

Temporal analysis of English and Spanish narratives

TERESA H. de JOHNSON

Saint Louis University, Saint Louis, Missouri 63103

DANIEL C. O'CONNELL

University of Kansas, Lawrence, Kansas 66045

and

EDWARD J. SABIN

Saint Louis University, Saint Louis, Missouri 63103

Two groups of 45 adolescent young men, native speakers of English and Spanish, respectively, retold a narrative after individual auditory and/or visual presentation of an 80-sec episode. Spanish stories were longer in mean number of syllables, but shorter in mean number of words, than were English stories; word length in syllables was shorter in English. In most respects, silent pauses were comparable across the two languages, but vocal hesitations varied considerably. Articulation and speech rates were much faster in Spanish than in English. Auditory presentation elicited the longest stories, visual presentation the shortest; audio-visual presentation, despite the visual enrichment of the stimulus, elicited stories of intermediate length.

The vast majority of temporal analyses of speech samples have been limited to one or another language; comparisons among various languages have not been common. Comparisons made with speech samples collected under comparable conditions and with a common standardized methodology have been even rarer.

Paradoxically, the earliest published pausological studies exemplify well such comparisons. Cattell (1885; 1886, p. 65) compared reading rates for six languages (English, French, German, Italian, Latin, and Greek) and concluded: "The rate at which a person reads a foreign language is proportional to his familiarity with the language. . . . The subject does not know that he is reading the foreign language more slowly than his own; this explains why foreigners seem to talk so fast." Wallin (1901) compared temporal aspects of poetry readings in several languages, but deemphasized comparisons among them.

More recently, Fónagy (1960) compared Hungarian with French and German poetry, but made measurements of pauses for the Hungarian poetry only. A study by Osser and Peng (1964, p. 124) "failed to reveal significant differences in speech rate between the two groups" in a cross-cultural study of speech rate between Japanese and Americans. Black, Tosi, Singh, and Takefuta (1966, p. 240) compared oral readings in Hindi, Spanish, and Japanese by native speakers with their readings of English and found that "the pauses used in native languages and English did not differ significantly." Hanley, Snidecor, and Ringel (1966, p. 105) found that

"male speakers of American English spend proportionately greater time in phonating than speakers of Japanese and Spanish," but not male readers. Hanley and Snidecor (1967, p. 146) found that in female readers "Spanish and Tagalog proved to have a relatively high phonation/time ratio (0.516-0.476) and American English and Japanese low ratios (0.442-0.376)."

Grosjean and Deschamps (1972, 1973, 1975) found that French and English speakers generally used the same amount of time in silent pauses, but in different patterns; with regard to vocal hesitations, French speakers produced almost as many drawls as filled pauses, whereas English speakers used more filled pauses by far in compensation for fewer drawls.

Barik (1973, 1975, 1977) compared French and English simultaneous translations with emphasis on consistent relationships across the languages. Unfortunately, the samples are extraordinarily small and not comparable across the two languages. A much more comprehensive sample, based on many hours of broadcast sampling in Spanish, Portuguese, French, Italian, English, and German, was collected by Glukhov (1975, p. 72), with the conclusion that "the specifics of the different languages have only a relatively slight influence on the univariate pause distribution function."

The comparisons of German and English by Kowal, O'Connell, O'Brien, and Bryant (1975) and O'Connell and Kowal (1972), summarized with additional Spanish comparisons in Sabin, Clemmer, O'Connell, and Kowal (1979), and the comparisons of Spanish and English by

Johnson (1978) have the added advantage of sharing identical methodological and measurement conventions with the present research.

It has been claimed that speech rate within a certain range is a language universal but that distributions of pauses are peculiar to specific languages (Osser & Peng, 1964, p. 124; O'Connell & Kowal, 1972, p. 163).

To what extent temporal phenomena such as speech and articulation rates, pausing, and other hesitation phenomena are constant or variable from one language to another is still a question requiring a great deal more empirical evidence. The present research is a modest effort to compare English and Spanish speech samples produced under specified conditions and measured in a standardized manner comparable with other recent research. It is hypothesized that phonetic, syllabic, syntactic, and lexical differences between the two languages will produce differences in the various temporal analyses, such as the patterning or distribution of pausing and other hesitation phenomena, evidence of constancies across languages to the contrary notwithstanding.

