Prosody in ʔayʔaǰuθəm: Downstep and reset as cues to constituency*

Marianne Huijsmans
University of British Columbia

Abstract: This paper explores prosodic constituency in ʔayʔaǰuθəm, an endangered Central Salish language. It compares the syntax-prosody mapping for pairs of sentences with the same string of lexical items but different syntactic structures. The results support the hypothesis that syntactic phrases map to prosodic phrases in ʔayʔaǰuθəm, with prosodic phrasing indicated by downstep and partial reset. A secondary finding of this study is that focus is associated with increased pitch movement on the head of the focused constituent.

Keywords: Salish, ʔayʔaǰuθəm, prosody, constituency, focus, downstep, reset

1 Introduction

In this paper, I examine prosody in ʔayʔaǰuθəm, a Central Salish language. This preliminary study focuses on the mapping between syntactic phrases (XPs) and prosodic phrases. Because most of the fluent elders are not literate in ʔayʔaǰuθəm, I use questions about pictures and direct translation as methods for eliciting target items. This generates a small, but controlled dataset for prosodic measurements. The results indicate that syntactic constituents map to prosodic constituents, with differences in syntactic structure correlating with differences in prosodic structure. Pitch is an important indicator of both prosodic constituency and focus with prosodic phrases indicated by partial pitch reset at the left edge and downstep within the phrase, while focused constituents are associated with a greater change in pitch.

2 Language background

2.1 Language affiliation and status

ʔayʔaǰuθəm is traditionally spoken in the Tla'amin (łaʔamin), Homalco (ʔop̓), Klahoose (t̓oq̓ʷ), and Comox (k̓omoks) First Nation communities in BC, Canada. These communities are the northernmost communities to speak a language belonging to the Central Salish branch of the Salish language family. ʔayʔaǰuθəm is severely endangered; according to the 2014 Report on the Status of B.C. First Nations Languages by the First Peoples' Heritage, Language and Cultural Council, there are 36 remaining fluent speakers.

* I am very grateful to everyone who shared their language with me, especially Phyllis Dominic and Joanne Francis, who provided the sound recordings for this analysis. č̓ɛč̓ɛhatanapɛ! I would also like to acknowledge the linguists who have worked on the language before me, Susan Blake, John Davis, and Honoré Watanabe, who have been very generous in sharing resources and ideas, and whose work I have heavily relied on. Finally, I would like to thank my committee who are always supportive and ready with good ideas and guidance: Henry Davis, Molly Babel, Michael Rochemont. Emily Elfner has been particularly influential on the direction and design of this project, and constantly patient and generous in answering all my questions about design and analysis. All errors are my own.
Contact info: mebhuijs@alumni.ubc.ca

In *UBC Qualifying Papers 4 (2015–2017)*,
University of British Columbia Working Papers in Linguistics,
Emily Sadlier-Brown, Oksana Tkachman, and Avery Ozburn (eds.). Vancouver, BC: UBCWPL, 2019.

2.2 Prosody

As for many endangered languages, the prosody of the language is under-documented and prosody above the word-level has not been previously examined. Word-level stress in ʔayʔaǰuθəm usually falls on the initial syllable (e.g. Blake 2000; Davis 1970; Watanabe 2003). Secondary stress is assigned according to a (left-aligned) trochaic pattern in words of three or more syllables (though certain types of suffixes, notably the stative suffix *-it*, disrupt this pattern. See Blake 2000; Watanabe 2003 for more detailed accounts). Pitch is an important means of indicating stress, with stressed syllables having audibly higher pitch than unstressed syllables (Blake 2000; Watanabe 2003); Watanabe (2003:26) also mentions duration as a cue to the contrastive secondary stress associated with stative forms.

The phonetic and phonological correlates of prosodic phrasing have not been previously described. This study explores two aspects of prosodic structure above the word: 1) whether there are prosodic constituents above the word level (and below the intonational phrase) marked by downstep and reset, 2) whether these prosodic constituents correspond to syntactic phrases. A secondary question in this investigation concerns whether there are intonational cues to focus, something which has been argued to be absent in other Salish languages (Benner 2006; Caldecott 2016; Koch 2008).

2.3 Syntax

I examine pairs of sentences, like those in (1) with the same string of lexical items, but different syntactic constituency. As is the case throughout the family, sentences are predicate-initial. Lexical items from different classes (verbs, nouns, adjectives) can function as the predicate (but see Davis and Matthewson 1999; Demirdache and Matthewson 1995 for arguments that Salish languages do distinguish lexical classes). A noun in predicate position may be preceded by a modifier, forming a complex nominal predicate (CNP), as in (1a). Another characteristic feature of Salish syntax is the availability of headless relative clauses (argued to be headed by a null pronominal in Davis 2003). The subject of (1a) is a headless relative clause *tə kʷanač ʔə tə θəkʷnačtən*, which roughly translates 'the ones sitting in the chair'. In (1b), the numeral functions as the predicate, followed by the DP modified by a relative clause.[1]

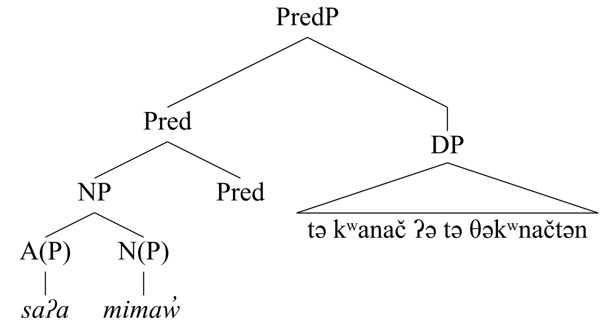
(1) a. [[$_{\text{CNP}}$ saʔa mimaẇ] [$_{\text{DP}}$ **tə** kʷanač [$_{\text{AdvP}}$ ʔə tə θəkʷnačtən]]]]
[[$_{\text{CNP}}$ two cat] [$_{\text{DP}}$ DET sit [$_{\text{AdvP}}$ OBL DET chair]]]
'The ones sitting on the chair are two cats.' (More colloquially: There are [$_{\text{FOC}}$ two cats] sitting on the chair.)

b. [[$_{\text{AP}}$ saʔa] [$_{\text{DP}}$ **tə** mimaẇ [$_{\text{CP}}$ kʷanač [$_{\text{AdvP}}$ ʔə tə θəkʷnačtən]]]]
[[$_{\text{AP}}$ two] [$_{\text{DP}}$ DET cat [$_{\text{CP}}$ sit [$_{\text{AdvP}}$ OBL DET chair]]]]
'The cats sitting on the chair are two.' (More colloquially: There are [$_{\text{FOC}}$ two] cats sitting on the chair.)

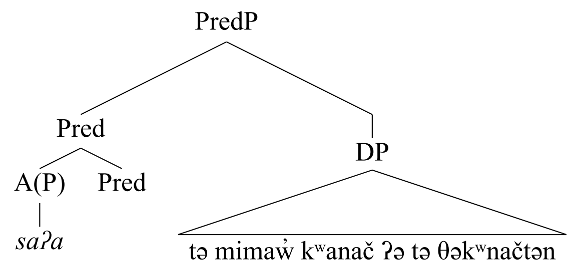
Assuming the minimum syntax necessary, this gives us the structures in (2) for the complex nominal predicate (2a) and numeral predicate (2b) structures, respectively.

[1] The glossing abbreviations used in this paper are as follows: 1 'first person', 3 'third person', CONJ 'conjunctive, CTR 'control transitive', DET 'determiner', ERG 'ergative' EXCL 'exclusive', IPFV 'imperfective', NEG, SBJ 'subject'. NTR 'noncontrol transitive', OBL 'oblique', PPL 'people', PL 'plural', PST 'past',STAT 'stative'.

(2) a.



b.



In careful speech, the structures of (1a) and (1b) are distinguished by the position of the determiner preceding the subject (bolded). When there is a complex nominal predicate, the determiner precedes the headless relative clause, while when there is a numeral predicate, the determiner precedes the subject DP that is modified by the headless relative clause. Determiners are frequently elided, however, making the two strings identical in terms of segmental content.

