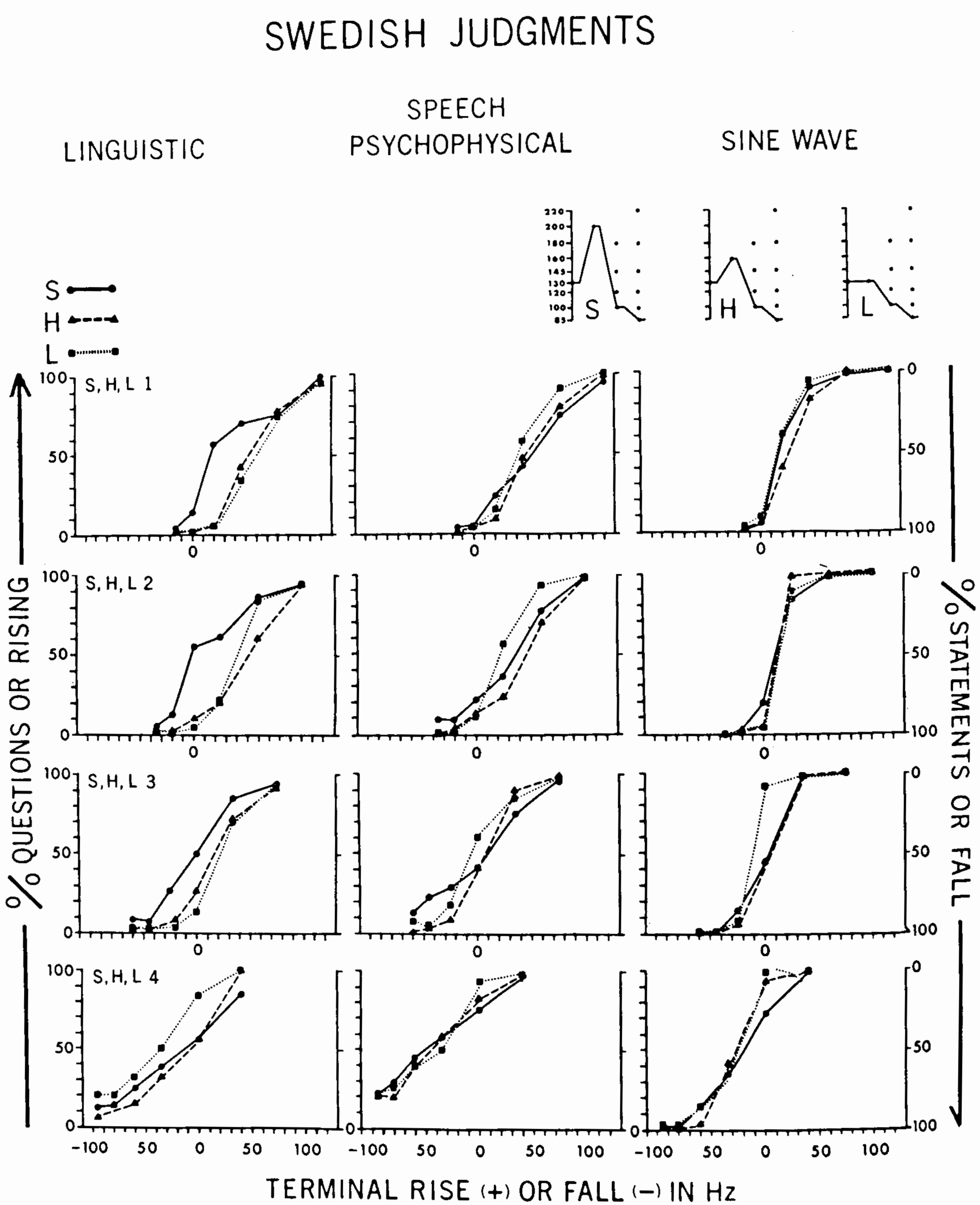


Fig. 4. Percentage of question or rise responses (left-axis) and statement or fall responses (right-axis) plotted as a function of terminal glide in Hz. Turning-point values are constant across rows and peak values are parameters of the curves. For Swedish subjects.



