Intonation Contours in St'át'imcets*

Martin A. Oberg
University of British Columbia

It has been observed in many languages that phrase level rises in fundamental frequency (F0) are associated with interrogative clauses. Recent work on Salish intonation has shown that yes/no questions in Squamish may not fit this generalization. However, no study has examined the contours associated with question words in Salish. This paper aims to fill this gap by examining F0 contours of question words in St'át'imcets (Lillooet), an Interior Salish language. This paper reports on two studies that compare the location of the F0 peak in stressed syllables of question words to non-question words. The results show that *swat*, 'who', has a later peak than the near homophone *zwat* 'to know'. A survey of other non-question words shows that the late pitch peak is uncommon in the language and suggests an association of a word level rising pitch contour with questions in St'át'imcets.

1 Introduction

This paper examines F0 contours in St'át'imcets (Lillooet), an interior Salish language. Recent work on Salish intonation has focused on yes/no questions. Jacobs (in prep) observered that in Sḵwx̱wu7mesh (a related Salish language) the sentence initial person marking clitic in declaratives has a higher F0 than in yes/no questions. This is contrary to the prediction made in Ladd's Strong Universalist Hypothesis (Ladd, 1981). Caldecott (p.c.) found that St'át'imcets has no final rise in either declaratives or yes/no questions as measured by comparing the max F0 of the last two syllables.

The present research aims to examine St'át'imcets *wh*-question intonation by comparing the shapes of syllable contours. Although conclusions about all *wh*-words are speculative, this paper shows that

*Collection of primary data was made possible by the resources available in field methods course work. I am also grateful for the effort and patience of our consultant, Laura Theverage.

swat, 'who', has a unique pitch contour not found on the near homophone *zwat*, 'know', or other words that occur in the same sentence position.

1.1 Background

Languages tend to differentiate sentence types by intonation. Hermann (1942) claims that interrogativity is signaled by a high pitch somewhere in the utterance. The Frequency Code (Ohala, 1983, 1984, 1994) describes pitch variation of sentence types as a result of ethological processes whereby low pitches are associated with dominance and assertions and higher pitched sounds are associated with submissiveness and questioning. Vaissiere (1995) uses this reasoning to posit a basic inventory of universal contours: a rise-fall typically associated with declaratives, and non-falling with questions. A sentence final rise in pitch associated with questions has been reported by many researchers. However, not all questions pattern this way. For example, English yes/no questions are often described as having a sentence final rise in pitch, but *wh*-questions typically have the opposite, final falling pattern associated with declaratives.

Other work has shown that important words of various sentence types tend to be associated with particular accents. Hedberg and Sosa (2002) found that the initial *wh*-question word is most often accompanied by a L-H* pitch accent, and the initial auxiliary in yes/no questions primarily receives a H* accent. This result is not to be confused with sentence final accents which typically differentiate yes/no questions and declaratives by either rising or falling pitch. However, this does not hold universally. A survey of 53 languages by Ultan (1978) found a slight preference for falling pitch on question words over high or rising pitch (52.1% vs 47.9%). This leaves the intonation patterns of newly studied languages open to empirical research. This present paper aims to add to the understanding of St'át'imcets intonation by examining *wh*-questions and looking at the shape of contours (rising vs. falling) instead of midpoint frequencies. Before explaining the measurements used in this study, I will first cover the syntactic issues relevant to the analysis.

1.2 Word order in St'át'imcets

St'át'imcets is a predicate initial language. There are two dialects: Upper, which has VOS word order, and Lower (spoken by the speaker in this study) which has VSO order. Predicates are sentence initial and take DP arguments. *Wh*-question words in St'át'imcets also function as predicates (see Davis et al. (1993) for arguments) and occupy initial position as seen in (1).

(1) Swat ku=az' ku=qmut[1]
who DET=buy DET=hat

'Who bought a hat?'

To claim that *wh*-words have a unique pitch contour, it must be shown that they behave differently than words of the same category. The first comparison tests whether *swat* behaves differently than other predicates. The near homophone *zwat*, '(to) know', was used to control for syntactic category, vowel, and immediate phonetic environment. If any word is to have a similar pitch contour, it would likely be a similar sounding word. However, unlike the *wh* use of *swat*, *zwat* in this position must have person agreement morphology (2) leaving the number of syllables on the target word uncontrolled.

