

## Phonetic realization of focus in Persian

VAHIDEH ABOLHASANIZADEH

*Shahid Bahonar University of Kerman*

SHOJA TAFAKKORI REZAYI

*Razi University, Kermenshah*

AZADEH SHARIFIMOQADAM

*Shahid Bahonar University of Kerman*

### Abstract

*The present study investigates how focus is phonetically realized in declarative and interrogative sentences in Persian. Focus is usually interpreted as contrastive focus in this language. That is, the contrasted element is chosen out of a closed set of entities and bears heavy stress. In this study, 12 native speakers of Persian recorded short declarative and interrogative sentences including Clitic Group and Phonological Word in neutral and focal conditions. The results show small acoustic differences in duration, intensity and spectral information between initial and final accented target words in neutral and focus conditions in Persian, by the side of substantial differences in  $f_0$ .*

**Keywords:** *focus, clitic group, phonological word, accent*

### 1. Introduction

One of the primary functions of prosody is to provide cues regarding the informational structure of discourse. Generally, words carrying new or important information in a given discourse are focalized in the utterance. Special weight can be given to any part of the utterance by using lexical, syntactic and intonation means. This is termed narrow focus as opposed to broad focus in which all parts of the utterance are given equal prominence

*Language Design* 14 (2013: 5-18)

(Ladd 1980). Contrastive focus, which is the main object of the study, is a subset of narrow focus whose function is to indicate an exclusive selection of an alternative out of a group of two or more possibilities. Focus for contrast is traditionally distinguished from focus for intensification, which is simply an equivalent means to using an intensifying adverb ( Hirst & Di Cristo 1988).

Focus, which is generally equivalent to emphasis, is a communicative function known to mainly manifest itself through  $f_0$  variations (Ladd, 1996 and references therein). According to Bolinger focus in English gives prominence to the syllables that are lexically stressed, primarily by assigning them a pitch accent. Many languages express focus in the prosodic phonology, and often the phonology and morphology appear to have arranged matters such that the focused part is more salient, and is characterized by greater pitch excursions (Gussenhoven, 2004). The present study is designed, as its primary goal, to reevaluate this view on the basis of Persian data. The secondary goal, which is necessitated by the first, is to understand the detailed  $f_0$  contours and their alignment with segmental materials as related to focus. We will start with a brief overview of how focus is treated in the American autosegmental-metrical (AM) framework. De Jong and Zawaydeh (2002) reasoned that stress and focus should expand the linguistically specified quantity differences but have no effect on or eliminate voicing differences, which is exactly what they found. While stress increased quantity differences, the difference between phonemically long and short vowels was larger in stressed syllables; the voicing effects were the same in stressed and unstressed locations. Here, just as voiceless stops in Arabic shorten previous vowels, English voiceless stops shortened previous vowels.

## **2. The AM theory**

In the AM theory, unlike the British tradition where nuclear tones are often described as contours, intonation is described in terms of two level tones—H and L (Ladd, 1996; Pierrehumbert, 1980). It is also assumed that pitch accents, which consist of either a single tone or two successive tones, are “phonologically located on metrically prominent syllables” (Pierrehumbert, 2000, p. 20). The AM theory assumes that there are two levels of phrasing in English intonation: the intermediate phrase and the full intonation phrase. An

intermediate phrase is marked by a phrase accent at its right (or left) edge (H- or L-), and an intonation phrase is marked by a boundary tone at its right edge (H% or L%).

One of the basic notions of the AM theory is that the components of intonation and their organization are defined primarily in form, while the meanings of these components are assigned only after their forms are established. The form of an intonation component involves two aspects. The first is its shape and relative height. In this theory, for example, a H is identified as corresponding to an  $f_0$  peak while a L an  $f_0$  valley. The second formal aspect of a tonal component is its relative prominence. This is mainly determined by auditory impression in both approaches (Cruttenden, 1997; Wightman, 2002).

### **3. 3. Focus as a communicative function**

From a functional perspective, focus refers to an emphasis on some part of a sentence as motivated by a particular discourse situation. For example, a narrow focus is put on “red” when “Your eyes are red” is said in response to “What’s wrong with my eyes?” In contrast, the focus will be on “eyes” when the question is “Is my nose red?” In this sense, focus is a discourse function P serving to highlight a particular piece of information against information already shared by the conversation participants (see Bolinger, 1972; van Heuven, 1994; Ladd, 1996; Xu , et.al, 2005 and Gussenhoven, in press). It should be noted that in this functional “definition” nothing is said about what the phonetic or phonological form of the highlighting is. Critically, once a word or syllable receives the nuclear accent, that accent is the only intonation component it carries. Furthermore, in both theories, by definition, there are no more prominence-related pitch events after the nuclear accent (Xu, et.al, 2005).

