

RISING INTONATION: NOT PASSED AWAY BUT STILL ALIVE

A reply to R. Geluykens

Anton BATLINER and Wilhelm OPPENRIEDER *

Received March 1988

In this reply to Geluyken's article 'Intonation and Speech Act Type' (this journal, Volume 11, 1987: 483–494) we want to argue that G.'s article suffers from methodological shortcomings and insufficient description of the data. Moreover, his conclusion that the role of rising intonation for the recognition of so-called 'Queclaratives' is overrated, is not backed up by his data. We try to find alternative explanations and outline the relevancy of rising intonation as one amongst other illocutionary force indicating devices.

1. Introduction

In his article '*Intonation and speech act type: An experimental approach to rising intonation in queclaratives*', (this journal, Volume 11, 1987: 483–494), Ronald Geluykens (henceforth: G.) concludes that "(...) the claim made in the literature that rising intonation is a deciding factor for the recognition of queclaratives [i.e., "utterances having the form of a declarative sentence but functioning as requests for information", p. 483] lacks any empirical support" (p. 491). Instead, he stresses the importance of pragmatic-contextual cues. In our reply, we want to argue that (1) important information about G's data is missing, (2) his conclusions are not backed up by his data, and (3) alternative explanations can be given. Finally, we want to show that the relevancy of rising intonation as an illocutionary force indicating device should neither be overrated nor underestimated. (From now on, we will speak of rising intonation as a 'cue' for sentence modality, instead of 'phonetic illocutionary force indicating device'.)

* Authors' address: A. Batliner and W. Oppenrieder, Institut für Deutsche Philologie, Universität München, Schellingstr. 3, D-8000 München 40, FRG.

2. Critique

2.1. Statistical approach

Any level of significance must be chosen in advance and is valid only as long as a single statistical test is applied to the data. Otherwise the α -error (error of the first kind, i.e. unjustified rejection of the null-hypothesis) will increase. It is not clear how many statistical tests G. carried out on his data (cf. p. 489 f.), but it must be some two-digit number. Strictly speaking, this approach jeopardizes his conclusions concerning the significance of his results, but of course his data can still be discussed on a purely descriptive basis.¹

2.2. Description of data

The description of G.'s experimental material is vexedly poor. Contours are represented schematically, but no exact information as to the time and frequency axis is given. G. refers globally to De Pijper (1983) "for more details" (p. 485). For the time axis, the position of the turning points in relation to the tonic is given, but the overall duration of the respective utterances and of the final rises can only be guessed at; cf. p. 486: "Total length of the utterances varied from 1660 up to 2020 msec." Without these details, we are at a loss to compute the slope of the rising contours; nevertheless, the slope can be an important factor that should be accounted for. Even worse: the only information given explicitly for the frequency axis is that on contour 1, (p. 486): "(...) a slight quarter tone dip just before the tonic (...)". If we assume that the vertical display of the contours on p. 487 is roughly proportional to a semitone scale, then the whole pitch range should be two semitones, as this quarter tone covers 1/4 of the whole range (Possibility 1). Of course this is a rather small range; cf. below. Another possibility could be that 'quarter tone' should read 'quarter range', cf. De Pijper (1983: 70); there the standard range of tone 1 is 9.6 semitones, of tone 2 14.4 etc. (Possibility 2 – De Pijper (1983: 70 f.)). Such ranges would make more sense (cf. below) but we simply don't know which information to rely on in our guesswork.

The stimuli used were schematically resynthesized (in analogy to Halliday's (1970) tone 1 to 4). This procedure can yield unnatural stimuli. Therefore, G.'s subjects "(...) were asked to mark a sentence as 'odd' if it sounded unnatural to them in some way; [since some utterances] received a markedly

¹ G.'s statistics could possibly be redeemed if we set the level of significance at 0.05 throughout, not at 0.001, as he does (p. 489), and if we further assume that he has not carried out more than 50 different tests on his data. In that case an adjusted level of significance of 0.001 can be computed using the formula $1-(1-\alpha)^{1/c}$ (c being the number of tests, in our case 50). But such a subsequent rescue operation is not exactly the right way of doing statistics.

high number of 'odd' judgments, they were excluded from our results" (p. 488). The question of naturalness might be a pivotal point, cf. below, but again, no exact information is given. What does 'markedly high' mean? And what about the other items? Did they receive a 'markedly low' number of 'odd' judgments, or just any number that was not 'markedly high'?

All the missing information should, and could have been given on just a few lines – even if the prospective readers of the *Journal of Pragmatics* might not be as interested in 'phonetic details' as the readers of e.g. the *Journal of Phonetics*.

