Speech Prosody 2016
31 May - 3 Jun 2106, Boston, USA

Focus marking and Pitch Register modification in Boro

Kalyan Das[1], *Shakuntala Mahanta*[1]

[1] Indian Institute of Technology Guwahati, India

d.kalyan@iitg.ernet.in, smahanta@iitg.ernet.in

Abstract

This paper describes the prosodic aspect of prominence in Boro, a tone language belonging to the Tibeto-Burman family. The results here describe three production experiments investigating the phonological properties of Boro words occurring in contexts like wide focus, contrastive focus, corrective focus and narrow focus with emphatic particles. Boro lexically distinguishes High and Low tones and they follow the pattern of right alignment. The experiments discussed here involve words with both H and L specification. The target words were placed in carrier sentences to elicit the focus conditions mentioned above. Ten speakers of Boro were asked to produce scripted sentences containing the target words. F0 normalized pitch curves and durational values of the target words were extracted with the aid of Prosody Pro (Xu 2013) in Praat. The pitch contours of the Intonational Phrases (IP) suggest focus marking with emphatic particles results in an H* associated to the particle itself and the pitch-range of the whole IP is raised. Both contrastive focus and corrective focus are expressed by compressing the duration of the target words and by lowering the register of the whole IPs. The paper presents this as evidence to show the discrete nature of pitch register modification in Boro.

Index Terms: contrastive focus, emphatic particles, pitch register

1. Introduction

Focus of an utterance generally represents a word or constituent that receives prominence [1]. Languages employ a variety of means for the expression of focus. It may be expressed in the syntax, by reserving a structural position for the focus constituent, like the end of the sentence or the position before the verb; or by using focus particles, which are placed in some structural positions before or after the focus constituent [1]. However, studies have found that prosody also conveys discourse level information as can be found in the discussion of focus by Gussenhoven [2]. Intonational Languages (e.g. English) conveys prominence by aligning a pitch accent --- a prominence lending tonal morpheme with the syllable in a word that bears primary stress [3]. The discussion on the strategies that tone languages use for highlighting information shows that focused component in an utterance is marked by variations in pitch range, length, intensity [4], phonological phrasing [5] and also by compressing the post-focus part of the utterance [6]. Works on non-African tone languages, research on the Bantu languages like Chichewa, Kinande, and Xhosa provide evidence for the fact that tone languages do make use of prosodic means to indicate focus. This article aims at a discussion of the tone that signals the information-structural category of focus in Boro.

Boro belongs to the Tibeto-Burman group of languages, and forms a branch along with Dimasa, Tiwa and Kokborok and the language is spoken mainly in the districts of Dhubri, Goalpara, Kokrajhar, Chirang, Baksa, Udalguri, Kamrup and Darrang in North Eastern State of Assam in India. Boro lexically distinguishes L and H tones [7] as shown below:

- gáo 'tear or split'
- gào 'shoot by arrow or gun'

The tone bearing unit (TBU) in Boro is the syllable. Brahma and Sarmah [8] present evidence to suggest that the High tone in Boro is pronounced with a rising contour when used in isolation. On the other hand the Low tone is pronounced in isolation with a falling contour. Das and Mahanta [9] on the other hand presents evidence showing that the Low tone in Boro is sometimes pronounced with low level pitch irrespective of the number of morae present. Thus words like lùŋ 'drink' and dɯì 'water' are pronounced with low level pitch. Tonal organization in Boro is such that sometimes only the second syllable may employ the lexically significant tone. While discussing tone assignment in disyllabic words, Sarmah [7] informs that the first syllable in all these cases has the default mid tone and the study also claims that Boro words must have only one tonal specification which is associated with the syllable on the right. The preceding syllable gets a mid tone.

This study analyses the prosodic aspect of contrastive focus, corrective focus and narrow focus marked with emphatic particles. Due to a possible conflict between focal tone and lexical tone, the TBUs have to accommodate only one of these while the other gets overridden. Prominence due to contrastive focus is expressed by a pitch accent in Intonation languages which involves longer duration and higher F0 values. Focus-Prominence-theory proposes that the expression of abstract focus prominence is language specific. The present study looks into the ways a tone language like Boro employs prosodic means for the expression of post-lexical pragmatic meaning like that of focus.