Kiss (1998) provides an elaborated definition of focus. She distinguishes between two types of focus and calls them as identificational focus and information focus. She suggests that identificational focus expresses exhaustive identification and moves into the specifier position of a functional projection, and on the contrary, information focus marks the non-

presuppositional nature of the information it carries and does not involve movement. According to Kiss every sentence must have an information focus, but the presence of an identificational focus is not obligatory in the sentence. Scabocsi (1997) argues the same lines for Hungarian. She states that a DP in derived focus position receives an exclusion-by-identification interpretation, while DP in situ does not.

Turning to Persian, identificational focus is usually interpreted as contrastive focus in this language. That is, the contrasted element is chosen out of a closed set of entities. By contrastive focus here we mean a focus that operates on a closed set of entities whose members are known to participants in the discourse. The following example illustrates a case of contrastive focus in Persian.

- (1) man be HASAN eteraz kard-am.  
I to H objection do.PST-3SG  
“I objected to Hasan.”  
(as opposed to someone else.)

The DP Hasan bears heavy stress, and receives contrastive reading. In the absence of a heavy stress, the DP in situ receives an interpretation showing information focus. Karimi (2005) believes that not only specific DPs, but also indefinite NPs along with Predictive bare nominals can function as contrastive focus in Persian, as illustrated respectively in (2) and (3). The reason is that nonspecific elements, when in focus, are interpreted as names of properties. In these cases the inherently nonspecific element denotes a referent that is a member of a relevant set, present in the discourse.

- (2) be Hasan ye xodkar dad-am  
To Hasan a pen give.PST-1SG  
“It was a pen( as opposed to a pencil) that I gave to Hasan.”  
(3) be Hasan xodkar dad-am  
To hasan pen give.PST-1SG  
“It was pen that I gave to Hasan.”

Kiss (1998) suggests that certain elements carry an inherent focus feature. These elements are only-phrases and wh-phrases. According to Karimi (2005) these phrases receive exhaustive focus interpretation in Persian. Persian allows two elements bearing contrastive focus in a single sentence only if at least one of them has inherent focus:

- (4) man be HASAN faqat ye xodkar dad-am.  
I to Hasan only a pen give.PST-1SG  
“It was to HASAN that I gave only pen.”

Miyagawa (1997), Karimi (1999c, 2003d), Kiss(2003) believe that focus movement is an instance of move triggered by a discourse- functional feature. It has also been argued that this movement places the focused element in the Spec of a functional focus phrase (Kiss 1994, 1998, 2003). In Persian contrasted elements bear heavy stress and optionally move into the Spec of FocP as in (5).

- (5) a. man ALI ro tu xiyabun did-am.  
I Ali SOM in street see.PST-1SG.  
“It was Ali whom I saw in the street.”  
b. [FocP ALIi ro [ man ti tu xiyabun did-am]].

Karimi and Taleghani (2003), and Karimi (2005) distinguish between two types of wh-phrases in Persian and state that just the wh-arguments rather wh-adjunts have focus interpretation. According to them wh-arguments can remain in situ and consequently receive information focus reading. Alternatively, they may move into the Spec position of FocP, where they get contrastive focus interpretation as (6) illustrates.

- (6) a. to ki ro tu xiyabun didi?  
You whom in street see-2SG  
“Whom did you see in the street?”  
b. [FocP ki ro to tu xiyabun didi?]

To sum up it could be said that the focused elements in Persian necessarily carry heavy stress, but their movement is optional, that is they can remain in

situ or move into the Spec of FocP. In the case of non wh-phrases even in both in-situ and derived positions, they carry contrastive reading, on the contrary as for wh-arguments, there is a subtle difference with regard to the type of focused readings, so that when they remain in-situ, they have information focus reading and alternatively, when they move into the Spec of FocP they get contrastive focus interpretation.

