Phonetica
Editor: E. ZWIRNER, Münster in Westfalen
S. KARGER – BASEL/NEW YORK (Printed in Switzerland)
SEPARATUM

Phonetica *11:* 175–185 (1964).

From the Haskins Laboratories, New York, N.Y.

An Experimental Study of Some Intonation Contours

By K. HADDING-KOCH* and M. STUDDERT-KENNEDY**

Introduction

Questions[1] are often said to be distinguished from statements by a terminal rise in fundamental frequency (f_0) as against a terminal fall. However, questions may also be distinguished by a comparatively high f_0 throughout the utterance (*Hermann*, 1942). Spectrographic analyses of Swedish speech have shown that, in this language, questions tend to be spoken on a higher f_0 than statements, usually ending in a moderate rise (*Hadding-Koch*, 1961). If four f_0 levels are postulated in the description, with arrows showing the direction of the terminal glide, the intonation contour[2] of a Swedish question, with a precontour marked before the stress, and the stress being indicated by two figures, could be described as 3 42↑[3] (the superscript 3 indicates the terminal level of the glide), or, if less "interested", as 2 32↑[3]. A typical statement, on the other hand, could be described as 2 31↓ or 2 21↓.

Similarly, a typical American English question is said to display a continuously rising contour that may be notated, for example, 2 23↑ or 2 23↑[4] (*Pike*, 1945; *Bronstein*, 1960). A typical American English statement may be notated exactly as in Swedish, 2 31↓.

* Also at Fonetiska Institutionen, Lund, Sweden.
** Also at Barnard College, Columbia University, New York.

[1] "Question" refers throughout to so-called Yes-No questions.

[2] The acoustic correlates of intonation are said to be changes in one or more of three variables: fundamental frequency, intensity, duration, with fundamental frequency being the strongest single cue (*Bolinger*, 1958; *Denes*, 1959; *Denes* and *Milton-Williams*, 1962). The present study is concerned with only one of these variables, fundamental frequency, and the term "intonation contour" refers to contours of fundamental frequency.

Herrn Professor Eberhard Zwirner zum 65. Geburtstag gewidmet.

However, polite statements in Swedish, though spoken on a lower frequency level than questions, quite often end with a rise. In American English also, terminal rises are reported to occur in statements (*Uldall, 1962*). In fact, *Uldall*, using synthetic speech, demonstrated that an utterance could have quite a large terminal rise and still be heard as a statement, if the rise was preceded by a high fall. If there was no pitch higher than the end point of the terminal rise, the utterance tended to be heard as a question.

These facts concerning both Swedish and American English suggest that not only the direction and range of the terminal glide, but the shape and level of the entire contour affect listeners' judgements (*Gårding* and *Abramson*, 1960; *Hadding-Koch*, 1961). The present experiment was designed to explore this notion in more detail by means of synthetic intonation contours, and to compare for Swedish and American listeners their preferred question and statement contours. In addition, as a partial check on the degree to which listeners could actually hear the detailed tonal movements involved, some purely psychophysical data were collected.

Method

The utterance *För Jane* [foe 'Jein] = *for Jane*, spoken on a monotone and in such a way as to be acceptable as Swedish to Swedes, as American to Americans was recorded on magnetic tape. From this recording forty-two different fundamental frequency contours, simulating Swedish intonation, were prepared by a procedure described below. The f_0 values were based on detailed spectrographic analyses of a long sample of the Swedish speaker's natural speech. The correspondences between level notation and fundamental frequency derived from this analysis are given in table I. As poles, a Swedish question contour, 2 42↑, and statement contour, 2 31↓, were used. In the present experiment the first number represents the level of the precontour on *För*, the second number the level of the intonation "peak", the third number the level of the "turning point" (Gårding, 1960) before the terminal glide. Between the poles of ideal question and statement, various f_0 values at peak, turning point and end point were introduced.

Table I

Correspondences between Level Notation and Fundamental Frequency in Cycles per Second (from *Hadding-Koch*, 1961)

Level	Fundamental frequency in cycles per second
4	370 and above
3	260–370
2	175–260
1	175 and below

Diagrams of the contours are reproduced above figures 1 and 2. All contours started at a fundamental frequency of 250 cps, sustained for 140 msec over *För.* They then rose to a peak of either 370 cps (the *S*, or superhigh series of contours) or 310 cps (the *H*, or high, series), dropped to one of three turning points: 130 cps (*S1* and *H1* series), 175 cps (*S2* and *H2* series), or 220 cps (*S3* and *H3* series), and then proceeded to one of seven end points between 130 cps and 370 cps. The rise and fall on either side of the peak lasted for 300 msec, the terminal rise or fall, from turning point to end point, lasted 200 msec. The actual contours were rounded at peak and turning point rather than pointed as in the schematic contours above the figures.

