

Temporal structure in speech and music: The timing of proparoxytones in Italian folksongs

Teresa Proto

Academy of Creative and Performing Arts, Leiden University

It has often been observed (Lerdahl & Jackendoff 1983, Repp 1991, Besson & Schoen 2001, Brown 2001, Gilbers & Schreuder 2002, Schreuder 2006) that language and music are structurally similar. Both display some kind of hierarchical organization, alongside a temporal and a melodic structure; moreover, they have a common timeframe of acquisition. Despite these similarities, however, little is known about their actual interaction in vocal music in general, and in individual singing idioms, in particular. Given the nature of the two systems involved, interactions would be expected at least at the following levels: in the melody, within the metrical / rhythmical organization, and in the temporal structure. In this paper, I tackle the issue of how the temporal structure of a language (in the specific case, Italian) is accommodated in the timing and rhythm of the music in songs.

I use 'temporal structure' to refer to the durational patterning of events, which is measured by the time intervals between successive onsets (IOI). If we consider duration alone, a rhythm can be described as a sequence of IOIs between notes (in the case of music) / vowels (in the case of speech). Perceptually, rhythm appears to be built on temporal structures that rely on longer intervals—, as opposed to shorter ones (Fraisse 1956). Longer intervals also play a key role in *grouping*, as segmentation tends to occur at events with relatively long IOIs.

One of the methods most effectively used to investigate the temporal organization of speech and music is the Normalized Pairwise Variability Index (henceforth nPVI). nPVI is an equation that measures the (average) degree of durational contrast between adjacent elements in a sequence, such as successive notes in a musical phrase, or successive vowels in a spoken sentence. The greater the contrast between adjacent durations in a sequence, the higher the nPVI value of that sequence. While this measure was developed in the field of linguistics (Grabe & Low, 2002; Low, Grabe, & Nolan, 2000; Ramus, 2002), it has also been applied to instrumental music, and to investigate the relations between linguistic and musical rhythm (Patel & Daniele 2003, Daniele & Patel 2013, Temperley 2017).

In Italian, like other so-called 'syllable-timed languages', the nPVI value is fairly low (Arvaniti 2012 reports a value of 48.5), meaning that sentences show little contrast in successive vowel durations. Similarly, Daniele & Patel (2013) found that the nPVI score of *instrumental* music has remained stably low over time for Italian composers compared to German/Austrian composers. Interestingly, this finding was not supported by research carried out on *vocal* music by Temperley (2017), who found comparable nPVI scores for German and Italian (52.5 and 54.3 respectively, for not anonymous songs).

While the nPVI has proved to be a powerful measure for determining durational variability at the level of the sentence and the musical phrase, I argue that it fails to capture within-word rhythmical relations in speech or singing. In this paper, I therefore take a different approach and investigate durations in a corpus of Italian folksongs, focusing mainly on the setting of words with antepenultimate stress. In doing so, I build upon work by Proto & Dell (2013) and Hayes & Kaun (1996). The latter formulated a

“rule schema for textsetting” (1996: 260), which states that the natural phonetic durations of syllables should be reflected in the number of metrical beats they receive. This rule schema seems to be at work in English folksongs, where syllables tend to receive a proportional number of metrical beats, that reflects their natural phonetic duration. Thus, for instance, the initial syllable of *city*, which is phonetically quite short, is set most comfortably to a single metrical position, whereas the long syllable *town* is preferably set to two positions. Proto & Dell (2013) found that in Italian folksongs there exists a special rhythmical configuration in which violating Hayes & Kaun’s rule would result in an ill-formed setting. This configuration involves setting the immediately post-tonic vowel of a proparoxytone word to a longer note than both neighbouring vowels.

In this study, I carry out a thorough examination of the rhythmical sequences in which words with antepenultimate stress occur in Italian folksongs. The inquiry has revealed that the durational contrast observed in speech between tonic and post-tonic vowels (and due to stress-conditioned or final lengthening, see Marotta 1985, D’Imperio and Rosenthal 1999) can be neutralized in singing (e.g. by setting the vowels to notes of equal duration); however, if some degree of durational contrast is inserted between the three final syllables of the word, the immediately post-tonic vowel may not receive extra metrical beats.

During the presentation these findings will be illustrated, and they will also be discussed in the light of the controversial issue whether the trochaic foot built from the lexically established stress in Italian is syllable-based or mora-based (Krämer 2009). The latter analysis has encountered more opposition due to the difficulty in accepting weight-sensitivity in a language, like Italian, without a length contrast in vowels. The data presented here (and the interpretation thereof) may provide further support for the mora-based account.

References

- A. Arvaniti. 2012. The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics*, 40, 351-373.
- M. Besson & D. Schoen. 2001. Comparison between Language and Music. *Annals of New York Academy of Sciences*. 232-258.
- J. R. Daniele & A. D. Patel. 2013. An Empirical Study of Historical Patterns in Musical Rhythm Analysis of German & Italian Classical Music Using the nPVI Equation. *Music Perception: An Interdisciplinary Journal* 31. 10-18.
- M. D’Imperio & S. Rosenthal. 1999. Phonetics and Phonology of Italian Main Stress. *Phonology* 16. 1–28.
- P. Fraise. 1956. *Les structures rythmiques*. Louvain: Editions Universitaires.
- D. Gilbers & M. Schreuder. 2002. Language and music in optimality theory. [ROA 571-0103]
- B. Hayes & A. Kaun. 1996. The role of phonological phrasing in sung and chanted verse. *Linguistic Review* 13. 243–303. DOI: 10.1515/tlir.1996.13.3-4.2
- M. Krämer. 2009. *The Phonology of Italian*. Oxford: OUP.
- F. Lerdahl, & R. Jackendoff. 1983. *A generative theory of tonal music*. Cambridge, Mass.: MIT Press.
- E. L. Low, E. Grabe & F. Nolan. 2000. Quantitative characterisations of speech rhythm: Syllable-timing in Singapore English. *Language and Speech*. 43. 377-401.
- G. Marotta, *Modelli e misure ritmiche. La durata vocalica in italiano*. Bologna: Zanichelli, 1985
- A. D. Patel & J. R. Daniele. 2003. An empirical comparison of rhythm in language and music. *Cognition* 87. B35-B45. DOI: 10.1016/S0010-0277(02)00187-7
- T. Proto & F. Dell. 2013. The structure of metrical patterns in tunes and in literary verse. Evidence from discrepancies between musical and linguistic rhythm in Italian songs. *Probus* 25. 105–138.
- F. Ramus. 2002. Acoustic correlates of linguistic rhythm: Perspectives. In B. Bell & I. Marlien (Eds.), *Proceedings of Speech Prosody, Aix-en-Provence* (pp. 115-120). Aix-en-Provence: Laboratoire Parole et Langage.

- B. H. Repp 1991. Some cognitive and perceptual aspects of speech and music. In *Music, Language, Speech and Brain. Proceedings of an International Symposium*, ed. by J. Sundberg, L. Nord and R. Carlson, 257–268. London: MacMillan.
- M. Schreuder 2006. Prosodic Processes in Language and Music. PhD diss., University of Groningen.
- D. Temperley. 2017. Rhythmic Variability in European Vocal Music. *Music Perception: An Interdisciplinary Journal* 35. 193-199.