ISCA Archive
http://www.isca-speech.org/archive

Speech Prosody 2012
Sixth International Conference
Shanghai, China
May 22-25, 2012

Functions of word order and intonation in information structuring in Puyuma

Anastasia Karlsson, Arthur Holmer

Centre for Cognitive Semiotics, Centre for Languages and Literature, Lund University, Sweden
Anastasia.Karlsson@ling.lu.se, Arthur.Holmer@ling.lu.se

Abstract

This paper presents the first results of analyzing the functions of intonation and word order in Puyuma. Puyuma is an endangered Austronesian language spoken in Taiwan. As other languages of this branch, Puyuma is verb initial and displays variations in word order, so that VOS, SVO and VSO are attested. The aim of our study is to explain these word variations. We investigate the connection between information structure, word order and intonation in three Formosan languages: Puyuma, Seediq and Bunun. We show that in Puyuma information structure is reflected in word order while the main function of intonation is to cue phrasing. The study is performed on semi-spontaneous speech.
Index Terms: intonation, word order, information structure, phrasing, Puyuma.

10.21437/SpeechProsody.2012-34

1. Introduction

Puyuma is an Austronesian language spoken in Taiwan, i.e. one of the languages often referred to as Formosan (although this is not a genetic grouping). We are working with the Nanwang dialect, spoken in Nanwang village. Though it is described as having less than 1000 speakers [1], the actual number is much lower. Today only people aged over 50 speak Puyuma as their first language. Puyuma is taught in school as a minority language and some younger Puyuma understand Puyuma but cannot speak it freely and do not use it. Most younger Puyuma have Mandarin Chinese as their first language, some others Minnan (Hokkien).

Puyuma is verb initial and displays a certain word order variation. The three word orders Verb Object Subject (VOS), Verb Subject Object (VSO) and Subject Verb Object (SVO) are attested:

'My father plants sweet potatoes.'

VOS: semarem dra bunga i namali
V O S
SVO: i namali semarem dra bunga
S V O
VSO: semarem i namali dra bunga
V S O

This variation was shown in [2] and replicated by one of the authors, Arthur Holmer, by collecting non-spontaneous material (in 2009). Speakers were asked to translate sentences from Chinese to Puyuma. Though word order variation was recorded, the reasons for it remained unclear. Our new research question is what conditions word order in Puyuma at clause level, at NP level and within relative clauses. To answer this question we perform an investigation by integrating syntax, information structure and intonation. In this paper we present first steps in our analysis.

1.1. First observations: phrasing

Intonation in Puyuma has not been studied. There is a short description of tonal contours in declaratives and interrogatives in a published Puyuma grammar [1]. Puyuma has no lexical tones. In our semi-spontaneous and spontaneous recordings of four speakers we observed systematic tonal patterns that reflect prosodic phrasing. Two places in a word are marked for intonation purposes [3]. Thus, almost every content word tends to build a prosodic phrase, analyzed as an accentual phrase (AP). AP is marked by a high boundary tone on the penultimate (H_{ap}). The word final syllable gets a high (H%) or a low (L%) boundary tone, indicating a major prosodic phrase (a tentative notion). The phonological transcription proposed here is still in progress. Nevertheless, our analysis of the function of the choice between high and low boundary tone is the result of a careful examination of a large set of data.

The most striking contrast between the choice of low and high boundary tone occurs in declaratives as opposed to polar (yes-no) interrogatives. In Puyuma, declaratives end with a high boundary, while interrogatives have a low boundary tone, which is a typologically unusual feature.

The contrast between low and high boundaries is found in other cases as well and thus is not only governed by the sentence type (see [3] for more details). Contexts in which differences in choice between low and high boundary tones occur are as follows:

High boundary tone	Low boundary tone
[declarative] *high*	[interrogative] *low*
[[S] *high* [VO]] *high*	[[S*topic*] *low* [VO]] *high*
[[Head *high* + Dep]] *high*	[[Dep *low* + Head]] *high*

An example of word order variation between Head and Dependent is Noun + Adjective [Head + Dep] vs. Adjective + Noun [Dep + Head] in Puyuma.

The choice between low and high boundary tones is governed by discourse. With a high boundary, the speaker signals that she is finished with the main information, while a low boundary means that information needs to be added. Thus, an interrogative does not add any new information; rather, this information has to be supplied by the answer. Thus, an interrogative is marked by a low tone, and a declarative by a high tone. Information relations within utterances are also marked by the same principle. For instance, topicalized elements (subject or object) do not contribute any new information and therefore are marked by a low boundary tone, which signals that the new information will follow.