The intonation was varied by means of the Intonator connected with the Vocoder at Haskins Laboratories, New York. The Vocoder first analyzes a speech sample in a bank of filters and then reconstitutes it in a simplified form on the basis of information obtained from the analysis (*Dudley*, 1939; *Borst* and *Cooper*, 1957). The fundamental frequency of the output is controlled by the Intonator, and may be varied independently of other characteristics of the speech sample. Thus, the same utterance may be given any desired number of different fundamental frequency patterns. Instructions to the Intonator are transmitted through photoelectric tubes responding to light reflected from a contour painted on an acetate loop. Also attached to the loop is a strip of adhesive magnetic tape bearing the speech sample to be processed. The loop is reeled past the photoelectric tubes of the Intonator and the magnetic input heads of the Vocoder. The outputs of the synthesizer in the present experiment were the forty-two stimuli previously described. These were recorded on magnetic tape and spliced into five different random orders, with a 5-second interval between stimuli and a 10-second pause after every tenth stimulus. They were presented to 25 Swedish and 24 American undergraduates in two counterbalanced sessions. In each session all five test orders were presented with a short pause between orders. In one session subjects were instructed to indicate for each stimulus whether it would be better characterized as a statement or a question (semantic judgement). In another session subjects were instructed to indicate for each stimulus whether it ended with a rising or a falling pitch (psychophysical judgement). Approximately half the subjects in each group (Swedish and American) made their psychophysical judgements first.

Results

In the semantic test, responses varied as a function of the fundamental frequency at all three of the variable points of the contours: peak, turning point, and end point.

Figure 1 presents the semantic data for the Swedish subjects (above) and American subjects (below) on the *S* series of contours (peak at 370 cps). Against the ordinate are plotted the percentages of question and statement responses. Against the abscissa are plotted the values of the terminal rise or fall in cycles per second of fundamental frequency (end point minus turning point): a negative value indicates a terminal fall, a positive value a terminal rise. Parameters of the curves are peak f_0 (370 cps) and turning point f_0 (130 cps for *S1*, 175 cps for *S2*, 220 cps for *S3*).

The effect of the terminal rise or fall is immediately obvious and very much as expected: for all three series the higher the terminal

