Journal of
Experimental Psychology

VOL. 97, NO. 2 FEBRUARY 1973

PERCEPTION OF READING RATE BY SPEAKERS AND LISTENERS

HARLAN LANE
Institute of Phonetics and Speech, University of Paris III, Paris, France

FRANÇOIS GROSJEAN [1]
Department of Anglo-American Languages, Literatures, and Civilizations, University of Paris VIII, Paris, France

The generalization that sensation grows more rapidly for the speaker than for his listener is supported by scaling autophonic and extraphonic reading rate, using the methods of magnitude production and estimation respectively. The obtained exponents of the two power functions 2.6 and 1.5, show that when a speaker doubles his reading rate he perceives a sixfold increase, whereas a listener perceives less than a threefold increase. This disparity, about the same size as that for voice level, indicates that the speaker's judgments of rate, like his judgments of level, are not based solely on the sound of his voice. When a speaker varies his rate of reading a known passage in order to produce a desired apparent rate, he primarily adds or subtracts pauses at strategic syntactic locations. It is the number of pauses, much more than the rate of articulation or the duration of the pauses, that determines the changes in overall rate.

When a speaker judges the characteristics of his own speech, the cues available include tactile and proprioceptive feedback, as well as bone-and-air-conducted sidetone. When he stops speaking and judges someone else's speech, he is deprived of all of these cues save the last and, as a listener, must found his judgments differently. Since the sensory characteristics of speaking and listening are so different structurally, it is not surprising to learn, from the accumulated experimental evidence, that they are quite different functionally (Lane, 1971b). As a rule, the speaker's perception of the properties of his own speech, his autophonic output, grows more rapidly than the corresponding physical magnitudes and more rapidly than his perception of these properties in the speech of another. The autophonic functions for voice level, vowel duration, and voice pitch are all steeper than the corresponding extraphonic functions (Lane, 1962).

In the present study, we turn to examine the speaker's perception of his own rate of reading against this background. Does apparent rate, like apparent level, grow as a power of physical rate and, if so, with what exponent? Does this autophonic scale have a slope greater than 1.0 (log–log coordinates), and is it steeper than the corresponding extraphonic scale? One more question helps to set the stage for this attempt to determine the autophonic scale of speech rate: To what extent does the speaker change his articulation rate and

[1] Requests for reprints should be sent to François Grosjean, Département d'Anglo-Américain, Université de Paris VIII, Route de la Tourelle, Paris 12, France.

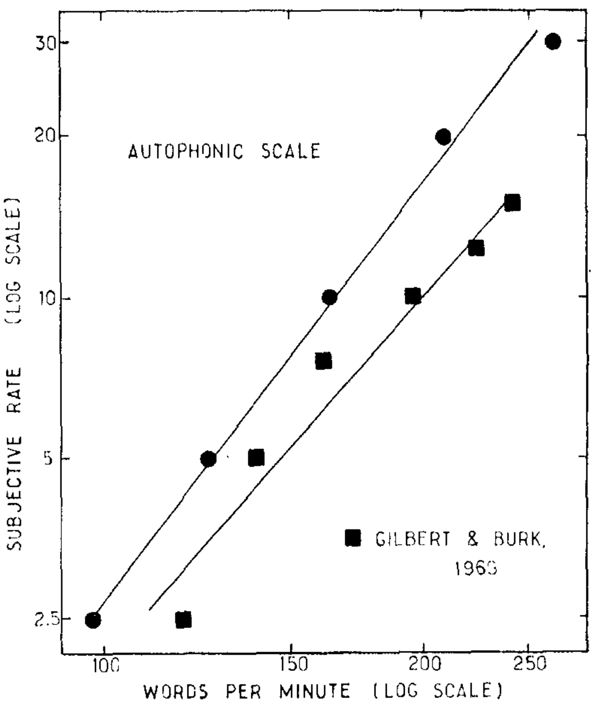

FIG. 1. Autophonic scale of reading rate. (Each circle represents the geometric mean of 96 magnitude productions, 8 by each of 12 speakers. The squares represent the arithmetic means of 288 fractionations or multiplications of reading rate, 24 by each of 12 speakers, in an experiment by Gilbert and Burk, 1969.)

to what extent the frequency and duration of his pauses, in order to produce an overall change in apparent rate of reading?