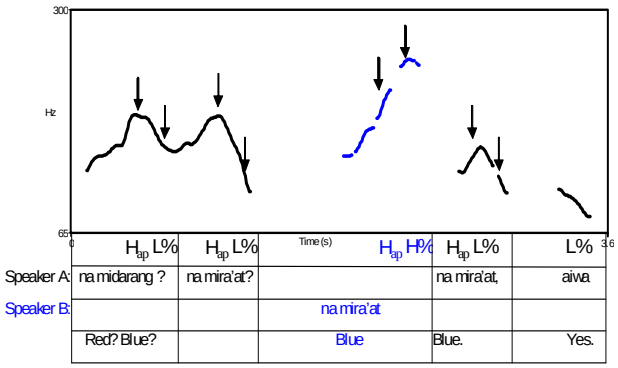

Figure 1: *Boundary tones in declaratives and interrogatives. An example from a spontaneous dialogue between two Puyuma speakers. Arrows show penultimate and final syllables.*

The aim of the present study is to find out whether word order and intonation interact with information structure in Puyuma, and, if so, how. Our previous observations on the types of boundary tones in Puyuma are the ground for the method chosen for this investigation.

2. Method

2.1. Speech material

Material for recordings was inspired by [4] though we modified stimuli to suit our research questions and local particularities. Stimuli used for the present study were aimed at triggering different information structures. Speakers were shown four pictures and asked to answer three questions for each picture. The three questions reflect three conditions: all new (Condition 1), given agent and new patient (Condition 2), new agent and given patient (Condition 3). Example:

Picture 1: A woman eating a banana and a man eating an apple.

Condition 1 question: What happens?
Condition 2 question: What is the woman eating and what is the man eating?
Condition 3 question: Who is eating the apple and who is eating the banana?

Answers were (for Picture 1 of four):

Condition 1: VOS:
maekan dra belbel na babayan
eat OBJ banana DET woman
“The woman is eating a banana”

Condition 2: TOP S V O:
na babayan i maekan dra belbel
DET woman TOP eat OBJ banana
“The woman, she is eating a banana”

Condition 3: SVO
na babayan maekan dra belbel
DET woman eat OBJ banana
“It is the woman who is eating banana”

TOP O CP S V
na belbel i amau na babayan na maekan
DET banana TOP CP DET woman DET eat
“The banana, it is the woman who is eating it”

Three speakers were recorded (all women), ages ranging from 57 to 77. All are speakers of the Nanwang dialect. Measurements of F0 were performed for two speakers, named Speaker A and Speaker P here.

2.2. Recording and analysis

The subjects were recorded in Taiwan, Nanwang village in June 2011 using a portable Edirol R-09 digital recorder and a lapel microphone. The utterances were digitized at 48 kHz sampling rate and 16-bit amplitude resolution and stored in .wav file format. Speaker A was recorded at her home in interaction with another speaker. Speaker P was recorded in a hotel room.

For acoustic measurements, the program *Praat* was used to display the waveform, spectrogram and fundamental frequency contour of each utterance. Minimum F0 values in the first syllable and maximum F0 values in the last syllable (or penultimate when the word is within topic (see section 1.1)) in each word were measured manually for each utterance. Function words are not included. In the Figures 4-7, the maximum F0 is shown with no consideration as to whether it occurred on the penultimate (in topicalized elements) or on the last syllable (elsewhere). By topicalized elements we mean elements which have been moved leftwards past the verb to clause-initial position, when this element is interpreted as GIVEN or as stage setting topic, rather than NEW (as in a cleft).

To be able to compare the two speakers, the measured F0 values in Hz were converted to semitones (St) and then normalized in accordance with [5]. A fixed semitone scale is used where the unit St is defined by:

$$St = 12[\ln(Hz/100)/\ln 2]$$

Normalization is performed by subtracting each subject’s average F0 in St (measured across production of each target sentence) from the individual St values. A total of 46 utterances for Speaker A and 29 utterances for Speaker P were measured.

3. Results

Figures 2 and 3 illustrate actual F0 courses in Condition 1 (VOS) and Condition 3 (SVO with S in focus). Plots showing the F0 measurement points in normalized semitones are presented in Figure 4 for both speakers for Condition 1, Figure 5 for Condition 2 and Figure 6 for Speaker A and Figure 7 for Speaker P for Condition 3. V = verb, S = subject, O = object, TOP S = topicalized subject, TOP O = topicalized object, CP= copula.