3 Theoretical background

I assume a mapping between syntax and prosody like that proposed in Match Theory, where prosodic structure reflects syntactic structure, but prosodic constraints may obscure the syntax-prosody relationship (e.g. Selkirk 2009, 2011).[2] This study focuses on the mapping between syntactic phrases (XPs) and prosodic phrases. I adopt the implementation of MATCH-PHRASE found in Elfner 2012:27–28, where a syntactic phrase exhaustively dominating a set of (phonologically overt) terminal nodes α should map to a prosodic phrase that exhaustively dominates the same set α of (phonologically overt) terminal nodes.[3] This results in a recursive prosodic structure where prosodic phrases may be embedded within prosodic phrases, mirroring recursive syntactic structure (e.g. Ito and Mester

[2] Match Theory involves indirect mapping, where the prosodic structure is not directly read from the syntax, but involves a level of prosodic representation that mediates between syntactic structure and phonetic realization. This is not crucial to formulating the hypotheses of this study, which could also be stated in terms of a direct mapping between syntax and prosody (e.g. Pak 2008; Wagner 2005, 2010, 2015).

[3] In Elfner (2012:28), this constraint is formulated in Optimality Theoretic terms, as a violable constraint:

(i) MATCH-PHRASE$_T$: Suppose there is a syntactic phrase (XP) in the syntactic representation that exhaustively dominates a set of one or more terminal nodes α. Assign one violation mark if there is no phonological phrase (ϕ) in the phonological representation that exhaustively dominates all and only the phonological exponents of the terminal nodes in α.

2013; Wagner 2005).

Where prosodic constituents are domains for downstep and reset, this means that downstep domains can be embedded within larger downstep domains (e.g. Berg, Gussenhoven, and Rietveld 1992; Féry and Ishihara 2010; Truckenbrodt 2002). This is schematized in Figure 1 from Féry and Ishihara (2010:5), where the vertical dimension represents pitch height. The horizontal lines represent the reference top line for each level of prosodic embedding. Here, three prosodic phrases (indicated by {} brackets) are grouped into a larger prosodic constituent. This larger prosodic constituent forms a downstep domain such that each prosodic phrase within it exhibits downstep relative to a preceding prosodic phrase also contained within it (the successively lowered reference line is shown in grey). Downstep occurs across the words within each prosodic phrase as well (the successively lowered reference line is shown in black). Partial reset occurs across the prosodic phrase boundaries, with the pitch at the left edge of a prosodic phrase reaching the reference top line established for that constituent (full reset would bring the pitch back to the top reference line for the sentence).

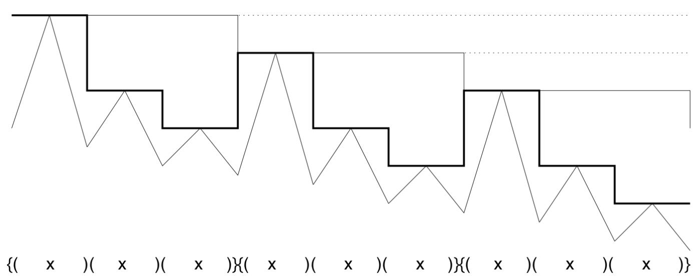


Figure 1: Downstep and partial reset across and within prosodic phrases with reference lines for each prosodic level, where the vertical dimension represents pitch height and the horizontal dimension has words (represented by () brackets) grouped into phrases (represented by {} brackets) across an utterance.

4 Predictions

Adopting this model for downstep and reset, we can predict contours for each of the sentences in (1). If syntactic phrases map directly to prosodic phrases, we predict contrasting prosodic structures for these two cases. Where *saʔa* 'two' and *mimaẇ* 'cat' form a complex nominal predicate, they belong to the same syntactic constituent and are grouped together prosodically. The headless relative clause that is the subject DP will also map to a prosodic phrase, giving a recursive prosodic structure (3a). In (3b), the numeral *saʔa* forms a predicate on its own, but does not map to a prosodic phrase since it is not phrasal; instead, it occurs at the left edge of the prosodic phrase that maps to the PredP in the syntax.[4] The subject DP *tə mimaẇ* maps to a prosodic phrase, and the embedded relative clause *kʷanač θəkʷnačtən* maps to a further embedded prosodic phrase.

[4] In these representations, I assume that syntactic phrases that are both minimal and maximal do not map to independent prosodic phrases.

(3) a. $(_{\phi}$ $(_{\phi}$ saʔa mimaw̓) $(_{\phi}$ tə kʷanač $(_{\phi}$ ʔə tə θəkʷnačtən)))
$[_{PredP}$ $[_{vP}$ two cat] $[_{DP}$ DET sit $[_{AdvP}$ OBL DET chair]]]
'There are two cats sitting on the chair.'

b. $(_{\phi}$ saʔa $(_{\phi}$ tə mimaw̓ $(_{\phi}$ kʷanač $(_{\phi}$ ʔə tə θəkʷnačtən))))
$[_{PredP}$ two $[_{DP}$ DET cat $[_{CP}$ sit $[_{AdvP}$ OBL DET chair]]]]
'There are two that are cats sitting on the chair.'

If prosodic phrasing is marked by downstep throughout a prosodic constituent and reset at the left edge of prosodic constituents, we predict different pitch contours for these two structures.[5] With the complex nominal predicate, there will be downstep between *saʔa* and *mimaw̓*, followed by partial reset for the embedded headless relative clause and downstep within the relative clause. The predicted prosodic contour is shown in Figure 2, where the downstep across prosodic phrases is represented in red, while downstep within the first prosodic phrase is represented in green.

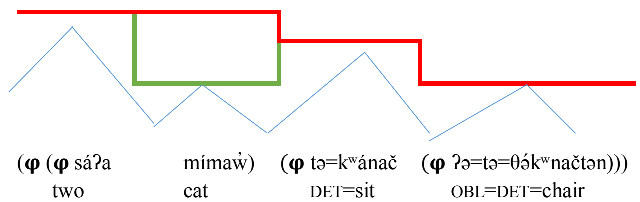


Figure 2: Predicted downstep and reset pattern for a sentence with a complex nominal predicate, where reference levels for successive downstep between prosodic phrases are represented in red and the reference level for downstep embedded within the initial prosodic phrase is represented in green. The vertical dimension represents pitch height and the horizontal dimension has words grouped into prosodic phrases (represented with () brackets) across an utterance.

With the numeral predicate, we predict partial reset on *mimaw̓*, since it is at the left edge of a constituent, but also downstep compared to the preceding word which is at the left edge of a less embedded prosodic phrase. The net result is shallower downstep, as shown in Figure 3 (see Féry and Ishihara 2010:22 for discussion of similar effects in Japanese).

The two structures focus different constituents and so are prompted in different contexts. In this paper, I adopt a notion of focus originally proposed in Rooth (1985), where focus evokes a set of alternatives to the focused constituent; the constituent replacing the wh-word in a wh-question is focused, for instance, because it picks out a single referent from set of possible referents that could be given in alternative answers. The initial position is typically associated with focus in Salish. Hence, a sentence like (1a) with a complex nominal predicate is felicitous in response to a question such as (4a) which prompts focus on the entire nominal. Similarly, the numeral predicate case focuses the numeral, so that (1b) occurs felicitously in response to a question inquiring 'how many' (4b). If a focused constituent is associated with prosodic prominence such as higher pitch or increased duration, we expect to find these prosodic cues on the nominal, the head of the focused constituent,

[5] Note that because the predictions regarding prosodic structures focus on left edges (i.e. partial reset at left edges), the relevant prosodic structure could also have been captured with alignment constraints (e.g. Selkirk 1986; Selkirk and Tateishi 1991; Truckenbrodt 1999). I adopt Match Theory because it allows a simpler implementation of the mapping between syntactic phrases and prosodic phrases for my purposes.

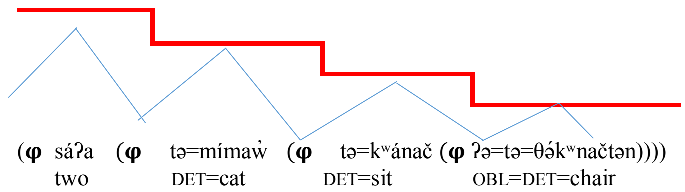


Figure 3: Predicted downstep and reset pattern for a sentence with a numeral predicate, where reference levels for successive downstep between prosodic phrases are represented in red. The vertical dimension represents pitch height and the horizontal dimension has words grouped into prosodic phrases (represented with () brackets) across an utterance

in sentences like (1a) and on the number in sentences like (1b).

(4) a. tam tə kʷanač ʔə tə θəkʷnačtən?
what DET sit OBL DET chair
'What is sitting in the chair?'

b. k̓ʷən tə mimaw̓ kʷanač ʔə tə θəkʷnačtən?
how.many DET cat sit OBL DET chair
'How many cats are sitting in the chair?'