METHOD

Subjects. Twelve graduate students, native speakers of English with no reported speech or hearing defects, served for about an hour in the experiment on autophonic rate. Four of them, plus four additional classmates, served in the subsequent experiment on extraphonic rate, lasting about a half hour.

Procedure. The method of magnitude production was used to determine the autophonic scale (Lane, Catania, & Stevens, 1961). The *S* was asked to read the experimental passage at normal rate. To the apparent rate of this reading, *E* assigned the numerical value 10. A series of values (2.5, 5, 10, 20, 30) was then named in irregular order, eight times each in all, and the speaker responded to each value by reading the passage with a proportionate apparent rate.

The method of magnitude estimation was used to measure extraphonic perception of speech rate. The *E* played a recording of the experimental passage read at moderate rate (162 wpm); to this standard he assigned the numerical value 10. Next he played 40 recordings of the passage, and the listener assigned to each a number proportional to its apparent rate. In all, five rates of reading by the same speaker were presented eight times in irregular order: 92, 131, 162, 255, and 360 wpm.

Materials. The experimental passage contained 51 words of text (including contractions), comprising 75 syllables: AS FAR AS I KNOW, I'M A FAIRLY NORMAL FIFTEEN YEAR OLD, NEITHER A COMPLETE PSYCHOLOGICAL CASE, NOR A CUT ABOVE THE OTHERS. I LISTEN TO RADIO LUXEMBOURG, MY HAIR FALLS FORWARD IN THE FASHIONABLE STYLE, AND I WEAR POLO NECK SWEATERS, BUT I DON'T CONSIDER MYSELF A GREAT POP FAN.

The magnitude productions were tape recorded (Uher 4000L) in an audiometric room. The tapes were played back into an oscillographic recorder (Mingograph EM 34) and the tracings measured with dividers to an accuracy of ± .02 sec. (paper speed, 25 mm/sec). These measures yielded, for each passage, the number and duration of the pauses and the runs (the stretches of speech between the pauses), with a pause defined as an interruption of the sound wave lasting more than .24 sec. The magnitude estimations were obtained from *S*s in the same audiometric room, listening to the tape recordings at comfortable level, with parallel headsets (Opelem-Socapex) supplied by an Opelem tape recorder (2M70).

RESULTS AND DISCUSSION

Autophonic perception of rate. The speaker's perception of his own reading rate is plotted in Figure 1 as a function of the rate actually produced. The results are well represented by a straight line in log–log coordinates. Consequently, the relation between perceived and actual rate can be considered a power function whose exponent is given by the slope of best fit to the geometric means of the data. The subjective scale of autophonic rate has a slope of 2.58 when determined by the method of magnitude production. Thus, autophonic rate joins the list of several dozen other continua on which psychological magnitude has been shown to be a power function of the stimulus (Stevens, 1962); and, more specifically, it joins the other autophonic scales indicating that the sensory mechanisms mediating the speaker's perception of his own speech amplify constant stimulus ratios into much larger constant subjective ratios. If we estimate the slope of the autophonic scale of reading rate to a first approximation at 2.6, then a subjective ratio of $\frac{1}{2}$ corresponds to a physical ratio of rate of $\frac{3}{4}$.

The most thoroughly studied of the other autophonic scales is certainly that for voice level, with exponent 1.2 (Lane, 1962, 1963; Lane et al., 1961; Lane, Tranel, & Sisson, 1970). The variability of magnitude productions of speech rate is compared to that for voice level in Table 1. Three sources of variability should be distinguished in both cases: (*a*) variability due to *S*'s choice of the standard 10, that is, his modulus; (*b*) variability due to *S*'s conception of a subjective ratio; and (*c*) variability due to different sense-organ operating characteristics. The first component of the total variance was removed to obtain the "corrected" estimates of population standard deviations shown in the table in the following way: A grand mean of the log vocal rates of the group was first computed. The mean of all of the log rates of a given *S* was then subtracted from the grand mean and the difference was added to each one of that *S*'s log rates. This operation left unchanged the slope of each *S*'s autophonic rate function, but it minimized the sum of the squared deviations of his productions around the regression line for the group. On the other hand, it should be noted that these standard deviations are unaffected by changes in the unit of measurement. Table 1 shows that almost half of the total variability obtained with the method of magnitude production is accounted for by differences among *S*s in their choice of modulus and that the productions of rate were less variable than those of level.

