Prosodic Focus in Vietnamese*

Stefanie Jannedy
Humboldt University of Berlin

This paper reports on pilot work on the expression of *Information Structure* in Vietnamese and argues that Focus in Vietnamese is exclusively expressed prosodically: there are no specific focus markers, and the language uses phonology to express intonational emphasis in similar ways to languages like English or German. The exploratory data indicates that (i) focus is prosodically expressed while word order remains constant, (ii) listeners show good recoverability of the intended focus structure, and (iii) that there is a trading relationship between several phonetic parameters (duration, f0, amplitude) involved to signal prosodic (acoustic) emphasis.

Keywords: Information Structure, Vietnamese, Focus, Perception (Statement-Question Matching)

1 Introduction

Mon-Khmer languages are known for the complexity of their tone system: lexical contrasts are marked by tonal (pitch) as well as laryngeal features (Yip, 1995). This interaction of voice quality and lexical tone also characterizes Vietnamese (Brunelle, 2003, 2006). Several more recent experimental studies have explored the perception of tone in the northern (Hanoi) and the southern (Saigon) Vietnamese dialect with six and five contrasting tones respectively, and have established that there is a higher and a lower pitch register (Brunelle, 2006;

* Many thanks are due to Tue Trinh and Phuong Ha for their valuable native linguist speaker judgments and for their patience during the recording sessions. I would also like to thank Philippa Cook (ZAS) and Anna McNay (HU) for comments on ongoing work and the participants of the 3rd Contrast Workshop at the ZAS for encouragement and positive feedback. Manfred Krifka and Bernd Pompino-Marschall have been incredibly supportive of this project, I thank them. I kindly thank Marc Brunelle (Univ. of Ottawa) for insightful comments on this paper and for discussions on the language. All shortcomings of this paper are my own.

Interdisciplinary Studies on Information Structure 08 (2007): 209–230
Ishihara, S., S. Jannedy, and A. Schwarz (eds.):

Michaud & Vu, 2004; Michaud, 2004; Michaud et al., 2006; Nguyễn & Edmondson, 1997; Brunelle & Jannedy, 2007). The f0-contours shown in Fig.1 are representative of the standard Hà Nội dialect. The only exception is the rising tone *sắc*, which is realized relatively low, a variant found in some young female Northerners. In the Hà Nội dialect, laryngealization is tone-medial in *ngã* (steeply rising f0 trajectory marked with "▲") and tone-final in *hỏi* and *nặng* (glottalization). The three tones with a laryngealized voice quality are represented by a dotted line. The *huyền* tone is partially breathy. The rising tone *sắc* is fully modal and usually rises from the bottom of the pitch range to the top. The three tones in the lower register are *hỏi, huyền* and *nặng*. The neutral tone is called *ngang* and remains fairly stable in pitch throughout.

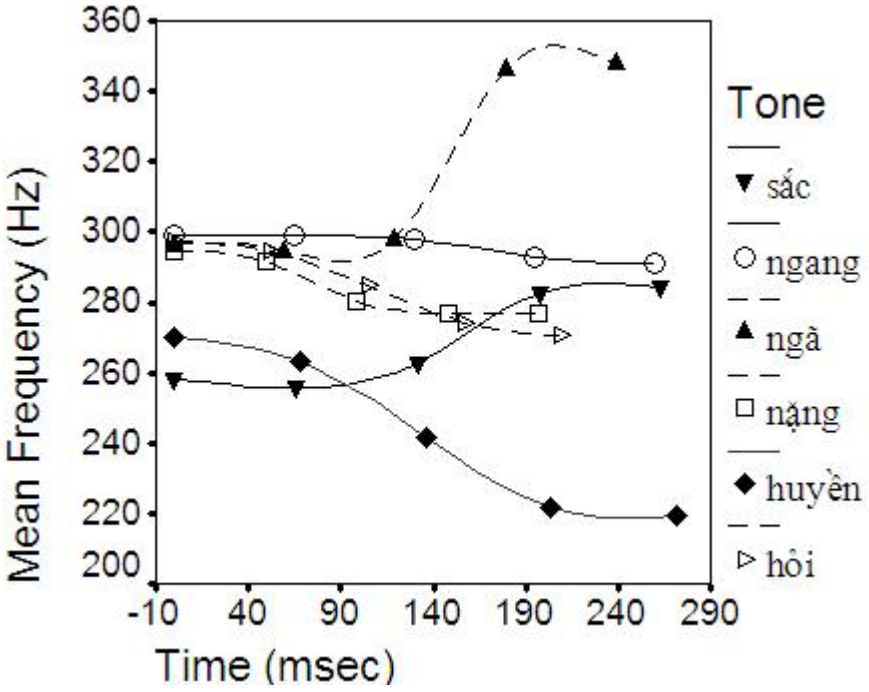

Fig 1.: Mean f0-contours (over five repetitions) for the six lexical tones of the Hà Nội dialect of Vietnamese as produced by a female speaker (used as stimuli in the experiment described in Brunelle & Jannedy, 2007).

Vietnamese is an isolating language, most words consists of mono-syllables. It is unclear though if syllables are the tone bearing units in Vietnamese (as is the case in Ewe, Hausa, Chicheŵa or Mandarin Chinese) or if moras are (as in Japanese or Thai, see Morén, 2003). Furthermore, it is remarkable that Vietnamese has no tone-sandhi rules, as we know them for languages such as Mandarin Chinese, Cantonese or Taiwanese. Tone-Sandhi refers to the changes in the values of lexical tones in the context of other tones. A well-known example from Mandarin Chinese is the change of a low-tone to a rising tone when it is followed by another low tone. No such consistent rules are known for Vietnamese and none of the standard grammar books on the language (Thompson, 1965; Nguyễn, 1997) make reference to it. There is also no phonological downstep: the successive lowering of high tones often observed in register tone languages. There may be other non-systematic intonational downtrends such as final lowering (the lowering of the pitch towards the end of an utterance or phrase) or declination (a decline of the f0 over the course of the utterance); however, with the exception of Dung et al. (1998), none of the grammars, offer somewhat systematic descriptions of intonational variation.

