

Pause perception: Some cross-linguistic comparisons

JOANN CHIAPPETTA, LAURA A. MONTI, and DANIEL C. O'CONNELL
Loyola University of Chicago, Chicago, Illinois

Native speakers of Italian and speakers of English who were not speakers of Italian listened to two passages of monologic Italian and reported occurrence and duration of pauses. Both groups underestimated frequency of occurrence and overestimated duration of pauses relative to measured occurrence and duration in both corpora. The nonspeakers' duration estimates were significantly different from those of the native speakers and were also closer to the actual durations. Interference and facilitation effects of language are discussed.

Historically, both psycholinguistic and perceptual research have until now largely neglected perception of pauses in spoken discourse. Kowal (1986) reviewed the early use of listeners' perception of pauses for the determination of pause occurrence and duration (Cowan & Bloch, 1948; Hahn, 1949; Newman & Mather, 1938). She also listed a number of later studies (Duncan, Rice, & Butler, 1968; Hannah & Engler, 1967; Mishler & Waxler, 1970; Pürschel, 1975; Tannenbaum, Williams, & Hillier, 1965; Taylor, 1969; Webb, 1969; Williams, 1970), as well as studies that have been limited to pauses of unusual duration (e.g., Maclay & Osgood, 1959) or have divided perceptually identified pauses into relatively short, moderate, and long durations (e.g., Clay & Imlach, 1971; Crystal & Quirk, 1964; Deese, 1984).

The justification for perceptual determination has often been implicit, but somehow related to convenience or lack of instrumental means of measurement (despite early availability of instrumentation; e.g., Snell, 1918a, 1918b; Wallin, 1901). Gülich (1970) went so far as to claim, "It would be inappropriate to want to correct the subjective auditory perception by means of an objective instrumental recording" (p. 277; our translation).

Her claim failed, however, to take into account the relationships of perceptually identified pauses to objective properties of spoken discourse. Rochester (1975/76) summarized the research on these relationships (Boomer & Dittman, 1964; Cowan & Bloch, 1948; Martin, 1970) as follows: "Long pauses are always detected and no further variables are needed for explanation, while detection of short pauses (50-200 in Cowan & Bloch's work; 50-110 msec in Martin's study) depends on linguistic cues" (p. 3). Butcher (1981) called attention to a play-off between objective duration and pause position in such perceptual judgments: "Whereas breaks between tone groups are not heard by 75% of listeners until they are approximately 220 ms long, breaks within tone groups

are heard by the same proportion of listeners when only 80 ms long" (p. 205). Esser (1977) added that individual perceptual style also has an influence. The matter is made even more complex by the fact that in some cases, only variability of syllabic duration, of acoustic energy, or of vocalic quality is found to be responsible for the perceptual report of pauses (see, e.g., Adams, 1979).

The task of identifying pauses perceptually is not without problems. Directing one's attention to individual pauses is quite different from directing one's attention to a meaningful message. It is quite possible that subjects' language habits cannot be suppressed and that they therefore interfere with or facilitate the task of perceiving (location or duration of) pauses.

A logical and straightforward test of this possibility involves the use of native speakers and nonspeakers of a given language to perceive pauses in recorded passages of that language. We hypothesized (1) that native speakers would perceive a higher proportion of shorter pauses than would nonspeakers, because of facilitation from language cues that would be unavailable to nonspeakers; and (2) that native speakers would overestimate duration, relative to both nonspeakers' estimates and objective measures of duration, because of interference from the same language cues.

METHOD

The subjects for the experiment were 7 native speakers of Italian and 12 speakers of American English who were not speakers of Italian. The nonspeakers were 7 women and 5 men, all undergraduate students at Loyola University of Chicago, averaging 20 years of age. The Italian speakers were 6 women and 1 man, ranging from 18 to 52 years of age; all had lived in the United States for at least 14 years.

Two comparable Italian language samples (A and B) were used. Both were radio homilies from Chicago station WEEF, used with the permission of the speakers. The cassette recordings were of media quality and were 227 and 183 syllables in length. Articulation rates (6.11 and 6.27 syllables/sec) and mean durations of pauses (477 and 648 msec) were similar, as were the numbers of pauses (20 and 21). All measurements were made by means of a Siemens Oscillomink L and an F-J fundamental frequency meter (Type FFM 6502), with a cutoff point of 130 msec for the minimum pause duration.