(2) Zwát-en=lhkan k=Elvis
know-DIR=1SG.SUBJ DET=Elvis
'I know Elvis'

Any differences observed in (2) could be viewed simply as a general difference between questions and declaratives. To address this, yes/no questions were examined to test whether predicates have a general question intonation. Yes/no questions are also predicate initial (3) allowing the same predicate to be examined in the same position.

(3) Zwat-en=lhkácw=ha k=Elvis
know-DIR=2SG.SUBJ=YNQ DET=Elvis
'Do you know Elvis?'

In these constructions, the yes/no question clitic *ha* follows the first word in the sentence. St'át'mcets stress follows a right-headed trochaic stress system (Roberts and Shaw, 1994) allowing (primary) stress to move from the first syllable to the third in (3). This study assumes that a primary/secondary stress distinction is not important. To motivate this, Caldecott (2006) found that the acoustic correlates of stress (mean F0, duration, and intensity) generally do not differentiate primary from secondary stress. In fact, out of four speakers, only one exhibited primary stressed vowels with significantly greater mean F0 than secondary stressed vowels. No other primary/secondary stress difference was observed.

Despite not being able to contrast *swat* and *zwat* without agreement morphology sentence initially, both are still stressed and, following Bolinger (1986, 15), are candidates to receive accents potentially associated with questions. One might look for an accent on the particle *ha*, however, stress is not allowed word finally (Eijk, 1997) preventing it from being able to carry an accent. This makes *zwat* the best candidate to receive an accent associated with an interrogative function. If there is a single contour associated with words in interrogative sentences, it would be observed for both yes/no and *wh*- questions. For *wh*-question words to

[1]All St'át'imcets sentences represent are written in practical orthography. They represent spoken forms which tend to collapse .

be considered to have a unique pitch contour, they need to behave differently than the initial predicate in declaratives (2), and predicates in yes/no questions. This would suggest separate contours for questions and declaratives as well as for each question type.

1.3 Contour Measurements

Returning now to the description of pitch contours, there are a few issues to address. First, that the location of the peak F0 is relevant to question intonation, and second, that the interval containing the peak affects the measurement of the peak location. One attribute used by researchers in studies of intonation is peak location. Researchers have found later F0 peaks to be associated with questions in a wide variety of languages. Gussenhoven and Chen (2000) observed that Chinese, Dutch, and Hungarian monolinguals judge sentences of an unknown language as questions when the accent syllable had a higher and later peak F0. House (2003) observed that Swedish listeners judge syntacticly ambiguous sentences more frequently as questions (rather than statements) when the final accent has a later peak and higher F0.

It is important to note that the location of peak F0 can be measured as an absolute time delay from some position (syllable or vowel onset), or the ratio over some interval (the entire syllable or just the vowel). Van Santen and Hirschberg (1994) measured absolute peak timing from syllable onset and relative position as a percent of the syllable rhyme (s-rhyme) "defined as the interval beginning with the start of the last sonorant in the onset (or vowel start if the onset has no sonorants) and that ends at the end of the last sonorant in the syllable," both of which yielded significant differences. This study measured the location of the peak F0, measured as a percentage (4) of the duration of the measurable pitch contour, corresponding to the s-rhyme.

$$(4) \quad PeakPercent = \frac{T_{maxFo} - T_{onset}}{T_{offset} - T_{onset}}$$

Using this formula, rising contours have a peak percent approaching 1, and falling contours near the beginning, PP = 0.

In addition to peak percent, which ignores F0 values, measurements were collected for maximum F0, F0 range, and final fall (5, 6)[2].

$$(5) \quad FinalFall(Hz) = F_{max} - F_{offset}$$

$$(6) \quad FinalFall(cents) = 1200 Log_2 \left(\frac{F_{max}}{F_{offset}} \right)$$

These measures help further differentiate contour shape and address the differences between correlates of question intonation. Higher F0 as proposed by Hermann and Ohala and later F0 peak as suggested by

Gussenhoven and House do not necessarily go hand in hand. For example, a word could have an earlier peak percent but higher F0, thus only supporting Hermanand Ohalaa. Measuring the final fall is a measure of peak stability. It is possible for two contours to have the same peak percent, but behave differently after the peak (either remain high, or fall). Such differences would be observed with this measure. The next two sections describe the experimental procedure for the studies.