Given the tonal complexity of the language and what has been stated in the sporadic reports published on tones, tone implementation and intonational emphasis, the question arises whether or not the language makes use of prosodic cues to signal information structural content or whether it needs to revert to other means such as the usage of particles or specialized syntactic positions to signal focus or topic. Occasional references to the use of prosodic means for emphasis and for phrasing can be found on some of the older, somewhat sparse, literature (Thompson, 1965; 1981; Nguyễn, 1990; Dung et. al. 1998).

> ”Heavy stress singles out the syllable or syllables of each pause group which carry the heaviest burden of conveying information. Weak stress accompanies syllables, which bear the lowest information-

> conveying load in the pause group. They often refer to things which have been brought up earlier or which are expectable in the general context. Other syllables are accompanied by medium stress."
>
> Thompson (1965:106)

Tran (1967:24) also describes intensity as one of the integral aspects of intonation in Vietnamese. Intonation contours are "superimposed on the basic tone system; they modify the pitch characteristics of the tones, but do not affect the tonemic contrast between them […] the basic intonation contours are intrinsically linked with the overall intensity patterns." Similarly, Michaud & Vu (2004) state: "Vietnamese also possesses intonational emphasis: as in many languages, the great variability observed in the realization of the lexical tones largely reflects the informational prominence of various syllables in the utterance..." and they conclude "[…] a stable correlate of emphasis is curve amplification, manifested [...] as an increased slope of F0 curve [...] or as F0 register raising."

The lack of detailed descriptions of phonetic or phonological properties of structuring or emphasizing information in Vietnamese is apparent. Evidence reported in the literature and our first pilot studies strongly suggest that Vietnamese shows properties that are often associated with intonational phrasing and prosodic prominence in intonation languages: it has pitch range effects of the same sort seen in the intonational marking of emphasis and it also has pausing and other rhythmic effects of the sort associated with intonational phrasing observed in English and German.

In studying prosodic prominences and the resulting pragmatic interpretation of *prosodic focus*, there are two over-arching questions that are more effectively responded to if they are addressed together. One question pertains to the mechanics of how the speaker imparts prominences to some parts of an utterance but not to others, while the other question addresses the listener's

interpretation of such prominences - i.e., the function of prosodic focus from the listener's point of view. A fundamental assumption in posing the first question is that the speaker has various methods at his/her disposal to make some part of an utterance prosodically more prominent than other parts. In English and languages like English, for example, one important means of making a particular word more prominent than surrounding words is to align a pitch accent — a prominence lending tonal morpheme — with the syllable in a word that bears primary stress. Most current accounts of prosodic focus in English recognize this mechanism of putting a constituent in prosodic focus, and in one particularly influential account, due to Selkirk (1984, 1995), this is the only mechanism recognized. Other accounts, however, suggest that other aspects of the tune also may play a role in imparting prominence. For example, the accented word that is the last accented material in its phrase is also aligned to another tonal morpheme, the phrase accent, which is simultaneously aligned to the end of the phrase as well. When it is followed immediately by the phrase accent, a pitch accent becomes the 'nuclear accent' in its phrase. In the account of Pierrehumbert (1980) and her colleagues (e.g., Beckman & Pierrehumbert, 1986; Beckman & Edwards, 1994), any nuclear accent is more prominent than all earlier, non-nuclear accents. (This is related to Ladd's (1980, 1996) notion of 'deaccenting', which says that an accented word can be made prominent if all following material is left unaccented, effectively positioning the nuclear accented word early in its phrase). The important point is that if word order remains constant and it can be observed that prosodic emphasis is being shifted from one constituent to another, a structure with an early prosodic prominence is cognitively more salient (due to the unaccented post nuclear tail) than a structure with a prosodic prominence late in the utterance (Beckman, 1996). This is probably due to the probability of distributions of early prominences versus late prominences in running discourse and the expectations that hearers have.

An equally fundamental assumption underlying the second question is that speakers use prosody and prosodic focus to facilitate and guide the hearer's understanding and comprehension of the message being conveyed at any particular time in a discourse. Thus, one of the uses of intonation is to guide the listener's interpretation of the utterance in relationship to the larger discourse context. Different intonational structures, then, are used to distinguish one discourse purpose, one extension of the current discourse state, from other possible moves in the mutual building of the discourse structure by the speaker and hearer, they are used to manage discourse content (Krifka, 2006). This function of intonation makes it difficult to test claims that two or more intonation patterns differ categorically.

This differs markedly from claims about the number of tones in contrast in languages such as Mandarin Chinese, Cantonese or Vietnamese, which can be tested by seeing whether the tune distinguishes one word from any other word that could have occurred in the same place. Listeners are generally very good at identifying which of two minimally contrasting words they heard. They are generally much less facile at identifying different discourse intentions, unless the differences also trigger a difference in truth conditions. One of the challenges for psycholinguistics, therefore, is to devise tasks that tap the listener’s competence in interpreting the intended discourse purpose rather than training listeners to attend to specific aspects of the signal. In studying the functions of prosodic focus, for example, the psycholinguist must find an experimental design that can be used to determine how exactly different prosodic manipulations contribute to the introduction of new entities or highlighting of old entities in the interpretation of the discourse purpose of an utterance.

2 Focus

The canonical word order in Vietnamese is SVO (Nguyễn , 1997; Thompson, 1965), and this structure is used consistently when answering any *wh*-focus alternative question (Krifka, 2006; 2007). That is, focus is always marked in situ for all sentence constituents. Consider the following example of a transitive sentence:

(1)

S	V	O
Phương	đi	xe đạp.
Phuong	ride	bicycle.

'Phuong is riding a bicycle.'