Instructions were given in the native language of the speaker (with some admixture of English in the case of the Italians). The nonspeakers

Address correspondence to Daniel C. O'Connell, Psychology Department, Loyola University of Chicago, 6525 North Sheridan Road, Chicago, IL 60626.

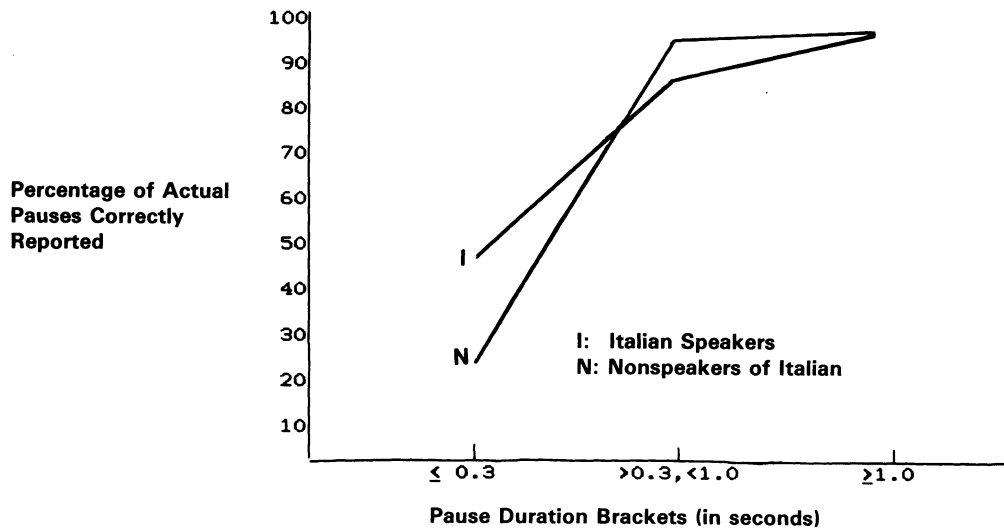


Figure 1. Percentage of actual pauses in combined Italian corpora (A and B) correctly reported by native speakers of Italian (I) and by nonspeakers of Italian (N), in three duration brackets: ≤ 0.3 sec; > 0.3 sec, < 1.0 sec; and ≥ 1.0 sec.

of Italian were tested as a group; the Italians were accessible only individually. Typed transcripts (without punctuation or capital letters except for proper names) were handed to the subjects, who were told, "A radio tape of the transcribed homily will be played for you. You have to do basically two steps: First, mark with a slash the pauses you hear and, second, estimate the duration of each pause you've marked (second to the nearest decimal point)." The homily was then played four times so that the subjects could review both identification and estimation a second time. The same procedure was repeated for the second homily without new instructions.

RESULTS

Reports of pauses at positions where pauses did not actually occur were nonsignificantly more frequent (for the combined corpora) among speakers of Italian (Group I) than among nonspeakers of Italian (Group N): median numbers of such false positives per subject were 4 and 2, respectively. More detailed analyses were carried out only for positions where pauses actually occurred. For the combined Italian corpora, the percentage of pauses correctly reported by Group I was 69.7% (200/287); the percentage for Group N was 74.9% (295/394).

The combined Italian corpora were then analyzed in three brackets of actual pause duration: ≤ 0.3 sec; > 0.3 sec, < 1.0 sec; and ≥ 1.0 sec (cf. Butcher, 1981, p. 205 f.). The brackets were intended to set off short, moderate, and long pauses. In the ≤ 0.3 sec bracket, the percentage of pauses correctly reported by Group I (46.0%; 29/63) was significantly greater than the percentage correctly reported by Group N (25.0%; 27/108) [$\chi^2(1) = 7.06$, $p < .01$]. In the > 0.3 sec, < 1.0 sec bracket, the percentage correctly reported by Group N (93.7%; 223/238) was significantly different from that reported by Group I (82.5%; 145/196) [$\chi^2(1) = 10.97$, $p < .001$]. Of all the actual pauses, 63.4% fell in this middle bracket. In the ≥ 1.0 sec bracket, there were no group differences:

the percentages of correctly reported pauses were Group I, 92.9% (26/28); Group N, 93.8% (45/48). The percentages for the two groups across the three brackets are shown in Figure 1. Mean actual and estimated durations and standard deviations of pauses in both Italian corpora for Groups I and N are given in Table 1. The means for estimated durations were calculated from the mean durations for subjects in each group for each position (i.e., corresponding to each position where an actual pause occurred). The rationale for this calculation and for the corresponding one-tailed t tests for related samples (Siegel, 1956) was dictated by our hypothesis that native speakers would overestimate duration relative to both nonspeakers' estimations and actual durations. For both Italian corpora, the durations estimated by Group I were significantly longer than the actual durations [$t(18) = 6.51$, $p < .0005$, and $t(20) = 8.74$, $p < .0005$, for A and B respectively]. Relative to the durations estimated by Group N, however, the durations estimated by Group I were significantly longer only for Italian corpus B [$t(20) = 9.05$, $p < .0005$]. We had not hypothesized that the durations estimated by Group N would also be significantly longer than the actual durations, but they were for both Italian corpora [$t(19) = 2.79$, $p < .01$, and $t(20) = 1.84$, $p < .05$, for A and B respectively].

Table 1
Mean Actual and Estimated Duration (in Seconds) and Standard Deviation (SD) of Pauses in Both Italian Corpora (A, B) for Italian Speakers (I) and Nonspeakers (N)

Corpus	Actual Duration		Estimated Duration			
	Mean	SD	I		N	
			Mean	SD	Mean	SD
A	0.48	0.29	0.84	0.31	0.76	0.34
B	0.65	0.33	1.08	0.31	0.72	0.29

DISCUSSION

It is clear from our findings that knowledge of the Italian language played an important role in these reports of pauses. In the longest duration bracket (see Figure 1), pauses were sufficiently salient by reason of their duration to be identified quite well by both native speakers and nonspeakers. In the middle bracket, however, that salience was reduced. Here the nonspeakers appear to have had an advantage in that they could attend to the physical stimulus without the compelling salience of meaning, whereas the language habits of the native speakers compelled them to attend to meaning, to the detriment of their perception and report of pause occurrence. We can refer to this as an *inhibitory effect* of the native language. In the case of the shortest duration bracket, the effect shifted in favor of the native speaker. In this bracket, the pause itself had diminished in salience to such an extent that the native speaker had to rely on expectations dictated by meaning. All else remained constant between the two groups, including the pause duration itself, intonation, syllable duration, level of acoustic energy, and vocalic quality. But the nonspeaker was almost entirely bereft of semantic and syntactic cues. In this case we can refer to a *facilitative effect* of the native language.

In view of the inhibitory and facilitative effects of the native language, the duration estimates also take on new significance. In the absence of any language cues, one could predict that the nonspeakers' duration estimates would be closer than those of the native speakers to the actual pause durations. Such was indeed the case. The nonspeakers' estimates were, of necessity, based on the physical stimulus, rather than on any additional expectations dictated by language cues.

The nonsignificant difference between native speakers and nonspeakers in false positives (reported occurrence of pauses in the absence of actual pauses) was in the same direction. It seems reasonable to expect that, with a more extensive sample, false positives would appear more frequently in the reports of native speakers than in the reports of nonspeakers, in consequence of the native speakers' language habits and the expectations dictated by their use.

To return now to the archival literature, we have confirmed Rochester's (1975/76) conclusion that "detection of shorter pauses . . . depends on linguistic cues" (p. 3). This dependence, however, is not absolute; the detection is simply facilitated by the addition of language cues to the cues supplied by the physical stimulus. We cannot accept Rochester's conclusion that "long pauses are always detected and no further variables are needed for explanation" (p. 3). Linguistic cues obviously made a notable difference in our middle duration bracket, and although there were no differences between groups in the longest duration bracket, both groups still failed to report some actual pauses (7.1% and 6.2% for Groups I and N, respectively).

The fact that language cues can be inhibitory as well as facilitative with regard to the reporting of occurrence and duration of actual pauses points up both the lawfulness and the complexity of these long-neglected perceptual phenomena. We would all be well advised not to heed Güllich's (1970) advice that it is "inappropriate to want to correct the subjective auditory perception by means of an objective instrumental recording" (p. 277; our translation). Our concern is not, in any event, for a correction, but for an explanation. The present research makes clear that both actual duration and estimated duration of pauses must be taken into account by a comprehensive explanation of pause occurrence and duration.