#### **4. Data collection and methodology**

The materials were presented to 12 educated native speakers in random order, blocked by intonation contour. Speakers were recruited from the University of Tehran and were aged between 26 and 37. They read each sentence twice and were freely allowed to repeat themselves if they thought they hadn't read a sentence correctly. The two best versions of each sentence by each speaker were selected; often these were the only two that were produced. In a few cases, we decided to discard utterances because of disfluencies, in which case all versions of the sentence were discarded. Of the  $2(\text{clitic}) \times 2(\text{voice}) \times 3(\text{focus}) \times 2(\text{mode}) \times 12(\text{speakers}) = 586$  utterances 572 were thus analyzed with the help of Praat [10]. We determined all segment boundaries in the target words, including that between stop closure and stop burst, for both voiced and voiceless plosives. In the case of voiced plosives, this meant that we had intervals of zero duration in a number of cases. Initial [t]'s were only measured for their bursts, as no reliable indication of the beginning of the closure was available. We measured  $f_0$ , segment durations, mean segment intensity levels, F1, F2 and F3 of the two vowels, and the Centre of Gravity of the stop bursts and [ʃ]. Subsequently, we averaged all values over the repetitions and performed analyses of variance (repeated measures) with STRUCTURE (clitic vs nonclitic), MODE (declarative vs interrogative), FOCUS (neutral, post-focal, focal), VOICE (voiced vs unvoiced), and SEGMENT. The levels for SEGMENT varied with the dependent variable. For duration, these included t-burst, [a], labial closure, stop burst, [e], [ʃ] and [e]; for formants and intensity these were the two vowels, and for Centre of Gravity the stop bursts and [ʃ].

## **5. Discussion**

The following analyses were performed to help us find the difference between focus and neutral. So a set of repeated measures ANOVAs were performed to analyze the importance of different phonetic cues for focus realization. We report the result for dependent variables such as duration, energy, formants as follows:

### *5.1. Duration*

The effect of Structure and mode on [t-burst] shows that its duration in the accented syllable is 8.59 ms longer than in the unaccented one and in the statements 6.76 ms longer than in the questions. The effect of voice on [ɑ] duration is 30.17 ms longer before the voiced consonant [b] than before the unvoiced [p].

The effect of STRUCTURE on [b,p] duration is due to 5.96 ms longer duration in the accented syllable than in the unaccented syllable and the effect of VOICE is due to 12 ms longer duration of voiceless consonant.

The effect of VOICE and MODE on [b,p -burst] shows that the duration of the burst of voiceless consonant is 34.40 ms longer than the duration of the burst of voiced consonant. The effect of MODE is due to 2.39 ms longer burst in the statements than in the questions.

The effect of VOICE on [e] duration is due to 16 ms longer duration after the voiced consonant.

The effect of FOCUS on the last [e] is due to 64.41 ms longer duration in the neutral condition than in the focal condition and the effect of MODE shows that duration of the last [e] is 108 ms longer in the questions than in the statements.

Table 1 The effect of structure, voice, focus and mode on the duration of segments.

	<i>Structure</i>	<i>Voice</i>	<i>Focus</i>	<i>Mode</i>
<b>T (burst)</b>	F=9.12 P=.012	NA	NA	F=8.65 P=.013
<b>a</b>	NA	F=125.28 P<.001	NA	NA
<b>b/p</b>	F=5.41 P<.01	F=14.01 P<.01	NA	NA
<b>b/p (burst)</b>	NA	F=166.28 P<.01	NA	F=10.27 P<.01
<b>e</b>	NA	F=38.79 P<.01	NA	NA
<b>f</b>	F=8.46 P<.01	F=39 P<.001	F=15.04 P<.01	F=11.53 P<.001
<b>e</b>	NA	NA	F=147.6 P<.01	F=108.54 P<.01

The segments in the stressed syllable are longer than in the unstressed syllable. Although the effect of focus on the duration of segments is not significant, segments in the focus conditions are slightly longer than in the neutral condition.

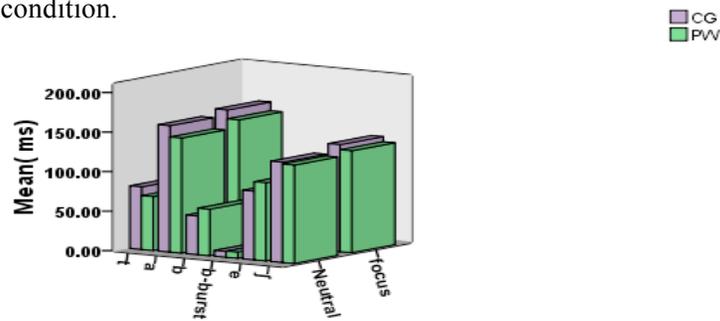


Fig 1. The mean duration of segments in Neutral and Focus conditions.