5 Methodology

5.1 Participants

I recorded two elders fluent in ʔayʔaǰuθəm who live in Vancouver. Both speakers are female and over 60 years in age. While both were born in Tla'amin (t̓ıšosəm), one grew up in Church House (ʔop̓) and the other in Klahoose (t̓oq̓ʷ), so that there is a possibility of small dialectal differences between the speakers. At this point, it is not known if there are prosodic differences between dialects, but there are no clearly audible differences that are perceptible to a non-native speaker.

5.2 Materials

In the previous section, I described how the two structures are prompted by different questions, invoking focus on different constituents. In order to elicit the two structures, I constructed a set of fifteen pictures. The first set contained five pictures of animals or people engaged in the same activity, as in (4). For these pictures, I would ask the consultant *tatam ga tiʔ?* 'what's going on there?'. This condition was included because this question is often answered with a complex nominal predicate structure when the noun is quantified in the answer, allowing comparison between broad focus intonation and narrow focus intonation, for the same complex nominal predicate structure.

The second set contained five pictures with a group of animals or people engaged in the same activity and a contrasting character engaged in a different activity. For these, I would ask the consultant which characters were engaged in one of the activities, e.g. *tam tə kʷanač ʔə tə θəkʷnačtən?* 'what is sitting on the chair?'.

Figure 4: Picture for ‘there are two cats sitting on a chair’

Figure 5: Picture for ‘[$_F$ two cats] are sitting on the chair’

Finally, the third set contained five pictures with a group of animals or people where most were engaged in one activity, but one was doing something else. For these, I would ask how many of the characters were engaged in the activity, e.g. *k̓ʷən tə mimaw̓ tə kʷanač ʔə tə θəkʷnačtən?* ‘how many cats are sitting on the chair?’.

The five sentences elicited are given in (5) (with the complex nominal predicate structure).

(5) a.

mus	mušmuš	tə	č̓iq̓-it
four	cow	DET	fence-STAT

‘Four cows are fenced (in the fence).’

b.

saʔa	mimaw̓	tə	kʷanač	ʔə	tə	θəkʷnačtən
two	cat	DET	sit-STAT	OBL	DET	chair

‘Two cats are sitting on the chair.’

Figure 6: Picture for ‘[F two] cats are sitting on the chair’

c. sisaya tumiš tə ʔəm∼ʔimaš ʔə tə θičim
two.PPL men DET PL∼walk OBL DET woods
‘Two men are walking in the woods.’

d. čalas mimaw̓ tə ʔił<iw>tan ǰanxʷ
three cat DET eat<PL> fish
‘Three cats are eating fish.’

e. sisaya tumiš tə ǰa∼ǰaq̓-at tə ǰaʔǰaʔ
two.PPL men DET PL∼fall-CTR DET tree
‘Two men are falling the tree.’

For the second task, I asked my consultants to translate sentences with contrastive focus on the DP (6a) or contrastive focus on the number (6b), where the difference in the contrast phrase makes the information structure explicit.

(6) a. saʔa mimaw̓ kʷ k̓ʷənəxʷan, xʷukʷt č̓an̓u
two cat DET see-NTR-1S.ERG.SBJ not.exist dog
‘I saw two cats, but no dogs.’

b. saʔa=ʔut kʷ mimaw̓ kʷ k̓ʷənəxʷan, xʷa čalas=as
two=EXCL DET cat DET see-NTR-1S.ERG.SBJ, NEG three=3CNJ
‘I saw just two cats, not three.’

The set of five target sentences is given in (7) in the numeral contrast condition.

(7) a. saʔa=ʔut kʷ mimaw̓ kʷ k̓ʷən-nəxʷ-an, xʷa čalas=as
two=EXCL DET cat DET see-NTR-1S.ERG.SBJ, NEG three=3CNJ
‘I saw just two cats, not three.’

b. saʔa=ʔut kʷ qigaθ kʷ k̓ʷən-nəxʷ-an, xʷa čalas=as
two=EXCL DET deer DET see-NTR–1S.ERG.SBJ, NEG three=3CNJ
‘I saw just two deer, not three.’

c. paʔa=ʔut ǰanxʷ kʷ məkʷ-t-at-uɬ, xʷa saʔa=as
two=EXCL fish DET eat-CTR-1PL.ERG.SBJ-PST NEG two=3CNJ
'We only ate one fish, not two'

d. paʔa=ʔut kʷ ʔapəls kʷ məkʷ-t-an-uɬ, xʷa saʔa=as
one=EXCL DET apple DET eat-CTR-1S.ERG.SBJ-PST NEG two=3CNJ
'I only ate one apple, not two.'

e. sisaya=ʔut kʷ tumiš kʷ ʔəm∼ʔimaš, xʷa čiɬaya=as
two.PPL=EXCL DET men DET PL∼walk NEG three.PPL=3CNJ
'Two men are walking, not three.'

Of course, it is necessary to be able to differentiate downstep and reset associated with prosodic constituency from any pitch movement associated with focus in order to answer the research questions of interest here (if there is prosody prominence associated with focus in ʔayʔaǰuθəm; Koch (2008) argues that prosody does not indicate focus in Nɬkepmxcín, and Benner (2006) and Caldecott (2016) also do not find prosodic effects of focus in SENĆOŦEN or St'at'imcets). However, the expected prosodic prominence (and corresponding pitch peaks) associated with focus is at least partially independent from the pitch reset predicted at constituent boundaries. For instance, in (8b), *saʔa mimaw̓* is focused, while *kʷanač* is not, so if there is higher pitch on *kʷanač*, it must be due to partial reset at the left edge of the headless relative clause, not focus prominence. Similarly, in (9b), the numeral is focused, while *mimaw̓* is not focused, but based on prosodic constituency, we expect *mimaw̓* to show partial reset. So, while prosodic effects associated with information structure may obscure the prosodic effects of constituency, the predictions for the pitch contour marking constituency are distinct from the predictions regarding focus prominence.

(8) a. tam tə kʷanač ʔə tə θəkʷnačtən?
what DET sit OBL DET chair
'What is sitting in the chair?'

b. ($_{\phi}$ ($_{\phi}$ saʔa mimaw̓) ($_{\phi}$ tə kʷanač ($_{\phi}$ ʔə tə θəkʷnačtən)))
[$_{\text{PredP}}$ [$_{\text{vP}}$ two cat] [$_{\text{DP}}$ DET sit [$_{\text{AdvP}}$ OBL DET chair]]]
'There are two cats sitting on the chair.'

(9) a. k̓ʷən tə mimaw̓ tə kʷanač ʔə tə θəkʷnačtən?
how.many DET cat DET sit OBL DET chair
'How many cats are sitting in the chair?'

b. ($_{\phi}$ saʔa ($_{\phi}$ tə mimaw̓ ($_{\phi}$ kʷanač ($_{\phi}$ ʔə tə θəkʷnačtən))))
[$_{\text{PredP}}$ two [$_{\text{DP}}$ DET cat [$_{\text{CP}}$ sit [$_{\text{AdvP}}$ OBL DET chair]]]]
'There are two that are cats sitting on the chair.'

5.3 Procedure

Recordings took place in the homes of the consultants. Consultants were recorded using lavaliere microphones and a Marantz recorder. For the first task, I recorded the 15 sentences several times with each consultant over several sessions in order to create a dataset with minimal pairs between the conditions for as many of the target items as possible. At first, the consultants were reluctant to use the numeral except in the answers to the *k̓ʷən* 'how many' questions, so I asked them whether I could use it in the answer to the *tam* and *tatam* 'what' and 'what's happening' questions. Both replied that it was possible, and started using the numerals (though they were still omitted on occasion). Where the answer to the *tatam* or broad focus question included a number, the sentence usually involved a complex nominal predicate construction. This seems to be the preferred construction for expressing quantification over entities, though it is possible to have a numeral modifying a DP in argument position (e.g. *tə saʔa mimaw̓* 'the two cats'). In order to have only one type of question salient at a time, I asked blocks of questions with the same information structure (as opposed to switching up types of questions, as I did in piloting the methodology, since this made the different questions all salient, so that the answers were more likely to vary in structure without being congruent with the question).

6 Analysis

The recordings were transcribed using ELAN and the annotations were exported as Praat textgrids. The tiers for words, syllables and vowels were hand-aligned in Praat. A pitch file was created in Praat for each target utterance (using the autocorrelation method described in Boersma 1993) with the frequency range set to 75 Hz–350 Hz (since both participants are over 60 years, so were unlikely to have F0 measurements above 350 Hz). Measurements below 90 Hz were double-checked and hand-corrected in cases of pitch tracking errors (octave errors were quite common for the elder of the two speakers, whose voice is quite creaky). A Praat script was used to extract pitch measurements at multiple points within each of the vowels. Twenty percent at the beginning and end of the vowels was excluded from measurement to avoid coarticulatory effects with the neighbouring consonants. In the results reported below, each pitch point for a vowel is an average over measurements taken at 40%, 50% and 60% of the duration of the vowel. The script also extracted measurements of duration for vowels, words and sentences.