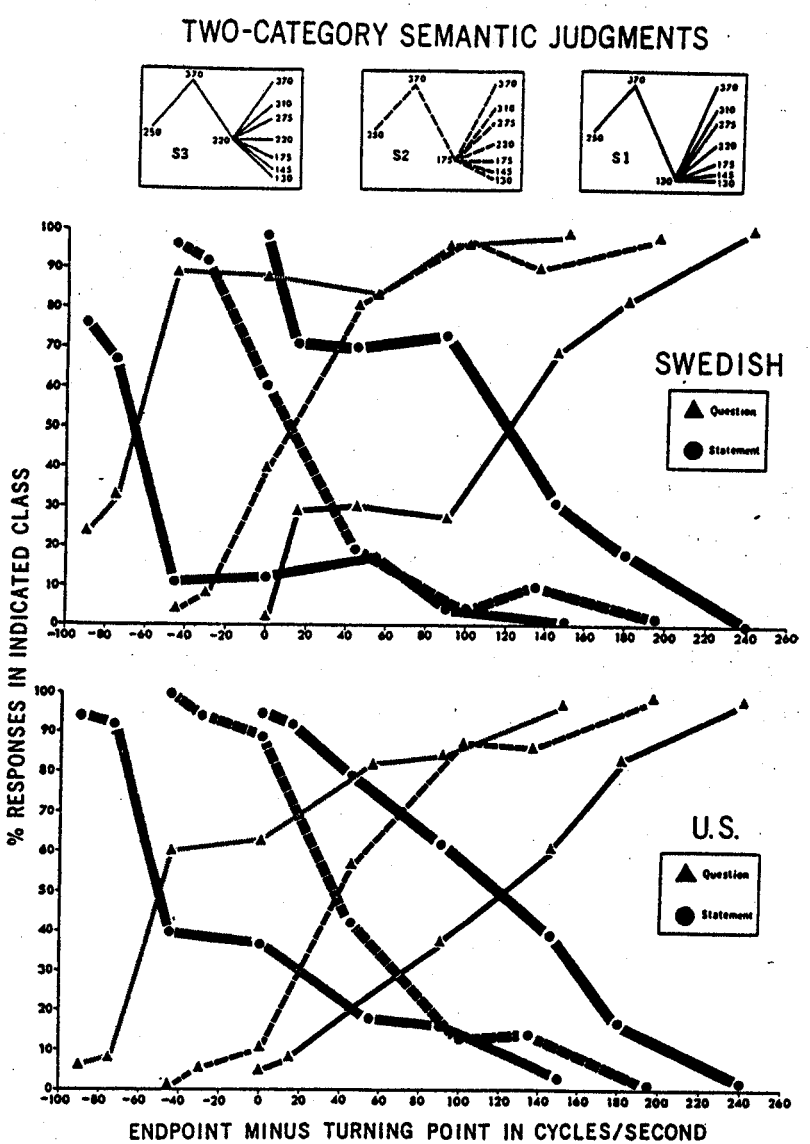

Fig. 1. Peak f_0 at 370 cps: percentage of statement and question responses as a function of the terminal rise (positive) or fall (negative) in cycles/second of f_0 (endpoint f_0 minus turning point f_0). Parameters of the curves are turning point f_0: 130 cps (*S1*), 175 cps (*S2*) and 220 cps (*S3*). The Swedish data are plotted above, the U.S. data below.

rise, the higher the percentage of question responses. Equally obvious is the effect of the fundamental frequency at the turning point.