2.3. Interpretation of results

G. uses three kinds of utterances: 'Y-utterances' like *You feel ill* are labeled 'question-prone': "(...) the verb 'feel' refers to the addressee's mental state, to which the speaker has no direct access (...) The utterance is therefore likely to be interpreted as a genuine request for information, despite its declarative form (...)" (p. 484). On the other hand, an 'I-utterance' like *I feel ill* is 'statement-prone': "Since the verb 'feel' in this instance refers to the mental state of the speaker rather than of the addressee, (...) it is very unlikely that someone would request information concerning his/her own mental state" (p. 485). An 'H-utterance' like *He is ill* is supposed to be rather neutral in this respect.

G. starts from the following basic hypothesis: "Cues of a pragmatic-contextual nature are the determining factor in recognizing an utterance with declarative form as a queclarative; the use of a rising intonation contour does not necessarily turn a declarative utterance into a question" (p. 484). In the experiment, Y-, I- and H-utterances with five different intonation contours (one of them being a fall, four of them being rises) had to be classified as questions or statements (p. 486 ff). The results seem to confirm the hypothesis: There is no marked difference between the five contours used (with one exception, cf. p. 490). Instead, there is a marked difference (we can't say 'significant', cf. above) between the question-prone Y-utterances on the one hand and the I- and H-utterances on the other. The former are much more often classified as questions, although not as often as G. expected (N.B.: 53%, i.e. chance level). Nonetheless, G. concludes: "The intonation contour appears to be of no importance for recognizing these utterances as questions" (p. 491). G. doesn't seem to recognize that the very fact – which he comments on himself, p. 491 – that even these Y-utterances with their clear-cut question-prone status were *not* recognized as questions, renders this general conclusion implausible. Such a conclusion would hold if the question-prone Y-utterances had got 'markedly high' question-responses, let's say 90% or more. As it stands, the only conclusion can be: We don't know whether G. used intonation contours that were appropriate for queclaratives, but as for

the contours used, a marked difference can be observed between Y-utterances on the one hand and I- and H-utterances on the other; this difference may be traced back to non-phonetic, pragmatic-contextual factors. G. states the right facts, but draws the wrong conclusion: "It has been shown that, *at least for the rising contours tested here*, the role of rising intonation in queclaratives is overestimated. Rising intonation does not contribute significantly to the recognition of queclaratives" (p. 493; our emphasis). This generalization from the sample to the population is simply a non sequitur. It is a misconception of rising contour as a binary all-or-none feature: not every rising F_0 -contour is a linguistically relevant rising intonation!

2.4. Coverage of phonetic literature

In his introduction, G. gives an account of the experimental studies on the issue; however, he doesn't quote such authors as, e.g., Studdert-Kennedy and Hadding (1973). These authors tested a variety of different intonation contours on the one word utterance *November* with Swedish and American English-speaking subjects. Although they found complex trading relations between the different parts of the contours, they conclude: "Terminal glide is the single most powerful determinant of linguistic judgments. None of the highly preferred question contours and few of the highly preferred statement contours (...) lack the appropriate terminal rise or fall" (Studdert-Kennedy and Hadding (1973: 305)). ('Highly preferred' means that at least 90% of the judgments were in a single category.) In *November* there is of course no structural, lexical, or pragmatic information about sentence modality whatsoever, and the subjects had therefore to rely entirely on the intonation contour, which obviously is the dominant factor. This result makes it appear even more mysterious that G.'s Y-utterances with their pragmatically-contextually defined question-proneness *and* a rising intonation should not yield clear-cut question-responses; this discrepancy should therefore have been commented on by G. (Note that only 15% of the Y-utterances with rising intonation have been judged as 'definite questions', all the others being classified as either 'more question than statement' (38%), 'more statement than question' (29%), or 'definite statement' (15%), cf. p. 488.) A reasonable conclusion would therefore run in the opposite direction of G.'s. An appropriate rising contour can be highly relevant; however, G.'s contours were simply not appropriate enough. Of course, we don't want to exclude the relevancy of other prosodic factors as "(...) pitch range, pauses, loudness, and tempo (...)" (p. 491). But as far as has been shown in the literature, these factors can perhaps, to a certain extent, alter the role of a rising intonation but never override it.