METHOD

Subjects were 90 young men: 45 native speakers of English from Saint Louis, Missouri, ages 15-5 to 18-4 (years-months) ($M = 17.2$, $SD = 8.2$ months), and 45 native speakers of Spanish from San Luis Potosi, Mexico, ages 15-3 to 18-9 ($M = 16.4$, $SD = 11.2$ months). Although sociocultural variables were not under study, care was taken that educational and socioeconomic backgrounds of the subjects were similar; both groups were from upper middle class, college preparatory, male-only high schools.

In order to differentiate performance characteristics due to task, three modes of presentation of the stimulus were used: audio-visual (film and sound track), visual (film only) and auditory (sound track only). Two identical films, one with English and one with Spanish sound track, were used: an 80-sec comic episode from "Clown Princes of Hollywood" (Blackhawk Films, Davenport, Iowa). Since each subject performed the same task, retelling the story in his own words, the modes of presentation served to differentiate the cognitive processing required by subjects to decode the story and recode it in their own words.

Verbatim transcripts of the voice recordings were prepared in conventional orthography (Spanish or English), but with no corrections, especially in words such as *entonces* (then), which was rendered in several versions: *'tonces*, *'tons*, *entons*. No punctuation was imposed on these transcripts.

The magnetic-tape voice recordings were transferred to a Brüel and Kjaer audio-frequency spectrometer (Type 2112) and level recorder (Type 2305), yielding a graphic record of acoustic energy in terms of amplitude over time. This permanent graphic record was used to objectively locate and measure the duration of silent pauses and other temporal responses. The minimum length of silent pauses was 270 msec.

The following response measures were studied: total time of stories, speech time (excluding silent pauses), story length in syllables, speech and articulation rates (syl/sec), phrase length (syl/silent pause), and percentage of silent pause time/total time. Four types of vocal hesitations were examined, including parenthetical remarks (e.g., "Well, you know"), false starts (e.g., "Driving the . . . he drove the car"), repeats (e.g., "The the man arrived"), and filled pauses (e.g., "uh, umm") as further specified elsewhere (Kowal, O'Connell, & Sabin, 1975; Sabin et al., 1979).

Frequency and length of silent pauses and vocal hesitations were adjusted per syllable in order to control for differences in sample length (number of silent pauses/syl, length of silent pauses/syl, number of vocal hesitations/syl, length of vocal hesitations/syl). The last two of these measures were indices of the number of occurrences of vocal hesitation *phenomena* per syllable and of the number of vocal hesitation *syllables* per syllable, respectively.

Each subject was seated in front of a 14 x 14 in. projection screen for individual viewing. Instructions were simply to watch and/or listen to the episode and then retell the story. Subjects were aware that their speech was being recorded.

RESULTS

Each of the response measures was subjected to a 2 by 3 (language by mode of presentation) analysis of variance. None of the interactions was significant. Since a number of response measures were involved, a conservative significance level ($p < .01$) was adopted. All results reported in this section were significant at or beyond that level.

Mean story length in syllables was the only response measure to yield significance of both main effects, mode of presentation ($F = 5.89$) and language ($F = 11.73$). Auditory presentation yielded the longest stories and visual the shortest, with audio-visual intermediate ($215 > 199 > 156$ syl). Spanish yielded longer stories than English ($215 > 165$ syl). The corresponding standard deviations were 68, 80, 68 and 85, and 57, respectively.

Mean total time and speech time were significant response measures ($F = 6.63$ and 5.15 , respectively) for the mode of presentation main effect only. Mean total time and speech time were both longest in the auditory presentation, shortest in the visual, and intermediate in the audio-visual ($56 > 50 > 41$ and $40 > 36 > 29$ sec, respectively). The corresponding standard deviations were 16, 18, 14 and 12, 14, 11, respectively.