No statistical analysis was undertaken due to the small size of the dataset, but the results were graphed using the ggplot package in R and inspected for shape and degree of variability. Results are reported for each speaker separately. Because comparing across conditions required exact minimal pairs, any tokens that did not have a minimal pair in the other conditions were excluded from the analysis.[6] For task 1, this meant excluding all tokens of one target item from PD and all tokens of three target items from JF, since substitution of lexical items or different morphological forms rendered the tokens not directly comparable. For the second task, none of the target items yielded minimal pairs for PD, but only one target item was excluded for JF. An additional five tokens were excluded for JF and one for PD because the morphological forms of the lexical items in these tokens did not match the other tokens of the target item. Finally, eight tokens for JF and four tokens for PD were excluded because of disfluencies such as pauses or stopping and repeating a word or syllable; for the younger speaker JF, utterances with pauses longer than 200 ms were excluded, but for the older speaker PD whose speech style is quite slow and involves frequent pauses above 200 ms (which are not perceived as disfluent) utterances were not excluded based on pause length, except where

[6] There are a few pairs that are not exactly minimal, where these still involve the same lexical items and number of syllables, differing only in the forms of the lexical items, such as for *ʔəmʔimaš* 'to walk (plural)' and *ʔiʔimaš* 'to walk (ipfv)'.

multiple such pauses occurred in a single utterance.

6.1 Results for task 1

6.1.1 Pitch (F0) measures for PD

The pitch tracks for each of the minimal pairs for task 1 (which involved the participants answering questions based on a series of pictures) for PD are represented in Figure 7, and the glosses provided in (10).[7]

(10) a. saʔa mimaẇ kʷanač θukʷnačtən
two cat sit-STAT chair

b. sisaya tumiš ʔəm∼ʔimaš/ʔi∼ʔimaš θičim
two.PPL men PL∼walk/IPFV∼walk into.woods

c. čalas mimaẇ ʔił<iw>tan/ʔi∼ʔiłtan ǰanxʷ
three cat eat<PL>/IPFV∼eat fish

d. sisaya tumiš ǰaq̓at ǰaʔǰaʔ
two.PPL men fall-CTR tree

The points in the graph (counting from left to right across the horizontal axis) correspond to consecutive vowels in the utterance, where each point represents measurements from the vowel at that syllable number within the utterance. Points connected by a line represent measurements taken from the vowels contained in the same token utterance and each vowel is labelled with the word that it appears in. Colours represent the three conditions. The numeral predicate condition (answers to the *k̓ʷən* 'how many' questions) is represented in red (NumPred), while the nominal focus condition (answers to *tam* 'what' questions) is represented in green (CNP-NF) and the broad focus condition (answers to *tatam* 'what's happening' questions) is represented in blue (CNP-BF).

In this data set, there were two main trends that differentiated the numeral predicate condition (NumPred) from the complex nominal predicate conditions (both with narrow focus (CNP-NF) and with broad focus (CNP-BF). In the numeral predicate cases, the distance in Hz between the mean pitch of the final syllable of the numeral and the mean pitch of the initial syllable of the nominal tended to be greater than for either of the complex nominal predicate conditions. The measurements for the individual tokens are laid out in Table 1. The numbers in the table show that for four of the five instances, the distance in Hz between the final syllable of the numeral to the initial syllable of the nominal is greater in the numeral predicate condition than for either of the other two conditions. The one counterexample is the second token of third item in the numeral predicate condition, which fails to show any reset for the noun.

Where the number is narrowly focused in the numeral predicate utterances, there is also a tendency for a greater change in pitch (F0) within the number, measured as the difference between the pitch of the initial and final syllable. Table 2 shows that the change of pitch from the initial syllable to the final syllable of the number in the numeral predicate condition (NumPred) is either equal to or

[7] Measurements were not taken for determiners, which are clitics, so they do not appear in the graphs and are not included in the glosses.

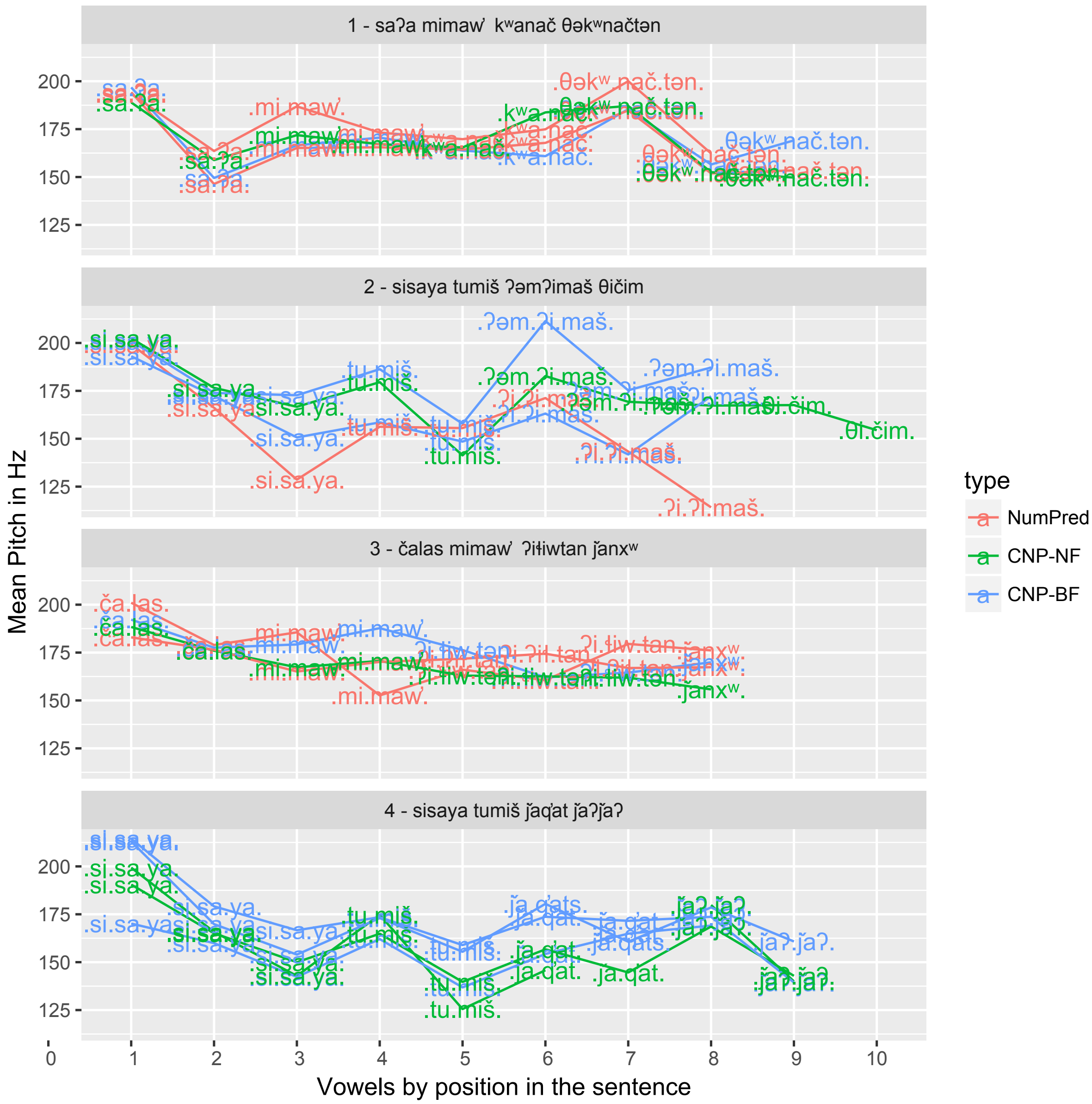


Figure 7: Pitch contours for PD for task1, plotted by the mean pitch (Hz) of vowels across each utterance and compared across conditions

greater than the change in pitch within the number for both of the complex nominal predicate conditions (CNP), with one counter example (the same token that was a counterexample to the above generalization).

Comparing the broad focus (CNP-BF) and narrow focus (CNP-NF) complex nominal predicate conditions, we observe a larger change in pitch (F0) on the nominal in the narrow focus condition. Again, the change in pitch is measured as the difference in pitch between the initial syllable and final syllable of the nominal. The measurements are given in Hz in Table 3.