### 5.2. Energy

An analysis of the variance (repeated measures) was performed on the energy of the segmented sections of the target words, with segments [a] and [e].

Table 2 The effect of structure, voice, focus and mode on the intensity of vowels.

<i>segment</i>	<i>Structure</i>	<i>Voice</i>	<i>Focus</i>	<i>Mode</i>
<b>a</b>	F=9.96 <i>P</i> =.010	NA	F=14.22 <i>P</i> =.003	NA
<b>e</b>	F=5.28 <i>P</i> =.042	F=6.52 <i>P</i> =.032	F=14.03 <i>P</i> =.003	F=6.63 <i>P</i> =.026

The result shows that the energy of vowel [e] in the accented syllable is 1.68 dB more than that in the unaccented syllable. In the focal condition, vowel [e] has 1.59 dB more energy than that in the neutral condition. In the questions, [e] has 1.64 db more energy than in the statements. The main effect for STRUCTURE ( $p=.010$ ) on vowel [a] is due to 2.97 dB more energy in the accented syllable and the main effect for FOCUS is due to 1.26 dB more energy in the focal condition (Fig 2).

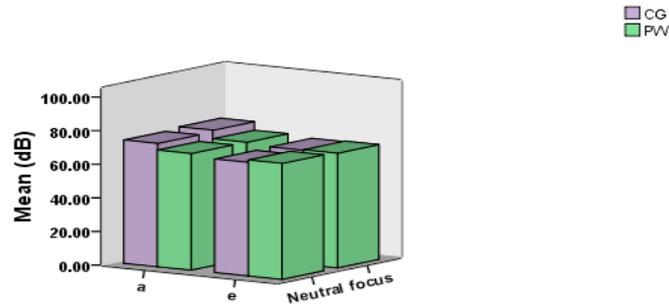


Fig 2. The mean energy of vowels [a] and [e] in the Focus and Neutral conditions.

### 5.3. Spectral information

In order to study the quality of vowels in the focus condition, we measured F1, F2 and F3 of vowel [a] and [e]. As for the difference between focus conditions ( $F[1,11]=6.84, p=.024$ ), we found that in the focal condition, F1 of [a] is 14.33 Hz higher than in the neutral condition. The main effect for STRUCTURE ( $F[1,10]=5.09, p=.045$ ) is due to 20.68 Hz higher F1 in the accented syllable than in the unaccented syllable. The main effect for STRUCTURE ( $F[1,11]=17.29, p<.01$ ) shows that F1 of the vowel [e] is 19.46 Hz higher in the accented syllable than in the unaccented syllable.

### 5.4. Fundamental Frequency

The interaction of vowel [e] STRUCTURE & FOCUS ( $F[1,11]=8.67, p<.01$ ) is significant. It shows that in the focal condition, vowel [e] in the accented and unaccented syllables has higher  $F_0$  than in the neutral condition. In the accented syllable,  $F_0$  of vowel in the focal condition [e] is 28 Hz higher than in the neutral condition.

Table 3 The effect of structure, voice, focus and mode on  $F_0$  (minimum, maximum and mean)

	<i>segment</i>	<i>Structure</i>	<i>Voice</i>	<i>Focus</i>	<i>Mode</i>
<i>Mean</i>	<i>a</i>	$F=26.57$ $P<.001$	$F=6.12$ $P<.01$	<i>NA</i>	<i>NA</i>
	<i>e</i>	$F=47.54$ $P<.001$	<i>NA</i>	$F=39.31$ $P<.001$	$F=31.38$ $P<.001$
<i>Max</i>	<i>a</i>	$F=22.32$ $P<.01$	<i>NA</i>	<i>NA</i>	<i>NA</i>
	<i>e</i>	$F=32.90$ $P<.001$	<i>NA</i>	$F=36.85$ $P<.01$	$F=24.64$ $P<.001$
<i>Min</i>	<i>a</i>	$F=26.09$ $P<.001$	<i>NA</i>	<i>NA</i>	$F=5.17$ $P<.01$
	<i>e</i>	$F=53.13$ $P<.001$	<i>NA</i>	$F=26.08$ $P<.001$	$F=29.28$ $P<.001$