Mean speech and articulation rates and length of vocal hesitations per syllable were significant response measures ($F = 25.93$, 58.64 , and 7.19 , respectively) for the language main effect only. Spanish yielded faster speech and articulation rates than did English ($4.24 > 3.66$, $6.08 > 5.00$ syl/sec, respectively) and greater length of vocal hesitations per syllable than English ($.102 > .071$ syl). The corresponding standard deviations were .59 and .46, .78 and .52, and .059 and .051, respectively.

In both the Spanish and English groups, all subjects used silent pauses, and the frequency in Spanish and English ($948 > 804$) was reasonably proportional to the total frequency of syllables in the Spanish and English samples, respectively ($9,672 > 7,340$). Mean length of silent pauses was comparable in Spanish and English (733 and 694 msec, respectively). Use of vocal hesitations, however, was considerably different in Spanish and English. The total frequencies are shown in Table 1.

Table 1
Total Number of Vocal Hesitations and Percentage of
Subjects Using Each Type in Spanish and English

Type of Vocal Hesitation	Spanish		English	
	N	%	N	%
Parenthetical Remarks	253	93	49	62
False Starts	143	84	89	78
Repeats	177	77	48	42
Filled Pauses	27	35	182	95

Despite the larger number of syllables in the Spanish sample, the total number of words was smaller than in English ($5,679 < 6,297$). Underlying this discrepancy was a shorter mean word length in English than in Spanish ($1.19 < 1.71$ syl/word).

DISCUSSION

The findings for mode of presentation are of interest apropos of encoding processes. The audio-visual mode unquestionably provided the richest array of stimulus information for encoding into a speech response mode. But it consistently elicited less response encoding in terms of syllables and time expended compared to the auditory mode. The auditory mode provided information already verbally encoded. The truncation of story length with the addition of visual to auditory mode indicates interference of some kind rather than simple redundancy of information. It is hardly surprising, on the other hand, that the purely visual mode elicited the shortest stories; a complete recodification of the nonverbal stimulus information was required. The absence of any interaction between mode of presentation and language main effects makes such an interpretation quite straightforward.

Variation of story length due to language is not quite as simple. Spanish stories had more syllables and fewer words than English stories. The paradox is due to a discrepancy in mean word length in syllables from Spanish to English, a fact which dramatically points up the impossibility of making use of words per minute as a comparative measure of rate across languages.

But the differential length of stories across Spanish and English also reflects a syllabic rate differential in both articulatory and speech rate measurements. Delattre's (1966) findings reflect an average syllable length in English of 242 msec, as compared with 227 msec in Spanish. In the present experiment, similar differences are reflected in faster Spanish articulation and speech rates relative to English.

The results summarized in Table 1 undoubtedly pose many questions regarding the cognitive processing underlying use of vocal hesitations and their expressive and rhetorical functions in discourse. A comparison of the parenthetical remarks and filled pauses proves particularly heuristic. Both the frequencies and the percentages of subjects contributing to those frequencies strongly suggest the use of certain parenthetical remarks in Spanish (e.g., *este*, *pues*, *bueno*) and filled pauses in English for equivalent functions. Examination of the transcripts confirms this suggestion; in context, for example, it is clear that the parenthetical remark "este" in Spanish frequently serves as a filler, just as "uh" does in English. The finding is strikingly parallel to that of Grosjean and Deschamps (1975): a compensatory use of draws instead of filled pauses in French and vice versa in English.

Another comparison can throw additional light on the time-gaining function of vocal hesitations. Given the faster speech and

articulation rates of the Spanish speakers, they clearly get more syllables for their money (in currency of time). But even here there is a compensating factor; almost twice as many syllables are expended by the Spanish speakers in vocal hesitations ($981 > 513$) as by the English speakers. The Spanish speaker spoke on the average 4.6 sec longer than the English speaker. On the whole, the gains and losses seem to indicate a very similar economy of time and effort in Spanish and English for the codification of the same story.