Table 1: Pitch (F0) change between the final syllable of the numeral (LastSyllNum) and initial syllable of the nominal (FirstSyllN) measured in Hz for PD. Sentence numbers correspond to the numbers given to the sentences in Figure 7; multiple lines of measurements occur corresponding to a single sentence number where there are multiple tokens of the sentence in a particular condition (where multiple tokens for a sentence in a particular condition appear in Figure 7).

	CNP-BF			**CNP-NF**			**NumPred**		
Sentence	LastSyllNum	FirstSyllN	Distance	LastSyllNum	FirstSyllN	Distance	LastSyllNum	FirstSyllN	Distance
1	149.5	166.56	17.06	158.84	171.87	13.03	146.35	165.18	18.83
							163.58	186.67	23.09
2	150.68	158.55	7.87	166.69	179.54	12.85	128.66	156.27	27.61
	172.71	186.31	13.6						
3	177.56	179.24	1.68	175.98	167.31	-8.67	178.94	185.54	6.63
							176.43	165.09	-11.34
Average			10.05			5.74			12.96
SD			6.75			12.48			15.67

Table 2: Change in pitch (F0) within the numeral, measured as the difference in Hz between the pitch of the initial syllable (Syll 1) and the final syllable (Syll 2 or Syll 3), for PD. Sentence numbers correspond to the numbers given to the sentences in Figure 7; multiple lines of measurements occur corresponding to a single sentence number where there are multiple tokens of the sentence in a particular condition (where multiple tokens for a sentence in a particular condition appear in Figure 7)

	CNP-BF				**CNP-NF**				**NumPred**			
Sentence	Syll 1	Syll 2	Syll 3	Distance	Syll 1	Syll 2	Syll 3	Distance	Syll 1	Syll 2	Syll 3	Distance
1	196.79	149.5		-47.29	188.93	158.84		-30.09	194.24	163.58		-30.66
									193.4	146.35		-47.05
2	201.04	173.92	172.71	-28.33	202.41	176.56	166.69	-35.72	199.34	166.16	128.66	-70.68
	193.05	172.34	150.68	-42.37								
3	192.19	177.56		-14.63	188.36	175.98		-12.38	201.03	178.94		-22.09
									182.94	176.43		-6.51
Average				-33.16				-26.06				-35.4
SD				14.73				60.96				23.73

Table 3: Change in pitch (F0) within the nominal, measured as the difference between the initial syllable (Syll 1) and final syllable (Syll 2) in Hz, for PD. Sentence numbers correspond to the numbers given to the sentences in Figure 7; multiple lines of measurements occur corresponding to a single sentence number where there are multiple tokens of the sentence in a particular condition (where multiple tokens for a sentence in a particular condition appear in Figure 7).

	CNP-BF			**CNP-NF**		
Item	Syll 1	Syll 2	Distance	Syll 1	Syll 2	Distance
PD						
1	166.56	170.57	4.01	171.87	167.12	-4.75
2	186.31	157.83	-28.48	179.54	141.32	-38.22
3	129.24	187.73	58.49	167.31	170.63	3.32
4	173.48	158.92	-14.56	174.71	125.53	-49.18
	162.07	136.82	-25.25	164.85	139.72	-25.13
Average			-1.16			-22.79
SD			35.68			22.06

6.1.2 Duration measures for PD

Since duration can play a role in both prosodic phrasing and focus prosody and this was a straightforward additional measure to take, the duration of each vowel across the sentences is also examined. The duration measures for each vowel across each sentence are plotted in Figure 8. The most notable finding is an increase in duration for the embedded predicate (the predicate in the relative clause). Note that the long duration of the vowels in *ǰaʔǰaʔ* 'tree' in the fourth panel must be treated with caution, as the glottal stop codas are generally not easily segmentable, but rather occur as creakiness on the vowel, giving the effect of a long creaky vowel.

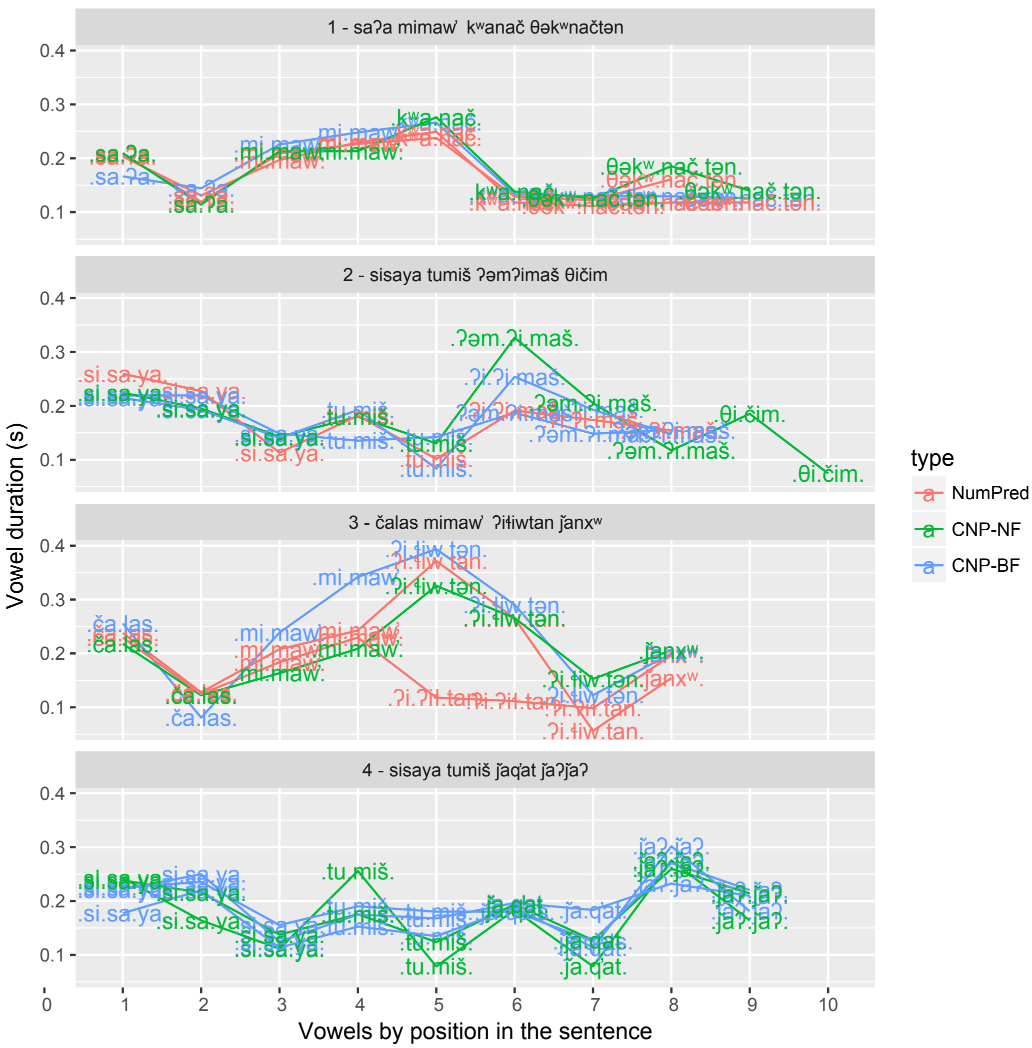


Figure 8: Duration of each vowel measured in seconds and plotted across each utterance for PD for task 1

6.1.3 Pitch (F0) measures for JF

For the other speaker, there was only one item (3) where there were exact minimal pairs across conditions, shown in Figure 9, and an additional item (subset according to whether the number was realized with two (2syll) or three (3syll) syllables) with minimal pairs between the broad focus (CNP-BF) and narrow focus (CNP-NF) complex nominal predicate conditions (1) and between the broad focus complex nominal predicate condition (CNP-BF) and numeral predicate condition (NumPred) (2). Glosses for these items are given in (11). What is notable here is the steep pitch movement on the number in the numeral predicate conditions. There is also steeper pitch movement on the nominal for the narrow focus complex nominal predicate condition (CNP-NF) condition compared to the broad focus nominal predicate condition (CNP-BF) (the fact that secondary stress in marked in the narrow focus condition on the three syllable nominal *təmtumiš*, but not in the broad focus condition, may also be an indication of prominence on the nominal in the narrow focus condition). These examples do not show a steeper reset between numeral and nominal for the numeral predicate condition compared to the other two conditions, unlike the results from the first speaker.[8]

(11) a. sisaʔa mimaw̓ kʷanač θəkʷnačtən
two.PPL cat sit-STAT chair

b. sisaya təm∼tumiš ʔəm∼ʔimaš/ʔi∼ʔimaš θičim
two.PPL PL∼man PL∼walk/IPFV∼walk into.woods

6.1.4 Duration measures for JF

With respect to duration, there was a tendency across utterances for increased duration on the embedded predicate just as for PD, as shown in Figure 10. Note that the longest duration is on the second syllable of *ʔəmʔimaš* ‘walk’ because the initial vowel is *ə* which tends to have shorter duration than other vowels (see e.g. Blake 2000).

