

Speech rhythm – intention or consequence? Cross-language observations on the hyper/hypo dimension

Diana Krull and Olle Engstrand

Department of Linguistics, Stockholm University

Formal and casual Swedish and Spanish were analyzed to replicate previous observations suggesting a temporal equalization of syllable-sized units in casual as opposed to elaborated speaking styles. Durations and standard deviations appearing in casual Swedish tended to resemble those in Spanish, and durational differences between unstressed and stressed units in casual Swedish were smaller than those observed in more elaborated Swedish. This corroborated the view that rhythmic patterns associated with stress- and syllable-timing may arise as consequences of more primary phonetic intentions rather than themselves representing such intentions.

1. Introduction

Last year, at FONETIK 2002, we reported some cross-language observations on speech timing in Swedish and Spanish (Engstrand & Krull 2002). Read and unscripted speech in Swedish and Argentinean Spanish were analyzed to test the hypothesis that temporal equalization of syllable-sized units typically occurs in casual as opposed to elaborated speaking styles. Results indicated 1) that durations and standard deviations appearing in unscripted Swedish tended to resemble those in Spanish, 2) that durational differences between unstressed and stressed units in unscripted Swedish were smaller than those observed in Swedish careful reading, and 3) that, in both languages and speaking conditions, the duration of syllabic units was partly determined by their size. These findings, however, were based on three speakers only. The purpose of the present paper is to see if the preliminary results could be replicated using more data. For convenience, the three original subjects are integrated in this report.

2. Methods

Unscripted monologues were produced by two male native speakers of Central Standard Swedish (JS and PL) and three male native speakers of Latin American Spanish (henceforth ‘Spanish’): one Argentinean (RF), one Mexican (FH) and one Guatemalan (JC). The read material consisted of Swedish and Spanish texts read in a formal speaking style. The Swedish texts were read by a female Central Standard Swedish speaker (GT), and by one of the male speakers (PL); the latter speaker, thus, produced both unscripted and read speech. Spanish texts were read by all three Spanish speakers who thus produced both unscripted and read speech. (Speakers JS, GT and RF were the subjects in Engstrand & Krull 2002.) All recordings were made in a sound-treated room using high quality professional equipment.

Parts of the digitized (16 kHz) material were labeled and analyzed using the Soundswell Signal Workstation. Utterances were segmented into syllable-sized units on the basis of auditory and spectrographic analysis. As explained in Engstrand & Krull (2002), these units were contoid-vocoid(-contoid) sequences reflecting opening-closing movements of the vocal tract. Single contoids were counted as unit onsets, and contoid clusters were counted as onsets if conforming to the 'sonority hierachy (Jespersen 1926, Engstrand & Krull 2001).

3. Results

Figure 1 depicts distributions of syllable-sized unit durations in Swedish speaker PL (left column) and Guatemalan Spanish speaker JC (right column). The durational range shown along the x axes is 0-600 ms in all four histograms. For both speakers, the top row represents careful reading and the bottom row represents unscripted speech. Units in prepausal position are removed throughout the data presentation. N is given for each case in Table 1.