2. Method

This study focuses on the behaviour of lexically marked morphemes when they occur in Intonational Phrases (IPs) expressing focus. Intonational Phrases in Boro corresponds to clauses. Contrastive focus was elicited by using the words: [bɔ́n] 'firewood', [daodɯí] 'egg', [zɯ̀o] 'ricebeer' and [bedɔ̀r] 'meat'. These words are placed in the carrier sentence frame in (1) for elicitation of contrastive focus. The subjects read out the sentences written in Devnagari script. They were also asked to pause for a moment before the beginning of the second clause (in Contrastive Focus sentences).

Subject$_i$ [x] Verb$_m$ but Subject$_i$ [y]$_{foc}$ Verb$_n$ (1)

Corrective focus is elicited in the form of an answer beginning with 'No' in response to a Yes/No question. The sentence frame in (2) was used for this. The speakers were asked to pause for a moment after the negative marker in the answer. The pause is intended to nullify any effect of the negative marker nɔŋá 'no' and the conjunction natʰaí 'but'

 doi: 10.21437/SpeechProsody.2016-177

which might result in a global lowering effect due to length. Kugler and Genzel's [10] experiment on the prosodic expression of focus in Akan has previously used pause as a deterrent to general declination effect on the later part of an utterance. Figure 1 presents an instance of the pause in the wave form extracted with the help of Praat. The target words used for this experiment were same as the ones used for eliciting contrastive focus.

Did Subject_i [x] Verb_m (2)
No, $\text{Subject}_{i/j}$ $[y]_{foc}$ Verb_m .

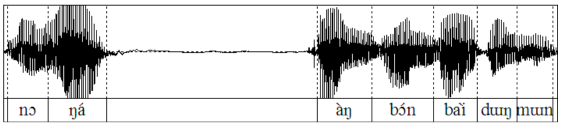

Figure 1: *Wave form extracted with the aid of Praat showing the amount of pause preceding the target sentences in the corrective focus conditions. Sentences expressing contrastive focus are also preceded by almost the same amount of pause.*

The intonational properties of the second clause in (1) and the answer in (2) is compared with that of an out-of-the-blue sentence expressing wide focus in the frame given in (3).

$\text{Subject}_{i/j}$ [y] $\text{Verb}_{m/n}$ (3)

Utterances recorded in the form of (3) are taken as the base line for comparing the effect of contrastive and corrective focus. The main prosodic parameters analyzed are F0 contours and duration.

The prosodic ramification of focus marking with emphatic particles are investigated by adding *bu* 'also' to focused constituents in sentences containing various combinations of lexical tones. As each Boro word can surface with one lexical tone, sentences containing a combination of three lexical tones are used for this experiment. The words used for this experiment are presented in Table 2.

2.1. Speech Material

The carrier sentences used to elicit the tonal properties of contrastive focus were constructed using a combination of the words given in Table 1. These sentences featured as the second clause following the conjunction natʰaí 'but' and also in an out-of-the-blue context for wide focus.

Table 1: *Sentence frame used for contrastive focus.*

àŋ 'I'	zù̀o 'ricebeer'	nagír-dɯŋ-mɯn 'search-Prf-Pst'
	bɔ́n 'firewood'	náŋ-gɯo-mɯn 'want-Fut-Pst'
bí-eɯ 'he-Nom'	bedòr 'meat'	bài-dɯŋ-mɯn 'buy-Prf-Pst'
	daodɯí 'egg'	

The prosodic properties of corrective focus was elicited using the same sentences as in Table 1 except that the verb zá-ji 'eat-Prsnt' was used instead of nagír-dɯŋ-mɯn 'search-Prf-Pst' and náŋ-gɯo-mɯn 'want-Fut-Pst'. These sentences featured after the negative marker *noŋá* 'No' in the data set.