6.2 Results for task 2

6.2.1 Pitch (F0) measures for JF for task 2

The second task (where the participants translated sentences with contrastive focus on the complex nominal predicate or on the numeral predicate) only resulted in minimal pairs for one of the speakers, JF. The contours for the sentences from this task are given in Figure 11,[9] and the glosses in (12). One complication with this dataset is that the particle *ʔut* often appears following the number in the numeral focus condition (giving the ‘only’ interpretation, see Huijsmans 2017). This introduces an

[8] In the middle panel of Figure 9, it appears that primary stress is falling on the second syllable of *təmtumiš* ‘men’ in the *k̓ʷən* condition. This pattern has been observed before for this speaker with words that have CəC∼ plural reduplication, but is surprising given the usual stress patterns of the language and is also not consistent across tokens.

[9] The range of the graphs for this speaker are not the same as for the PD. This is because the two speakers have different vocal ranges, with PD have a lower range and creakier vocal quality and JF having a higher pitch range. Since the data from each speaker is not directly compared, I allowed the graphs to reflect this difference between the speakers.

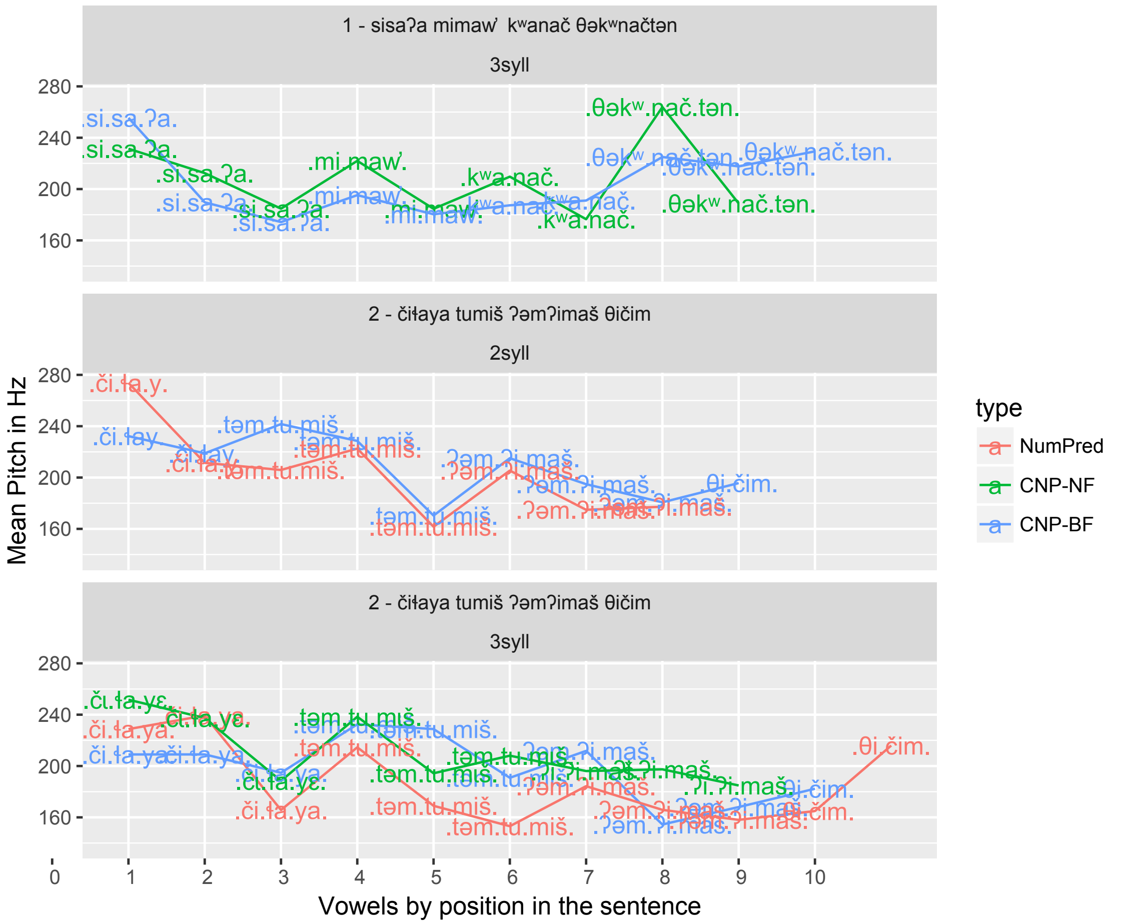


Figure 9: Pitch contours for JF for task1, plotted by the mean pitch (Hz) of vowels across each utterance and compared across conditions

extra syllable into the numeral focus tokens, and may extend the number of syllables of the numeral, depending on the phonological domain in which this clitic is parsed.[10] For the purposes of this paper, I do not include the clitic in measurements of the pitch of the initial prosodic word (the number), since this would make the data difficult to compare across the conditions.

(12) a. čalas/saʔa=ʔut mimaw̓ k̓ʷən-əxʷ-an-uɬ
three/two=EXCL cat see-NTR–1S.ERG.SBJ-PST

b. saʔa/saʔa=ʔut qigaθ k̓ʷən-əxʷ-an-uɬ
two/two=EXCL deer see-NTR-1S.ERG.SBJ-PST

[10] When the predicate is a monomoraic root, a second-position clitic will be footed with the predicate to create a bimoraic foot, as pointed out in Blake (2000). At this point, we do not know for sure what domain clitics belong to when their host is not prosodically 'deficient' (i.e. their host forms a bimoriac foot on its own), nor whether all the second-position clitics behave the same in this regard. See Kroeber (2002) for further discussion of the properties and distribution of clitics in ʔayʔaǰuθəm.

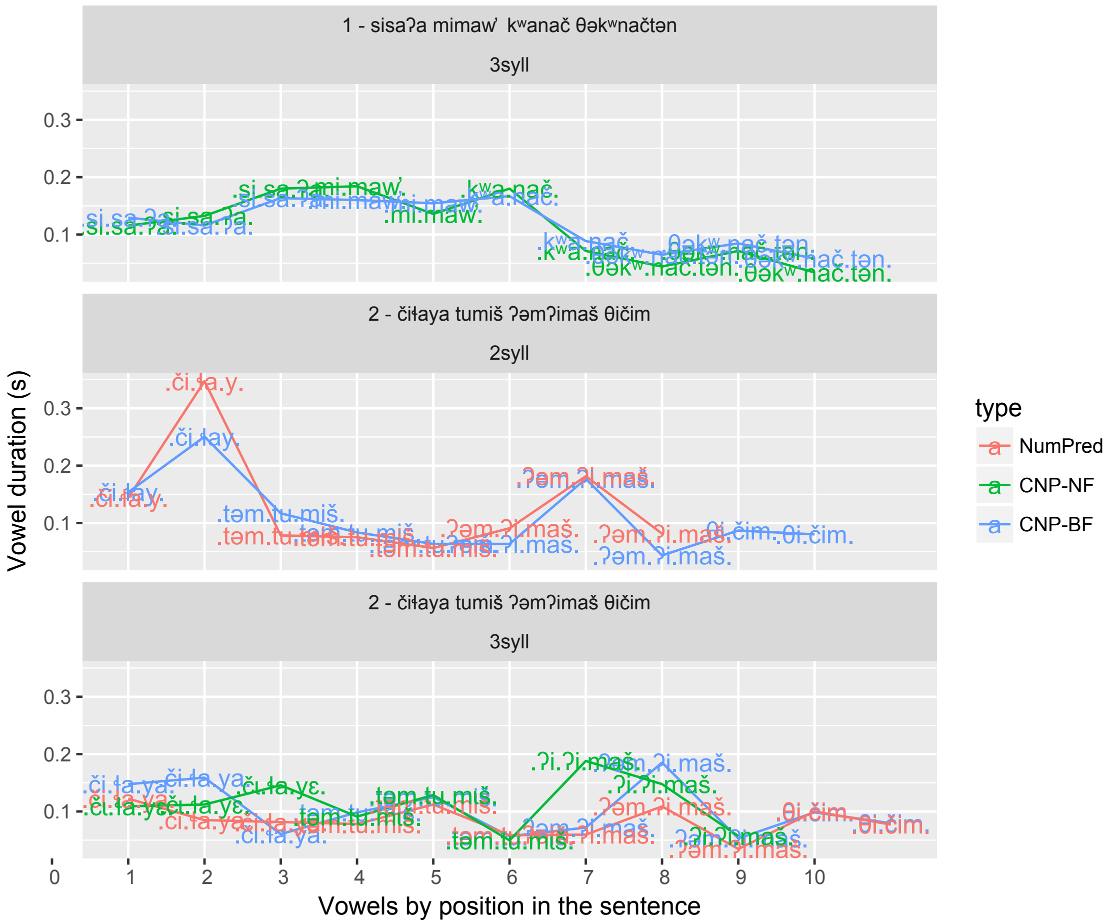


Figure 10: Duration of each vowel measured in seconds and plotted for each utterance for JF for task 1