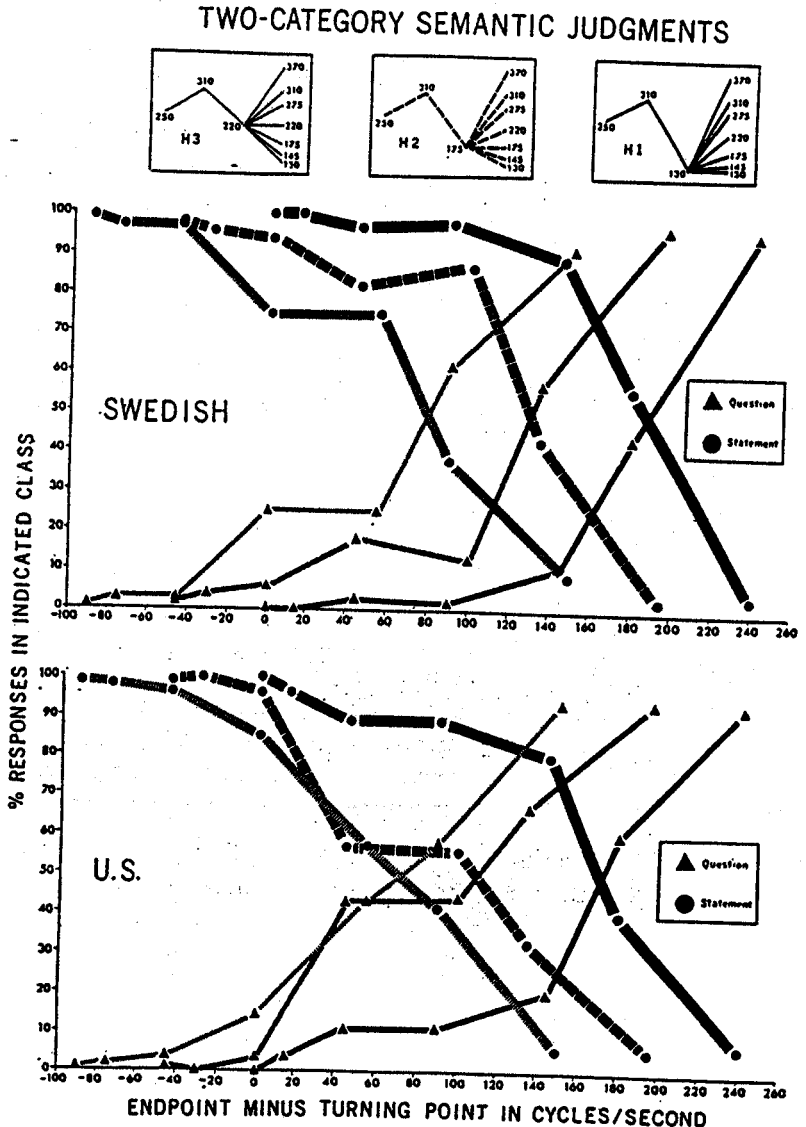

Fig. 2. Peak f_0 at 310 cps: percentage of statement and question responses as a function of the terminal rise (positive) or fall (negative) in cycles/second of f_0 (endpoint f_0 minus turning point f_0). Parameters of the curves are turning point f_0: 130 cps (*S1*), 175 cps (*S2*) and 220 cps (*S3*). The Swedish data are plotted above, the U.S. data below.

For purposes of comparison we may consider the so-called points of subjective equality, that is, the indifference points at which subjects' responses cross over from predominantly statements to predominant-

ly questions. For the Swedish subjects we find the crossover in the *S1* series at a final rise of 120 cycles, in the *S2* series at a final rise of 12 cycles, and in the *S3* series at a final fall of 65 cycles. Thus, the f_0 value of the turning point may quite override the effect of the terminal rise or fall. For example, a terminal fall of 45 cycles is heard as a statement 96 % of the time when the turning point is at 175 cps, but as a question 89 % of the time when the turning point is at 220 cps. Similar effects are present in the American data. But the Americans display some preference for statements over questions. As compared with the Swedes they require somewhat smaller terminal falls to be sure they hear statements, somewhat larger terminal rises to be sure they hear questions.

In the *H* series (peak f_0: 310 cps) the number of questions heard again increases with the f_0 value at the turning point, although less markedly. Figure 2 presents the data for the Swedish subjects (above) and the American subjects (below). Here the groups differ little in their question curves. But the Swedes display a preference for statements, particularly in the *H2* and *H3* series.

If the *S* and *H* series are compared (figs. 1 and 2), it appears that more questions are heard in the *S* series and more statements in the *H* series. In other words, a rise in f_0 at the peak – exactly as a rise in f_0 at the turning point – is accompanied by an increase in the number of questions heard. Figure 3 facilitates this comparison by displaying the Swedish data for the *S2* and *H3* series on the same axes. For example, a stimulus with a final rise of 100 cycles is heard as a question 96 % of the time when the peak is at 370 cps (*S*), but as a statement 89 % of the time when the peak is at 310 cps (*H*). In other words, the lowered peak overrides the effects of a substantial final rise, and induces a virtual reversal of the response distributions.

Turning, finally, to the results of the psychophysical tests, in which subjects were asked to indicate whether the contours ended with a rising or a falling pitch, we find greater subject uncertainty but the same general effects as have been described for the semantic tests. Figure 4 compares the American psychophysical and semantic data for the *H1* series of contours: the two sets of data are nearly identical. Contours identified as statements tend to be heard as having a terminal fall (even when, in fact, the final contour is rising), while contours identified as questions tend to be heard as having a terminal rise. On other series, the agreement is not always so marked. For example, figure 5 displays the American psychophysical and

