PRODUCTION AND PERCEPTION OF FOCUS IN SÜMI

Amos Teo, Melissa Baese-Berk

University of Oregon
ateo@uoregon.edu, mbaesebe@uoregon.edu

ABSTRACT

This paper investigates the homophony/polysemy between a morphological agentive marker and a contrastive focus marker in Sümi, a Tibeto-Burman language of Northeast India. Both are realized by a phrasal suffix *-no* that attaches to grammatical subjects, but the interpretation of the suffix varies by clause type. The present study examines whether transitive and intransitive subjects in contrastive focus receive any special prosodic marking that is recognizable to native listeners. The study has implications for understanding the development of agentive/focus marking in Sümi, as well as other languages of the Himalayas, and in New Guinea and Australia where similar homophony/polysemy between agentive and focus markers has been found.

Keywords: prosody, focus, perception, differential case marking, Tibeto-Burman

1. INTRODUCTION

1.1. Differential subject marking

Differential subject marking (DSM) refers to a system of case marking that does not merely encode the grammatical relation of "subject", but also semantic and pragmatic information, which include: animacy of the referent, contrastive focus etc. These systems are considered "partial and probabilistic" [3], since speakers appear to have some freedom in whether or how they mark a noun for case, without changing the representational meaning of an utterance. Research has increasingly invoked information structure or management to explain some of the triggers of DSM. For example, the ergative in Jingulu also appears to mark discourse prominence [6]. Yet, despite the appeal to information structure, few studies examine co-occurrences of DSM with prosodic patterns, which cross-linguistically are relevant to the realization of information-structural categories [2]. Some notable exceptions are work on intonation and case marking in Burmese [4] and Jaminjung [7]. Even so, no work has investigated whether speakers of such languages use prosodic cues to help in the interpretation of these case markers, or if they rely more on top-down information, e.g. the type of sentence in which the marker appears. In the present study, we examine the co-occurrence of DSM with prosody in perception and production of an under-studied language.

1.2. Language background

Sümi is a Tibeto-Burman language spoken mainly in Nagaland, North-East India by an estimated 104,000 speakers. The language has a system of DSM, whereby certain grammatical subjects in Sümi are "optionally" marked by a phrasal suffix *-no*.

The function of *-no* depends on the sentence type in which it occurs. In verbless sentences, as in (1), *-no* is obligatory only when there is narrow focus on the subject. In transitive sentences (2 participants), as in (2), *-no* is also obligatory, but it simply marks the agent/doer of the action, whether the agent is in narrow focus or not. In intransitive sentences (1 participant), as in (3), *-no* is optional, and its use is often associated with narrow focus, though this depends on the speaker.

(1) *Atsü-no akijeu.* /àtsɨ no àkìʒèū/
'The dog (not anything else) (is) bigger.'
(2) *Atsü-no awu ha cheni.* /àtsɨ no àwù hā tʃènī/
'The dog is chasing a chicken.' or
'The dog (not anything else) is chasing a chicken.'
(3) *Atsü(-no) zü ani.* /àtsɨ (no) zɨ̀ ànī/
'The dog is sleeping.' or
'The dog (not anything else) is sleeping.'

Sümi also has three contrastive tones distinguished in production by F_0 height [8]. The suffix *-no* itself is not specified for lexical tone and since it occurs at the right-edge of the phrase, it is a potential site for intonational tones. However, given the widespread use of F_0 for lexical differentiation in Sümi, it was unclear if native listeners would rely on prosodic differences in perception. We were therefore hesitant to begin with a resource-intensive study of focus production and decided to first run a perception experiment with stimuli produced by a non-naïve native speaker. Specifically, we asked:

(1) Are Sümi listeners able to use prosodic cues to distinguish between the agentive function vs. the narrow focus function of the suffix *-no* in transitive vs. intransitive clauses?

(2) Does sentence type (transitive vs. intransitive) affect listeners' interpretation of the suffix *-no* as agentive (i.e. marking a doer of an action)?

We tested the hypothesis that agentive and narrow focus *-no* were distinguishable by prosodic cues in this perception task. If agentive and narrow focus *-no* were homophonous to listeners, we expected only sentence type to affect its interpretation, with listeners more likely to rate verbless and intransitive sentences with *-no* as having narrow focus than transitive sentences. If the suffixes were not homophonous, we expected that listeners would rate sentences that had been uttered with narrow focus prosody as having narrow focus. However, an interaction with sentence type was also possible, with prosodic cues only affecting listeners' interpretation of transitive sentences, since *-no* is optional in intransitive sentences, and previous language consultants had associated its appearance with narrow focus.

