



Contents lists available at ScienceDirect

Journal of Phonetics

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

Research Article

Acoustic correlates of focus in Marathi: Production and perception

Preeti Rao^{a,*}, Niramay Sanghvi^a, Hansjörg Mixdorff^b, Kamini Sabu^a^a Department of Electrical Engineering, Indian Institute of Technology Bombay, India^b Beuth University Berlin, Germany

ARTICLE INFO

Article history:

Received 12 December 2016

Received in revised form 12 June 2017

Accepted 14 July 2017

Available online 4 August 2017

Keywords:

Acoustic prominence

Focus

Trade-offs in acoustic correlates

Fujisaki modeling

Marathi prosody

ABSTRACT

Focus or prominence is an important linguistic function of prosody. The acoustic realisation of prominence in an utterance, in most languages, involves one or more acoustic dimensions while affecting one or more words in the utterance. It is of interest to identify the acoustic correlates as well as their possible interaction in the production and perception of focus. In this article, we consider the acoustics of focus in Marathi. Previous studies on Hindi, the more researched member of the Indo-Aryan family, have reported that the well-known rising F_0 pattern on non-final content words in an utterance becomes hyper-articulated when the word is in focus. The associated F_0 excursion, duration and intensity increase and are accompanied by post-focal compression of pitch range. A preliminary goal of the present study was to verify whether Marathi exhibits similar behaviour. We used Subject–Object–Verb (SOV) structured utterances with elicited focus on each word by 12 native Marathi speakers. We observed that each narrow focus location is accompanied by a distinct set of local and global acoustic correlates in F_0 , duration and intensity which closely parallel previous observations on Hindi. F_0 cues were also examined via the accent command amplitudes of the Fujisaki model. F_0 range, duration and intensity were found to vary significantly with focus condition prompting a study to examine their relative importance in the perceptual judgement of focus. Perception testing with synthetically manipulated utterances revealed that duration cues are interpreted in a categorical manner, relatively uninfluenced by the pitch cues. Only when duration is ambiguous, does the on-focus F_0 cue appear to play a role. An explanation for this may lie in the normal F_0 -rise characteristic of the content words in Marathi, making pitch a less dependable functional cue for focus.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

Focus refers to the marking of new information in an utterance over given information by emphasizing it in some way. Focus may be of different widths with “broad focus” indicating all new information, and “narrow focus” indicating that one word or phrase at a specific location in an utterance is emphasized over all others. Focus is typically achieved via the manipulation of prosody although in some languages it may be realised by changing word order (Swerts, Krahmer, & Avesani, 2002). The phonetic marking of focus itself is based on a common set of acoustic attributes although the precise cues, and the relative importance of each, are language dependent (Koreman et al., 2009). Although prominence is typically associated with a word, the acoustic cues may be described at the word level or sub-word levels such as syllable

or even vowel nucleus (Rosenberg & Hirschberg, 2009). Experimental studies as early as 1958 on English demonstrated that manipulating duration, fundamental frequency (F_0) and intensity influenced listeners’ judgement of stress on a syllable (Fry, 1958). Although segmental quality can also be a factor in stress or prominence perception due to possible hyper-articulation (Cho, Lee, & Kim, 2011), nearly all languages use one or more of the following acoustic dimensions of speech: F_0 , duration and intensity. Systematic variations of these properties at the word or syllable level are utilised both in the production and perception of focus. Accentuation of a segment is often accompanied by the deaccentuation of other segments leading to further enhancement of the prominence effect. In stress accent languages, it is the lexically stressed syllables that undergo modification under changing focus (Bolinger, 1986). The acoustic correlates of focus in a particular language and their dependence on the location and type of focus are of interest both from a scientific viewpoint as well as for speech technology applications such as automatic speech

* Corresponding author.

E-mail address: prao@ee.iitb.ac.in (P. Rao).

understanding, computer-aided pronunciation training and text-to-speech (TTS) synthesis. In the latter, accurate models for prosody can serve to generate more intelligible and natural speech.

There have been numerous studies about the phonetic realisation of focus in several languages using elicited utterances by native speakers. Breen, Fedorenko, Wagner, and Gibson (2010) studied the prosodic realisation of information structure in American English, as a representative West-Germanic language. Based on utterances elicited in response to *wh*-questions, the acoustic properties of narrow focus were measured in different sentential locations. It was observed that focused material is produced with longer duration, higher *F0* and greater intensity than non-focused material. The strongest indicators for discriminating the different focus conditions were word duration, word maximum intensity, word mean and maximum *F0*. In an earlier study specifically on English declarative intonation, Xu and Xu (2005) reported that the pitch range of the focused word is expanded, the pitch range of the post-focus items, if any, is compressed and lowered, and the pitch range of the pre-focus items, if any, remains neutral.