The sentences for investigating the nature of focus expressed with the help of emphatic markers are formed using the frame in Table 2. The emphatic marker *bu* 'also' was added to all the constituents in the resultant sentences. This yielded a total number of 32 sentences

Table 2: *Sentence frame for focus with emphatic particles*

Subject	Object	Verb
bikʰunzú-a mother-in-law-Nom	*daodɯí* 'egg'	*zá-dɯŋ-mɯn* eat-Prf-Pst
àŋ I	*bedòr* 'meat'	*bài-dɯŋ-mɯn* buy-Prf-Pst

2.2. Participant and Recording

Ten male speakers of Boro between 22-28 years of age who are born and raised in their native villages were asked to produce the sentences used for the first two experiments. Four of the same group of speakers participated in the third experiment. The experiments were carried out in a quiet environment in Bhatipara and Bashbari villages. Five of the participants were from Bashbari village in Parbatjhora sub-division of Kokrajhar district of Assam in India and the other five speakers were the residents of Bhatipara village situated at the border between Bilashipara subdivision in Dhubri District and Kokrajhar Districts

For the first two experiments the speakers produced 5 iterations of each of the sentences and only the middle three have been considered for the present experiment. However the recording of one speaker for contrastive focus had to be avoided later due to noise. The third experiment contained 5 iterations of each of the sentences by the 4 speakers. An Edirol Roland R-09HR with its inbuilt microphone was used for the recordings. The recordings were digitized at a sampling frequency of 44.1 kHz and 32 bit resolution.

2.3. Measurement Criteria

Each iteration of the individual sentences was first extracted and saved as separate wave files using the speech analysis software- Praat 5.3.04_win32 [11]. Individual sound files of the sentences were further segmented into the syllable level, with the four target words segmented into the phoneme level, and Praat TextGrid files were created for each of the sentences for acoustic analysis. The segmented files were processed with a Praat Script [ProsodyPro] for obtaining measurements of F0 and duration. Time-normalized process in ProsodyPro takes into account 10 F0 points considering each interval as the temporal domain of normalization and also allows averaging of F0 contours across repetitions as well as speakers [12]. The script also provides values for meanF0, maxF0, minF0, duration and so on. The averaged time-normalized F0 values of all the iterations of each sentence were plotted as line graphs in order to observe the difference between the pitch contours expressing various types of focus. The raw meanF0 and duration values for each domain extracted with the help of ProsodyPro were normalized for speaker specific effects by using z-score normalization [13][14]. The z-score normalized meanF0 and duration values $p_z(n)$ were manually calculated by using the equation in (4). In this equation, μ represents mean of meanF0 or duration and σ stands for standard deviation and $p(n)$ is for meanF0 or duration for each domain. Repeated measures style of analysis of variance (ANOVA) was conducted in SPSS with fixed or independent factors focus (wide/contrastive, corrective) and token (1/2/3), and random factor speaker. Further details are provided in the following sections where results of the observations of the pitch contours and statistical analysis are presented.

$$p_z(n) = \frac{p(n) - \mu}{\sigma} \quad (4)$$

3. Result

3.1. Prosody of Contrastive Focus

Evaluation of the averaged normalized pitch contours of the sentences expressing wide focus and contrastive focus conditions revealed that contrastive focus (CF) lowers the pitch of the target word. This also leads to the lowering of the other tones in the vicinity. Figure 2 compares the difference between averaged time normalized F0 contours of wide focus and contrastive focus renditions of *àŋ [bɔ́n]$_{foc}$ bài-duɯŋ-muɯn* 'I bought firewood'. It can be seen that the pitch range of the sentence is lowered when it expresses contrastive focus. It is also found that the z-score normalized meanF0 values of the vowel of the TBUs averaged for all the tokens expressing contrastive focus significantly differ from the F0 value of it with wide focus. A Repeated Measured ANOVA was performed in SPSS to explore the effect of Focus on meanF0 (m) of the target tone-bearing vowels. The results have shown that the z-score normalized meanF0 for the vowel in the second syllable of bedòr 'meat' in wide focus condition (m = 0.382 hz, n=54) differs significantly from its occurrence in contrastive focus condition, $F(1, 8) = 29.93$, $p < 0.05$, m = -0.897 hz. There was also no significant interaction between the meanF0 of the repetitions and the two focus conditions, $F(5, 40) = 3.46$, $p > 0.05$. Results for the vowel in bɔ́n 'firewood' (n=81) have shown a significant effect of focus on meanF0, $F(1, 8) = 89.14$, $p < 0.05$ with the averaged value for wide focus being 0.743 hz and for contrastive focus it's being -0.560 hz. There was also no significant interaction between the meanF0 of the repetitions and the two focus conditions, $F(8, 64) = 1.35$, $p > 0.05$. The meanF0 for the vowel in zɯ̀o 'ricebeer' (n=108) significantly differs between its occurrence in the two focus conditions (m=0.539 hz for wide focus and m= -0.581 hz for contrastive focus) discussed here, $F(1, 8) = 406.2$, $p < 0.05$. No significant interaction between the meanF0 of the repetitions and the two focus conditions, $F(11, 88) = 1.14$, $p > 0.05$, were found. Contrastive focus also significantly lowers the meanF0 (m=0.749 hz for wide focus and m=0.043 hz for contrastive focus) of the vowel of the second syllable of daoduɯí 'egg', $F(1, 8) = 34.5$, $p < 0.05$. There was also no significant interaction between the meanF0 of the repetitions and the two focus conditions, $F(2, 16) = 1.24$, $p > 0.05$.

