[1]Note two errors on p. 651: "sentences, 2, 12, and 15" should read "sentences 2, 11, and 14"; and the table captions, which are identical, should contain "naive subjects" only for Table II and "experimenters" only for Table III.

[2]The comments in this paragraph and the parenthetical comment in the preceding paragraph were not contained in the original reviews. Note that the first lines in the formulas for $m$ and $r(x,y)$ should be enclosed in square brackets.

Lieberman, P. (1967). *Intonation, Perception, and Language* (MIT, Cambridge, MA).

Lieberman, P., Katz, W., Jongman, A., Zimmerman, R., and Miller, M. (1985). "Measures of the sentence intonation of read and spontaneous speech in American English," J. Acoust. Soc. Am. 77, 649–657.

Maeda, S. (1976). "A characterization of American English intonation," Ph.D. thesis, MIT (unpublished).

# Reply to Bruno H. Repp

Philip Lieberman, William Katz, Allard Jongman, Roger Zimmerman,[a] and Mark Miller
*Brown University, Providence, Rhode Island 02912*

(Received 1 July 1985; accepted for publication 3 July 1985)

The reader will readily note that much of Repp's criticism [J. Acoust. Soc. Am. 78, 1114–1116 (1985)] reflects his contention that there has been a procedural transgression on the part of the authors and the Society that establishes a "dangerous precedent." Since Repp provides a partial account we feel that it is first necessary to address this issue.

PACS numbers: 43.70.Fq, 43.71.Es

## I. PROCEDURAL ISSUES

In his abstract and introduction Repp suggests that a conventional review process was initiated, and expresses surprise that the Society found it necessary to authorize an independent review (Repp, 1985). In Repp's reply to our first submission, he informed us that based on the reactions of three reviewers, of whom one suggested outright rejection, he would not consider the paper for publication. In the next paragraph of his letter he stated,

> "Since I am just about to leave for the ASA meeting in Norfolk, I have decided not to delay my response any further and thus not to read the paper."

In other words, reviewers had been selected and their comments had been accepted without the former Associate Editor's ever having read the paper. It is Repp who was attempting to introduce a "dangerous precedent" in the peer review process. A minimal requirement for a responsible, informed editor is that he/she read the manuscript. Not having done so, it is puzzling that Repp felt competent to select reviewers, much less make a decision without considering the comments of the reviewers in relation to the paper's content. On further inquiry, we were informed that the three reviewers were proponents of the viewpoints that we were calling into question in our paper. The three reviewers furthermore did not, as Repp insinuates, independently and unanimously note the supposed flaws that Repp cites. This unanimity follows from Repp's correspondence with the reviewers he selected, after we responded to their initial comments. These factors led to the Society's eventual decision to override Repp's decision and submit the article to an independent review process which resulted in the paper's being accepted.

[a]Now at Voice Processing, Cambridge, MA.

## II. CONTENT-RELATED ISSUES

### A. Experiment 1: Testing the visual abstraction procedure

Repp states that by, "failing to distinguish between inter- and intrasubject consistency" we do not test the type of "consistent" response with which an "experienced researcher" might apply the visual abstraction procedure criteria. Repp misses the point of this experiment entirely; he also appears to be unaware of one of the basic principles on which the scientific method rests—replication. The visual abstraction method, if it is to be considered an objective metric, must demonstrate a low degree of intersubject variability (otherwise different researchers employing this technique would end up with disparate measurements of $F_0$ linearity, slope, etc. for the same data). It is possible that intrasubject variation might be significantly lower than the intersubject variation. However, this would be irrelevant to the basic question. Without the possibility of replication, objective inquiry in a scientific mode is impossible. The data of experiment 1 show that different people fitting straight lines to $F_0$ contours using the visual abstraction procedure, fit lines that have very different slopes. Examination of the slopes in Tables II and III shows that the visual abstraction procedure does not generate replicable data. The mean declinations noted by Maeda (1976) for his three speakers were 21.0, 28.4, and 32.2 Hz. The standard deviations noted in these tables are comparable to these values.

It is unclear what Repp is attempting to prove when he states that, "these individuals (the authors) are largely inexperienced, with the exception of the senior author" and that, "no indication of the nature and amount of training of the four co-authors is given." We remind Repp that the task here involves fitting a straight line to a series of points; it is most

unlikely that anyone needs years of speech research in order to accomplish this task. Similarly, the inclusion of a control group was *not* "to deny the role of skill and experience in speech research," but to insure that the researchers did not unknowingly respond in a manner that would produce high variability. The data show that the researchers were as good as or better than the control group—we did not bias our data.