TABLE 1

STANDARD DEVIATIONS (IN DB.) OF VOCAL LEVELS[a] AND READING RATES OBTAINED BY MAGNITUDE PRODUCTION

Criterion value	Vocal level (n = 72)[b]		Reading rate (n = 96)[c]	
	Uncorrected	Corrected	Uncorrected	Corrected
2.5	5.2	3.4	2.6	1.9
5	6.3	3.3	1.9	1.3
10	6.6	2.6	1.2	.5
20	6.9	2.8	1.2	1.1
30	6.7	4.0	1.7	1.6

Note. The *SD*s were corrected for the component of variability due to *S*'s choice of modulus and for bias due to sample size. The *SD*s are expressed in db. to facilitate comparison.
[a] The *SD*s of vocal levels were taken from Lane, Catania, and Stevens (1961).
[b] 72 = 24 *S*s × 3 Magnitude Productions/*S*.
[c] 96 = 12 *S*s × 8 Magnitude Productions/*S*.

When the autophonic scale of speech rate is constructed from fractionations and multiplications of reading rate, the results are reasonably consistent with those obtained by magnitude production in this study. Gilbert and Burk (1969) had 12 *S*s read a 55-word passage at speeds $\frac{1}{4}$, $\frac{1}{2}$, and $\frac{3}{4}$ normal, and $\frac{1}{4}$ and $\frac{1}{2}$ greater than normal. When their results are plotted in log–log coordinates (Figure 1), the straight line of best fit has a slope of 2.3. On the one hand, the use of fractionation below the standard and multiplication above is expected to steepen the psychophysical function (cf. the results for halving and doubling of loudness, Poulton & Stevens, 1955). On the other hand, curiously enough, these *E*s played *S*'s standard back to him after each production, which may have encouraged him to found his judgments of rate more heavily on auditory impressions than he would otherwise have done; that is, the rate of growth of autophonic changes may have been attenuated by the more moderate growth of extraphonic changes. The fractionation data are probably more in line with the autophonic scale of speech rate than it already appears for yet another reason. The lowest reading rates are associated with the largest variance (Table 1) and the greatest skewness. Since Gilbert and Burk used the arithmetic mean as a measure of central tendency, the rates at the lowest fractionation values and therefore the slope of the psychophysical function may be underestimated. In fact, if the arithmetic mean had been used in the present study instead of the geometric mean, a slope of 2.50 and not 2.58 would have been obtained for the autophonic scale.

Extraphonic perception of rate. The extraphonic perception of reading rate is described in Figure 2. The magnitude estimates are well fit by a straight line in log–log coordinates, with slope 1.5 (the mean and *SD* of the eight individual slopes are 1.5 and .37). The figure also shows the transformed category estimates of reading rate from an experiment in which a 55-word passage recorded at eight differ-