2 Study 1

2.1 Materials

English translations of target sentences (e.g. 1, 2, 3) were printed on note cards and given to the consultant to translate to St'át'imcets. The sentences used for this study were elicited in prior work and known to have the desired translation. A total of 27 utterances were used for analysis. This included 15 *wh*-questions, 8 yes/no questions, and 4 declaratives.

2.2 Procedure

A native St'át'imcets speaker (female and also fluent in English) was instructed to read each card to herself and then ask the question in St'át'imcets. No time requirements were set and she proceeded at her own pace. Recordings were made on a Marantz PMD660 digital recorder at 44.1 kHz and transferred to digital storage for analysis by PRAAT, version 4.4.30 (Boersma and Weenink, 2001). The beginning and end of the measurable F0 contour corresponding to the s-rhyme (van Santen and Hirschberg (1994)), which included /w/ and /a/, was marked for each word to be automatically processed by a script.

2.3 Results

Figures 1 - 3 show an example demarcation and pitch contour for sentences used in the analysis.

Figure 4 plots a linear interpolation of the average F0 at %5 intervals. From visual inspection, one can observe that *swat* tends to increase throughout most of its duration whereas *zwat* rises to a maximum F0 roughly halfway through and then falls.

Swat has a significantly higher peak F0 compared to *zwat* in yes/no questions (2323Hz vs 195Hz) $T(21)=4.486$, $p < .001$. Peak percent

[2]Measuring an interval in cents is sensitive to the *perceptual* difference between two frequencies whereas measuring in Hz captures the *physical* difference between the rates of vibration. For example, the interval between 100Hz and 200Hz sounds the same as a 200Hz to 400Hz interval (they are both octaves – a doubling of F0). The sizes of the intervals in Hz (100Hz and 200Hz, respectively) are different, but both equal 1200 cents. This can be calculated using the formula: $c = 1200\text{Log}_2(A/B)$.