c. saʔa/paya=ʔut ǰanxʷ məkʷ-t-at-uɬ
two/one=EXCL fish eat-CTR-1PL.ERG.SBJ-PST

d. saʔa/paya=ʔut ʔapəls məkʷ-t-an-uɬ
two/one=EXCL apples eat-CTR-1S.ERG.SBJ-PST

As shown in Table 4, the change in pitch (F0) (distance in Hz between the pitch of the initial syllable and pitch of the final syllable) within the numeral tends to be greater in the numeral predicate condition (NumPred) than the complex nominal predicate condition (CNP).

There is also an effect of greater reset between the numeral and the nominal in the numeral predicate condition, measured as the difference in Hz between the pitch (F0) of the final syllable of the numeral and the initial syllable of the nominal, as shown in Table 5; however, in both conditions, there was not always pitch reset on the noun (giving rise to the negative values in the table).

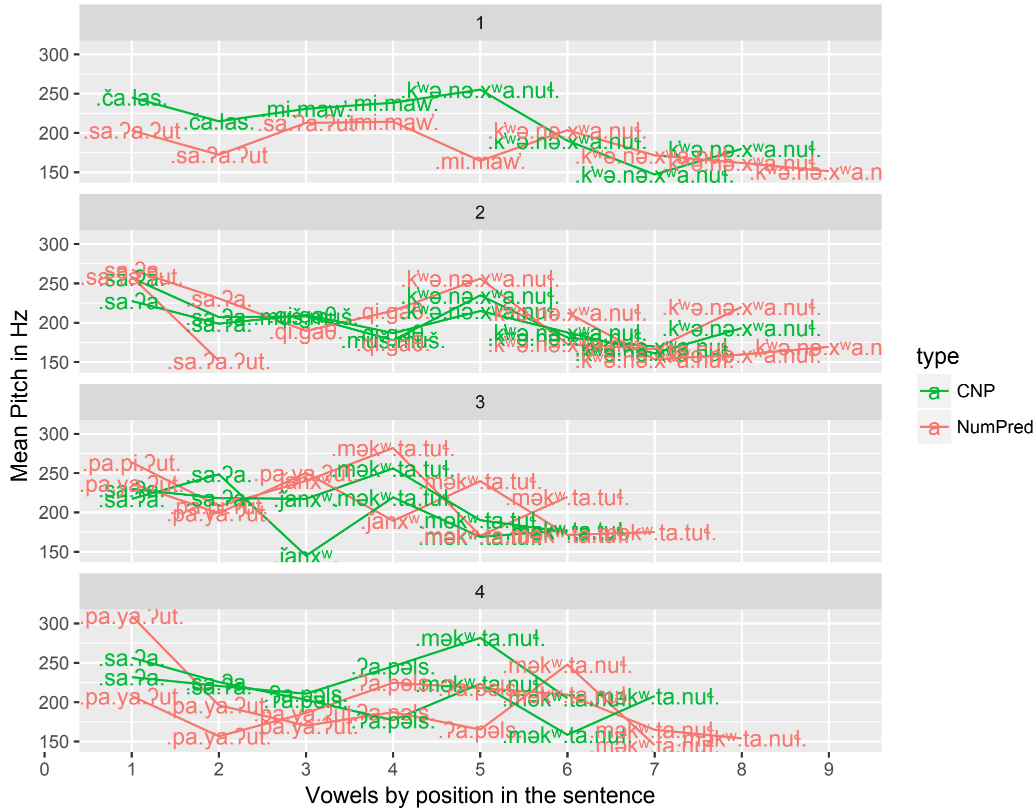


Figure 11: Pitch contours for JF for task 2, plotted by the mean pitch (Hz) of vowels across each utterance and compared across conditions

Table 4: Change in pitch within the numeral, measured as the difference in Hz between the pitch (F0) of the initial syllable (Syll 1) and final syllable (Syll 2), for JF for task 2. Sentence numbers correspond to the numbers given to the sentences in Figure 11; multiple lines of measurements occur corresponding to the same sentence number where there are multiple tokens of the sentence in a particular condition (where multiple tokens for a sentence in a particular condition appear in Figure 11).

	CNP			**NumP**		
Sentence	Syll 1	Syll 2	Distance	Syll 1	Syll 2	Distance
1	245.21	214.85	-30.36	203.27	172.27	-31
2	256.9	207.12	-49.78	276.22	230.78	-45.44
2	227.66	198.8	-28.86	258.66	151.6	-107.06
3	229.9	218.14	-11.76	263.63	207.95	-55.68
3	217.24	248.49	31.25	236.76	199.05	-37.71
4	256.55	225.59	-30.96	309.69	196.1	-113.59
4	231.89	220.63	-11.26	207.91	156.63	-21.68
Average			-18.82			-63.12
SD			25.67			33.33

Table 5: Change in pitch (F0) between the numeral and the nominal, measured as the difference in Hz between the pitch of the final syllable of the numeral (LastSyllNum) and the initial syllable of the nominal (FirstSyllNom), for JF for task 2. Sentence numbers correspond to the numbers given to the sentences in Figure 11; multiple lines of measurements occur corresponding to a single sentence number where there are multiple tokens of the sentence in a particular condition (where multiple tokens for a sentence in a particular condition appear in Figure 11).

	Complex nominal predicate			**Numeral predicate**		
Sentence	LastSyllNum	FirstSyllNom	Distance	LastSyllNum	FirstSyllNom	Distance
1	214.85	230.7	15.85	172.27	214.3	42.03
2	207.12	208.11	0.99	230.78	189.64	-41.14
2	198.8	210	11.2	151.6	174.32	22.72
3	218.14	217.37	-0.77	207.95	240.22	32.27
3	248.49	144.99	-103.5	199.05	189.45	-9.6
4	225.59	203	-22.59	196.1	187.25	-8.85
4	220.63	210.78	-9.85	156.63	224.67	68.04
Average			-4.2425			10.915
SD			40.85			37.03

6.2.2 Duration measures for JF for task 2

As for task 1, the most significant finding with respect to duration was that there tends to be an increase in duration within the embedded predicate, as shown in Figure 12. The increased duration does not always correspond to the initial syllable of the embedded predicate, however, which is *ə* for these predicates (see Blake (2000) for a discussion of the properties of *ə*, including its lack of weight).

7 Discussion

Overall, the findings support the hypothesis that there are prosodic phrases that map to syntactic constituents in ʔayʔaǰuθəm. In section 4, we predicted that if there are prosodic phrases that map to syntactic phrases, there would be a difference in the prosodic contour associated with the numeral predicate structure and complex nominal predicate structure. Since there is a syntactic constituent boundary between the numeral and the nominal in the numeral predicate structure, there should be a prosodic constituent boundary triggering partial reset on the nominal, while with a complex nominal predicate, there is no constituent boundary between the nominal and the numeral, so there will be downstep between the numeral and nominal with no partial reset.