The relative constancy of silent pauses across Spanish and English is in accord with other recent research findings. The hypothesized variations in patterning and distribution were found to be characteristic of vocal hesitations, particularly in the functional trade-off between filled pauses in English and certain parenthetical remarks in Spanish. These variations underline the necessity for detailed analysis of each type of hesitation in research across languages and the corresponding inadvisability of grouping all vocal hesitations under a single response category.

REFERENCES

- BARIK, H. C. Simultaneous interpretation: Temporal and quantitative data. *Language and Speech*, 1973, 16, 237-270.
- BARIK, H. C. Simultaneous interpretation: Qualitative and linguistic data. *Language and Speech*, 1975, 18, 272-297.
- BARIK, H. C. Cross-linguistic study of temporal characteristics of different types of speech materials. *Language and Speech*, 1977, 20, 116-126.
- BLACK, J. W., TOSI, O., SINGH, S., & TAKEFUTA, Y. A study of pauses in oral reading of one's native language and English. *Language and Speech*, 1966, 9, 237-241.
- CATTELL, J. MCK. Ueber die Zeit der Erkennung und Benennung von Schriftzeichen, Bildern und Farben. *Philosophische Studien*, 1885, 2, 635-650.
- CATTELL, J. MCK. The time it takes to see and name objects. *Mind*, 1886, 11, 63-65.
- DELATRE, P. A comparison of syllable length conditioning among languages. *International Review of Applied Linguistics*, 1966, 4, 183-198.
- FÓNAGY, I. Die Redepausen in der Dichtung. *Phonetica*, 1960, 5, 169-203.
- GLUKHOV, A. A. Statistical analysis of speech pauses for Romance and Germanic languages. *Soviet Physics. Acoustics*, 1975, 21, 71-72.
- GROSJEAN, F., & DESCHAMPS, A. Analyse des variables temporelles du Français spontané. *Phonetica*, 1972, 26, 129-176.
- GROSJEAN, F., & DESCHAMPS, A. Analyse des variables temporelles du Français spontané: II. Comparaison du Français oral dans la description avec l'Anglais (description) et avec le Français (interview radiophonique). *Phonetica*, 1973, 28, 191-226.
- GROSJEAN, F., & DESCHAMPS, A. Analyse contrastive des variables temporelles de l'Anglais et du Français: Vitesse de parole et variables, composantes, phénomènes d'hésitation. *Phonetica*, 1975, 31, 144-184.
- HANLEY, T. D., & SNIDECOR, J. C. Some acoustic similarities among languages. *Phonetica*, 1967, 17, 141-148.
- HANLEY, T. D., SNIDECOR, J. C., & RINGEL, R. L. Some acoustic differences among languages. *Phonetica*, 1966, 14, 97-107.
- JOHNSON, T. Lara, Lorena and Erika: Monolingual and bilingual development at age four. *1977 Mid-America Linguistics Conference*. Columbia: University of Missouri, 1978. Pp. 105-112.
- KOWAL, S., O'CONNELL, D. C., O'BRIEN, E. A., & BRYANT, E. T. Temporal aspects of reading aloud and speaking: Three experiments. *American Journal of Psychology*, 1975, 88, 549-569.

- KOWAL, S., O'CONNELL, D. C., & SABIN, E. Development of temporal patterning and vocal hesitation in spontaneous narratives. *Journal of Psycholinguistic Research*, 1975, 4, 195-207.
- O'CONNELL, D. C., & KOWAL, S. Cross-linguistic pause and rate phenomena in adults and adolescents. *Journal of Psycholinguistic Research*, 1972, 1, 155-164.
- OSSER, H., & PENG, F. A cross-cultural study of speech rate. *Language and Speech*, 1964, 7, 120-125.
- SABIN, E. J., CLEMMER, E. J., O'CONNELL, D. C., & KOWAL, S. A pausological approach to speech development. In A. W. Siegman & S. Feldstein (Eds.), *Of time and speech*. Hillsdale, N.J: Erlbaum, 1979. Pp. 35-55.
- WALLIN, J. E. W. Researches in the rhythm of speech. *Studies from the Yale Psychological Laboratory*, 1901, 9, 1-142.

(Received for publication February 8, 1979.)