The results for duration of vowels expressing contrastive focus shows that Boro exercises significant durational reduction of the vowels of the TBUs getting contrastive focus. Table 3 presents the results of repeated measures ANOVA done on the z-score normalized duration value for the vowels. It can be seen from the results that compression of duration of the target words is another way of employing contrastive focus in Boro.

3.2. Prosody of Corrective Focus

Evaluation of the pitch contours for the sentences expressing corrective focus (CrF), in comparison to their occurrence with wide focus, has revealed that pitch range plays a discrete phonological role in expressing corrective as well as contrastive focus. Figure 3 below presents the averaged time normalized pitch contours of *bí-eu [daoduɯí]$_{foc}$ zá-ji* 'He has eaten egg' pronounced in both wide focus and corrective focus contexts. It can be seen that the pitch contour for the corrective focus expression remains lower till the end. Similar comparison of the time normalized pitch contours for the other sentences has revealed that corrective focus is consistently expressed by lowering the pitch register.

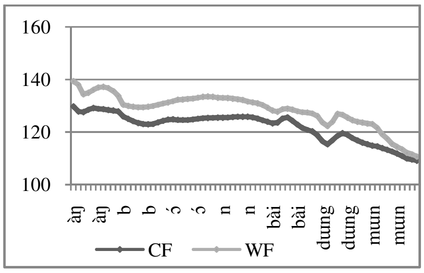

Figure 2: *Averaged time normalized pitch contour of àŋ bɔ́n bài-duɯŋ-muɯn 'I bought firewood' with wide focus (grey line) and contrastive focus (black line) (n=27)*

Table 3: *Durational difference between the occurrence of the vowels in the TBUs of the target words in Wide Focus and Contrastive Focus conditions. Interaction is not-significant for all the levels.*

Vowel	Focus	Mean	Significance
ɯ̀o	WF	3.165 ms	F(1, 8) = 70.480, p= 0.000, n= 108
	CF	2.529 ms	
ɔ́	WF	1.988 ms	F(1, 8) = 39.936, p= 0.000, n= 81
	CF	1.686 ms	
ò	WF	1.515 ms	F(1, 8) = 35.577, p= 0.000, n= 54
	CF	1.233 ms	
ɯí	WF	2.330 ms	F(1, 8) = 25.588, p = 0.001, n=27
	CF	2.040 ms	

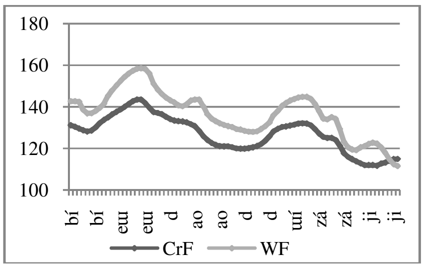

Figure 3: *Averaged time normalized pitch contour (n=30) of bí-eu [daoduɯí]$_{foc}$ zá-ji 'He has eaten egg' with wide focus (grey line) and corrective focus (black line).*

Evaluation of duration of the TBUs of the vowels of target words in the corrective focus and wide focus contexts has revealed that the duration is reduced when they express corrective focus. A repeated measures Anova test (on z-score normalized duration values) done in SPSS has revealed that the duration of the vowels is significantly reduced when they express corrective focus. Table 4 explains this.