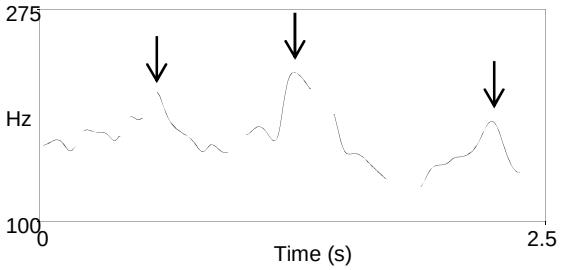

Figure 2: *Illustration of tonal boundaries in Condition 1 (all new). Word order is VOS. Arrows show high boundary tones on last syllables of V, O, and S. The utterance is [midinkesan]V [dra idus]O [na mainayan]S "A man is holding a spoon" (as an answer to question "What do you see in the picture?"). Speaker P.*

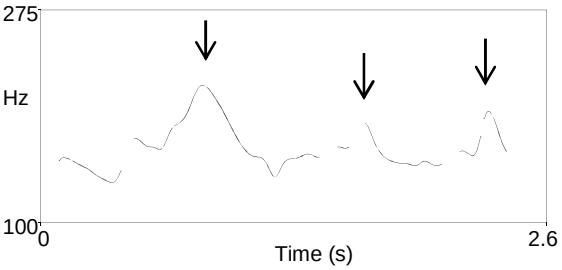

Figure 3: *Illustration of tonal boundaries in Condition 3 (new subject). Word order is SVO. Arrows show high boundary tones on last syllables of S, V, and O. The utterance is [na mainayan]S [midinkesan]V [dra idus]O "It is the man who is holding a spoon" (as an answer to question "Who is holding the spoon?"). Speaker P.*

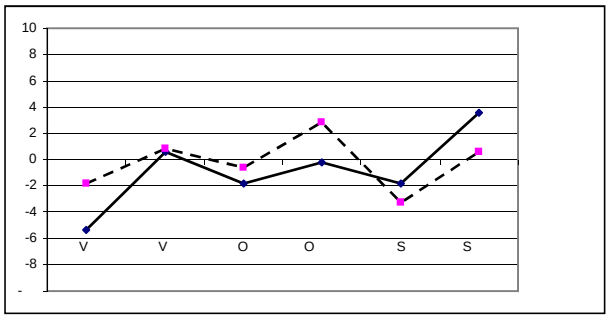

Figure 4: *Normalized F0 measurement points in semitones for Speaker A (plain line) and Speaker P (dotted line). Condition 1: all new. The word order displayed is VOS. Two measurement points per word are shown.*

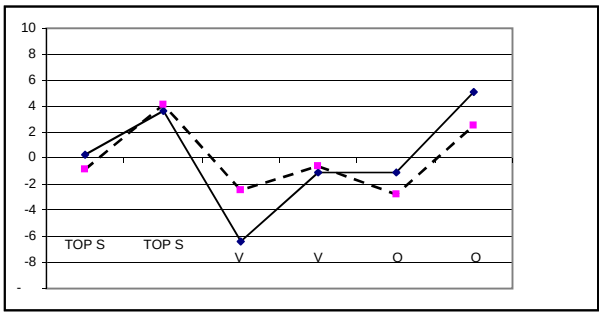

Figure 5: *Normalized F0 measurement points in semitones for Speaker A (plain line) and Speaker P (dotted line). Condition 2: given agent (S) and new patient (O). The word order displayed is TOP S + VO. Two measurement points per word are shown.*

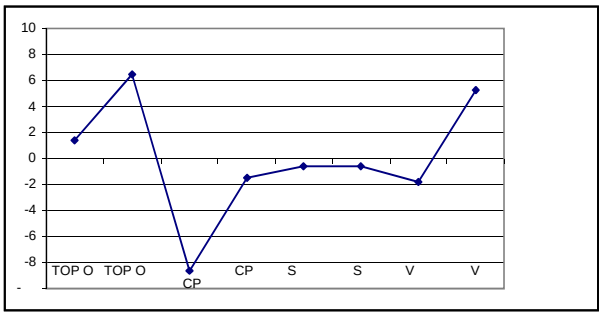

Figure 6: *Normalized F0 measurement points in semitones for Speaker A. Condition 3: new agent (S) and given patient (O). The word order displayed is TOP O + CP S V. Two measurement points per word are shown.*

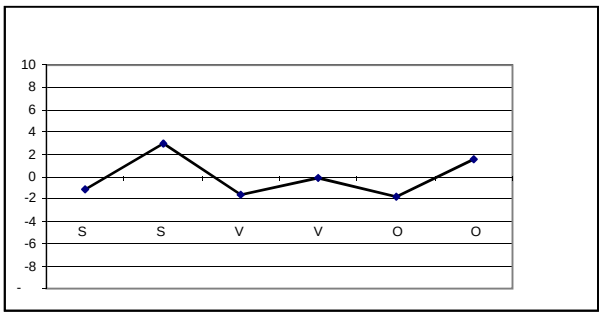

Figure 7: *Normalized F0 measurement points in semitones for Speaker P. Condition 3: new agent (S) and given patient (O). The word order displayed is SVO. Two measurement points per word are shown.*

3.1. Word order

Condition 1, "all new", triggers VOS (illustrated in Figure 4) or (in a few cases) S TOP + VO word order with both speakers. Condition 2, "given agent and new patient" triggers TOP S + VO (Figure 5), i.e. the agent is topicalized. Condition 3, "new agent and given patient" triggers SVO with Speaker P (Figure 7). The two main word order patterns for Speaker A in Condition 3 are clefting of the new agent CP S + VO and topicalization of the given patient TOP O + CP SV (Figure 6).