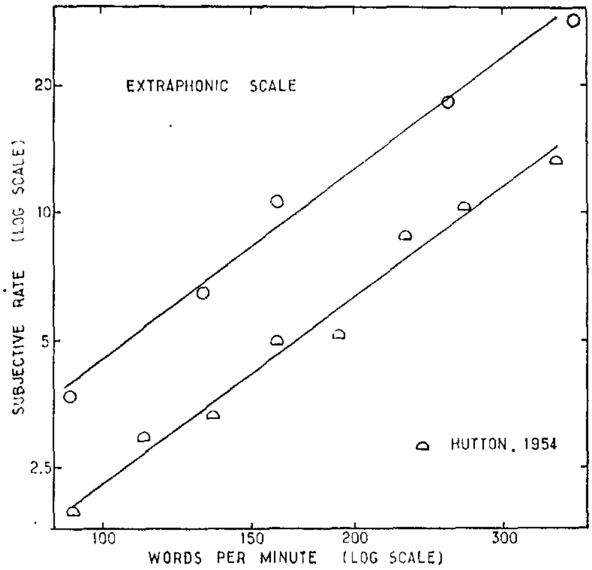

FIG. 2. Extraphonic scale of reading rate. (Each circle represents the geometric mean of 40 magnitude estimations, 5 by each of 8 listeners. Also plotted are transformed average category estimates of rate assigned by 10 listeners [Hutton, 1954].)

ent rates was presented to 10 listeners. Hutton (1954) reported that $G = 11.97 \log R - .46$, where G is a category estimate from 1 to 9, and R is the rate in words per minute. Schneider and Lane (1963) have shown that when a category scale is a simple logarithmic function of a physical continuum, as in this case, then it is related to the ratio estimation scale of the same continuum by: $\log M = n/b(G - a)$, where M is a magnitude estimate, a and b the intercept and slope, respectively, of the category scale, and n the exponent of the ratio scale. The excellent fit of the transformed category estimates of rate to the extraphonic function obtained in the present study confirms that the category scale of rate is concave downward when plotted against the ratio scale, as is expected for a prothetic continuum (Stevens & Galanter, 1957).

The disparity between the slopes of the autophonic and extraphonic functions (Figures 1 and 2), which was displayed by each of the four *S*s who served in both experiments as well as by the groups, is quite instructive. It means that when a speaker doubles his reading rate he perceives a sixfold increase, whereas a listener perceives less than a threefold increase. It is noteworthy that this ratio, approximately 2:1, between the sensory dynamics of speaking and listening is similar to that obtained for voice level (Lane et al., 1970). The generalization that sensation grows more rapidly for the speaker than for the listener seems to be borne out.

The disparity between the perception of spoken rate and heard rate also shows that the speaker does not rely exclusively on his hearing in judging his rate. This was expected. It is clear that, as far as judging his voice level is concerned, the speaker does not rely on his hearing. The autophonic scale of voice level is the same for normal speakers, for those who are experimentally deafened with intense masking noise, and for those who are deaf from birth (Lane, 1963). All three populations claim they are judging vocal effort when judging their own voice level, and indeed, the autophonic level of a speech sound and the work done on the air in producing it are proportional; they both grow as the 1.2 power of sound pressure (Ladefoged & McKinney, 1963). Further evidence that autophonic judgments are not loudness judgments is obtained whenever speakers vary their levels in order to match changes in the loudness of a criterion stimulus, to compensate for changes in sidetone loudness, or to maintain intelligibility despite increasing noise (Lane et al., 1970). All of these tasks confirm the 2:1 disparity in the operating characteristics of the underlying sensory mechanisms. Finally, the finding that judgments of one's own reading rate must depend more on interoceptive cues than on hearing was expected in the light of the evidence that it is these cues and not sidetone that are critically involved in the servomechanism control of speech (Lane, 1971a).

Rate as a complex variable. The autophonic scale plotted in Figure 1 shows the changes in the speaker's perception of his rate associated with changes in the rate of reading a passage. These latter changes, however, are the sum of changes in the time spent articulating and in the time spent pausing, and it is natural to wonder about the relative contribution of these two variables to the autophonic perception of rate. Since the exponent of a power function $y = x^m$ is unchanged under the similarity group transformation, $x' = cx$, the slope of the autophonic scale will not change whether it is plotted against articulation time, pause time, or their sum (or words per minute) *if* articulation time is a constant fraction of the total time. Figure 3 shows, however, that the articulation time ratio (= phonation time ratio) is not a constant. The percentage of the time spent articulating seems to grow about linearly with the rate, reaching nearly 100% in the present experiment. The whole picture does not agree with that painted by Gilbert and Burk (1969) who thought that "the results for w.p.m. rate indicate a significant serial increase as subjects progress from the slowest to the fastest rates studied, p.t.r. [phonation time ratio] on the other hand being altered only for establishing a difference between the faster and slower rates [p. 201]." Average articulation time ratios are less in the latter experiment and in Hutton's (1954) experiment (data reported for one *S*), but then the passage employed had four sentences compared to two here, and hence two more places for "full stops." Another factor probably contributing to reduced articulation–time ratios in these two experiments is the duration criterion for a pause; reported in neither study, it was probably shorter than in the present experiment.