Table 4: *Durational difference between the occurrences of the vowels in the target words in Wide Focus and Corrective Focus condition. Interaction is not-significant for all the levels.*

Vowel	Focus	Mean	Significance
ɯ̀o	WF	2.829 ms	F(1, 9) = 59.350, p = 0.000, n= 60
	CrF	2.326 ms	
ɔ́	WF	1.980 ms	F(1, 8) = 38.919, p = 0.000, n = 30
	CrF	1.561 ms	
ò	WF	1.420 ms	F(1, 9) = 44.074, p = 0.000, n= 60
	CrF	0.983 ms	
ɯí	WF	2.210 ms	F(1, 9) = 62.414, p = 0.000, n = 90
	CrF	1.835 ms	

3.3. Prosody of Focus with Emphatic Particles

Averaged normalized pitch contours for the sentences with the emphatic particle have shown that *bɯ* 'also' surfaces with a high F0 and also the sentences with this particle are pronounced with a higher pitch range. Figure 4 presents a comparison between a sentence with the emphatic particle *bɯ* 'also' and the related wide focus sentence without the particle. It can be seen on the lower panel of Figure 4 that the emphatic particle is pronounced with a higher F0. At the same time, the sentence with the emphatic particle is pronounced with a higher pitch range.

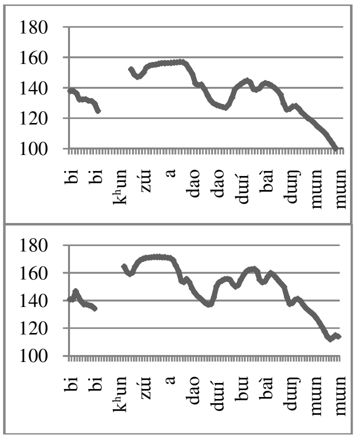

Figure 4: *Averaged time normalized pitch contours (n=20) of bikʰunzúa daoduí baìduŋmun 'Mother-in-law bought egg' on upper and of bikʰunzúa daoduíbɯ baìduŋmun 'Mother-in-law bought egg also' on the lower panel.*

A similar pattern of overall pitch range difference and higher F0 on the emphatic particle are noticed in the other instances of the occurrence of *bɯ* 'also' in the data set. The meanF0 of the emphatic particle in Figure 4 is compared with that of the preceding syllable bearing the lexical tone. The result has shown that the meanF0 for the emphatic particle *bɯ* 'also' (158.39 hz, n=20) is higher than that of the second syllable of *daoduí* 'egg' (149.58, n=20). This kind of pitch prominence on the emphatic particle provides acoustic evidence for forwarding a H* pitch accent on the particle itself. Table 5 compares the average pitch of each of the syllables (n = 20) in the sentence *bikʰunzúa daoduí baìduŋmun 'Mother-in-law bought egg'* pronounced with wide focus and with the emphatic particle *bɯ* 'also' occurring with all the three constituents in it. It can be seen that the occurrence of the emphatic particle with the subject or object or the final verb results in higher average pitch for each of the syllables in the sentence. This kind of acoustic evidence suggests that focus with emphatic particles in Boro is expressed by H* pitch accent on the emphatic marker along with overall pitch range raising.

Table 5: *Z-score normalized meanF0 (n=20) of each of the syllables (except the first one) are higher when the sentence is pronounced with the emphatic particle attached to any one of its constituents.*

Syllable	Wide Focus	Subject	Object	Verb
		+ bɯ	+ bɯ	+ bɯ
bi	-0.912	-0.856	-0.876	-0.818
kʰun	9.921	9.947	9.778	10.012
zú	9.718	9.756	9.983	10.125
a	9.912	9.969	10.298	10.475
dao	8.770	8.893	9.079	9.284
duí	9.138	9.243	9.246	9.626
baì	9.143	9.355	9.539	9.655
duŋ	7.988	8.339	8.342	8.632
mun	7.066	7.258	7.351	7.858