In the accented syllable, the mean, minimum, and maximum of Vowel [e] is 41.05 Hz, 36 Hz and 35 Hz higher than in the unaccented syllable and in the

focal condition, the mean, Minimum and maximum of vowel [e] is 16 Hz, 12 Hz and 17 Hz higher than in the neutral condition. in the interrogatives, the mean, minimum and maximum of vowel [e] is 33 Hz, 28.59 Hz and 35 Hz higher than in the declaratives.

In the accented syllable, the mean, minimum and maximum of Vowel [a] is 44 Hz, 35 and 42 Hz higher than in the unaccented syllable and in the focal condition, the mean of vowel [a] is 6 Hz higher than in the neutral condition. In the interrogatives, the minimum of vowel [a] is 15 Hz higher than in the declaratives.

Phonological focus on voicing did not increase the size of voicing effects, while phonological focus on quantity did increase the size of quantity effects. The  $f_0$  peak of a word is consistently higher under a narrow focus than in the neutral-focus sentence. At the same time, the general locations of the  $f_0$  peaks are largely the same with or without narrow focus. This is seen in the graphs in Fig. 3.

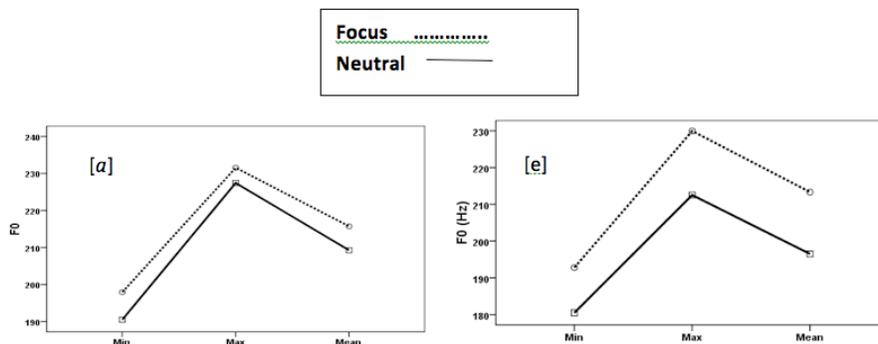


Fig 3. The  $F_0$  minimum, maximum and mean of vowels [a] and [e] in Focus and Neutral conditions.

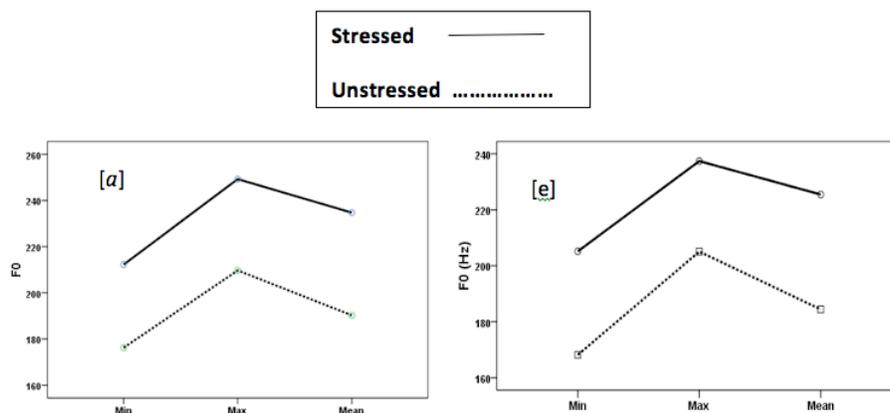


Fig 5. The  $F_0$  minimum, maximum and mean of vowels [a] and [e] in stressed and unstressed position.

## 6. Conclusion

Through the examination of detailed  $f_0$  contours in short Persian declarative and interrogative sentences with different focus conditions, we found that the pitch range of the focused item is expanded. The results show the effect of contrastive focus on  $F_0$  excursion size and accented vowel duration and energy by comparing the measurement points in two conditions: 1) contrastive focus condition and 2) noncontrastive focus condition. It can clearly be seen that under contrastive focus, the accented vowel becomes longer and the rise size and energy becomes larger as well. The duration and  $F_0$  attributes of contrastive focus that have been established by several researchers are largely corroborated here (e.g. Couper-Kuhlen 1984, Cooper & al. 1985).