We elicited replies to focus alternative questions asking for sentence focus (a), subject focus (b), object focus (c), verb focus (d), and VP focus (e) from two native speakers of Hà Nội Vietnamese. A sample paradigm is shown below. (Also see the appendix).

(2) a.	Chuyện gì vậy?	**What is happening?**
	[**Phương đi xe đạp**]$_F$	[**Phuong is riding a bicycle**.]$_F$
b.	Ai đi xe đạp?	**Who** is riding a bicycle?
	[**Phương**]$_F$ đi xe đạp.	[**Phuong**]$_F$ is riding a bicycle.
c.	Phương đi gì?	**What** is Phuong riding?
	Phương đi [**xe đạp**.]$_F$	Phuong is riding a [**bicycle**.]$_F$
d.	Phương làm gì với xe đạp?	What **is** Phuong **doing** with the bicycle?
	Phương [**đi**]$_F$ xe đạp.	Phuong [**is riding**]$_F$ the bicycle.
e.	Phương làm gì vậy?	**What is** Phuong **doing**?
	Phương [**đi xe đạp**.]$_F$	Phuong [**is riding a bicycle**.]$_F$

In each panel in Fig. 2, we have bracketed the particular part of the utterance that was in focus.

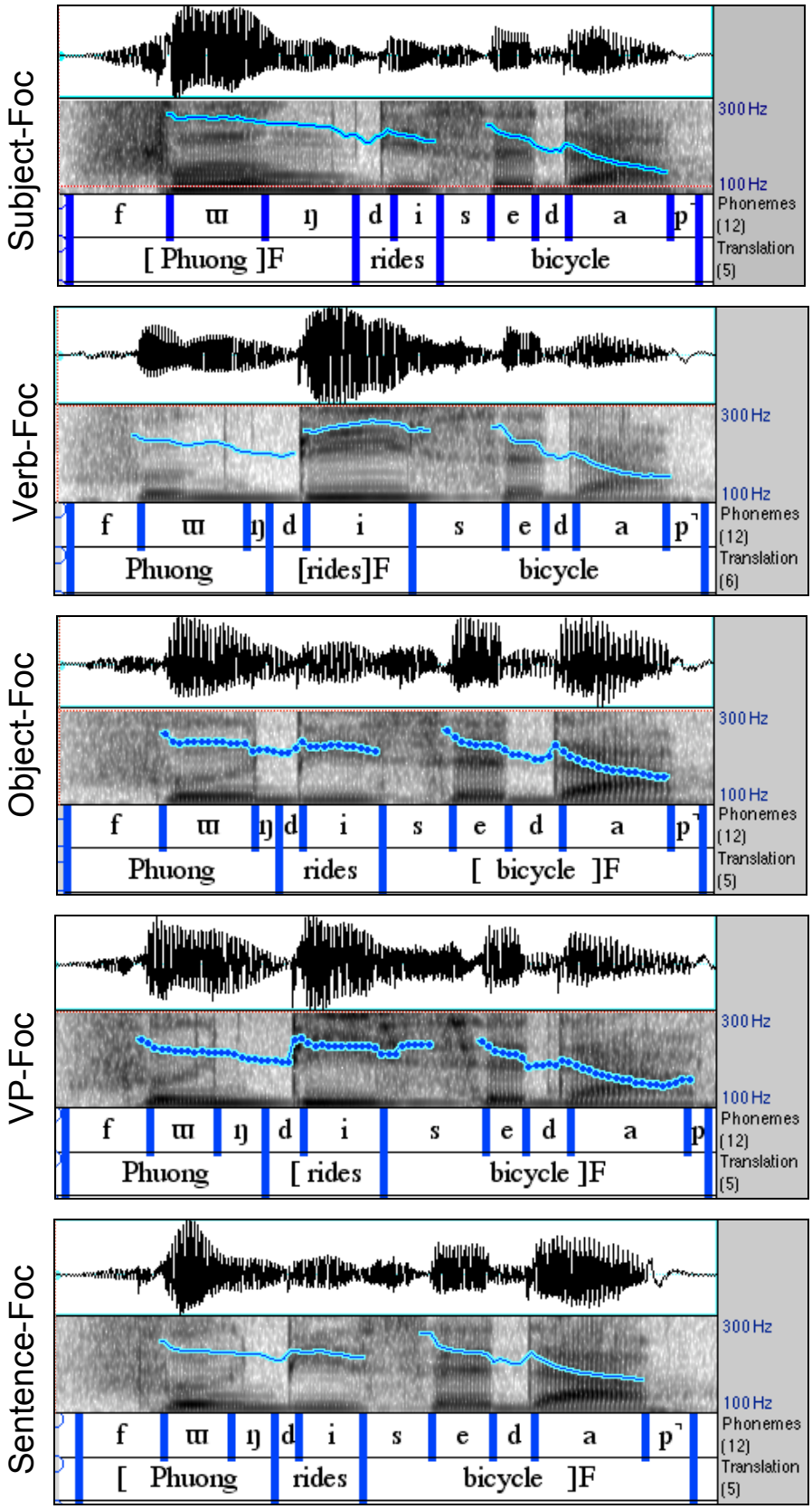

Fig. 2: Spectrogram, waveform and f0 display of five segmented and annotated replies to *wh*-focus alternative questions for speaker 1.

Most importantly, it should be noted that word order remained constant and hence, any kind of contrast between the five kinds of focus condition is expressed prosodically. All f0-curves are plotted on the same pitch range (100Hz to 300Hz) and all sentences are lexically identical, thus we can visually compare these patterns. There appear to be differences in the amplitude (a raw acoustic measure of the strength or volume of a signal) of the signal, as is clearly visible in the waveform (upper display) of each panel. According to native speaker intuitions, amplitude (measured in decibel [dB]) does play a role in Vietnamese to express acoustic emphasis. The intensity of the signal is defined as "average rate of flow of energy per unit time per unit area", measured in watts per cm^2 (Poser, 2002). And loudness in turn, is a perceptual response to the physical property of intensity. That is, roughly speaking, the psychological percept of amplitude is loudness. Note that in the subject focus (Sub-Foc) case, the vowel in the name *Phương* has a particularly great amplitude, visible especially in contrast to the verb focus (V-Foc) case where the vowel in the verb *đi* has the greatest amplitude. In the verb phrase focus (VP-Foc) case, both the verb and the object appear to have a greater amplitude, while in the object focus (O-Foc) panel, there does not seem to be a clear picture with regard to the differentials in amplitude of the signal.