REFERENCES

- ADAMS, C. (1979). *English speech rhythms and the foreign learner*. The Hague: Mouton.
- BOOMER, D. S., & DITTMAN, A. T. (1964). Hesitation pauses and juncture pauses in speech. *Language & Speech*, **63**, 215-220.
- BUTCHER, A. (1981). *Aspects of the speech pause: Phonetic correlates and communicative functions*. Kiel, West Germany: Institut für Phonetik der Universität Kiel.
- CLAY, M. M., & IMLACH, R. H. (1971). Juncture, pitch, and stress as reading behavior variables. *Journal of Verbal Learning & Verbal Behavior*, **10**, 133-139.
- COWAN, J. M., & BLOCH, B. (1948). An experimental study of pauses in English grammar. *American Speech*, **23**, 89-99.
- CRYSTAL, D., & QUIRK, R. (1964). *Systems of prosodic and paralinguistic features in English* (= Janua Linguarum, series minor, **39**). The Hague: Mouton.
- DEESE, J. (1984). *Thought into speech: The psychology of a language*. Englewood Cliffs, NJ: Prentice-Hall.
- DUNCAN, S., JR., RICE, L. N., & BUTLER, J. M. (1968). Therapists' paralinguistic features in peak and poor psychotherapy hours. *Journal of Abnormal Psychology*, **73**, 566-570.
- ESSER, U. (1977). *Sprachpsychologische Untersuchungen zum Pausenverhalten in der spontanen Sprachproduktion*. Unpublished doctoral dissertation, Karl-Marx-Universität, Leipzig.
- GÜLICH, E. (1970). *Makrosyntax der Gliederungssignale im gesprochenen Französisch*. München: Fink Verlag.
- HAHN, E. (1949). An analysis of the delivery of the speech of first grade children. *Quarterly Journal of Speech*, **35**, 338-343.
- HANNAH, E. P., & ENGLER, L. (1967). Juncture phenomena and the segmentation of a linguistic corpus. *Language & Speech*, **10**, 228-233.
- KOWAL, S. (1986). *Zur Sprachproduktion von Politikern*. Unpublished manuscript.
- MACLAY, H., & OSGOOD, C. (1959). Hesitation phenomena in spontaneous speech. *Word*, **15**, 19-44.
- MARTIN, J. G. (1970). On judging pauses in spontaneous speech. *Journal of Verbal Learning & Verbal Behavior*, **9**, 75-78.
- MISHLER, E. G., & WAXLER, N. E. (1970). Functions of hesitations in the speech of normal families and families of schizophrenic patients. *Language & Speech*, **13**, 102-117.
- NEWMAN, S., & MATHER, V. G. (1938). Analysis of spoken language of patients with affective disorders. *American Journal of Psychiatry*, **94**, 913-942.
- PÜRSCHEL, H. (1975). *Pause und Kadenz: Interferenzerscheinungen bei der englischen Intonation deutscher Sprecher*. Tübingen: Max Niemeyer Verlag.
- ROCHESTER, S. R. (1975/76). Defining the silent pause in speech. *Journal of the Ontario Speech & Hearing Association*, **8**, 1-4.
- SEGEL, S. (1956). *Nonparametric statistics: For the behavioral sciences*. New York: McGraw-Hill.
- SNELL, A. L. F. (1918a). An objective study of syllabic quantity in English verse. *Publication of the Modern Language Association*, **33**, 396-408.
- SNELL, A. L. F. (1918b). Pause: A study of its nature and its rhythmic function in verse, especially blank verse. In F. N. Scott (Ed.), *Contributions to rhetorical theory* (Vol. 8). Ann Arbor, MI: Ann Arbor Press.
- TANNENBAUM, P. H., WILLIAMS, F., & HILLIER, C. S. (1965). Word predictability in the environments of hesitation. *Journal of Verbal Learning & Verbal Behavior*, **4**, 134-140.
- TAYLOR, I. (1969). Content and structure in sentence production. *Journal of Verbal Learning & Verbal Behavior*, **8**, 170-175.
- WALLIN, E. W., JR. (1901). Researches in the rhythm of speech. *Studies from the Yale Psychological Laboratory*, **9**, 1-142.
- WEBB, J. T. (1969). Subject speech rates as a function of interviewer behavior. *Language & Speech*, **12**, 54-67.
- WILLIAMS, F. (1970). Psychological correlates of speech characteristics: On sounding disadvantaged. *Journal of Speech & Hearing Research*, **13**, 472-488.

(Manuscript received for publication September 12, 1986.)