Acoustic properties of prominence marking in German and Norwegian were compared by Koreman et al. (2009) to find that the former used *F0* and intensity cues while Norwegian relied on duration cues. In a more recent study on German, perceptual prominence of a word was observed to be strongly correlated with *F0* transition, syllable duration, maximum intensity and mean harmonics-to-noise ratio (Mixdorff et al., 2015). Unlike English, German and Greek, Georgian is found to use duration to signal focus, with focus showing very little influence on intonation (Skopeteas & Féry, 2010). Turkish uses duration and intensity to cue focus (Ipek, 2011). It was reported that while on-focus pitch variation was absent in Turkish, narrow focus in sentence initial position elicited post-focal pitch compression. A perception experiment with the same recorded utterances across different focus locations showed that, in fact, initial focus was recognized most accurately by listeners. In a tone language like Mandarin, on-focus increase in *F0*, duration and intensity have been observed across dialects with post-focal compression of intensity and pitch ranges playing a prominent role with Beijing Mandarin speakers (Chen, Wang, & Xu, 2009). Korean also shows both narrow focus and post-focal effects that involve all the acoustic parameters (Lee & Xu, 2010).

Most studies investigating the acoustic correlates of focus have been based on production, i.e. by using measurements on elicited speech by native speakers. More definitive conclusions can be expected from the systematic testing of focus perception by native listeners provided suitable stimuli are available. The relationship between variations of stimulus parameters and perceived prominence can reveal the importance of the different acoustic cues provided they can be teased apart, which is usually possible only with synthesized stimuli. There have been a few studies of this type on intonation based features that have attempted to relate *F0* variation to perceived prominence (Gussenhoven, Repp, Rietveld, Rump, & Terken, 1997; Gussenhoven & Rietveld, 1988; Ladd & Morton, 1997; Terken, 1989; Terken & Hermes, 2000). The role of other prosodic parameters—mainly intensity and segmental durations—has been studied much less. Vainio and

Järvikivi (2006) in the course of studying the dependence of prominence perception on tonal features in Finnish, found that intensity has an effect and that speakers could exploit a perceptual trade-off between pitch height and intensity increase. That languages differ in the degree to which they exploit duration, *F0* and intensity in production but also to some extent in perception was recognised in the course of a recent large study on information structure across languages (Andreeva, Barry, & Koreman, 2015).

Among Indo-Aryan languages, Hindi has been studied relatively extensively for its prosody, as also specifically for the prosody of prominence. Hindi is a head-final (SOV) language with relatively free word order. While particles (equivalent to the English words “only”, “also” or “even”) can serve to mark focus, this is not necessarily the case (Patil et al., 2008). Several studies on Hindi intonation have noted that while declarative sentences follow the universal pitch down-step pattern, the individual (non-final) content words each show an *F0* rise from left to right across the word (Harnsberger, 1996; Moore, 1965; Pandey, 2015; Rajendran & Yegnanarayana, 1996). Harnsberger (1996) assigned an LH intonation pattern to all but phrase-final words in his study with 3 native speakers, and observed that syllable weight had no role to play in this. Similar observations by several others have led to the view that each content word constitutes a separate phrase (Féry, 2010; Patil et al., 2008). Several other Indian languages share this characteristic (Féry & Gisbert Fanselow, 2008; Hayes & Lahiri, 1991; Khan, 2007, 2016) including Marathi (Rao & Srichand, 1996). The property has also been exploited for word segmentation of continuous speech in Hindi (Rajendran & Yegnanarayana, 1996; Rao & Srichand, 1996). The *F0* rise characteristic may mean that the more significant *F0* changes would necessarily occur towards the end of the word which raises the question of the “lexically stressed” syllable within a word. Hindi and Marathi are considered syllable-timed languages and lexical stress per se is not a well understood aspect. However there have been some commentaries that assign lexical stress based on syllable weight in Hindi (Dyrud, 2001) and Marathi (Dhongde & Wali, 2009) although the acoustic correlates of stress are unknown or even considered non-existent as in some other South Asian languages (Khan, 2016). Ohala (1986) suggests that stress is weaker than in English but only one syllable can be made prominent in a word.

Given the word-level *F0* rise characteristic, focus, in Hindi, has been shown to bring in additional prosodic effects, namely (i) exaggeration of the *F0* rise on the focus word by way of greater excursion (Dyrud, 2001; Harnsberger 1996, 1999; Moore 1965; Patil et al., 2008). (ii) compression of the post-focal pitch range although rising pitch accents are still realised on the post-focal content words (Patil et al., 2008). Pre-focal elements do not undergo modification. Genzel and Kügler (2010) found that contrastive focus in Hindi is linked with an increase in duration of the stressed syllable as well as increased *F0* rise across the word that came as much from lowering the L tone as raising the H tone in the hammock shape. Puri (2013) found the main acoustic correlates of focus in Hindi by bilingual speakers (of Indian English) to include increased duration, as well as an increase in *F0* excursion, both on the focused element, and post-focal reduction in

Download English Version:

<https://daneshyari.com/en/article/5124090>

Download Persian Version:

<https://daneshyari.com/article/5124090>

[Daneshyari.com](https://daneshyari.com)