Journal of Phonetics 67 (2018) 21–33

Contents lists available at ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/Phonetics

Research Article

Tonal and morphophonological effects on the location of perceptual centers (p-centers): Evidence from a Bantu language

Kathryn Franich

Department of Linguistics, University of Chicago, 1115 E. 58th Street, Rosenwald Hall, Room 224, Chicago, IL 60637, USA

ARTICLE INFO

Article history:
Received 9 May 2017
Received in revised form 2 November 2017
Accepted 7 November 2017

Keywords:
Rhythm
Speech perception
Speech timing
Tone
Morphology
Phonology

ABSTRACT

Perceptual centers (or 'p-centers') correspond to the perceptual moment of occurrence of a syllable or word, and are crucial in the perception of speech rhythm. A metronome alignment task was used to investigate how tone and prenasalization—two elements which affect speech timing and which also interact acoustically—influenced p-center location in Medʉmba, a Grassfields Bantu language. Plain CV words bearing low tones were found to have p-centers which were later (farther from consonant releases and closer to vowel onsets) than those bearing high tones, but the observed effect was not present in prenasalized words. We attribute this difference to the effects of tone depression and slope leveling in prenasalized forms. While prenasalization generally led to earlier p-centers (mirroring effects found for onset clusters in other languages), forms with morphologically-derived prenasal onsets behaved more like plain CV forms, suggesting that nasal prefixes do not contribute to p-center timing. Our findings for derived prenasal sequences parallel similar articulatory findings for languages with simplex onset coordination, where consonant 'clusters' actually behave as separate timing units.

1. Introduction

Rhythmic timing in speech has been of longstanding interest to linguists, though identifying precisely which elements in speech should be expected to behave as rhythmic, and in what way, remains a problem. Early approaches to rhythm analysis proposed a central role for *isochrony*, or equal timing between elements in speech, forming the basis of the distinction between hypothesized rhythm classes such as stress-timed and syllable- or mora-timed (Abercrombie, 1965, 1967; Bloch, 1950; Pike, 1945). Acoustic evidence for isochrony has been notoriously weak, however, with most studies providing little evidence for equal spacing of stress beats, syllables, or moras in any language under investigation (Bolinger, 1965; Dauer, 1983, 1987; Delattre, 1966; Roach, 1982; Shen & Peterson, 1962). This fact has led some to abandon the notion of isochrony altogether in characterizing speech rhythm, focusing instead, for example, on crosslinguistic differences in patterns of durational variability of consonant and vowel intervals in the speech stream (Dellwo, 2006; Grabe & Low, 2002; Ramus, Nespor, & Mehler, 1999). While these approaches have provided some empirical support for distinctions in rhythm class between languages (or at least for a continuum of rhythm types), different duration-based metrics have sometimes led to distinct rhythmic categorizations of the same language, or lack of a clear categorization for some languages. Furthermore, many factors, such as the type of speech elicited (e.g. isolated sentences versus conversational speech), have been found to influence results (see Arvaniti, 2009, 2012 for comprehensive overviews). Crucially, these approaches also fail to address the strong intuition among listeners of different languages that speech sounds isochronous (Lehiste, 1977). Indeed, a growing body of research shows that *perceived* isochrony in speech is beneficial for speech processing (Brown, Salverda, Dilley, & Tanenhaus, 2011; Dilley & McAuley, 2008; Dilley, Mattys, & Vinke, 2010, 2012). To better understand the facilitative role of perceived isochrony in speech processing, a clearer picture of how it relates to different acoustic and structural factors in language will be necessary.

The connection between perceived isochrony and phonetic properties has received a great deal of attention within studies of *perceptual centers*, or 'p-centers', which reflect the perceptual moment of occurrence of a syllable and serve as the locus of perceived isochrony (Fowler & Tassinary, 1981; Morton, Marcus, & Frankish, 1976; Rapp, 1971). P-center effects were first described as such by Morton et al. (1976) in relation to a

E-mail address: kfranich@uchicago.edu
URL: http://home.uchicago.edu/kfranich/

https://doi.org/10.1016/j.wocn.2017.11.001
0095-4470/

memory experiment examining whether number recall could be improved if digits were presented at evenly-spaced temporal intervals in a list. In this experiment, it was found that digits manipulated into isochronous sequences based on onset of acoustic energy were not perceived to sound evenly-spaced, and in fact sounded quite arhythmic to listeners. When listeners were asked to adjust syllables to sound more regularly-timed, the adjustments made suggested that p-centers were tied neither exclusively to the syllable onset nor the vowel onset, but somewhere in between; the authors concluded that it must correspond to some function relating the two.