The correct picture of amplitude may be confounded in the O-Foc example due to the fact that the Vietnamese word *xe đạp* is a compound which requires emphasis on the second syllable in order to be interpreted as a compound (cf. Dung et al., 1998:399). Ingram & Nguyễn (submitted) find task related differences in the emphasis patterns in compounds (naming task versus reading task). In more formal settings such as the reading task, they find more reflexes of compound final emphasis than in the naming task. They attribute these to formality or register differences. Our data was elicited in a question-

answer paradigm which could potentially be construed as a casual conversation and thus, as non-formal.

The three simple transitive SVO test sentences used in the perception study are listed below. The focus conditions are the same as in example (2) above (see the Appendix for an explicit listing of the tested utterances). Note that the sample sentence in (3a) is specified for the neutral tone, the level tone *ngang*, with exception of the last syllable, which carries the *nặng* (final laryngealization) tone. We deliberately selected a tonal specification that has the potential for rises and falls during the course of the utterance so that we may explore the potential variation of the f0 range imposed under different focus conditions.

(3) a. Phuong is riding a bicycle. Phương đi xe đạp.
b. Lan is drinking coffee. Lan uống cà-phê.
c. Men is drinking water. Mến uống nuốc.

The sentence in (3b) has a neutral tone on the Subject, a rising tone on the verb (*sắc*) and a falling tone *huyền* on the first syllable of the compound *cà-phê* and a neutral tone again on the final syllable, while the sentence in (3c) is specified lexically throughout with the modal rising tone *sắc.*

Note though that the three utterances above are specified differently for lexical tone. The first sentence type *Phương đi xe đạp.* is lexically specified throughout with the level tone while the third sentence *Mến uống nuốc.* has all rising tones. The third sentence *Lan uống cà-phê.* combines neutral, rising and falling lexical pitch patterns. These few examples already show the complex interplay between lexical tone on the one hand and intonational requirements to signal information structure on the other hand.

The graphs in Fig. 3 show stylized f0 contours, generated by logging the maximum F0 during a labeled interval, that is, during a phoneme. These individual points were plotted and the lines between the points are interpolations rather than actual f0-trajectories. Note further that Vietnamese has complex vowel sounds such as <ướ> that are considered monophthongs rather than diphthongs.

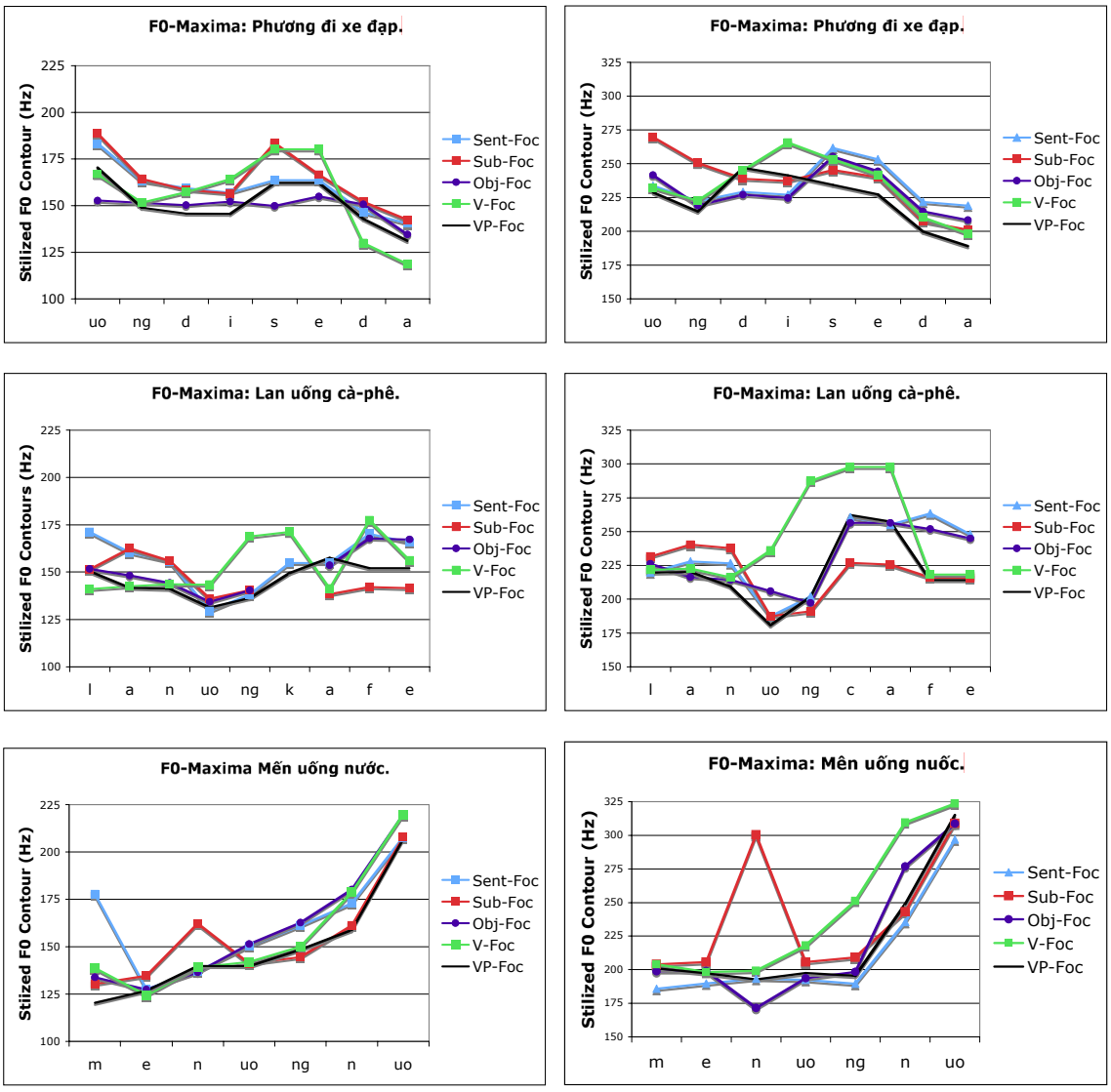

Fig. 3: Stilized F0 Contours (interpolations between the maximum f0 value of each labeled phoneme).