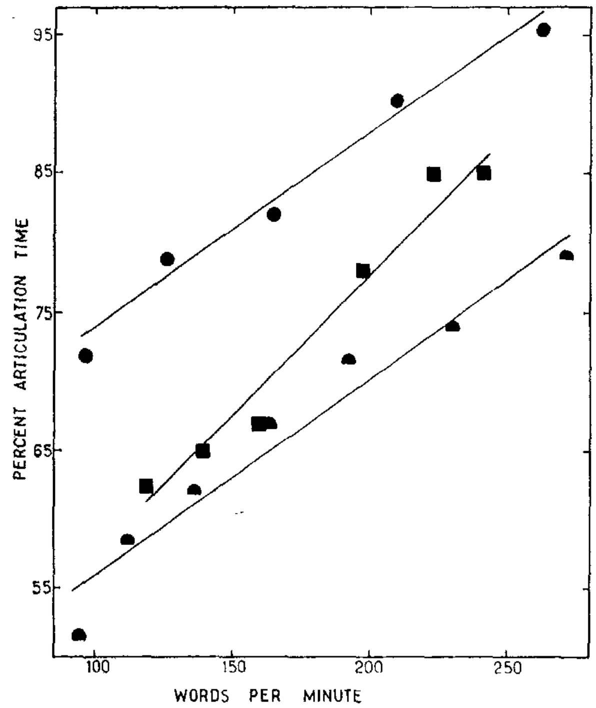

FIG. 3. Percentage of time spent articulating as a function of the overall rate. (Each circle represents the articulation–time ratio [×100] averaged across the 12 speakers. The squares represent the articulation–time ratios for the fractionations and multiplications of reading rate obtained by Gilbert and Burk, 1969. The half circles show articulation–time ratios of category productions by 1 speaker [Hutton, 1954].)

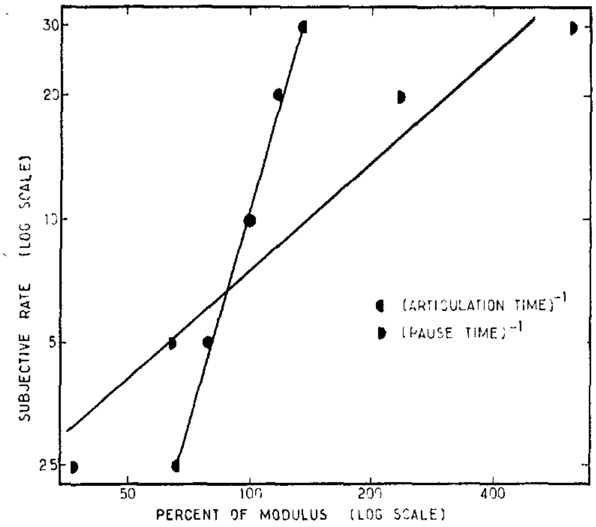

FIG. 4. Autophonic judgments of reading rate as a function of articulation time and pause time that comprise the physical rate. (For convenience, the reciprocals of these two components of the measures plotted in Figure 1 are shown, expressed as a percentage of their average values at the modulus.)