## 7. References

- BOLINGER, D. L. (1972). "Accent is predictable (if you're a mind reader)". *Language*, 48, 633–644.
- COOPER W. E., EADY S. J. & MUELLER P. R. (1985). "Acoustical aspects of contrastive stress in question-answer contexts". *JASA* 77, 2142-2156.
- COUPER-KUHLEN, E. (1984). "A new look at contrastive intonation" in Watts R.J. & Weidmann U. (eds), *Modes of interpretation: essays presented to Ernst Leisi on the occasion of his 65th birthday*, Tübingen, Gunter Narr, 137-158.
- CRUTTENDEN, A. (1997). *Intonation* (2<sup>nd</sup> ed.). Cambridge: Cambridge University Press.
- CRYSTAL, D. (1969). *Prosodic systems and intonation in English*. London: Cambridge University Press.
- HIRST, D. & DI CRISTO, A. (1998), (eds.) *Intonation systems: a survey of twenty languages*, Cambridge, Cambridge University Press.
- DE JONG, K.J & ZAWAYDWH, B. A. (2002). "Comparing stress, lexical focus, and segmental focus: patterns of variation in Arabic vowel duration". *Journal of Phonetics*, 30, 53-75.
- LADD D. R. (1980). *The structure of intonational meaning: evidence from English*, Bloomington, Indiana University Press.
- GUSSENHOVEN, C. (2004). *The phonology of tone and intonation*. Cambridge University Press, Cambridge UK.
- GUSSENHOVEN, C. (in press). "Types of focus in English". In: D. Bring, M. Gordon, & C. Lee (Eds.), *Topic and focus: Intonation and meaning. Theoretical and crosslinguistic perspectives*. Dordrecht: Kluwer.
- KARIMI, S. (2005). *A minimalist approach to acrambling: evidences from Persian*, The Hage, Mouton.
- KARIMI, S. (2003). "Focus Movement and the Nature of Uninterpretable Features". In *Formal approaches to function in grammar*, A. Carnie, H. Harley and M. Willie (eds.), 297-306. Amsterdam/Philadelphia: John Benjamins.
- KARIMI, S. & AZITA T. (2003). "Looking alike is not being alike: a story of argument and adjunct wh-constructions in Persian". Paper Presented At the Western Conference on Linguistics (WECOL), Tuscon, Arizona.

- KARIMI, S. (1999). "Is Scrambling as Strange as We Think?", *MIT Working Papers in Linguistics*, 33: 159-190.
- KISS, E. K. (2003). "Argument scrambling, operator movement, topic movement in Hungarian". In *Word order and scrambling*, S. Karimi (ed.), 22-43. Oxford/Berlin: Blackwell Publishers.
- KISS, E. K. (1994). "Scrambling as the base-generation of Random Complement Order". In *Studies on scrambling*, N. Corver, and H. van Riemsdijk (eds.), 221-300. Berlin/ New York: Mouton de Gruyter.
- LADD, D. R. (1996). *Intonational phonology*. Cambridge: Cambridge University Press.
- MIYAGAWA, S. (1997). "Scrambling as an Obligatory Movement". In the *Proceedings of Naznan University International Symposium on Japanese Language Education and Japanese Language Studies*, 81-92. Naznan University, Nagoya.
- PIERREHUMBERT, J. (1980). *The phonology and phonetics of English intonation*. Ph.D. dissertation, Massachusetts Institute of Technology.
- PIERREHUMBERT, J. (2000). "Tonal elements and their alignment". In M. Horne (Ed.), *Prosody: theory and experiment* (pp. 11-36). London: Kluwer Academic Publishers.
- VAN HEUVEN, V. J. (1994). "What is the smallest prosodic domain?" In P. A. Keating (Ed.), *Papers in laboratory phonology* (pp. 76-98). Cambridge: Cambridge University Press.
- WIGHTMAN, C. W. (2002). "ToBI or not ToBI". In *Proceedings of the first international conference on speech prosody*, Aix- en-Provence, France (pp. 25-29).
- XU, Y. & XU, C. 2005. "Phonetic realization of focus in English declarative intonation". *Journal of Phonetics*, 33, 159-197.