The three graphs on the left show the stylized f0-curves from the male speaker whereas the three graphs on the right show the stylized f0-curves for the same utterances but for the female speaker. Note that we have avoided to plot the initial or final voiceless obstruents in the utterances as f0 cannot be cleanly logged during these sounds. Each line in a graph represents one repetition of the five focus conditions the utterance was produced in. Despite the range of variation observable, there are also commonalities: for example, the subject-focus and the verb-focus utterances appear to have rather pronounced f0-maxima rather early in the utterance, while sentential or object-focus utterances show pitch excursions later, towards the end of the utterances.

For the all rising contour (bottom panel), we can observe the general tendency of a low onset of the contour and a relatively steep final rise, whereas the all neutral contour (top panel) displays a final fall and much less overall variation in the f0 from the onset of the utterance to the end. The tonal contour displayed in the bottom panel appears much less consistent in terms of an overall tendency of the f0 contour throughout the utterance. These observations however can only be viewed as general tendencies, the amount of data is not sufficient enough to make more generalizable statements about the interaction of lexical tone and phrasal tone requirements.

2.1 Perception test

The test material was recorded in a *wh*-question-answer paradigm from a male and a female native speaker of the northern dialect of Vietnamese. While the questions and replies were presented in writing, both speakers were present for the recordings and prompted each other with the questions, they were rendered as quasi-spontaneous rather than read. For each focus condition and sentence type, we elicited one through three tokens of which both speakers selected their “best” renditions.

To understand and evaluate the listener's competence in interpreting the intended discourse purpose of an utterance, we wanted to test whether the *wh*-focus alternative question was recoverable from the reply utterance presented out of context. Six native listeners of Vietnamese, naïve as to the purpose of the experiment, aged between 21 and 26, participated in a short forced-choice identification perception task. The test data consisted of three sentence types that were each elicited in five focus conditions and spoken by our two native speakers (3 x 5 x 2 = 30 test sentences).

These 30 test sentences were played five times each (in randomized order) to each of the six listeners that participated. The sounds were presented over *Sennheiser* headphones and were called up by a script in Praat. The listeners were asked to match each heard utterance back to one of the five questions that were visually displayed to them on a computer screen.

Thus, we elicited 900 responses in total (30 sentences x 5 repetitions x 6 listeners = 900). That is, a total of 180 responses were collected for each of the five focus conditions tested (900 items in perception test / 5 focus conditions = 180 items per focus condition). A summary of the data and responses is provided in Table 1.

	Stimulus -Type				
response	Sub-Foc	V-Foc	O-Foc	VP-Foc	S-Foc
Subject	142 (78.89)	4 (02.22)	3 (01.67)	7 (03.89)	14 (07.78)
Verb	5 (02.78)	135 (75.00)	10 (05.56)	34 (18.89)	7 (03.89)
Object	11 (06.11)	15 (08.33)	94 (52.22)	34 (18.89)	33 (18.33)
Verb Phrase	9 (05.00)	21 (11.67)	33 (18.33)	46 (25.56)	56 (31.11)
Sentence	13 (07.22)	5 (02.78)	40 (22.22)	59 (32.78)	70 (38.89)
Grand Total	180 (100%)	180 (100%)	180 (100%)	180 (100%)	180 (100%)

Table 1: Number of responses in five categories per stimulus type (raw numbers and percentages).

A chi-square test on the raw counts of the observed data was significant (χ^2= 998.47, df = 16, p<.001), indicating that the listeners did not match answer utterances randomly to questions. That is – despite the word order remaining constant in all five focus conditions – the prosody helps to disambiguate and lets listeners correctly match answers to questions. In fact, as Fig. 4 shows, listeners identified the subject-focus, verb-focus and object-focus questions that matched the utterances they heard, quite well. There are less reliable patterns in the VP and sentential focus condition. However, results indicate that even in these conditions, listeners responded above chance level (20%).

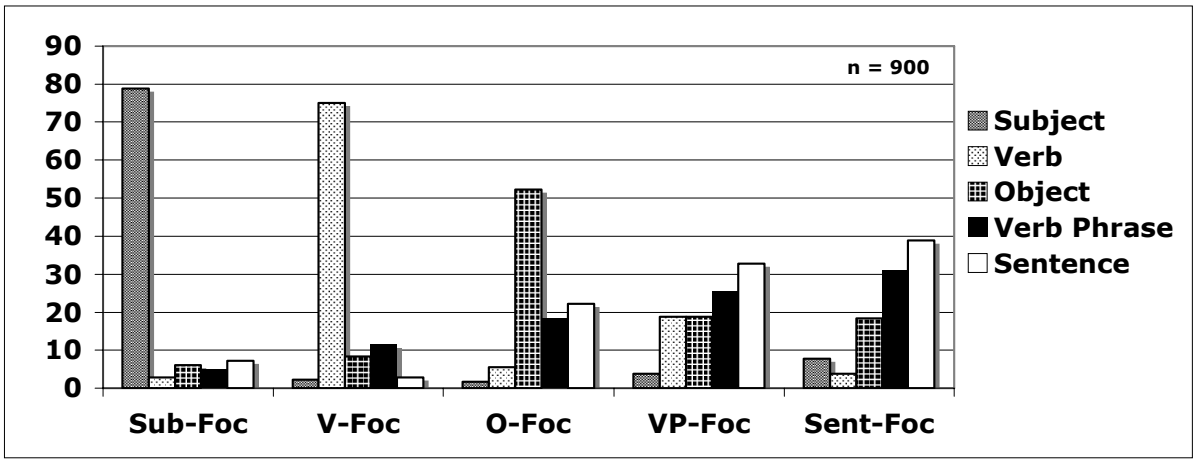

Fig 4: Visualization of the data (in %) presented in Table 1.