The speaker in fact changes articulation time much less than pause time to achieve a given increase in apparent rate. When autophonic judgments are plotted against these two component variables of physical rate in log–log coordinates, the slopes of the straight lines of best fit are − 3.4 and − 0.9, respectively. For convenience, Figure 4 shows the reciprocals of these two functions, with articulation time and pause time expressed as a percentage of their average values at the average modulus (165 wpm), specifically, 15.4 and 3.4 sec. (82% articulation time). In terms of articulation rates, the modulus, 4.8 syllables/sec, was bracketed by a range of only 3.4 syllables/sec. Apparently, the speaker who wishes to double his apparent

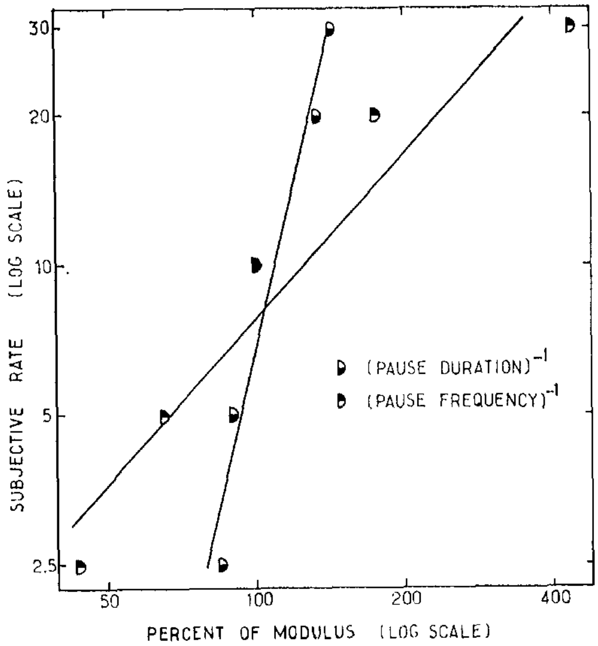

FIG. 5. Autophonic judgments of reading rate as a function of pause duration and pause frequency, which comprise total pause time (Figure 4). (For convenience, the reciprocals of these two components of the measures plotted in Figure 4 are shown, expressed as a percentage of their average values at the modulus.)

rate simply spends half as much time pausing.

Does he do this by pausing half as often or by cutting the durations of his pauses in half—or both? The former, according to Figure 5, is correct. On the average, twice the apparent rate was associated with about half the number of pauses but 85% of the average pause duration. The

TABLE 2

PERCENTAGE OF PAUSE SLOTS FILLED BY TWO OR MORE SPEAKERS AS A FUNCTION OF RATE OF READING THE EXPERIMENTAL PASSAGE AND NATURE OF THE GRAMMATICAL JUNCTURE

Pause slot following	No. of pause slots[a]	Speech rate (wpm)				
		97	126	165	210	263
Sentence[b]	1	100	100	100	100	100
Clause	6	100	100	100	100	50
Content word	23	91	65	13	9	0
Function word	20	60	10	0	0	0
Internal syllable	24	8	4	0	0	0
Overall	74	57	34	14	12	5

[a] All but the last of the categories are mutually exclusive.
[b] There is no pause slot after the last word of the passage.

slopes of the straight lines of best fit in log–log coordinates are, respectively, − 1.1 and − 4.3. Incidentally, the slope relating autophonic rate to total pause time (Figure 4) follows necessarily from the slopes reported in Figure 5 for pause frequency and duration and the fact that pause time is their product. In general, if $y = x^a$ and $y = z^b$, then $y = (xz)^{ab/a+b}$.

The nearly linear reduction in pause frequency with increasing autophonic rate (Figure 5) is not accomplished, of course, by indiscriminate suppression of pauses. Goldman-Eisler (1968) states that, with passages read aloud, pauses should be found after phrases, clauses, and at the end of sentences. Table 2 shows the percentage of all pause slots filled, as a function of the nature of the grammatical juncture and the rate of reading. At the modulus rate (165 wpm), her prediction is indeed confirmed, with the addition of two pauses after content words not located at clause boundaries: GREAT and POP. These exceptions are most likely caused by the particular stress distribution over the expression A GREAT POP FAN. At the next highest speed, the pause after GREAT is dropped, while all the others are retained. At the very highest speed, only the sentence-final and three of the clause-final pauses are retained, namely those before the constructions beginning NEITHER . . . NOR, AND, and BUT. When the reader gives magnitude productions of rate below his modulus, he prefers to add pauses after content words, inserting them after function words and after syllables within a word only as a last resort. Two content words are never followed by pauses: YEAR in FIFTEEN YEAR OLD, and POLO in POLO NECK SWEATERS. Clearly, stress is the conditioning factor, as it may have been in creating an occasional pause after A in I'M A FAIRLY NORMAL FIFTEEN YEAR OLD; the other occurrences of A are never followed by pauses, nor are those of THE and TO.