Other work has tried to identify more precisely the acoustic bases for the p-center effect. Marcus (1981) used a similar design—termed the *rhythm adjustment method* (see Villing, Repp, Ward, & Timoney, 2011 for further details)—where subjects evaluated temporal regularity between a set of alternating base and test sounds and adjusted test sounds such that they sounded more evenly timed with base sounds. Marcus found that a CV syllable's p-center tended to occur close to the vowel onset, but was influenced both by the duration of its onset consonant and, to a lesser degree, by its vowel and coda consonant durations. Most notably, increased onset consonant duration led to a shift in p-centers away from the vowel onset and to a point earlier in the syllable. Similar results were found using the same method by Pompino-Marschall (1989) and by Harsin (1997). Results of all of these studies also pointed to the importance of the role of the distribution of acoustic energy at critical frequency bands in determining p-center locations, a possibility which was tested explicitly in perception experiments and supported by Howell (1984, 1988), Scott (1998), Scott and Howell (1992).

Though the p-center effect is characterized as a perceptual phenomenon, various speech production studies examining speech timing have found similar results to those found in perception studies. Allen (1972) found that, when instructed to tap 'on the beat' of designated syllables within a sentence, subjects located their taps just before vowel onsets, and the precise duration between tap and vowel onset varied as a function of the duration of the prevocalic consonant. Similar results were found by Rapp (1971) when having subjects repeat noncewords to a regular metronome pulse (a method we refer to as the *metronome asynchrony procedure*). Fowler and Tassinary (1981) also had subjects repeat syllables to a metronome pulse, this time examining effects of different consonant clusters on p-center alignment. They found that, while speakers tended to align utterances such that vowel onsets occurred with the metronome pulse, as the number of segments in the syllable onset increased, the pulse occurred farther from the vowel onset, occurring instead somewhere within the onset consonant cluster. This finding was replicated by Šturm and Volín (2016) for the Czech language.

Additional studies have investigated the possible role of other aspects of phonology and syllable structure in affecting p-center location. Cooper, Whalen, and Fowler (1986), using the rhythm adjustment method, showed that onset cluster effects on p-centers are not influenced by the category of the segments involved, but rather depend exclusively on the duration of the cluster sequence as a unit. Further work has investigated contributions of the syllable rime to p-center location. Cooper et al. (1986, 1988) showed that duration of the syllable rime exerted a small but consistent influence on the location of p-centers, but not nearly as strong an effect as had been found with syllable onsets. The authors conclude from this asymmetry that syllable structure is an important factor in determining the effect of a particular segment on p-center timing. Similar results were obtained by Šturm and Volín (2016), who found that vowel duration and coda duration had a far weaker effect on p-center timing than onset duration in Czech.

Browman and Goldstein (1988) compare p-center results to those found with respect to an articulatory phenomenon known as the *C-center effect* (short for 'consonant center'). This refers to the strong tendency for the timing of onset consonantal gestures to correspond as a unit to other articulatory landmarks such as the target of the vowel gesture and the acoustic release of the syllable coda,[1] such that the midpoint of the onset consonant or cluster of consonants remains stably timed with these landmarks. Just as the p-center of a syllable moves consistently away from the syllable's vowel onset as onset duration increases, so, too, does the C-center. Thus, it would appear that p-centers may be tied to articulatory events, as has also been proposed by Fowler (1979, 1983). Also similar across the two phenomena is the fact that both appear far more dependent on properties of syllable onsets than syllable codas. These results are interpreted within the framework of Articulatory Phonology (Browman & Goldstein, 1990, 1998, 2000) as evidence that onset consonantal gestures, at least in English, are coupled *in-phase* with the following vowel, whereas codas are coupled 180° *anti-phase* to the vowel gesture. Thus, when the structure of the onset is changed through the addition of a consonant, the coordinative structure of the onset + vowel portion is affected as a whole, whereas addition of a coda consonant affects coda timing, but not vowel timing (at least not to a large degree).

Further work has set out to explicitly tie the p-center effect to articulatory events, but results have been mixed. de Jong (1994), by examining stimuli extracted from an articulatory database, found that while articulatory gestures predicted p-center location just as well as (if not better than) acoustic events, no single articulatory landmark (e.g. tongue tip minimum position or jaw maximum position) acted as the sole correlate of p-center location. Rather, the author concluded that p-centers correspond to a complex of articulatory events in the syllable. Likewise, Patel, Naito, and Löfqvist (1999) examined both acoustic and kinematic data in relation to p-center location, finding that no one landmark from either type of measure was exclusively tied to p-center location.

Despite the elusive nature of acoustic and articulatory landmarks associated with p-centers, the phenomenon remains remarkably robust across speakers of a language, and even across languages. Hoequist (1983), using the rhythm adjustment method, found that speakers of English, Spanish, and Japanese all tend to align p-centers around the onset of the vowel in monosyllables. Barbosa, Meireles, and Vieira (2005), using the metronome asynchrony method, found similar results for Brazilian Portuguese. Crosslinguistic results have revealed some interesting language-specific behavior of

[1] Note that different studies have used articulatory landmarks other than the release of the syllable coda as anchor points for measuring C-center effects; see Tilsen (2012) for an overview.