2. EXPERIMENT

2.1. Participants

10 participants, 5 male and 5 female, took part in the perception experiment. They were recruited from the Institute of Chartered Financial Analysts of India (ICFAI) in Dimapur, Nagaland. All participants were native Sümi speakers. They were all between the ages of 20-25 years, with no hearing difficulties reported.

2.2. Materials

For the experiment, three sentence types in two focus conditions were studied. The sentence types were:

Transitive:	_____*-no ha cheni.*
	'_____ is chasing.'
	(object not explicitly mentioned)
Intransitive:	_____*-no zü ani.*
	'_____ is sleeping.'
Verbless:	_____*-no akijeu.*
	'_____ (is) bigger.'

The transitive sentences were still considered grammatical when the object was not mentioned. The verbless sentences were used as a control group, since it was only possible to interpret them as having narrow focus.

The focus conditions were broad (non-narrow) focus on the sentence vs. narrow focus on the subject. Our language consultant was asked to respond to a set of pictures with two different questions, using the same sequence of words for each focus condition. The questions used were:

Broad:	*Kiu shi ani kea?*
	'What is happening?'
Narrow:	*Khu no ha cheni/zü ani/akijeu?*
	'Who is chasing/sleeping/bigger?

12 lexical nouns referring to animals found in Nagaland were used. These were balanced for tone on the final syllable, with four nouns ending with Low tone, four ending with Mid and four ending with High. In addition, two nouns, one ending with Low and one with Mid tone, were used for training purposes. All words were expected to be known by native speakers.

All audio stimuli were produced by Dr Salome Kinny, the main language consultant for the project. The recordings were done using a Tascam DR-100MK-II and head microphone in a quiet room with the lead researcher present.

The visual stimuli that accompanied the written Sümi question prompts were illustrated by Mr Obeto Kinny, who is a member of the Sümi community. He was asked to draw pictures of animals in a style that would be recognizable to people in Nagaland. The same pictures were then used in the perception experiment.

2.3. Procedure

The perception experiment was run in PsychoPy (v3.0) [5], with pre-recorded audio instructions and written instructions in Sümi. Participants listened to the stimuli using Sony MDR7506 headphones in a quiet room.

The experiment was divided into two parts. In the first part, participants were told that they would see some pictures and that a speaker would describe them in Sümi. Sometimes, the speaker would be emphasizing who was doing the action; other times, the speaker would be emphasizing the action. Effort was made to use language similar to what previous language consultants had used to describe the two functions of *-no*. Participants had to decide what the speaker was emphasizing in a four-alternative forced choice task and responded by pressing one of four keyboard keys, depending on whether they thought the speaker was emphasizing the actor or the action, with the middle two options allowing them to indicate uncertainty. After training without feedback, they were presented with 48 target stimuli (transitive and intransitive sentences). All visual stimuli featured a pair of animals.

In the second part of the experiment, the participants were told that they would still see pictures of the same animals in pairs but this time, the speaker would state that one animal was the bigger one. The participants had to decide whether the speaker was emphasizing which of the two was bigger and respond with either 'yes' or 'no'. After training without feedback, they were presented with 12 target stimuli (verbless sentences).

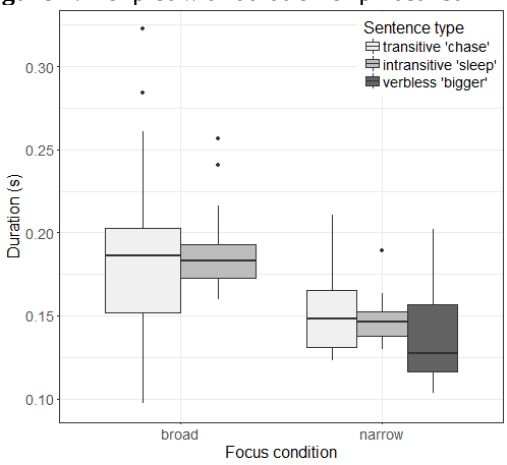
3. RESULTS

3.1. Analysis of production stimuli

A number of acoustic measures were done on the production stimuli to test the effect of focus condition and/or sentence type: (a) duration of suffix *-no*; (b) F_0 across *-no*; (c) duration of last syllable of noun preceding *-no*; and (d) F_0 over the final syllable of the intransitive and transitive sentences, which both end with the same morpheme *-ni* 'present tense'. The first two measures were done because the suffix, which is not specified for lexical tone, was identified as a potential location for prosodic events, similar to the Japanese particles *wa* and *ga* [1], [9]. The third measure was done because the last syllable of nouns is where the main tonal contrast is typically found in Sümi nouns and a potential site for prominence marking. The fourth measure was to look for evidence of post-focal F_0 compression.