In this regard Repp asserts, based on his understanding of Fig. 3, the "possibility of deliberate violation of rules." The problem is that the criteria are so broad and the relationship between an actual $F_0$ contour and the hypothetical declination line so tenuous, that different people will draw different lines. Repp's insinuation concerning "deliberate violation of the rules" is offensive and is, in fact, not borne out by Fig. 3. Readers of our paper will note that the seven naive subjects whose lines are shown in the bottom graph in Fig. 3 also vary in the way that they "apply" the rule. The problem rests in the inherent nature of the subjective procedure that has been the basis for many declination studies—looking at an $F_0$ plot and drawing a straight line that attempts to connect "valleys" that supposedly can be connected by a line. If proponents of the declination theory wish to have their data accepted it is incumbent on them to demonstrate that their basic data—lines visually fit to $F_0$ contours—can be replicated. No such test has ever been reported by the proponents of declination theories.

### B. Experiment 2: Quantitative least-squares-fit all-points line.

Repp would be well advised to consult the introductory statistics texts which he refers to in his gratuitous comment concerning our "general unfamiliarity with statistical procedures." The reader will note that Repp confuses the "number of points" that lines were fit to, with "degrees of freedom" used in establishing significance for a Pearson correlation value. Repp's questions concerning how "lines were fit to single data points, and how fits to two points could yield less than perfect correlation" demonstrate that Repp is unfamiliar with this statistical procedure. He does not realize that the degree of freedom, *df*, in Table IV is equal to the number of data points minus 2, i.e., $n - 2$. The lines that Repp believes were fit to "single data points" were fit to three, those that Repp believes were fit to "two points" were fit to four.

Repp also does not appear to understand how to make use of statistical methods, nor does he appear to have looked at the data that are presented in our paper. Obviously, a "correlation of constant magnitude will be increasingly significant as the number of degrees of freedom increases." However, this does not mean as Repp apparently believes, that calculations based on a large number of data points will "naturally lead to more significant correlations" than calculations based on a small number of data points. The data of Table IV, for example, show that topline for read sentence number 2 for five data points achieves a greater degree of significance than the all-points line for the 89 data points of read sentence number 5. In short, the alleged "confound" which Repp points to does not exist. Rather, an empirical test using an adjusted degree of freedom scale demonstrated that the all-points rms fit line is, in the majority of cases, the most linear of the three line-fitting procedures tested.

Finally, Repp appears to have lost sight of the basic claim of the declination theory, that the valleys are constrained to a linear relationship from which $F_0$ variations keyed to various linguistic and affectual phenomena deviate. In other words, the valleys are supposed to provide a linear base line from which $F_0$ deviations occur. If all the points of the $F_0$ contour show the same variation from linearity as the valleys, than the valleys have no special value in describing the $F_0$ contour. The statistical procedure thus correctly notes that a better fit to the linearity hypothesis exists when all the points of an $F_0$ contour have the same level of linearity, than when only three or four do. The data derived in experiment 2, the correlation coefficients, and levels of significance noted in Table IV, show that this is the case.

### C. Experiment 3: Spontaneous versus read speech breath-group versus declination

Repp's comment concerning our "sweeping conclusions" is odd. Our claims are modest and preliminary. We note that, "The data that we have discussed show that a great deal of variability exists; however one characterizes the overall $F_0$ contour." We also note that, "The data of experiment 3 show that sentences derived from spontaneous speech do not show as consistent a pattern of declination as do similar sentences read by the same speakers. It is evident from the data for spontaneous speech that terminals are generally falling for simple declarative sentences, and that the nonterminal parts of the $F_0$ contours have less declination than do those derived from read speech. This is consistent with the claims of the breath-group theory. It is also germane to note that the data demonstrate that neither declination or breath-group theories of intonation adequately describe all the sentences of the corpus." The reader can contrast our conclusions with the claims of declination theories like Pierrehumbert (1979) and Maeda (1976) which claim that a downwards slope of the $F_0$ contour is a necessary and universal acoustic cue that segments the flow of speech into sentences. In this regard it is significant that Maeda and Vaissiere have independently concluded that declination does not characterize $F_0$ contours derived from the spontaneous utterances of speakers of French (Vaissiere, 1985).

Maeda, S. (1976). "A characterization of American English intonation," Ph.D. thesis, MIT (unpublished).

Pierrehumbert, J. (1979). "The perception of fundamental frequency declination," J. Acoust. Soc. Am. 66, 363–369.

Repp, B. H. (1985). "Editorial comment and critique of 'Measures of the sentence intonation of read and spontaneous speech in American English' [J. Acoust. Soc. Am. 77, 649–657 (1985)," J. Acoust. Soc. Am. 78, 1114–1116.

Vaissiere, J. (1985). Personal communication.