4. Discussion

The discrete nature of pitch register modification in Hausa and Jita has been previously discussed in Inkelas and Leben [15] and in Downing [16]. Borras-Comes, Vanrell and Prieto [17] show how pitch range differences can express discrete linguistic distinction in Catalan. Results presented in the previous sections show that pitch register modification in Boro is grammatically encoded as this modification is related to a specific way of expressing focus. Focus marking with emphatic particles results in two distinct prosodic modifications. Firstly an H* is associated to the particle itself and secondly the pitch-range of the whole IP is raised. On the other hand, both contrastive focus and corrective focus are expressed by compression of the duration of the focused words. A second prosodic means used for contrastive and corrective focus is the phonological feature of lowering the register of the IPs expressing these two kinds of focuses. Morphological elements marked with the property of pitch accents has been already discussed in studies like Gordon [18] where evidence has been presented in support of morpholexical pitch accents in Chickasaw. The emphatic particles in Boro also seem to contribute a similar property to the intonation structure of the language. However, this accentual property of the emphatic particles surface only when they are attached to non-final constituents. The final occurrences of the emphatic particles do not result in H* pitch accent. It results only in the raising of the pitch register of the IP. The non-final occurrences of the emphatic particles result in an additional feature of an H* pitch accent. Contrastive and corrective focuses lower the pitch height of all the non-final lexical tones leading to the overall lowering of the pitch register. However the final lexical tone does not undergo much lowering and this can be noticed in both Figure 2 and 3. This kind of lowering of the overall pitch of the utterance can be explained by introducing a register Low tone which Boro uses as a prosodic means for marking contrastive and corrective focus. On the other hand focus marking with the help of emphatic particles is prosodically marked by a register High tone which raises the overall pitch of the whole IP.

5. References

[1] C. Gussenhoven, *The Phonology of Tone and Intonation.* Cambridge: Cambridge University Press, 2004.

[2] C. Gussenhoven, "Focus, mode and the nucleus," *Journal of Linguistics*, vol. 19, no. 2, pp. 377–417, 1983.

[3] S. Jannedy, "Prosodic focus in Vietnamese," *Interdisciplinary studies on information structure,* vol. 8, pp. 209-230, 2007.

[4] Y. Xu and C. X. Xu, "Phonetic realization of focus in English declarative intonation," *Journal of Phonetics,* vol. 33, no. 2, pp. 159-197, 2015.

[5] L. J. Downing, "The prosody and syntax of focus in Chitumbuka," *ZAS papers in linguistics,* vol. 43, pp. 55-79, 2006.

[6] Y. Xu, "Effects of tone and focus on the formation and alignment of f0 contours," *Journal of phonetics,* vol. 27, no. 1, pp. 55-105, 1999.

[7] P. Sarmah, *Some Aspects of the Tonal Phonology of Bodo.* M Phil dissertation, Hyderabad: Central Institute of English and Foreign Languages, 2004.

[8] Brahma, A. and Sarmah, P., "An Acoustic Study of Boro Tones", Power Point Slides, NEILS6, Retrieved from http://sealang.net.neils/, 2012

[9] K. Das and S. Mahanta, "Tonal Alignment and Prosodic Word domain in Boro," *The 5the International Symposium on Tonal Aspects of Languages.* ISCA. (forthcoming)

[10] F. Kugler and S. Genzel, "On the prosodic expression of pragmatic prominence: The case of pitch register lowering in Akan," *Language and Speech,* vol. 55, no. 3, pp. 331-359, 2012.

[11] P. Boersma and D. Weenik, "Praat: doing phonetics by computer (Version 5.3.04_win32) [Computer programme]. Retrieved on January 12, 2012, from http://www.praat.org/Bradley 1982. 2012.

[12] Y. Xu. "ProsodyPro—A tool for large-scale systematic prosody analysis," pp 7-10, 2013.

[13] P. J. Rose, "Considerations on the normalization of the fundamental frequency of linguistic tone," *Speech Communication.* Vol. 6, no. 4, pp. 343-352, 1987.

[14] P. J. Rose, "How effective are long term mean and standard deviation as normalization parameters for tonal fundamental frequency," *Speech Communication.* Vol. 10, no. 3, pp. 229-247, 1991.

[15] S. Inkelas and W. Leben, "Where phonology and phonetics intersect: the case of Hausa intonation," *Papers in Laboratory Phonology I: Between the grammar and the physics of speech,* pp 17-34, 1990.

[16] L. J. Downing, *The tonal phonology of Jita.* Vol. 5. Lincom GmbH: Munchen, 1996.

[17] J. Borras-Comes, M. del M. Vanrell, and P. Prieto, "The role of pitch range in establishing intonational contrasts," *Journal of the International Phonetic Association.* Vol. 44, no. 01, pp. 1-20, 2014.

[18] M. Gordon, "An Autosegmental/metrical model of Chickasaw intonation," *Prosodic typology- the phonology of intonation and phrasing,* pp. 301-339, 2005.