While we did not find overall higher pitch on the nominal in the numeral predicate condition, we did find a tendency for a steeper reset to the nominal for both speakers. We also found some effect of embedded CP boundary across the conditions, but the marking of this boundary was more complex than expected: while in some cases it had a pitch peak predicted for partial reset, in other cases it seemed to be marked primarily by increased duration. Further research should investigate how pitch and duration interact in marking constituent boundaries, and whether longer duration is always associated with CP boundaries in particular. If this turns out to be the case, it would motivate mapping CP to a different type of prosodic constituent, an IP, with distinct phonetic correlates (thanks to Natalie Weber, p.c., 2017 for pointing this out), a move entirely consistent with the proposed

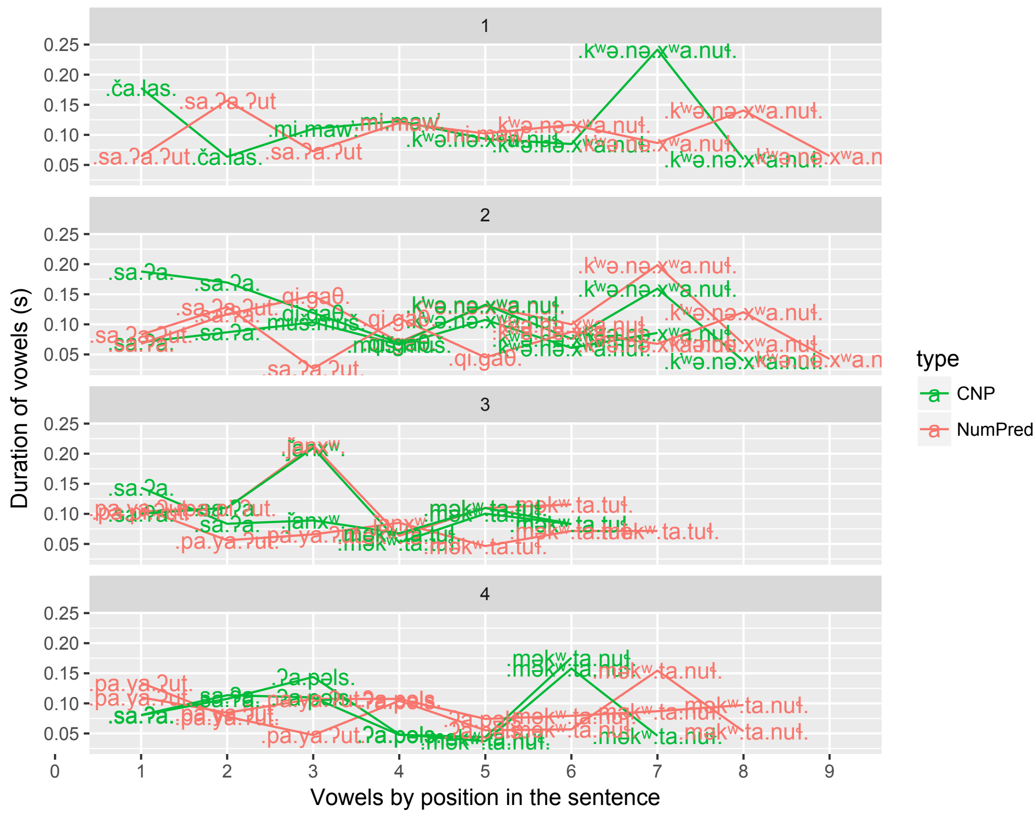


Figure 12: Duration of each vowel in seconds plotted across each utterance for JF for task 2

matching constraints in Match Theory (Selkirk 2009, 2011).

The pitch contours also showed greater change in pitch on focused constituents, suggesting that focus as well as constituency plays a role in determining the pitch contour of a sentence. For both speakers and both tasks, there was greater change in pitch within the numeral in the numeral predicate condition (where the numeral was focused either in the answer to a *k̓ʷən* 'how many' question or contrastively focused), than in the complex nominal predicate condition(s). In the first task, for both speakers, there was also greater change in pitch within the nominal where the nominal was the head of a focused complex nominal predicate (where the nominal predicate was focused in answer to a *tam* 'what' question), than when the complex nominal predicate occurred in the broad focus condition (there was no broad focus condition in the second task, so this comparison was not possible).

8 Conclusion

The results from this study are necessarily tentative given the small data set and variability within it. However, it provides useful groundwork for future investigation into the prosody of ʔayʔaǰuθəm, giving preliminary evidence both for syntax-prosody mapping at the phrasal level, and for prosodic effects of focus. These findings show similar constituency effects to those described in Koch (2008) for Nłkepmxcín, but differ in finding prosodic effects of focus where Nłkepmxcín does not mark focus prosodically (but note that Koch's methodology is quite different, since he works with a much larger corpus of recordings, but does not compare only directly minimal pairs); it also differs from the description of SENĆOŦEN prosody in Benner (2006) and the findings of Caldecott (2016) for St'át'imcets (though again the methodology is quite different), which both point to these languages lacking prosodic prominence marking focus. Although we cannot be sure that the different findings are not due to different methodologies, it is quite probable that there is variation within the Salish language family with respect to whether focus is indicated with intonation. Future work should examine syntax-prosody mapping for a wider range of structures, as well as the role of duration in marking constituency, particularly at CP boundaries. The findings also indicate that the prosodic effects of information structure merit further attention, so it would be useful to examine whether given material is de-accented, for example, as well as whether different types of focus (e.g. corrective, contrastive) have different intonation profiles.

References

Benner, Allison. 2006. Prosody of SENĆOŦEN. Paper presented at the 41st International Conference on Salish and Neighbouring Languages.

Berg, Rob van den, Carlos Gussenhoven, and Toni Rietveld. 1992. *Downstep in dutch: Implications for a model*, 335–359. Cambridge, MA: Cambridge University Press.

Blake, Susan. 2000. On the distribution and representation of schwa in sliammon (salish): Descriptive and theoretical perspectives. Doctoral Dissertation, University of British Columbia, Vancouver, BC.

Boersma, Paul. 1993. Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. *Proceedings of the Institute of Phonetic Sciences* 17:97–110.

Caldecott, Marion. 2016. St'át'imcets intonation contours: a preliminary study. *Canadian Journal of Linguistics* 61:119–155.

Davis, Henry. 2003. Mind the gap: on plural agreement and A' extraction in St'at'imcets (Lillooet Salish). In *Papers for the 38th International Conference on Salish and Neighbouring Languages*, 23–46. Vancouver, BC: UBCWPL.

Davis, Henry, and Lisa Matthewson. 1999. On the functional determination of lexical categories. *Revue québécoise de linguistique* 27:29–69.

Davis, John. 1970. Some phonological rules in Mainland Comox. Master's thesis, University of Victoria, Victoria, BC.

Demirdache, Hamida, and Lisa Matthewson. 1995. Syntactic categoreies in St'at'imcets(Lillooet Salish). In *Papers for the 30th International Conference on Salish and Neighbouring Languages*, 69–75.

Elfner, Emily. 2012. Syntax-prosody interactions in Irish. Doctoral Dissertation, Univervisity of Massachusetts at Amherst, Massachusetss.

Féry, Caronline, and Shinichiro Ishihara. 2010. *How focus and givenness shape prosody*, 36–63. Oxford, UK: Oxford University Press.

Huijsmans, Marianne. 2017. The scalar exclusive operator ʔut in ʔayʔaǰuθem: 'no more than' and no exceptions! In *UBC Qualifying Papers 4*. UBCWPL.

Ito, Junko, and Armin Mester. 2013. Prosodic subcategories in Japanese. *Lingua* 124:20–40.

Koch, Karsten. 2008. Intonation and focus in nłeʔkepmxcin (thompson river salish). Doctoral Dissertation, UBC, BC.

Kroeber, Paul. 2002. Lexical clitics in two Salish languages. Ms., Indiana University.

Pak, Marjorie. 2008. The postsyntactic derivation and its phonological reflexes. Doctoral Dissertation, Massachusetts Institute of Technology, Boston, Massachusetts.

Rooth, Mats Edward. 1985. Association with focus. Doctoral Dissertation, UMass Amherst, Massachusetts.

Selkirk, Elisabeth. 1986. On derived domains in sentence phonology. *Phonology Yearbook 3* 371–405.

Selkirk, Elisabeth. 2009. On clause and intonational phrase in japanese: The syntactic grounding of prosodic constituent structure. *Gengo Kenkyu* 136:35–73.

Selkirk, Elisabeth. 2011. *The syntax-phonology interface*, 435–483. Oxford, UK: Blackwell Publishing.

Selkirk, Elisabeth, and Koichi Tateishi. 1991. *Syntax and downstep in japanese*, 519–543. Dordrecht: Kluwer.

Truckenbrodt, H. 2002. Upstep and embedded register levels. *Phonology* 19:77–120.

Truckenbrodt, Hubert. 1999. On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry* 30:219–256.

Wagner, Michael. 2005. Prosody and recursion. Doctoral Dissertation, MIT, Boston, MA.

Wagner, Michael. 2010. Prosody and recursion in coordinate structures and beyond. *Natural Language and Linguistic Theory* 28:183–237.

Wagner, Michael. 2015. *Phological evidence in syntax*.

Watanabe, Honoré. 2003. *A morphological description of Sliammon, Mainland Comox Salish, with a sketch of syntax*. ELPR Publications Series AZ-040. Osaka: Osaka Gakuin University.