

Pause reports for spontaneous dialogic speech

LORI A. FRIEDMAN and DANIEL C. O'CONNELL
Georgetown University, Washington, D.C.

Native speakers of English ($N = 40$) reported occurrence and estimated duration of pauses for English and German spontaneous dialogue. Occurrence was underestimated significantly more for the German than for the English. Duration was underestimated significantly more for the English than for the German, and for positions where no actual pauses occurred (false positives) than for positions where there were actual pauses. The results for the English sample only are compared with similar estimates based on ethnomethodological transcriptional procedures ($N = 7$). Both methods are found to yield similarly nonveridical reported occurrence and estimated duration.

Recent research (Carpenter & O'Connell, 1988; Chiappetta, Monti, & O'Connell, 1987; Duez, 1985; Gardner, McMillan, Townsend-Handscorn, Barrett-Bates, & O'Connell, 1990; Monti, O'Connell, & Chiappetta, 1990; Stuckenberg & O'Connell, 1988) has challenged the veridicality of reported occurrence and estimated duration of pauses in oral discourse. In general, both overestimates and underestimates of duration occur, and both false positives (report of pauses where no actual pauses occur) and false negatives (failure to report actual pauses) are frequent.

There remain, however, numerous questions about these phenomena. The research methods have complicated this uncertainty in that Duez (1985) has used a method other than the remainder of the research listed, and the Monti et al. (1990) research was concerned not with actual oral discourse as stimulus, but rather with the surmising of pauses that would be appropriate were a printed text actually to be spoken. Several native languages have been studied, but, to date, no spontaneous dialogic speech has been investigated (except by Duez with the alternate method and only for French). The research of Gardner et al. (1990) did indeed investigate dialogic speech, but the dialogue was according to dramatic convention rather than spontaneous.

Another recent research development has involved the questioning of psycholinguistic and ethnomethodological transcriptional methods for oral discourse. Deese (1984) has found all such methods, whether used for research purposes or for official or legal records, "to be inaccurate to varying degrees" (p. 21). More specifically, for the purpose of recording pause occurrence and duration, these methods have been found to be inadequate (O'Connell & Kowal, 1990a, 1990b). To date, there has been no direct comparison of an experimental method and a transcriptional method of identifying pauses.

The present research extends the evidence regarding reported occurrence and estimated duration of pauses to spontaneous dialogue. It is hypothesized that native speakers of American English will be both more accurate

and more efficient in reporting occurrence of pauses in English than in reporting pauses in German. It is also hypothesized that there will be a difference in the estimated duration across the two languages. The limitations of extant evidence necessitate that this hypothesis be non-directional. Finally, it is hypothesized that estimated duration at false positive positions will be shorter than estimated duration at actual pause positions.

METHOD

Stimulus materials consisted of an excerpt from an English television interview of former President Reagan by Dan Rather and an excerpt from a German television interview of former East German Secretary General Krenz by Fritz Pleitgen. Table 1 lists the descriptive statistics for these two speech samples. These were derived 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 measurable pause duration. In both the English and the German samples, the interviewer's articulation rate was faster than the interviewee's (English, 6.42 > 5.61 syl/sec; German, 6.06 > 6.00 syl/sec). In both languages, the interviewee spoke more syllables than did the interviewer (English, 2.57 times as many; German, 2.34 times) and used more ontime (i.e., time between onset and offset of speech sound) to do so (English, 2.94 times as much; German, 2.36 times).

Native-English-speaking American subjects (21 men, 19 women; mean age = 19) were volunteer participants in the research. Only 1 woman had any appreciable fluency in German, and her data revealed no notable divergence from the others.

Instructions were presented as a typed cover sheet to the English and German texts as follows: "The general objective of this study is to discover the nature of pauses and their role in language. In an effort to

Table 1
Descriptive Statistics for English and German Passages

Statistic	Passage	
	English	German
Total time (msec)	96,306	60,156
% of offtime (pause time)/total time	26.62	13.81
Speech rate (syl/sec total time)	4.27	5.10
Articulation rate (syl/sec ontime [see text])	5.82	6.02
Mean pause duration (msec)	475	462
Syl/word	1.44	1.62
Syl/phrase (per pause)	7.21	12.28
Time (msec)/phrase (per pause)	1,248	2,074
Words/phrase (per pause)	4.94	7.44

Correspondence should be addressed to Daniel C. O'Connell, Department of Psychology, Georgetown University, Washington, DC 20057.

do so, we shall examine your responses of pauses with regard to location and duration. In all, you will be presented with two recorded passages: one in English and one in German. In addition, you will be provided with a typed replication of the passage on which you will record the location and the duration of each pause. You will be presented with each passage four times. For the first two presentations of the passage, please mark with a slash (/) where you believe a pause to exist. Keep in mind that we define a pause as complete silence. For the second two presentations, please record your estimation of the pause duration (i.e., 0.0 seconds). To avoid confusion, we will instruct you immediately before each presentation of the passage."