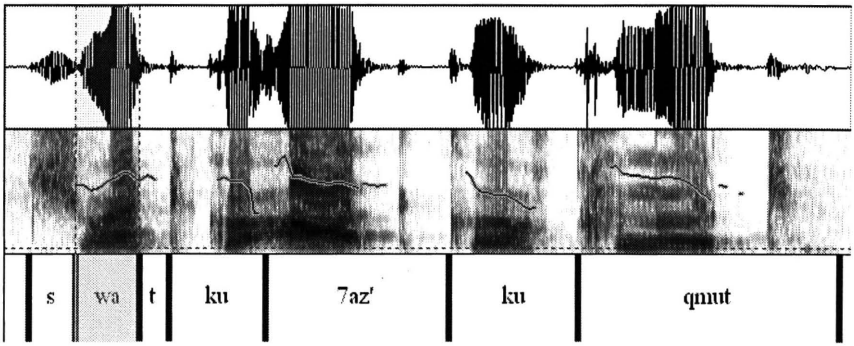

Figure 1: Pitch contour of a *wh*-question with *swat*

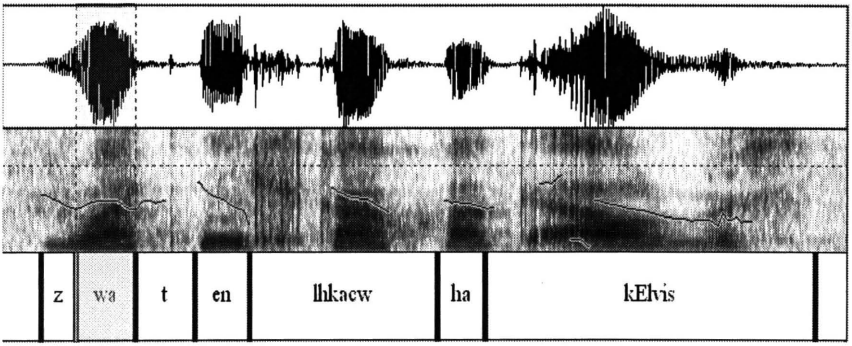

Figure 2: Pitch contour of question with *zwat*

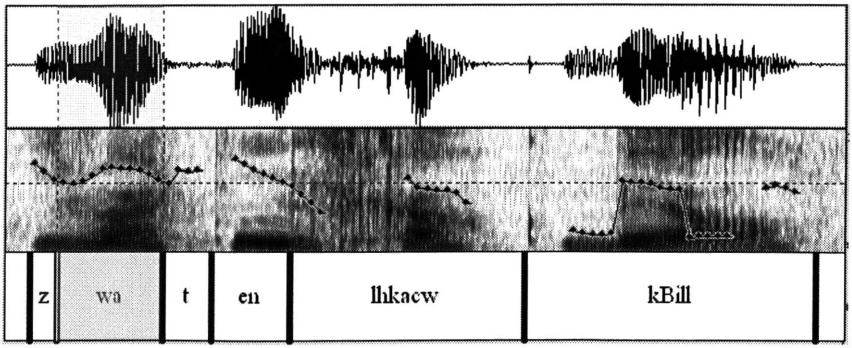

Figure 3: Pitch contour of declarative with *zwat*

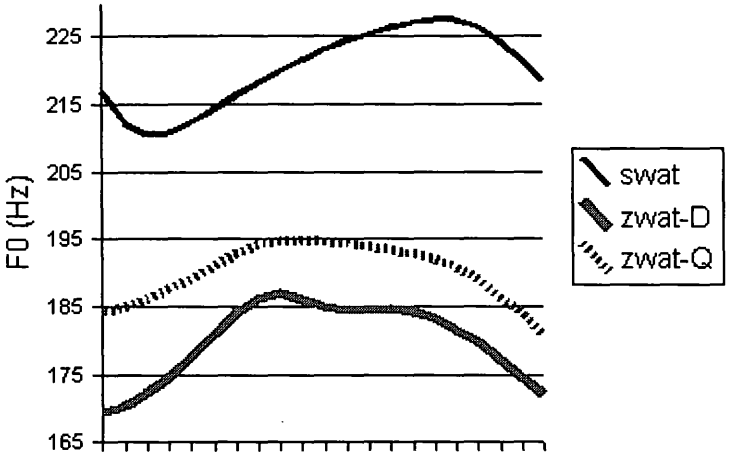

Figure 4: F0 contours of *swat* and *zwat*

values were also significantly different (.79 for *swat* and .46 for *zwat*) T(21)=7.425 $p < .001$. The overall change in F0 from maximum to minimum was not significantly different suggesting that it is not the amount of change in pitch that is important, but instead the location of the peak. Elicitations for *zwat* in declarative sentences had only started (N=4), but these preliminary results suggest that it also differs from *swat* (*zwat*-D PP = .53).

Computing an relative interval size from the initial F0, allows for an easy visual comparison of the relative contours. Figure 5 plots mean interval size from initial F0 to each measured frequency measured in cents. From this visual representation one can see that the fall after the peak is significantly greater ($p < .001$) for both *zwat* in questions (147cents) and declaratives (156cents) compared to *swat* (63.2cents). However, a significant correlation with peak percent (Pearson's r = -.62) shows that these measures are not easily separated.

Another characteristic that is easier to see when the contours are graphed in cents is that there appears to be a common inflection point located 25% through the contour where the *slope* starts to decrease. To illustrate this, one can estimate the slope at each point by calculating the slope between the preceding and following points using the equation:

$$(7) \quad slope_i = \frac{Fo_{i+1} - Fo_{i-1}}{2}$$

This formula estimates the slope *at* each point, not *between* points, and no slopes are calculated for end points. A peak in *slope* marks where the increasing F0 slows down. A peak (or trough) in F0 only occurs where the

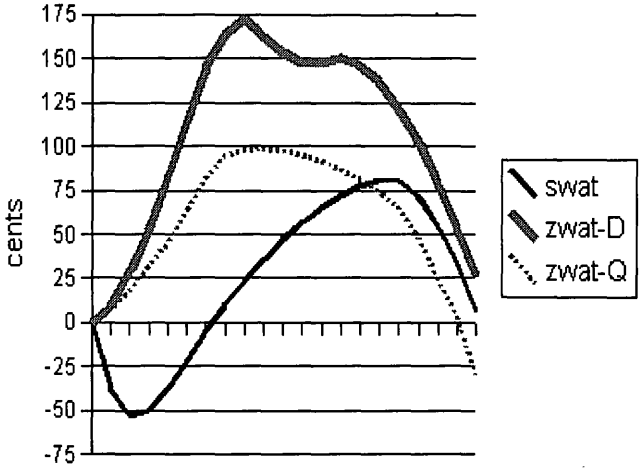

Figure 5: Normalized contours of *swat* and *zwat* in cents

slope *equals* 0. Figure (6) plots the resulting graph. The infection point, where the slope is maximized and starts to get smaller, has been circled.