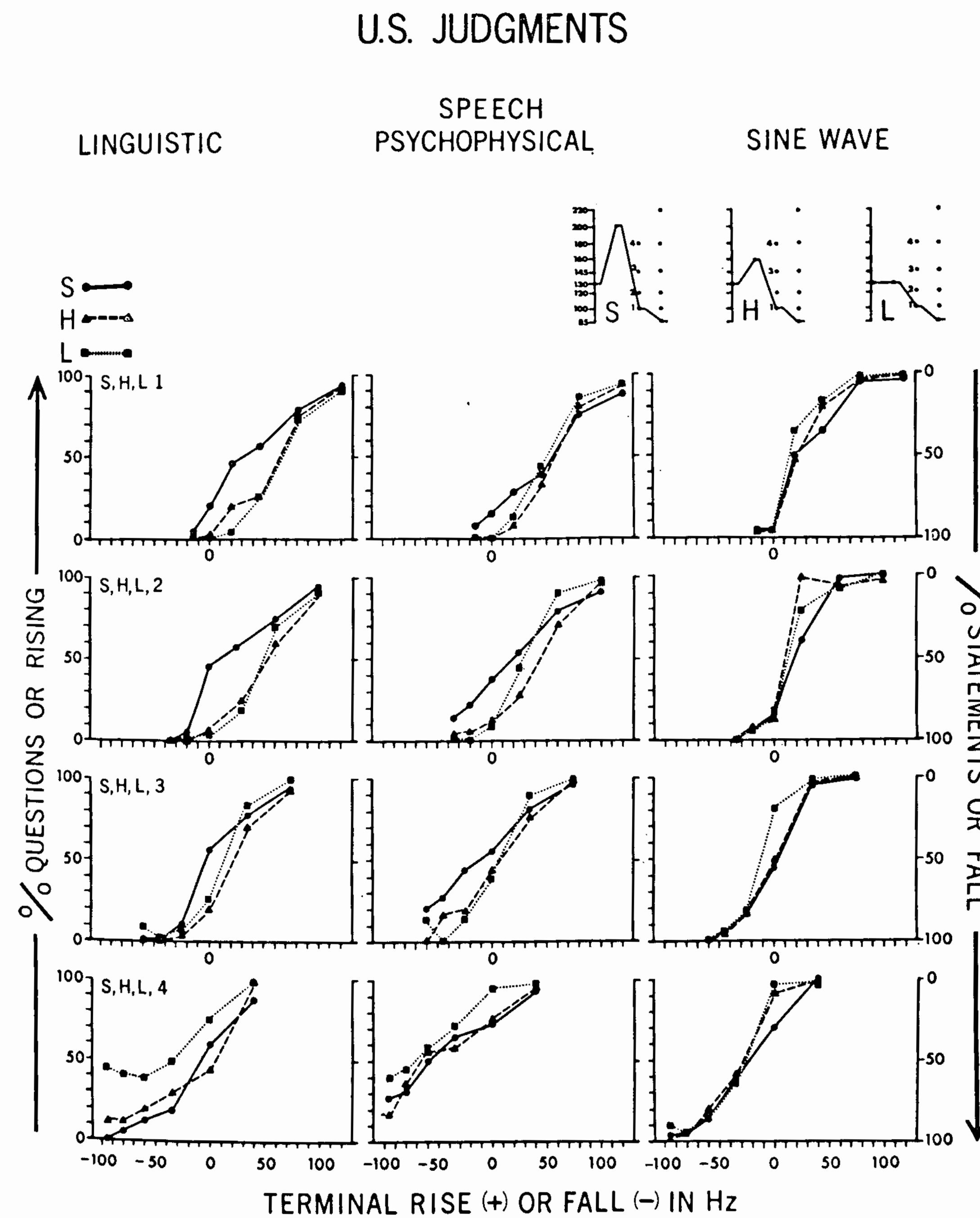


Fig. 5. As for Figure 4, for the American subjects.