3. Alternative explanations

Anybody who has ever tried to (re-)synthesize F_0 -contours will agree that strange things can happen in such an endeavor. As already mentioned, the naturalness tests, as carried out by De Fijper (1983) and G., can possibly help decide whether or not the test utterances conform to reasonable standards of English, but they cannot decide whether it is an acceptable English question in general, or queclarative (out of context) in particular. An appropriate test should ask: "Is this utterance a good candidate for a queclarative?" Or one should combine the utterances with a context that strongly favors a queclarative interpretation and ask: "How well do utterance and context go together?" (As to this kind of test, cf. Oppenrieder (1988).) A plausible explanation of G.'s result is therefore that his contours were simply unable to indicate a good queclarative. It is therefore very likely that such contours lose some or all of their weight as speech act indicating devices. For the subjects, the other indicating devices, such as the lexical material and the position of the finite verb, are then more reliable. Their judgments will therefore reflect the statement-proneness of the grammatical form of the sentences ('second' position of the finite verb) – this tendency weakened sometimes by the use of the subject pronoun *you* –, whereas the role of an inadequate intonation contour would be reduced, as compared to its role in the interpretation of normal, non-resynthesized and non-stylized speech.

If possibility 1, above, applies (range of two semitones), then the rising terminal contour is simply too small and we haven't got questions but 'tentative statements' (cf. G.'s own explanation on p. 492). However, if possibility 2 obtains (range of 9.6–14.4 semitones), then the rising terminal contours should be prominent enough, i.e. something else must have gone wrong. What exactly, cannot be decided without close inspection of the material, the experimental design, etc.²

To sum up, the only sound proof of G.'s claim that rising terminal contours are not a deciding factor would be provided by an experiment in which Y-utterances got about 90% or more question-responses *and* I- and H-utterances with the same F_0 -contours at the same time got very few question-responses. In a further experiment, these utterances should then come out as naturally sounding queclaratives, not only as naturally sounding English utterances in general. (Our alternative claim is that with an appropriate F_0 -contour that clearly indicates queclaratives, both Y-utterances and I-/H-utterances would get roughly the same question-responses. But this is a matter of empirical verification or rejection.)

² Maybe the steep final rises are too short; as far as we can guess, they have a duration of ca. 80 msec – comparable rises in De Fijper's (1970: 140 ff.) corpus have a duration of ca. 200 msec. In this case, a rise of ca. 12 semitones might be perceived mainly not as a linguistically relevant intonation contour, but rather as a final 'yodeling' disturbance.

4. Some further remarks on the role of rising intonation

We do not want to give the impression that in our opinion, a terminal rising contour is necessarily the only cue for questions. It is one amongst other cues (phonetic, syntactic, contextual, lexical, and pragmatic), but it is very likely the most important *phonetic* cue. In normal (face-to-face) communication, a plethora of possible cues are available, and we fully agree with G. that in such a situation, it is not settled yet whether queclaratives with rising terminal contour are more frequent than those with falling terminal contour. But in a more restricted communication situation, not all different kinds of cues are available. In this sense, experiments represent an even more restricted situation than do, e.g., phone conversations: Only phonetic information, no contextual information whatsoever is available. In such a situation, an appropriate terminal rise should be the decisive cue. Still, we might be in agreement with G. if we say that the results of a perception experiment cannot simply be extrapolated to real life situations; hence it is not surprising that G. found *only* 33% utterances with rising intonation in a corpus of 60 queclaratives (p. 483). But suppose now we can turn the tables: If G. had found *up to* 33% utterances with rising intonation in a production corpus, then this could be taken as a strong argument for the claim that rising intonation as a cue for queclaratives is not a mere experimental artifact or a phonetician's pipe dream.³

We can, therefore, with some slight (but decisive) modifications subscribe to G.'s conclusion in his abstract on p. 483: "Provided pragmatic cues are sufficiently strong to determine speech act status, rising intonation *is shown to be* [should read: *might be*] virtually without impact; if, on the other hand, pragmatic cues do not favor any particular speech act type, intonation *may, but need not, act as a* [should read: *can and possibly must be the only*] cue for determining question-status" (our emphasis).

References

- De Pijper, Jan Roelof, 1983. Modelling British English intonation. Dordrecht: Foris.
Halliday, M.A.K., 1970. A course in spoken English: Intonation. Oxford: Oxford University Press.

³ Following G. there seems to be little reason for speakers of English to use a redundant rising contour if they can get along just as well, and with less articulatory effort, by using a falling contour. In Finnish e.g. – a language which does not use rising intonation in questions, nor for other grammatical purposes – questions are always signalled morphologically, but this is not the case in English (and German, Swedish, Danish etc.). G. seems to doubt whether intonation should have any linguistically relevant function at all; cf. pp. 484, 491. However, it would not make much sense if such a strong and prominent formal contrast were not used by English speakers to indicate functional differences – at least in situations where other cues are not available.

- Oppenrieder, Wilhelm, 1988. Intonatorische Kennzeichnung von Satzmodi. In: Hans Altmann, ed., *Intonationsforschungen*. Tübingen, Niemeyer. pp. 169–205.
- Studdert-Kennedy, Michael and Kerstin Hadding, 1973. Auditory and linguistic processes in the perception of intonation contours. *Language and Speech* 16: 293–313.