The subjects were then taken through the four steps with an example ("To err is human/to forgive divine") and given an explanation of our use of "0.0 seconds," namely, to avoid suggesting any specific long or short bias to the subjects regarding their duration estimates.

The subjects then turned the page to the texts of the English and German samples, as first the English and then the German were played four times. The subjects were run in two groups. The texts were typed margin to margin without punctuation or capitalization (except for initial capital letters for German nouns as required by the language) on $8\frac{1}{2} \times 11$ in. paper. Filled pauses were noted as "UH" in English and "ÄH" in German.

The experimental design therefore consisted of a 2×2 factorial, with between-subject (men, women) and within-subject (English, German) variables. For estimated duration only, position (actual pauses, false positives) became an additional within-subject factor.

It is not feasible to subject ethnomethodological transcriptional procedures to group experimentation. Essentially, the method requires that a transcriber be entirely self-paced and make as many passes through the material (or parts thereof) as needed. Hence, 7 tutorial students (4 men, 3 women; mean age = 21) were supplied with a copy of the ethnomethodological conventions (Atkinson & Heritage, 1984, pp. ix-xiv), and an audio cassette of the English passage. The use of a stopwatch was left to their option. They were asked to follow the conventions as meticulously as possible. No other instructions were given. Time spent on the transcription ranged from about 3-8 h.

RESULTS

The ANOVA for estimated duration yielded two main effects, language and position (in all cases, $df = 1, 38$). Mean estimated duration was significantly shorter for English than for German ($348 < 386$ msec; $SDs = 258$ and 280 ; $F = 4.32$, $p < .05$). Both estimates were underestimates of the respective actual durations (475 and 462 msec). Mean estimated duration was also significantly shorter for false positives than for actual pauses ($327 < 407$ msec; $SDs = 248$ and 285 ; $F = 11.95$, $p < .01$). Neither the main effect for gender nor any of the interactions were significant.

Accuracy was defined as the percentage of actual pauses reported. Since the means for the men and the women were almost identical, a simple one-tailed t test for correlated samples was calculated for the English and German percentages. The English reports were significantly more accurate than were the German [$64.5\% > 29.6\%$; $SDs = 11.6$ and 17.2 ; $t(39) = 13.63$, $p < .0005$]. A similar index of efficiency was defined as the percentage of correct reports/(correct reports + false positives). For this measure, the English reports again were significantly more efficient than were the German reports [$87.0\% > 48.2\%$; $SDs = 7.3$ and 19.2 ; $t(39) = 12.78$, $p < .0005$].

The major comparison made possible through the addition of the ethnomethodological data is the percentage of

accuracy of reported occurrence of pauses with this procedure over all transcribers (59.2%). These have been broken down according to duration brackets (plus the otherwise excluded, unreliably measurable pauses < 130 msec in duration) in Figure 1 and compared with the corresponding percentages yielded from the experimental procedures.

Instead of an underestimate, the ethnomethodological transcriptions yielded an overestimated mean duration ($611 > 475$ msec; $SD = 357$). This mean was contributed to by only 4 of the transcribers, since one of the legitimate options of the conventional transcriptional procedure is to note pauses without measurement thereof.

DISCUSSION

The Gardner et al. (1990) research differed from the present study only in that the dialogue was a dramatic convention rather than truly spontaneous. In that study, the comparable English-speaking subjects reported 82.5% of the actual pauses in the English sample and 78.5% in the German sample. Since both listening and reading are habits of searching for meaning, the complexity, unpredictability, and rapidity of the spontaneous dialogue in the present experiment may arguably be accepted as having contributed to the lower corresponding percentages (64.5% and 29.6%) of accuracy. The articulation rates, for example, in the previous study were 4.54 and 3.30 syllables per second (syl/sec) for the English and German samples, respectively, whereas the corresponding articulation rates in the present samples were 5.82 and 6.02 syl/sec—well over 2 syl/sec faster in both cases.

This argument is consistent with recent findings by Ulasevich, Leaton, Kramer, Hieke, and O'Connell (1990). The English and German samples in that study were much more highly structured (poetry rather than spontaneous dialogue), were monologic, and the respective articulation rates were very slow (3.20 and 3.47 syl/sec). The obtained percentages of actual pauses reported were correspondingly high (93.5% and 82.0%). At the other extreme, a stimulus-degradation condition with the same poems yielded surmises (based on unpunctuated texts without poetic format) of only 40.0% and 21.9% of the actual pauses in the same English and German poems (Monti et al., 1990).