Since word order has remained constant, the difference between the focus conditions has to be marked prosodically. However, precisely what parameters (duration, f0, intensity, vocal effort) or what combination thereof are modified is less clear at this point. Considering the VP-Focus and Sentential-Focus conditions, it appears that listeners have a general preference for less marked questions such as those asking for a broader focus constituent such as Sentence focus. Since this study is based on only a relatively small amount of exploratory data, we cannot make further claims about this observation at this stage.

2.2 F0 & duration

Since there is no morphological focus marker in Vietnamese and given the good level of recoverability of the subject, verb, and object focus questions in our question-answer pairing test, there must be something distinguishing these morphosyntactically identical utterances. To make some of these prosodic patterns that listeners probably attend to 'visible', we time-normalized the fundamental frequency contours for each focus condition and calculated the mean over three repetitions of the sentence. For time normalization of the fundamental frequency contour, each labeled interval (in this case, phonemes) is divided into the same number of points (in this case 10). Time normalization allows for a direct comparison of differences in the f0 per labeled interval (see Xu, 1999). Note that in the graph below, the initial obstruent [f] and the final obstruent [p] are omitted from the plot. It is notable that the f0 – on average - is highest during the unrounded high back vowel [ɯ] in the subject focus condition, whereas it is highest during the vowel [i] in the verb focus condition.

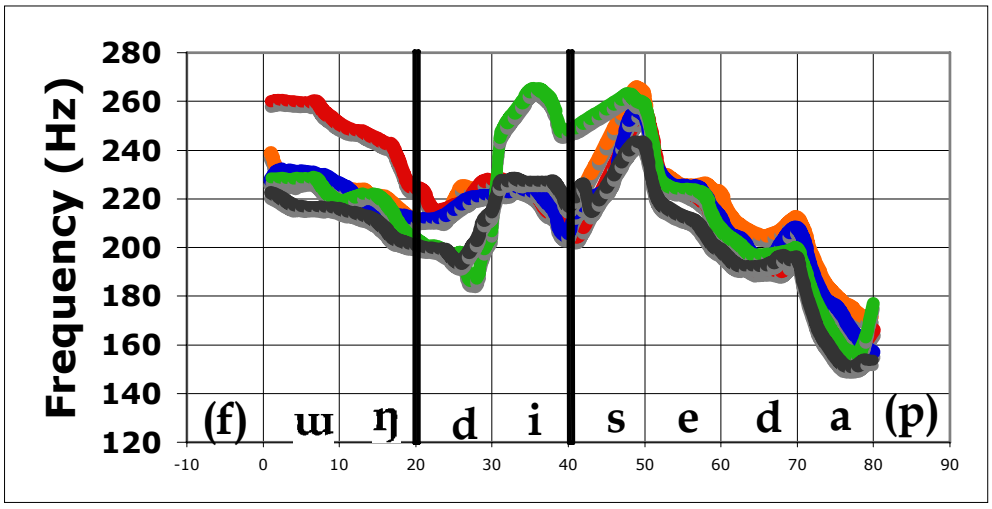

Fig. 5: Plot of the mean (n=3 per focus condition) of time normalized f0-contours for the five focus conditions as produced by our female speaker.

The representation of the data in Fig. 5 is based on actual f0-trajectories whereas the representations in Fig.3 are interpolations between measured f0-maxima. The type of representation below is preferred to evaluate f0-contours, however, in the absence of enough data to generate means, the graphs in Fig. 3 give decent approximations of the overall f0 patterns found in the data. Thus, it appears that local changes in the f0 as we know them from stress accent languages such as English and German, appear to play a role in the expression of focus in Vietnamese. We are reluctant at this point to call these local prominences 'accents' as this term has a specific meaning in the literature. Rather, we term them accentual prominences that are clearly visible for the subject and verb focus conditions.