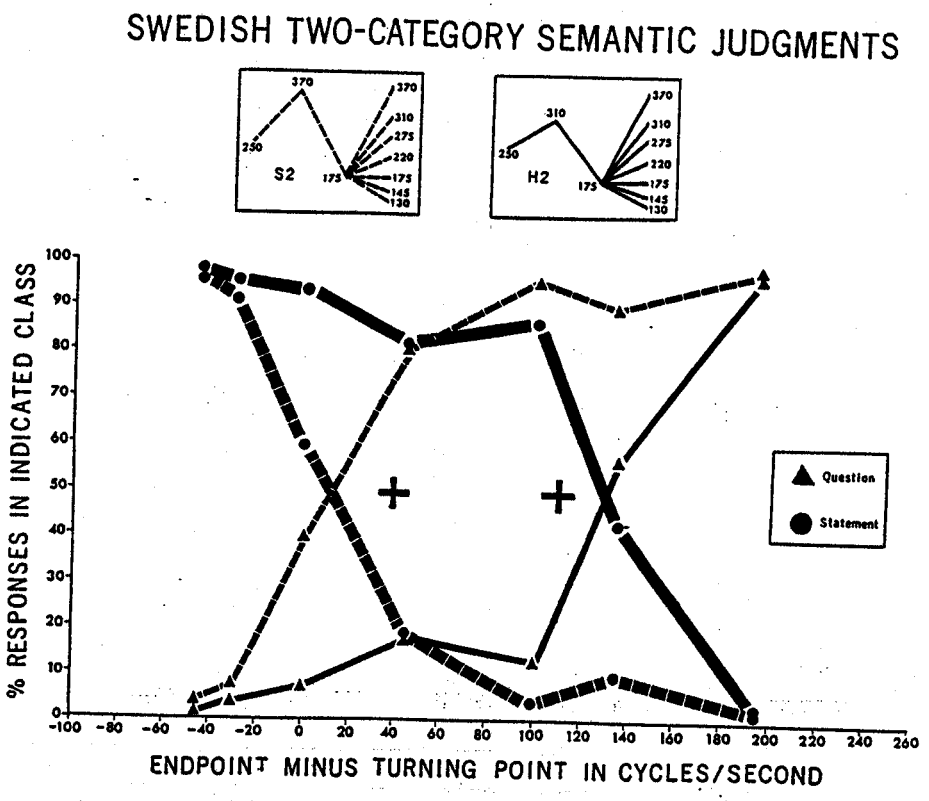

Fig. 3. Percentage of statement and question responses as a function of the terminal rise (positive) or fall (negative) in cycles/second of f_0 for Swedish subjects. The curves for two values of peak f_0 [370 cps (*S*) and 310 cps (*H*)], with turning point f_0 constant at 175 cps, are compared. The crosses indicate the points of subjective equality for the U.S. subjects in the *S2* (left) and *H2* (right) series.

semantic data from the *S3* series. Here, as is generally true, the psychophysical judgements are more uncertain than the semantic – particularly for the contours displaying terminal falls. Nonetheless, there is still remarkable agreement between the two sets of curves.

Discussion

The results confirm what naturalistic observation and some previous experiments have already suggested: that listeners may make use of the entire f_0 contour in identifying questions and statements. Not only terminal rise or fall, but also preceding peak and turning point are relevant. These three variables interact in a manner that

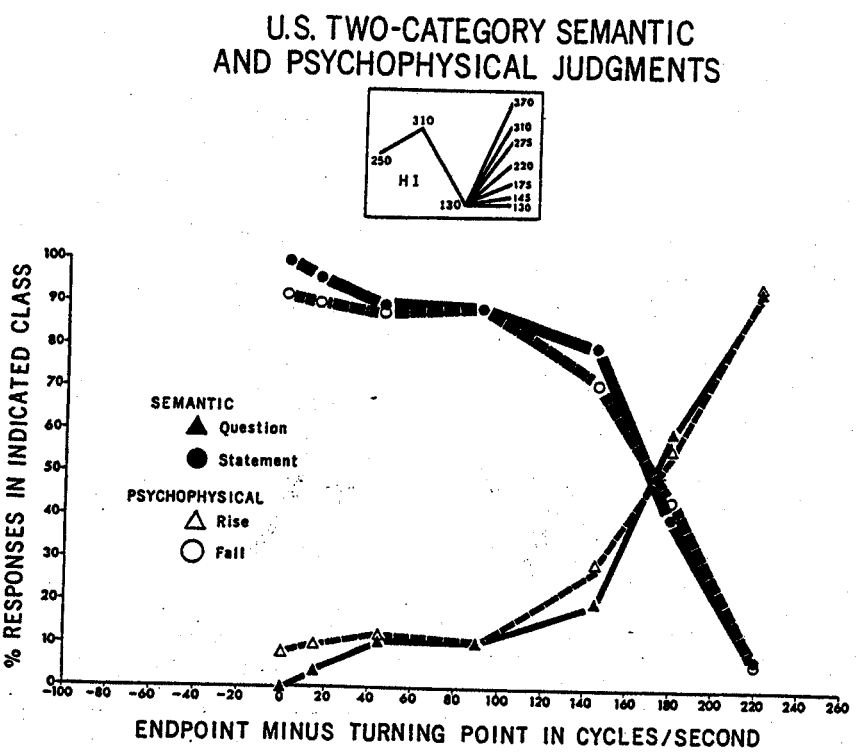

Fig. 4. Percentage of statement and question responses (semantic: solid line) and of rise and fall responses (psychophysical: hatched line) as a function of terminal rise (positive) or fall (negative) in cycles/second of f_0. Data from U.S. subjects on the *H1* stimulus series.

cannot be easily described. But, in general, for a given f_0 at the other two points, an increase in f_0 at the third point leads to an increase in the number of questions heard.