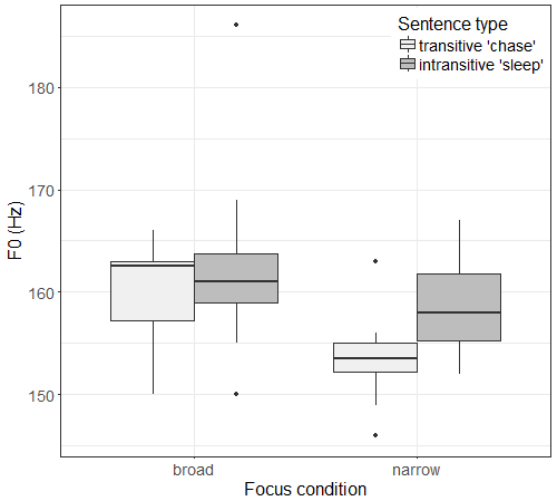
Of these measures, only (a) duration of *-no* and (d) F_0 over the final syllable were affected by focus condition and/or sentence type. Figure 2 shows that the duration of the vowel of *-no* was shorter in narrow focus than in broad focus, regardless of sentence type, although there was more variance in the transitive sentence under broad focus. A significant difference was found between the two focus conditions, $F(1,65) = 24.673$, $p < .001$, but no effect was found for sentence type, $F(2,65) = .471$, $p = .63$; or interaction with sentence type, $F(1,65) = .013$, $p = .91$.

Figure 2: Boxplot with duration of phrasal suffix *-no*



There was also some evidence for post-focal F_0 compression: F_0 at the midpoint of the final syllable of the sentence was lower in narrow focus than in broad focus, though the difference was larger in the transitive sentences than in the intransitive ones, as shown in Figure 3. A significant difference was found between the two focus conditions, $F(1,52) = 13.147$, $p < .001$, as well as between the sentence types, $F(1,52) = 5.804$, $p = .02$; but no interaction effect was found, $F(1,52) = .811$, $p = .37$.

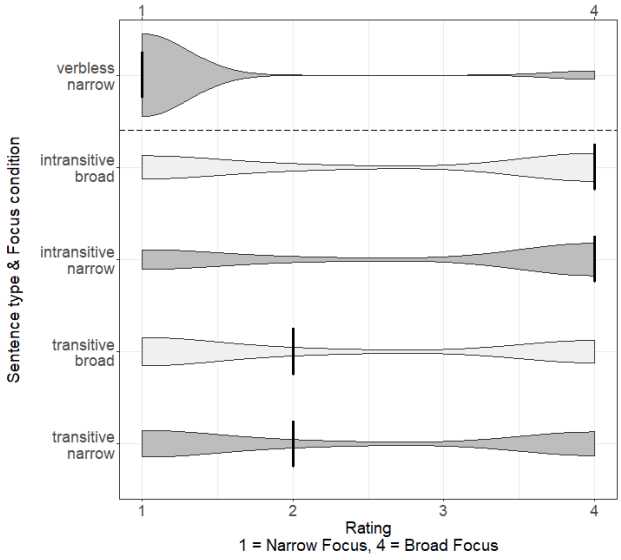
Figure 3: Boxplot of F_0 at vowel midpoint of final syllable of sentence



3.2. Perception experiment results

The results of the perception experiment were converted to a 2-point scale for verbless sentences and a 4-point scale for transitive and intransitive sentences, where "1" corresponds to a narrow focus interpretation and "4" to a broad focus interpretation. Figure 4 presents a violin plot showing that sentence type affected listeners' interpretation of sentences with the suffix *-no* more than any prosodic cues associated with broad vs. narrow focus. Listeners rated verbless sentences as having narrow focus. Listeners tended to rate transitive sentences with *-no* as having narrow focus on the subject; and intransitive sentences with *-no* as having broad focus.

Figure 4: Violin plot with rating of phrasal suffix *-no* in different sentence types and focus conditions. Crossbars indicate the median score.



We analyzed the responses to the transitive and intransitive sentence stimuli using a mixed effects model with sentence type, focus condition, and the interaction between sentence type and focus as fixed factors and participant as a random effect. The results support the picture presented above: only sentence type is a significant predictor of rating, ($\chi^2(2) = 10.143$, $p = .006$). On the other hand, focus condition was not a significant predictor of rating ($\chi^2(2) = 2.061$, $p = .357$), nor was the interaction between sentence type and focus condition ($\chi^2(2) = .394$, $p = .530$). These results reflect the trend we see above where we find listeners more likely to rate the transitive sentences as having narrow focus on the subject; and intransitive sentences as having broad focus. Estimates and *t*-values from the best fitting model are presented in Table 1.

Table 1: Estimates and *t*-values for best fitting model for interpretation score.