The inflection point for *zwat* in declaratives is slightly earlier, but as mentioned before, this condition had limited tokens. Still, the majority of the decrease in slope does occur at the same time intervals. This inflection may be tied to the pronunciation of the /w/, however, this is beyond the scope of this paper. For present purposes, it suffices to point out the similarities in the contours where they exist.

The boxed portion of the slope contours is another shared characteristic: a rapid F0 decrease during the last 20% of the contour, perhaps showing the transition to the syllable final /t/. If we consider only the middle portion of the contour (after the inflection at 25% and before the rapid fall during the last 20%) where there is less influence by surrounding articulations, we can highlight the differences of peak location (figure 7).

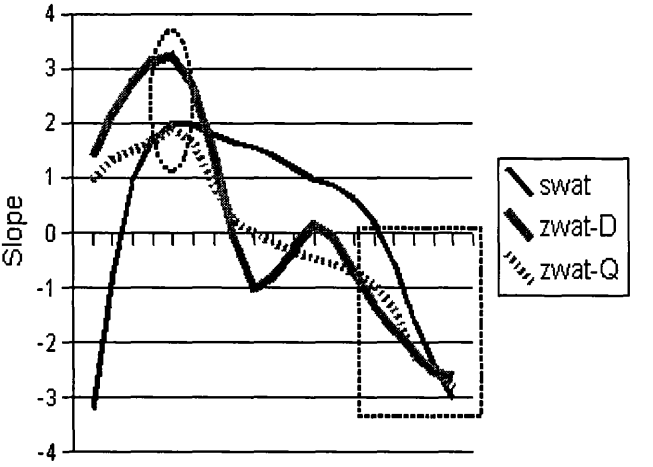

Figure 6: Estimated slope of contours

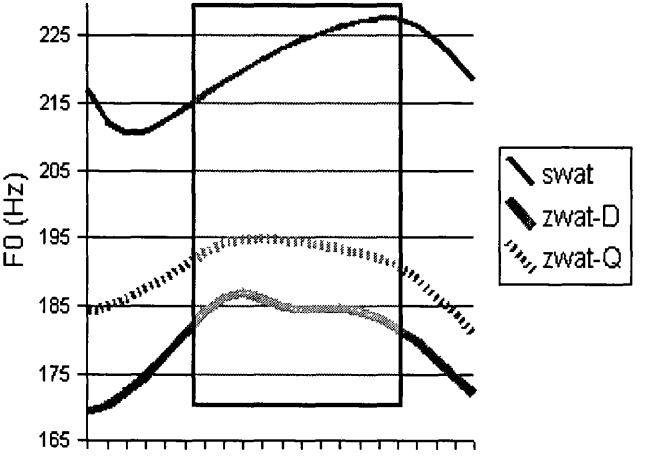

Figure 7: F0 contours of *swat* and *zwat*

The results in this section are summarized in Table 1 which shows mean values for word type. Means significantly different from from those measured for *swat* are marked with *s.

Table 1: Mean results for word type

Word	Peak F0 (Hz)	Peak Percent	Final Fall (Hz, cents)	F_{0MAX}-F_{0MIN} (Hz, cents)
swat	233	.79	8.3, 63.2	23, 176
zwat-Q	195***	.46***	16**147***	21, 193
zwat-D	194***	.53	17*, 156***	19, 183

significance level: ** < .01; *** < .001

2.4 Discussion

Significant differences of peak percent suggest that the pitch contour observed on *swat* is different from the non-*wh* word *zwat*. An additional difference in peak F0 was observed for *swat*. These can be easily seen in figure (7). Both results confirm cross-linguistic associations of high F0 and later peaks with questions. However, this only applies to *wh*-words; the predicate *zwat* behaved similarly in both declarative and yes/no question situations. It is unclear at this point whether there is a difference predicates in yes/no and declarative sentences.

However, the comparison with *zwat* was not representative of all non *wh*-questions, and the comparisons with declaratives was not complete. Given that the immediate phonetic environment was nearly identical, the evidence here suggest that these differences are more likely due to the *swat*'s use as a *wh*-word. A follow up study examined the F0 contours of a larger selection of sentence initial verbs (verbs with richer phonetic environments) to test if the observed contour for *swat* is distinct from non *wh*-words. These comparisons are described in section 3.