The significant difference between the mean estimated duration at false positive and actual pause positions ($327 < 407$ msec) is hardly surpris-

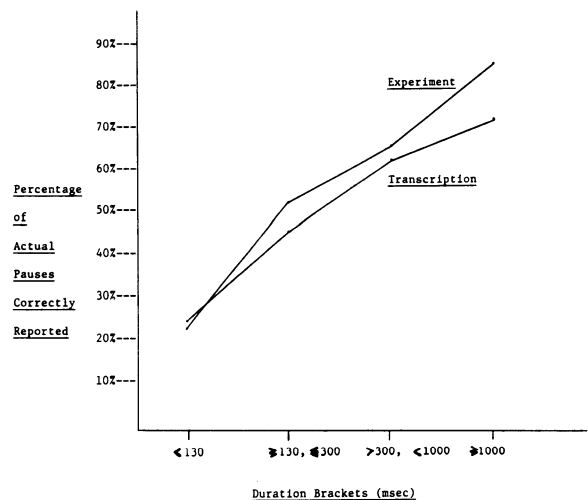


Figure 1. Percentage of actual pauses correctly reported in the experiment and by the ethnomethodological transcriptional procedure within four duration brackets, for the English passage only.

ing, since the false positive positions involved no reliably measurable silences at all but were the result of other stimulus cues misconstrued as pauses by the subjects. The difference between the mean estimated duration for the English and the German sample (348 < 386 msec), however, is the more impressive in that the actual difference in duration, although slight, was in the other direction (475 > 462 msec). That both estimates were underestimates of the respective actual pause durations was interesting as well, and it is speculated that such underestimates are a result of the rapid articulation rate found in the stimulus materials (5.82 and 6.02 syl/sec in the English and German samples, respectively) and of a consequent general perception of rapid-paced speech.

It is clear from all these data, as from the other recent research, that subjects are unable to deal with the temporal properties of either their native language or of a language in which they have little or no proficiency. The processing demands become particularly difficult (and correspondingly frustrating) with the complexity, unpredictability, and rapidity of spontaneous dialogue and especially with the shift to a non-native language.

It has been claimed recently (Klos & Ellgring, 1987) that temporal properties can most properly be identified and transcribed by stopwatch transcriptional procedures: "The manual method of identifying speech pauses longer than 390 msec is reliable and yields results as good as or better than the electronic method" (p. 64; our translation). The authors acknowledge, however, that the advantage for the manual method holds only when recordings are of poor audio quality and only for the longer pauses.

There is no evidence in the present research to confirm such a claim. In fact, only 59.2% of all actual pauses were correctly transcribed by such procedures. Figure 1 (which deliberately incorporates even the unreliably measurable pause positions < 130 msec in duration and otherwise in the experiment considered as false positives) clearly shows that the self-pacing and corresponding unlimited use of time do not increase the accuracy of pause reports at all. In none of the duration brackets is there any superiority for the ethnomethodological transcriptions over the experimental data, even though the transcribers all used more than 10 times as much time for their transcriptions as was spent on the experiment. The additional time does not alter the perceptual habits consequent upon knowledge of language.

Both the experimental and the ethnomethodological components of the present study indicate that a veridical transcription of the temporal properties of spontaneous dialogue cannot be obtained without careful

physical measurement of the real time, particularly in pauses of sub-second magnitude.

REFERENCES

- ATKINSON, J. M., & HERITAGE, J. (1984). *Structures of social action: Studies in conversation analysis*. Cambridge: Cambridge University Press.
- CARPENTER, S., & O'CONNELL, D. C. (1988). More than meets the ear: Some variables affecting pause reports. *Language & Communication*, 8, 17-27.
- CHIAPPETTA, J., MONTI, L. A., & O'CONNELL, D. C. (1987). Pause perception: Some cross-linguistic comparisons. *Bulletin of the Psychonomic Society*, 25, 103-105.
- DEESE, J. (1984). *Thought into speech: The psychology of a language*. Englewood Cliffs, NJ: Prentice Hall.
- DUEZ, D. (1985). Perception of silent pauses in continuous speech. *Language & Speech*, 28, 377-389.
- GARDNER, J. C., McMILLAN, H. S., TOWNSEND-HANDSCOMB, D., BARRETT-BATES, R., & O'CONNELL, D. C. (1990). Reporting pauses in dramatic dialogue. *Bulletin of the Psychonomic Society*, 28, 167-170.
- KLOS, T., & ELLGRING, H. (1987). Manuelle versus elektronische Analyse von Sprechpausen. *Zeitschrift für experimentelle und angewandte Psychologie*, 34, 64-71.
- MONTI, L. A., O'CONNELL, D. C., & CHIAPPETTA, J. (1990). "Surmising" pauses. *Journal of Psycholinguistic Research*, 19, 43-55.
- O'CONNELL, D. C., & KOWAL, S. (1990a). A note on time, timing, and transcriptions thereof. *Georgetown Journal of Languages & Linguistics*, 1, 203-208.
- O'CONNELL, D. C., & KOWAL, S. (1990b). Some sources of error in the transcription of real time in spoken discourse. *Georgetown Journal of Languages & Linguistics*, 1, 453-466.
- STUCKENBERG, A., & O'CONNELL, D. C. (1988). The long and the short of it: Reports of pause occurrence and duration in speech. *Journal of Psycholinguistic Research*, 17, 19-28.
- ULASEVICH, A., LEATON, B., KRAMER, D., HIEKE, A. E., & O'CONNELL, D. C. (1990). *American and Japanese subjects' reports of pause occurrence and duration*. Unpublished.

(Manuscript received November 28, 1990.)