Fixed effect	Estimate	Standard error	*t* value
Sentence type	.283	.160	1.776
Focus condition	.075	.160	.470
Sentence-Focus Interaction	.142	.226	.628

4. DISCUSSION

The results of the study show that the agentive and narrow focus forms of *-no* are not distinguished by listeners via prosodic cues in this task. Despite the presence of acoustic differences in the stimuli, i.e. a shorter suffix duration and post-focal F_0 compression in the narrow focus condition, listeners relied only on sentence type to interpret the function of the suffix.

An unexpected finding was that listeners tended to rate intransitive sentences with *-no* as having broad focus, and not narrow focus, given that the suffix is not obligatory on intransitive subjects and was treated as a focus marker in these sentences by previous language consultants. It was similarly unexpected that listeners tended to rate transitive sentences as having narrow focus, instead of broad focus.

Here, we consider the possibility of a task effect because the method involved playing sentences that included the verb. In natural speech, speakers can unambiguously achieve narrow focus by producing the subject noun phrase alone, so the inclusion of the verb in the stimuli may have led listeners to rate the intransitive sentences as having broad focus. This effect may have also been present in the transitive stimuli but was mitigated by the omission of the grammatical object.

Nevertheless, the presence of such a task effect does not negate our main findings. In fact, the interpretation of the intransitive sentences as having broad focus may also be driven by language change. The language consultants who would interpret *-no* in intransitive sentences as a narrow focus marker were often older than many of the experiment participants, and they also came from more rural areas. It is therefore possible that we are seeing a semantic shift in progress, whereby younger urban speakers are treating *-no* less like an agentive/focus marker and more like a grammatical subject marker. This idea is supported by discussions with younger speakers, as well as work-in-progress looking at inter-speaker variation in the use of *-no* in video description tasks.

Overall, the findings support the broader view that speakers of some languages with DSM use top-down information, such as sentence type, to interpret the case markers. However, we are hesitant to generalize to all other languages with DSM, because our results may be due to Sümi being a tonal language, in which F_0 is already used for lexical differentiation. We would therefore welcome similar research on non-tonal languages with DSM that have richer inventories of intonational units.

In terms of future work, we are seeking more participants for this study. Given limited resources in the field and the results of our current study, we feel that a large production study would only be worth doing if we can identify other possible contexts in which prosodic cues might play a role in disambiguating the functions of case markers, e.g. counter-expectation. Similarly, if listeners had shown sensitivity to prosodic cues in this task, we would consider manipulating the acoustic characteristics of the stimuli to identify which cues were most salient. We would also like to advocate for the inclusion of such perception experiments as part of the repertoire of tools available to linguistic fieldworkers, who are usually not native speakers *or* listeners of the languages they work on. Such experimental work can reveal insights that would not be found by working with a few language consultants or even by doing a production study with many speakers.

5. SUMMARY

We have shown evidence that although the phrasal suffix *-no* is produced with different prosodic cues in its agentive and narrow focus functions, the forms are homophonous to native Sümi speakers in perception. This adds to the view that the interpretation of differential case markers in Sümi is driven not by differences in the forms of the case markers, but by the use of the case markers in certain sentence types.

6. REFERENCES

[1] Finn, A. 1984. Intonational Accompaniments of Japanese Morphemes *wa* and *ga*. *Language and Speech* 27(1), 47-57.

[2] Lambrecht, K. 1994. *Information structure and sentence form*. Cambridge: Cambridge University Press.

[3] McGregor, W. 2010. Optional ergative case marking systems in a typological-semiotic perspective. *Lingua* 120(7), 1610–1636.

[4] Ozerov, P. 2014. Pragmatics and information structure apart. Differential Case Marking in Burmese. Paper presented at the workshop *Case and agreement: between grammar and information structure*, Hebrew University Jerusalem, 13 January.

[5] Peirce, J.W. 2007. PsychoPy - Psychophysics software in Python. *J Neurosci Methods* 162(1-2), 8-13.

[6] Pensalfini, R. 1999. The rise of case suffixes as discourse markers in Jingulu: a case study of innovation in an obsolescent language. *Australian Journal of Linguistics* 19(2), 225-240.

[7] Schultze-Berndt, E. 2016. Agents in focus. "Optional" ergativity in Jaminjung and information structure. Paper presented at the Linguistics Association of Great Britain meeting, University of York, 7 September.

[8] Teo, A. 2014. *A phonological and phonetic description of Sumi, a Tibeto-Burman language of Nagaland.* Canberra: Asia Pacific Linguistics.

[9] Venditti, J. 2000. *Discourse structure and attentional salience effects on Japanese intonation.* (Doctoral dissertation). The Ohio State University, Columbus, OH.