To sum up, two factors seem to govern linguistic judgments of intonation contours such as those of this study: fundamental frequency at the peak and perceived terminal glide. The entire contour is then interpreted as a unit with these factors in some serial, weighted combination: if a low peak is heard, listeners tend to interpret the utterance as a statement, unless they also hear a large terminal rise; if a high peak is heard, listeners tend to interpret the utterance as a question, unless they also hear a large terminal fall.

However, this account is not complete. Peak and terminal glide are not simply and always additive in their effects. A contour with a steady rise from precontour to endpoint, may require a relatively small terminal rise to be heard as a question (see, for example, the linguistic judgments of contour *L*3). Here, it is the overall sweep of the pattern that determines the judgment rather than the frequency levels of particular segments of the contour.

Finally, we note that the main drift of the data was remarkably similar for the two language groups. The differences were small. For example, the Swedish preference for an overall high contour on questions, suggested by naturalistic observation, was reflected in the Swedish subjects' consistently giving more question responses than the Americans to contours of the *S*2, *S*3, *S*4 series. And the wider use of a steadily rising contour for questions in American English seems to be reflected in the tendency for Americans to hear the contours of the L4 series as questions, even when terminated by a strong fall. However, the main results of the study are consistent with a tendency, reported for many languages, to signal statements with a low contour, questions with a high.

*Studdert-Kennedy*:
*Haskins Laboratories, New Haven*
*also Graduate Center and Queens College*
*City University of New York*
*Hadding*: *Lund University, Sweden*

## REFERENCES

Cooper, F.S.
1965 "Instrumental Methods for Research in Phonetics", in *Proceedings of the Vth International Congress of Phonetic Sciences*, Münster, 1964 (Basel, Karger) pp. 142-171.
Hadding-Koch, K. and M. Studdert-Kennedy
1963 "A Study of Semantic and Psychophysical Test Responses to Controlled Variations in Fundamental Frequency", *Studia Linguistica* XVII-2:65-76 (Lund).
1964 "An Experimental Study of Some Intonation Contours", *Phonetica* 11:175-185.
1965 "Intonation Contours Evaluated by American and Swedish Test Subjects", in *Proceedings of the Vth International Congress of Phonetic Sciences*, Münster, 1964 (Basel, Karger) pp. 326-331.



## DISCUSSION

JAMES (Toronto)
Did you notice any discrepancy between the actual physical reality of the contours, that is, the direction of the intonation patterns as revealed by instrumental analysis, and the perception revealed by the judgment of the listeners?

STUDDERT-KENNEDY
Examination of the curves will show many examples of non-veridical judgments of the terminal glide.

COLLIER (Belgium)
Do the results of these experiments merely corroborate those of your previous investigation or have they also led to an advanced insight into the perceptual and (or) linguistic relevance of the contours used?

STUDDERT-KENNEDY
The results confirm those of our previous study to a considerable extent. At the same time they permit us to resolve certain ambiguities. We are now able to say that the turning-point effect is primarily auditory, while the peak effect is primarily linguistic.

VANDERSLICE (New York)
I am worried about the effect of forced-choice testing using only the two linguistic response categories 'question' and 'statement'. I would claim that English, at least, has intonations corresponding to each of these, and also a fall-rise contour which can be used for either in the right circumstances, and many of the stimuli you used seem to be excellent realizations of this contour.

STUDDERT-KENNEDY
We certainly do not take the categories 'question' and 'statement' to be exhaustive. We were not attempting to botanize American English and Swedish intonation contours, but to estimate the relative importance of the terminal glide and other sections of the contour in linguistic judgments and, more particularly, to examine the auditory basis for these judgments. The restriction of judgments to two categories was therefore largely one of experimental design intended to increase the reliability of the results. If we recall that each subject made only five judgments of each contour under each experimental condition, it is evident from the smooth, monotonic curves that we did achieve a high degree of inter-subject reliability.