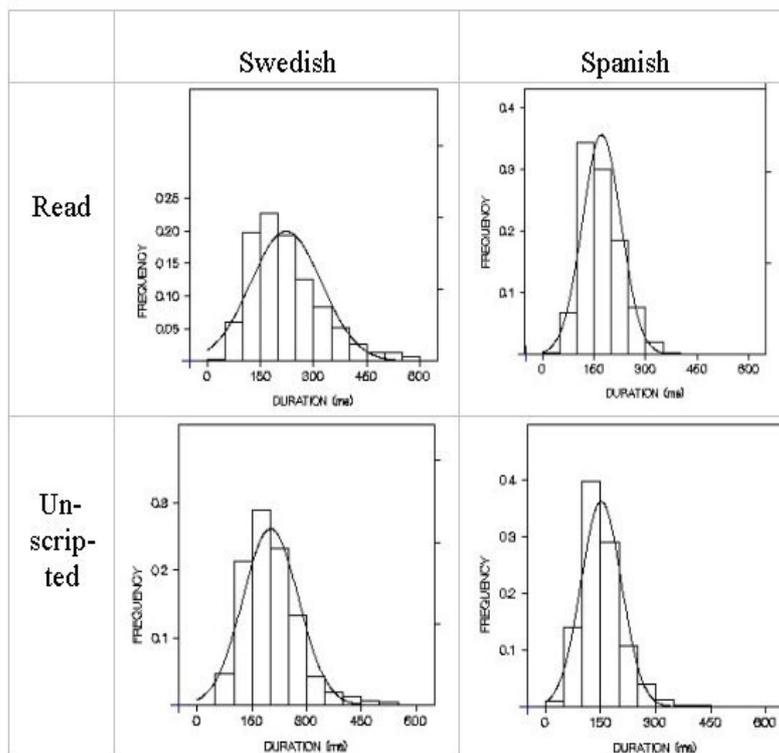


Figure 1. Syllable-sized unit durations (ms) in read and unscripted Swedish, and read and unscripted Spanish. A normal distribution is superimposed on each histogram.

For the Swedish data, the frequency distribution for the read condition is broader and flatter than that for the unscripted condition. It is also clear that the Spanish distributions are narrower than the Swedish distributions in both conditions. However, the distribution for the Swedish unscripted condition bears a closer similarity to the Spanish distributions than does that for the Swedish read condition. These patterns are in accordance with the observations previously made in Engstrand & Krull (2002).

Table 1. Durations (ms) for syllable-sized units in read and unscripted Swedish and Spanish.

| Language | Speaker | Condition | Mean | Std | N |
|-----------|---------|------------|------|-----|-----|
| Swe | GT | Read | 200 | 86 | 350 |
| Swe | JS | Unscripted | 178 | 62 | 306 |
| Swe | PL | Read | 223 | 102 | 455 |
| Swe | PL | Unscripted | 201 | 76 | 414 |
| Spa (Arg) | RF | Read | 156 | 59 | 167 |
| Spa (Arg) | RF | Unscripted | 155 | 51 | 287 |
| Spa (Mex) | FH | Read | 173 | 58 | 139 |
| Spa (Mex) | FH | Unscripted | 172 | 67 | 171 |
| Spa (Gua) | JC | Read | 172 | 56 | 337 |
| Spa (Gua) | JC | Unscripted | 152 | 55 | 523 |

Table 1 shows mean unit durations and standard deviations for the read and unscripted conditions in both languages. The durations pertaining to Swedish are consistently greater than those pertaining to Spanish. However, durations for the Swedish unscripted condition tend to fall inbetween those for the Swedish read condition and those for the Spanish read and unscripted conditions. An effect of speaking condition on the dispersion about the means is also evident such that standard deviations for the Swedish unscripted condition are clearly smaller than those for the read condition. The Spanish data display no such consistent effect.

Table 2. Mean durations (ms) for stressed and unstressed syllable-sized units in different speaking conditions in Swedish and Spanish. The rightmost column shows the difference between stressed and unstressed unit durations in the respective languages and speaking conditions (read and unscripted).

| Language | Condition | Speaker | Stress | Mean | Std | N | +/- stress |
|-----------|-----------|---------|--------|------|-----|-----|------------|
| Swe | Read | GT | y | 261 | 76 | 153 | 107 |
| | | | n | 154 | 60 | 197 | |
| Swe | Unscr | JS | y | 226 | 56 | 115 | 76 |
| | | | n | 150 | 47 | 191 | |
| Swe | Read | PL | y | 301 | 104 | 154 | 118 |
| | | | n | 183 | 74 | 301 | |
| Swe | Unscr | PL | y | 244 | 74 | 170 | 43 |
| | | | n | 201 | 76 | 413 | |
| Spa (Arg) | Read | RF | y | 198 | 64 | 58 | 64 |
| | | | n | 134 | 41 | 109 | |
| Spa (Arg) | Unscr | RF | y | 170 | 58 | 122 | 26 |
| | | | n | 144 | 42 | 165 | |
| Spa (Mex) | Read | FH | y | 221 | 54 | 43 | 70 |
| | | | n | 151 | 45 | 94 | |
| Spa (Mex) | Unscr | FH | y | 220 | 61 | 53 | 70 |
| | | | n | 150 | 58 | 118 | |
| Spa (Gua) | Read | JC | y | 212 | 46 | 108 | 59 |
| | | | n | 153 | 50 | 228 | |
| Spa (Gua) | Unscr | JC | y | 189 | 58 | 168 | 55 |
| | | | n | 134 | 44 | 355 | |