3 Study 2

This follow up was designed to test whether the rising contour observed for *swat* is distinct from other contours in sentence initial position. The extent to which the contour is not observed in a larger selection of words will support the hypothesis that the *wh*-word *swat* has a unique contour.

3.1 Methods

This study added 41 sentences to the analysis: 30 were taken from elicitation environments, and 11 from spontaneous speech in the

form of a story. The first word from each sentence was analyzed in the same manner as described in the previous section.

Of the words analyzed, 3 were predicates (8), 6 were conjunctions *nilh t'u7* (9), 10 were the cleft introducer *nilh* (10), 3 were quantifiers (11), and 19 were the auxiliary *wa7*, 10 of which were in yes/no questions (12), 6 were in declaratives followed by person pronouns clitics (13), and 3 functioned as imperfective to main verbs (14).

(8) Áolsem k=John
Sick DET=John
'John got sick'

(9) Nilh t'u7 tsícw=kalh múta7 ni=n-snúk'w7=a
FOC=just go=1PL.POSS CONJ DET=1SG.POSS-cousin=EXIS
'I went with my cousin.'

(10) Nilh s=Monique ti=áz'=a ti=lawíw's=a
FOC NOM=Monique DET=buy=EXIS DET=dress=EXIS
'It was Monique that bought a dress'

(11) Tákem swat wa7 zwat-en-táli k=Lisa
All who IMPF know-DIR-TOP DET=Lisa
'A lot of people know Lisa'

(12) Wa7=ha zwat k=Lisa
IMPF=YNQ know DET=Lisa
'Is Lisa known?'

(13) Wá7=lhkan ka=paqwalíkst=a
IMPF=1SG.SUBJ OOC=read=OOC
'I know how to read'

(14) Wa7 zwát-en-as John ni=n-snúk'w7=a
IMPF know-DIR-3ERG John DET=1SG.POSS-friend=EXIS
'John knows my friend'

3.2 Results

Figure 8 plots a linear interpolation of the average F0 at %5 intervals for each type of word, only non-falling contours are highlighted. Most words in this study were accompanied with a falling pitch contour obviously different from *swat*. The category of roots had a rise similar to *zwat* in study 1(peak percent = .35). Since the majority of the contours exhibit a falling pattern, peak percent is close to 0. *nilh t'u7 (2)* has a higher peak percent than expected for a falling contour, but this is due to one token reaching a max F0 at the end of the contour. Significant differences in peak percent and final fall were found for nearly every word

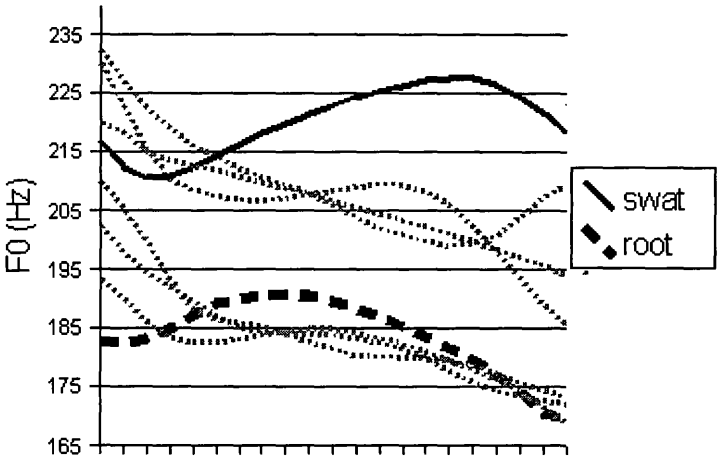

Figure 8: F0 contours sentence initial words

type. Table 2 shows mean values for Initial F0, peak percent, and final fall. *Swat* has been included for reference and *s denote significant differences from its mean values.