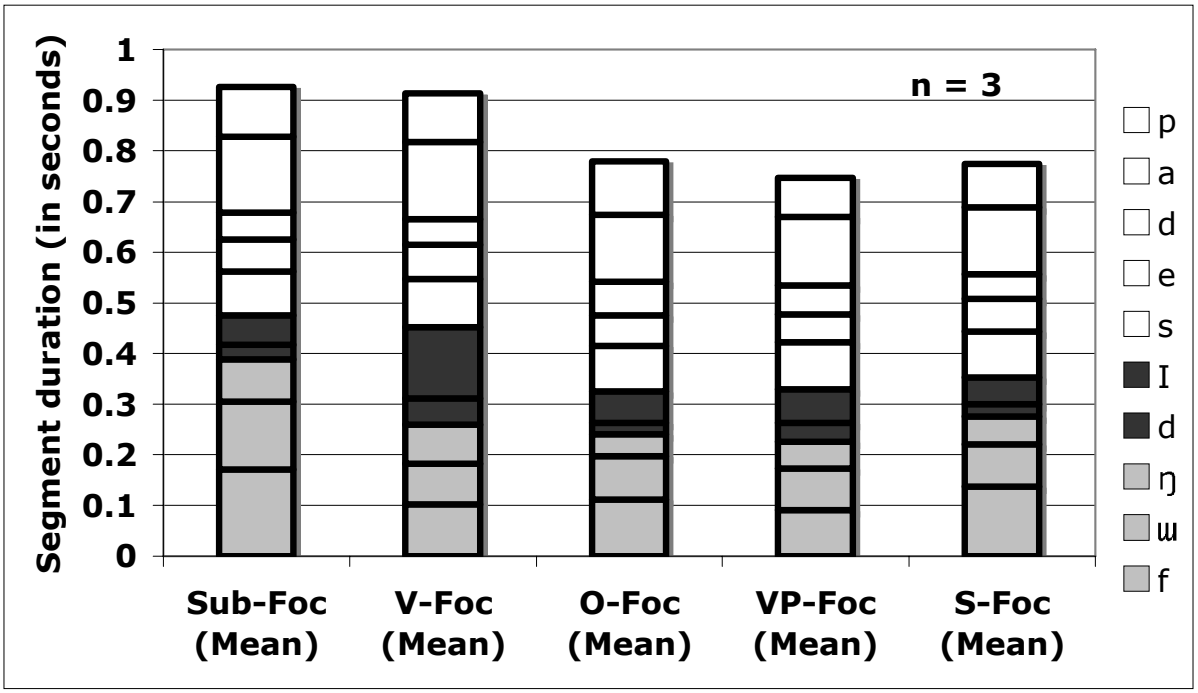

Fig. 6: Duration (in seconds) of each segment in the sentence "Phương đi xe đạp" based on three tokens rendered by one speaker.

None of the other focus conditions appear to have such a distinct pattern, not even the object-focus, even though the object focus reply was reliably matched to the object focus wh-question. Thus, we suspect an interaction of prosodic

parameters to play a role in the interpretation of focus conditions. For example, also note the durational differences between the five focus conditions, displayed in Figure 3. This graph is also only based on three utterances, thus, there is room for variability with the inclusion of more data.

Nevertheless, it appears that there is justification for speculating that durational cues such as the overall length of the utterance or the duration of subcomponents of the utterance (such as the subject (light grey shading in the first bar) or the duration of the verb (dark grey shading in the V-Foc condition) serve as cues to classification and interpretation.

Given the limited amount of data that the f0 and duration observation (Figures 5 and 6) is based on, we need to treat these results with caution but they can nevertheless be taken as an initial indicator that the interaction of prosodic factors does contribute to the encoding of focus conditions in Vietnamese. This said, given that word order remains constant and that no morphological markers are used to indicate focus, we claim that focus is exclusively prosodically (phonologically) marked in Vietnamese, through a combination of different prosodic parameters, including f0, duration and amplitude.

Even though object focus can only be realized in-situ in Vietnamese, there are non-canonical OSV sentences in Vietnamese. According to our informants, though, these are non-felicitous replies to object focus questions. Instead, they claim, OSV utterances must be interpreted as contrastive topic (Jannedy & McNay, 2007).

3 Information Structure

Based on our fieldwork notes and the small amount of data that we have collected so far, we have provided an overview of some general patterns that we have observed in our pilot data on the expression of focus in Vietnamese. The

results from the perception study show that listeners are generally quite able to detect the contextual meaning of the message (information structural content rather than just lexical content), that is, they are performing rather well, matching statements back to questions. That is, the generally, questions are well recoverable from the answer utterances, despite the range of variability observed in the actual renditions of the statements. This indicates to us that information structural content is consistently encoded via prosody. As the amount of data is too limited to conduct greater scale statistical analyses, we would like to conclude with some summary remarks on the descriptive patterns and observed tendencies that we found in on the Vietnamese data.

In summary, we find that focus in Vietnamese is exclusively expressed through phonology and prosody while the canonical word order must remain in tact. We have observed trading relationships between f0, duration and amplitude and possibly spectral tilt (voice quality) to mark emphasis, but how and in what context which parameters are used, remains unclear as of now. There also appear to be interactions between the lexical tonal specifications of utterances and the more global intonational requirements that an utterance must have to satisfy information structural requirements. Further, whether or not the different means that Vietnamese utilizes to signal emphasis are functionally equivalent or contrast with one another in any meaningful way or if they are socially distributed remains to be investigated. Naturally, these claims have to be tested against larger amounts of data collected from more speakers and under a greater variety of syntactic constructions and variability of tonal co-occurrences.

Appendix: Corpus for Perception Test

3 sentence-types in 5 focus conditions:

1.	Chuyện gì vậy?	(What's happening?)	[Phương đi xe đạp.]$_F$
2.	Ai đi xe đạp?	(Who is riding a bicycle?)	[Phương .]$_F$ đi xe đạp.
3.	Phương đi gì?	(What does Phương ride?)	Phương đi [xe đạp.]$_F$
4.	Phương làm gì với xe đạp?		
	(What does Phương do with the bicycle?)		Phương [đi]$_F$ xe đạp.
5.	Phương làm gì vậy?	(What does Phương do?)	Phương [đi xe đạp.]$_F$
6.	Chuyện gì vậy?	(What's happening?)	[Lan uống cà-phê.]$_F$
7.	Ai uống cà-phê?	(Who is drinking coffee?)	[Lan]$_F$ uống cà-phê.
8.	Lan uống gì?	(What does Lan drink?)	Lan uống [cà-phê.]$_F$
9.	Lan làm gì với cà-phê?		
	(Was macht Lan mit dem Kaffee?)		Lan [uống]$_F$ cà-phê.
10.	Lan làm gì vậy?	(What does Lan do?)	Lan [uống cà-phê.]$_F$
11.	Chuyện gì vậy?	(What's happening?)	[Mến uống nước.]$_F$
12.	Ai uống nước?	(Who is drinking water?)	[Mến]$_F$ uống nước.
13.	Mến uống gì?	(What does Mến drink?)	Mến uống [nước.]$_F$
14.	Mến làm gì với nước?		
	(Was macht Mến mit dem Wasser?)		Mến [uống]$_F$ nước.
15.	Mến làm gì vậy?	(What does Mến do?)	Mến [uống nước.]$_F$

References

Beckman, M. E. (1986) *Stress and non-stress Accent.* Foris Publications Holland, Dordrecht, the Netherlands.