As to the actual f_0 values preferred by listeners for the question and statement contours, the Swedish results agree with the predictions based on spectrographic analysis. The contours yielding the greatest proportion of question responses (*S3*) had a peak at 370 cps and a turning point at 220 cps. This contour yielded a high proportion of question responses even when there was a moderate terminal fall. But the preferred question contour was 2 42 plus a final rise. Similarly, the contours yielding the greatest proportion of statements (*H1*) had a peak at 310 cps and a turning point at 130 cps. Here, a large proportion of statement responses occurred even when there was a considerable terminal rise (cf. Introduction). But

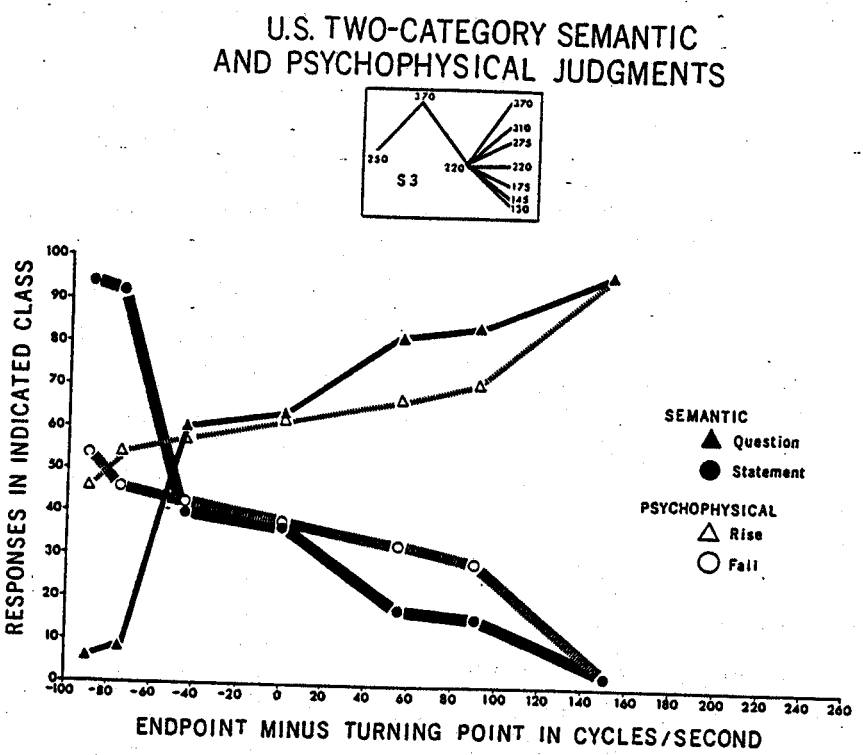

Fig. 5. Percentage of statement and question responses (semantic: solid line) and of rise and fall responses (psychophysical: hatched line) as a function of terminal rise (positive) or fall (negative) in cycles/second of f_0. Data from U.S. subjects on the *S3* stimulus series.

the preferred statement contour was 2 31. Since the turning point f_0 (130 cps) was the lowest f_0 in the experiment, no final fall could occur with the *H1* contour.

As was stated earlier, Swedish and American English are said to have similar typical statement contours, but different typical question contours. As to questions, the data of the present experiment do not contradict this. For, although both groups selected a typical Swedish question (*S3*, 2 42↑) as their preferred question contour, the Americans did require a higher terminal rise to reach complete agreement on their question responses than the Swedes: lacking the typical continuously rising question of American English (2 23↓), they gave more weight to the terminal glides than did the Swedes. However, they also gave more weight to the terminal glides in the preferred statement series (*H1*). This suggests that the two

groups may differ in their preferred statement as well as in their preferred question contours. Further experiments designed to examine in more detail the differences between the two groups are in progress.