The results for pause frequency taken together with those for pause location (Figure 5 and Table 2) lead to the following conclusion. Although rate in wpm is a convenient variable for comparisons across studies carried out with diverse materials and methods, it turns out that a particular component of rate accounts for most of the variance in magnitude productions; this rate component is the number of pauses. When the speaker varies his rate of reading a known passage, he primarily adds or subtracts pauses of largely the same duration at strategic syntactic locations; he alters articulation rate and pause duration much less. In so doing, the speaker seems to be guided not only by autophonic sensation but also by several other variables that remain to be disentangled: the volume of air in his lungs, mechanical limits on the rate of articulator displacement, the punctuation of the passage, its stress distribution, and the demands of intelligible communication (Grosjean, 1971; Lane, 1971a; Lane, Grosjean, LeBerre, & Lewin, in press).

REFERENCES

GILBERT, J. H., & BURK, K. W. Rate alterations in oral reading. *Language and Speech*, 1969, **12**, 192–201.

GOLDMAN-EISLER, F. *Psycholinguistics: Experiments in spontaneous speech.* London: Academic Press, 1968.

GROSJEAN, F. Le rôle joué par trois variables temporelles dans la compréhension orale de l'anglais, étudié comme seconde langue. Unpublished doctoral dissertation, Université de Paris VIII, 1971.

HUTTON, C. L. A psychophysical study of speech rate. Unpublished doctoral dissertation, University of Illinois, 1954.

LADEFOGED, P., & MCKINNEY, N. P. Loudness, sound pressure, and subglottal pressure in speech. *Journal of the Acoustical Society of America*, 1963, **35**, 454–460.

LANE, H. L. Psychophysical parameters of vowel perception. *Psychological Monographs*, 1962, **76** (44, Whole No. 563).

LANE, H. L. The autophonic scale of voice level for congenitally deaf subjects. *Journal of Experimental Psychology*, 1963, **66**, 328–331.

LANE, H. L. The Lombard Sign and the role of hearing in speech. *Journal of Speech and Hearing Research*, 1971, **14**, 677–709. (a)

LANE, H. L. Production et perception de la parole: Rapports et différences. *Phonetica*, 1971, **23**, 94–125. (b)

LANE, H. L., CATANIA, A. C., & STEVENS, S. S. Voice level: Autophonic scale, perceived loudness, and effects of sidetone. *Journal of the Acoustical Society of America*, 1961, **33**, 160–167.

LANE, H. L., GROSJEAN, F., LEBERRE, J., & LEWIN, K. Properties of foreign language utterances that control their comprehension. *International Journal of Psycholinguistics*, in press.

LANE, H. L., TRANEL, B., & SISSON, C. Regulation of voice communication by sensory dynamics. *Journal of the Acoustical Society of America*, 1970, **47**, 618–624.

POULTON, F. C., & STEVENS, S. S. On the halving and doubling of the loudness of white noise. *Journal of the Acoustical Society of America*, 1955, **27**, 329–331.

SCHNEIDER, B., & LANE, H. L. Ratio scales, category scales, and variability in the production of loudness and softness. *Journal of the Acoustical Society of America*, 1963, **35**, 1953–1961.

STEVENS, S. S. The surprising simplicity of sensory metrics. *American Psychologist*, 1962, **17**, 29–39.

STEVENS, S. S., & GALANTER, E. H. Ratio scales and category scales for a dozen perceptual continua. *Journal of Experimental Psychology*, 1957, **54**, 377–411.

(Received January 26, 1972)