Beckman, M. E. (1996) The Parsing of Prosody. *Language and Cognitive Processes,* 11 (1/2), 17-67.

Beckman, M.E. & Edwards, J. (1994) Articulatory Evidence for Differentiating Stress Categories. In *Phonological Structure and Phonetic Form. Papers in Laboratory Phonology III*, Keading, P.A. (ed.) Cambridge University Press.

Beckman, M. E., & J. B. Pierrehumbert (1986) Intonational Structure in Japanese and English. *Phonology Yearbook* 3:255--309.

Brunelle, M. (2003). Coarticulation Effects in Northern Vietnamese. *Proceedings of the 15th International Congress of Phonetic Sciences (ICPhS),* Barcelona**:** 2673-2676.

Brunelle, M. (2006) Tone Perception in Vietnamese Dialects. Presentation given at the *TIE-2 Conference* at the ZAS (Berlin), Sept. 2006.

Brunelle, M & Jannedy, S. (2007) Social Effects on the Perception of Vietnamese Tones. Accepted to *ICPhS 2007* in Saarbrücken.

Dung, B.T., Huong, T. T. & Boulakia, G. (1998) Intonation in Vietnamese, in D. Hirst & A. Di Cristo (eds.), *Intonation Systems: A Survey of Twenty Languages.* Cambridge University Press, Cambridge.

Ingram , J. & Nguyễn, T. (under review) Stress, tone and word prosody in Vietnamese compounds. Submitted to *Journal of the Acoustical Society of America.*

Jannedy, S. & Fiedler, I. (manuscript) Prosody of Focus Marking in Ewe. Humboldt University of Berlin.

Jannedy, S. & McNay, A. (2007) Contrastive Topic Marking in Vietnamese – Prosody, Word Order, And Morphology. Paper presented at the *3rd Workshop on Contrast: Towards a Closer Definition.* Zentrum für Allgemeine Sprachwissenschaft, Berlin).

Krifka, M. (2006) Notions of Information Structure. In Féry, C., Fanselow, G. & Krifka, M. (eds.) *Interdisciplinary Studies on Information Structure* (ISIS) 06 (pp. 13-54). Potsdam: Universitätsverlag Potsdam.

Krifka, M. (2007) The Semantics of Questions and the Focusation of Answers. In Lee, Ch., Gordon, M. & Büring, D. (eds.) *Topic and Focus*. Dordrecht, Springer, 139-150.

Ladd, R. D. (1980) *The Structure of Intonational Meaning.* Indiana University Press, Bloomington.

Ladd, R. D. (1996) *Intonational Phonology.* Cambridge University Press.

Michaud, A. (2004) Final Consonants and Glottalization: New Perspectives from Hanoi Vietnamese. *Phonetica* 61: 119-146.

Michaud, A. (2006) Replicating in Naxi (Tibeto-Burman) an Experiment Designed for Yorùbá: An Approach To 'Prominence-Sensitive Prosody' vs.

'Calculated Prosody'. Proceedings of *Speech Prosody 2006*, pp. 819-822, Dresden, 2-5 May 2006.

Michaud, A. & Vu, T.N. (2004) Glottalized and Non-Glottalized Tones under Emphasis: Open Quotient Curves remain stable, F0 curve is modified. Proceedings of the *Speech Prosody 2004*, pp. 745-748, Nara, Japan,.

Michaud, A., Vu Ngoc, T., Amelot, A. & Roubeau, B. (2006) Nasal release, nasal finals and tonal contrasts in Hanoi Vietnamese: an aerodynamic experiment , *Mon-Khmer Studies*, 36.

Morén, B. (2003) The Mora is the Tone Bearing Unit in Thai. Presentation at the Annual Meeting of the Linguistics Society of America, Atlanta, USA.

Nguyễn Đ.-H (1990) *Vietnamese*. London Oriental and African Language Library. John Benjamins, Amsterdam and Philadelphia.

Nguyễn, V. L. & Edmondson, J. (1997) Tones and Voice Quality in Modern Northern Vietnamese: Instrumental Case Studies. *Mon-Khmer Studies* 28: 1-18.

Pham, A. (2003) The Key Phonetic Properties of Vietnamese Tone: A Reassessment. Paper published at the *Proceedings of the 15th International Conference of Phonetic Sciences* (ICPhS).

Pierrehumbert, Janet. (1980) *The Phonology and Phonetics of English Intonation*. Ph.D. dissertation, MIT.

Poser, B. (2002) Amplitude, Intensity & Loudness (manuscript). Downloadable at: *www.ling.upenn.edu/phonetics/docs/Amplitude.pdf*

Selkirk, E. O. (1984) *Phonology and Syntax: The Relation between Sound and Structure.* Cambridge, MA: MIT Press.

Selkirk, Elisabeth O. (1995). Sentence prosody: Intonation, stress and phrasing. In *Handbook of Phonological Theory*, ed. John Goldsmith, pp. 550–569. Cambridge, MA: Blackwell.

Thompson, Laurence C. 1965. *A Vietnamese Reference Grammar.* University of Washington Press, Washington. (2nd edition, 1987, University of Hawai'i Press, Honolulu).

Tran, H. M. (1967) Tones and Intonation in South Vietnamese. Series A - Occasional Papers #9, *Papers in Southeast Asian Linguistics* No.1. Nguyễn,

D. L., Trễn, H. M. & D. Dellinger (eds.). Canberra, Linguistics Circle of Canberra.

Xu, Y. (1999). Effects of Tone and Focus on the Formation and Alignment of F0 Contours. *Journal of Phonetics* 27: 55-105.

Yip, M. (2002) *Tone.* Cambridge University Press.

Stefanie Jannedy
Humboldt Universität zu Berlin
SFB 632 „Informationsstruktur“
(Location: Mohrenstr. 40-41)
Unter den Linden 6
10099 Berlin
Germany
jannedy@ling.ohio-state.edu