3.2. Tonal contours

As can be observed from the figures, each syntactic phrase is marked by a prosodic boundary on its right edge. The boundaries are signaled by a tonal rise. This is realized on the final syllable of the final lexical word in each syntactic phrase. When a word is topicalized, the final boundary tone is a fall instead of a rise, and it is preceded by a high phrase accent on the penultimate. Obligatory boundaries are found in two positions. Thus, the utterance final word always gets a high boundary tone. When a syntactic unit is moved to the left it is marked by a boundary tone which gets higher F0 values than a boundary tone that marks a syntactic unit in the same position in all-new condition VOS with no left-dislocation (compare Figure 4 and 7 Speaker P).

4. Discussion

Our study shows that information structure is mainly reflected by word order in Puyuma. The "all new" condition triggers the default word order VOS. When the subject is in focus it is moved to initial position, yielding SVO word order (Figure 7) or topicalization of the given object and clefting of the focused subject (Figure 6). When the object is in focus it is not moved to initial position, instead the subject (which is given in this Condition) is topicalized. Thus, object is not moved to be focused.

Intonation is used to mark phrase boundaries, and their functions are conveyed by the choice between low or high boundary tone. Besides the phrasing tones, we do not find any extra pitch accents added to mark a focused word. Thus, we find rising or falling boundary tones at the end of the phrases.

Previously [3] we found that each word is phrased as an accentual phrase cued by a rising F0 on its penultimate. In the present investigation we could show that tonal contours also reflect syntactic phrases and we find rising boundary tones at the right edge of each syntactic group. The overlapping between prosodic and syntactic phrasing is quite high, but in many cases, especially in very fast speech, syntactic phrases can be clumped together prosodically into one prosodic phrase. However, there are two cases which are always marked by tonal boundaries. First, an utterance is always marked by a final high boundary tone. It is realized as a fall from a high level and is truncated when the coda is voiceless. Second, when a syntactic unit is left-dislocated, it obligatorily gets a boundary tone. In this case the tone is high, similar to the utterance final tone or low, preceded by a high tone in the penultimate, when the word is in topic position. Topic is marked by the optional topic marker *i* in Puyuma. Given the optionality of the topic marker, the low boundary often becomes the only means of contrasting TOP S + VO and SVO constructions.

5. Conclusions

Even though we cannot draw any strict conclusions about the role of prosody in information structuring in Puyuma from the presented material, we have evidence for an interaction between information structure, syntax and intonation.

Information structure is primarily reflected by word order in Puyuma. Intonation serves mainly to signal prosodic boundaries in Puyuma. We find a fairly high overlap between syntactic phrases and prosodic grouping but only two positions are obligatorily marked by prosodic boundaries, namely the end of an utterance and the end of a left-dislocated syntactic unit. By this, there is an indirect relation between prosodic phrasing and information structure in Puyuma: information structure > syntax > prosodic phrasing. Intonation alone is not used for conveying information structure in Puyuma.

6. Acknowledgements

The work reported in this paper has been carried out within the research project *Information structure, prosody and the origins of Formosan relativization*, funded by the Swedish Research Council. We further gratefully acknowledge financial support from the Centre for Cognitive Semiotics at Lund University, and from the Ministry of Education, Taiwan.

7. References

[1] Teng, Stacy Fang-Ching. A reference grammar of Puyuma. Canberra: Pacific Linguistics, 2008.

[2] Huang, Mei-Jin. Beinan-yu cankao yufa. [Reference grammar of Puyuma]. Taipei: Yuanliou, 2000.

[3] Karlsson, A. and Holmer, A., "Interaction between word order, information structure and intonation in Puyuma", in Mitsuaki Endo [Ed], Papers in Austroasiatic and Austronesian Linguistics, 28-38, 2011.

[4] Skopeteas, Stavros et al, Questionnaire on Information structure. Online: http://opus.kobv.de/ubp/volltexte/2007/1241/ . 2006.

[5] Fant, G. and Kruckenberg, A., "Analysis and synthesis of Swedish prosody with outlooks on production and perception", in Fant, G., Fujisaki, H., Cao, J. and Xu, J., [Eds], From traditional phonology to modern speech processing, 73-95, Foreign Language Teaching and Research Press, Beijing, 2004.