Table 2: Mean results for word type

Word	Peak F0 (Hz)	Peak Percent	Final Fall (Hz, cents)	F_{0MAX}-F_{0MIN} (Hz, cents)
swat	233	.79	8.3, 63	23, 176
nilh t'u7	278	.26***	360***, 52.5***	57.3 ***, 397**
cleft	232	.07***	381***, 46.3***	46.5*, 383***
quant	222	.13***	259***, 30.5***	30.9, 263
wa root	209	.01***	418***,45.7***	46.6*, 428**
waQ	201	.10***	306***, 31.8***	35*, 336**
wa=pn	192*	.13***	269***, 27.8***	28.3, 276
root	190**	.35***	169**, 17.1*	21.5, 210

significance level: * < .05; ** < .01; *** < .001

3.3 Discussion

The pitch contours of the words in initial position examined in this section are significantly different from that of *swat* as measured by peak percent and final fall. By and large, the most common contour is a exhibits a falling pattern. There were, however, some slight similarities between *swat* and other word types. The second *nilh t'u7* patten shares a late rise with *swat*, but differs in that it ends high. The *nilh* cleft has the

same fall-rise-fall pattern, but has an overall falling trend observed in the early peak percent (.07) and significantly larger final fall. This broader selection of words is further evidence that the rising contour associated with *swat* is unique to its status as a *wh*-word.

4 Conclusion

This study found the contour of the *wh*-word *swat* to be different from that of other sentence initial words. This included a near homophone *zwat*, the much used imperfective *wa7* in declarative and yes/no questions, the focus cleft introducer *nilh*, the conjunction *nilh t'u7*, and quantifiers, all of which were significantly different from *swat* in peak percent, peak F0, and final fall. These measures quantify the rising pitch contour observed for *swat* and show that the other words have a rise-fall contour or one that generally falls.

In the second study reported here, *swat* no longer has the highest peak F0. The second *nilh t'u7* conjunction has a much greater peak F0 and the *nilh* cleft rivals the peak F0 of the *wh*-word. This does not have to take away from the association of high pitch and questions. There are other intonational uses of pitch (focus in this case). This observation actually strengthens the case for late pitch peaks as a cue for *wh*word intonation. There is still no conclusive description of St'át'imcets yes/no questions, however, this study suggests that an overall higher F0 may be a factor.

This study marks the beginning of the phonetic analysis of St'át'imcets *wh*-questions. Additional analysis of other *wh*-words is needed to test whether the findings presented here are part of a general *wh*-word phenomenon. Studies of sentence wide intonation will be able to test whether St'át'imcets *wh*-questions have the same 'tune' as declaratives as has been observed in other languages.

References

Boersma, P., and D. Weenink. 2001. Praat, a system for doing phonetics by computer. *Glot International* 5:341–345.

Bolinger, D. 1986. Intonation and its Parts: Melody in Grammar and Discourse.

Davis, H., D. Gardiner, and L. Matthewson. 1993. A comparative look at WH-questions in Northern Interior Salish. *28th International Conference on Salish and Neighboring Languages* 28.

Eijk, J. 1997. *The Lillooet Language: Phonology, Morphology, Syntax.* UBC Press.

Gussenhoven, C., and A. Chen. 2000. Universal and language-specific effects in the perception of question intonation. *Proceedings of the 6 thICSLP* 91–94.

Hedberg, N., and J.M. Sosa. 2002. The Prosody of Questions in Natural Discourse. *Proceedings of Speech Prosody 2002* 375–378.

Hermann, E., H. Kees, and L. Morsbach. 1942. *Probleme der Frage*. Vandenhoeck & Ruprecht.

House, D. 2003. Perceiving question intonation: the role of pre-focal pause and delayed focal peak. *Proc 15th ICPhS* 755–758.

Ladd, D.R. 1981. On Intonational Universals. *The Cognitive Representation of Speech* 389–397.

Ohala, JJ. 1983. Cross-language use of pitch: an ethological view. *Phonetica* 40:1–18.

Ohala, JJ. 1984. An ethological perspective on common cross-language utilization of F0 of voice. *Phonetica* 41:1–16.

Ohala, J.J. 1994. The frequency code underlies the sound-symbolic use of voice pitch. *Sound Symbolism* 325–347.

Roberts, T., and P.A. Shaw. 1994. Optimality in the St'at'imcets (Lillooet Salish) Stress System. *Canadian Linguistic Association, Calgary* .

van Santen, J.P.H., and J. Hirschberg. 1994. Segmental effects on timing and height of pitch contours. *Proceedings of the International Conference on Spoken Language Processing* 2:719–722.

Ultan, R. 1978. Some general characteristics of interrogative systems. *Universals of human language* 4:211–248.

Vaissiere, J. 1995. Phonetic explanations for cross-linguistic prosodic similarities. *Phonetica* 52:123–130.

Martin Oberg
oberg.m.a@gmail.com