Table 2 presents mean durations for stressed vs. unstressed syllable-sized units in the two languages and speaking conditions. As also observed in Engstrand & Krull (2002), the

difference between the read and unscripted conditions is stress-dependent. For Swedish subject PL, for example, the stressed units in the read and unscripted conditions are 301 and 244 ms, respectively, and the corresponding figures for the unstressed units is 183 and 201 ms (i.e., the stressed units tend to be shorter than the unstressed units, a tendency also seen in Argentinean Spanish speaker RF). In Spanish, the pattern is similar but less extreme. For JS, for example, the stressed units in the Spanish read and unscripted conditions are 212 and 189 ms, respectively, whereas the corresponding figures for the unstressed units are 153 and 134 ms.

The differences between stressed and unstressed syllable durations in the respective languages and speaking conditions are summarized in the rightmost column of Table 2. The stress effects in the Swedish read condition (GT: 107 ms and PL: 118 ms) agree with the durational effect of stress observed in stress-timed languages (Eriksson 1991), whereas the durational quantum added by stress in the remaining conditions (unscripted Swedish as well as read and unscripted Spanish, range: 26-76 ms) would be more typical of syllable-timed languages.

4. Summary and conclusions

The present data have largely corroborated the preliminary conclusions drawn in Engstrand & Krull (2002). Thus, temporal equalization of syllable-sized units appears to take place in casual speech. Also, the previous observation that durational differences between unstressed and stressed units were smaller in unscripted Swedish than in careful reading was confirmed by the new data. Thus, the unscripted speaking condition was less typically stress-based than was the read condition. However, the final observation made in Engstrand & Krull (2002), that of a linear relationship between number of segments and duration of syllable-sized units, has not yet been put to test with the new data. Nevertheless, the new data have strengthened the general impression that speaking style may have a considerable effect on temporal aspects of stress- and syllable-timing. In conclusion, then, the original hypothesis, that the temporal aspects of speech rhythms can be explained as consequences of more primary phonetic intentions rather than themselves representing such intentions, remains a viable one (cf. Dauer, 1983; Strangert, 1985; Engstrand 1987; Eriksson, 1991).

5. References

- Dauer, R. (1983) Stress-timing and syllable-timing reanalyzed, *Journal of Phonetics*, 11, 51-62.
- Engstrand, O. (1987) Durational patterns of Lule Sami phonology, *Phonetica*, 44, 117-128.
- Engstrand, O. & Krull, D. (2001) Simplification of phonotactic structures in unscripted Swedish, *Journal of the International Phonetic Association*, 31, 41-50.
- Engstrand, O. & Krull, D. (2002) Duration of syllable-sized units in casual and elaborated speech: Cross-language observations on Swedish and Spanish. *Papers from FONETIK 2002, TMH-QPSR*, 44, 69-72.
- Eriksson, A. (1991) Aspects of Swedish speech rhythm, *Gothenburg Monographs in Linguistics*, 9, Department of Linguistics, University of Gothenburg.
- Jespersen, O. (1926) *Lehrbuch der Phonetik*. 4. Auflage, Leipzig & Berlin: Teubner.
- Strangert, E. (1985) Swedish speech rhythm in a cross-language perspective, *Acta Universitatis Umensis, Umeå Studies in the Humanities*, 69.