Finally, the psychophysical data perhaps throw some light on the process by which the f_0 values at peak and turning point exert their influence on listeners' semantic judgements. These data show that listeners were unable to follow the terminal glide with anything like the precision that might have been predicted from simple pure tone pitch discrimination (*Stevens* and *Davis*, 1938): psychophysical judgements were influenced by peak and turning point f_0 very much as semantic judgements. Insofar as semantic and psychophysical judgements agree (as in fig. 4), it would seem that listeners may have been using the *perceived* direction of the terminal glide rather than its physically measured direction to make their semantic decisions. The role of peak and turning point f_0 would then seem to lie in their effect on the perception of the terminal glide. On the other hand, insofar as the psychophysical data display greater uncertainty than the semantic (as in fig. 5), the peak and turning point f_0 values would seem to exert an independent influence on the semantic judgement, presumably in some weighted combination with the perceived terminal glide. Further experiments, including a study of the perception of non-speech control signals, have been designed to study these relations more fully and are now in progress.

Acknowledgement. We should like to acknowledge that the research reported here was supported in part by a grant from the National Science Foundation, Washington, D.C.

Summary

The intonation contour of a single utterance was systematically varied and presented for judgement to Swedish and American subjects. In counterbalanced sessions, subjects were asked to classify the contours as (1) questions or statements (semantic judgement), (2) having terminal rises or falls (psychophysical judgement). The data from the two groups of subjects and from the two types of judgements are described and compared.

Eine experimentelle Untersuchung von Intonationskonturen

Zusammenfassung

Die Intonationskontur einer einzelnen Äußerung wurde systematisch verändert und schwedischen und amerikanischen Versuchspersonen zur Beurteilung dargeboten.

In zwei sich das Gegengewicht haltenden Sitzungen sollten die Versuchspersonen entscheiden, ob die Konturen 1. Fragen oder Feststellungen wären – semantische Beurteilung –, und ob sie 2. am Ende steigen oder fallen – psychophysische Beurteilung. Die so gewonnenen Daten werden beschrieben und verglichen.

Etude expérimentale des contours d'intonation

Résumé

Le schéma intonatif d'un énoncé a été transformé de façon systématique et présenté pour appréciation à des sujets suédois et américains. Dans des séances de présentation organisées de façon homogène, les sujets devaient décider si les contours qui leurs étaient proposés étaient:

1° interrogatifs ou énonciatifs – appréciation sémantique – et

2° montants ou descendants – appréciation psychophysique.

Les données ainsi recueillies ont été soigneusement décrites et comparées les unes aux autres.

References

Bolinger, D. L.: A theory of pitch accent in English. Word *14:* 109–149 (1958).

Borst, J. M. and *Cooper, F. S.:* Speech research devices based on a channel vocoder (abstract). J. acoust. Soc. Amer. *29:* 777 (1957).

Bronstein, A. J.: The pronunciation of American English: An introduction to phonetics (Appleton-Century. Crofts, Inc., New York 1960).

Denes, P.: A preliminary investigation of certain aspects of intonation. Language and Speech *2:* 106–122 (1959).

Denes, P. and *Milton-Williams, J.:* Further studies in intonation. Language and Speech *5:* 1–14 (1962).

Dudley, H.: Remaking speech. J. acoust. Soc. Amer. *11:* 169–175 (1939).

Gårding, E.: A study of the perception of some American English intonation contours, Paper read before 75th Meeting Mod. Lang. Ass. Amer., Philadelphia, 28 December (1960).

Gårding, E. and *Abramson, A. S.:* A study of the perception of some American English intonation contours, Haskins Labs. Quarterly Progress Report *34,* New York, June (1960).

Hadding-Koch, K.: Acoustico-phonetic studies in the intonation of Southern Swedish (Gleerups, Lund 1961).

Hermann, E.: Probleme der Frage. Nachrichten von der Akademie der Wissenschaften in Göttingen *3–4:* 1942.

Pike, K. L.: The intonation of American English (Univ. of Michigan, Ann Arbor 1945).

Stevens, S. S. and *Davis, H.:* Hearing, its psychology and physiology (John Wiley & Sons, New York 1938).

Uldall, E. T.: Attitudinal meanings conveyed by intonation contours. Language and Speech *3:* 223–234 (1960). – Ambiguity: question or statement? or "Are you asking me or telling me?". Proc. IV Int. Congr. Phon. Sciences; pp. 779–783 (Mouton & Co., The Hague 1962).

Authors' addresses: Kerstin Hadding-Koch, Fonetiska institutionen, Kävlingevägen 20, *Lund* (Sweden) and Michael Studdert-Kennedy, Haskins Laboratories, 305 E. 43rd St., *New York,* N.Y. 10017 (USA)