Prosody in Ktunaxa Interrogatives: An Initial Examination of Acoustics and Perception*

Brian Diep
University of British Columbia

Chenxi Xu
University of British Columbia

Molly Babel
University of British Columbia

M. Ryan Bochnak
University of British Columbia

Abstract: Across languages, prosody is often used to differentiate between different phrase types such as questions and statements. Although universal tendencies exist, there remains cross-linguistic variation in prosody for differentiating interrogatives and declaratives. The current work focuses on documenting prosody in Ktunaxa, an understudied and severely endangered language isolate. We conduct an acoustic analysis of Ktunaxa speech samples in order to quantify fundamental frequency (f0) contour differences in Ktunaxa utterances. We find that polar questions in Ktunaxa have higher overall f0 than that of declaratives, whereas wh-questions and tag questions display distinct patterns possibly related to focus. We further conduct a series of perception experiments to understand whether Ktunaxa speakers are either implicitly or explicitly aware of these differences and use prosodic cues alone to identify utterance type. We find evidence that Ktunaxa speakers indeed can use prosody to classify utterances at better-than-chance levels.

Keywords: prosody, perception, interrogatives, delcaratives, Ktunaxa

1 Introduction

Intonation refers to "the use of suprasegmental phonetic features to convey 'postlexical' or sentence-level pragmatic meanings in a linguistically structured way" (Ladd 2008:6). According to Ladd (2008), the heart of intonation lies in *pitch*, which is a listener's interpretation of the rate of vibration of the vocal folds, manifested as the fundamental frequency (f0) in the acoustic signal, and in *relative prominence* (or focus), with focus frequently marked by pitch accents. A higher pitch has been reported to mark interrogativity cross-linguistically, while where and how it is manifested varies across languages (Haan 2001; Ohala 1983). Ktunaxa is an Indigenous language spoken by the Ktunaxa people who traditionally reside in southeastern British Columbia, southwestern Alberta, western Montana, and parts of Washington and Idaho (First Peoples' Language Heritage Language and Cultural Council 2016). Questions are marked by the morpheme *k* (which is also the subordinate complementizer) in Ktunaxa across subtypes (Morgan 1991; Sandoval and Zhou 2023). Compared with declaratives (1), polar questions (2) feature the presence of the *k* morpheme (usually sentence-initial), while the indicative *-ni* suffix is absent.[1].

* Many thanks to our Ktunaxa consultants Violet Birdstone and Dorothy Alpine, and our participants whose participation made this work possible. We also thank members of the 2023—24 Field Methods course at UBC for comments, suggestions, and encouragement on this ongoing work.

Contact info: brdiep@mail.ubc.ca, xcx23xcx@mail.ubc.ca, molly.babel@ubc.ca, ryan.bochnak@ubc.ca

[1] We use the standardized Ktunaxa orthography (Kootenai Culture Committee 1999), where ɬ = [ɬ], ȼ = [t͡s], and V· = long vowel. Glosses are as follows: A.DEM = anaphoric demonstrative; ADV = adverbializer/preverb marker; COMP = complementizer; EVID = evidential; HAB = habitual; IND = indicative; M.DEM = medial demonstrative; OBV = obiative; POSS = possessive; PROG = progressive.

In *Proceedings of the International Conference on Salish and Neighbouring Languages 59*,
D. K. E. Reisinger, Laura Griffin, Ella Hannon, Gloria Mellesmoen, Sander Nederveen, Bruce Oliver, Julia Schillo, Lauren Schneider, Bailey Trotter (eds.). Vancouver, BC: UBCWPL, 2024.

(1) Sił ʔi·ku**ł**ni ka·pis Ȼan.
s-ił ʔi·ku**ł**-ni ka·pi-s Ȼan.
PROG-ADV drink-IND coffee-OBV John
'John is drinking coffee.'

(2) **K**sił ʔi·kuł ka·pis Ȼan?
k=s-ił ʔi·kuł ka·pi-s Ȼan
COMP=PROG-ADV drink coffee-OBV John
'Is John drinking coffee?'

However, how intonation plays a role in marking interrogativity in Ktunaxa remains unclear. Although a previous study investigating the prosodic realizations of focus in Ktunaxa found a direct mapping between higher pitch and focus among declaratives (McClay 2017), the interaction between prosody (including intonation) and interrogatives has not been studied. Moreover, the acoustic features (e.g., f0) are not equivalent to their psychoacoustic realizations (i.e., perception) (e.g., Ladd 2008), leading to the necessity of conducting perceptual research along with acoustic analysis to understand how users of auditory-vocal languages actually use language. However, how intonational cues are utilized during speech perception in Ktunaxa remains a question. Therefore, the present study serves as a novel and preliminary investigation of the intonation in Ktunaxa interrogatives, focusing on both acoustic analysis and its perceptual interpretations.

2 Background

2.1 Intonation and interrogativity

In line with the Universalist Hypothesis, Haan (2001) proposed that high pitch is associated with interrogativity cross-linguistically. The high pitch in interrogatives can be realized in different forms, including a final rise, a globally higher register, a raised nuclear accent peak, or simply less declination (Haan 2001). Several factors could affect such realizations, among which are cross-language differences and cross-question-type differences. On the one hand, the same type of question may have different types of pitch patterns across languages. For example, polar questions are commonly realized as a final rise in English and Cantonese, a high-pitched beginning in Finnish or Western Arabic, a globally higher pitch in Swedish, Chickasaw, or Mandarin, a raised nuclear accent peak in Bengali, and less pitch declination in Danish or French (Haan 2001; Thorsen 1980; Xu and Mok 2012). On the other hand, different types of questions may also have different types of pitch patterns, even within the same language. For instance, although polar questions in English frequently feature a final rise, Hirschberg (2000) and Hedberg et al. (2004) found that over 80 percent of wh-questions (i.e., questions containing question words such as *what*, *where*, *who*) had a final fall in their corpora of spontaneous speech in American English.

Moreover, the degree of importance of intonation in marking interrogativity is also influenced by the semantics and syntax of the questions. According to the Functional Hypothesis proposed by Haan (2001), the intonational properties of questions will be more salient in the absence of lexical and/or syntactic markers. For instance, in an investigation of Dutch interrogatives, Haan (2001) found that the high and rising pitch was most evident in questions without lexical and syntactic markers, and the least salient in questions marked by both semantics (i.e., a question word) and syntax (i.e.,

inversion). Consistent with this theory, Caldecott (2016) also reasoned a lack of prosodic correlates of interrogatives in SENĆOŦEN, a Salish language, by the presence of syntactic and morphological cues.

2.2 Perception of interrogative intonation

Given the acoustic mapping between a high pitch and interrogativity, another question may arise, i.e., whether such mappings can be perceived and used by language users, especially when the intonational cues are acoustically subtle with the presence of syntactic and/or semantic cues. Research conducted on questions in different languages with different high-pitch realizations has generally suggested that such prosodic cues in marking interrogativity are perceivable. For instance, after applying low-pass filters that filter out the semantic information of the stimuli, Medrano-Miller (2022) coded different types of sentences and successfully identified yes/no questions with an accuracy rate of 84 percent in Secwepemctsín, a Salish language. When conducting perception tasks among native Cantonese and Mandarin speakers, Xu and Mok (2012) also found that the categorization accuracy of low-pass-filtered utterances was above the chance level, although lower than the perception of unfiltered utterances. In the perception of Tashlhiyt Berber, an Afroasiatic language in which yes/no echo questions have a globally higher pitch and wider pitch span, speakers identified intonation contours in higher pitch registers more quickly and more often as questions than contours in a lower register (Roettger and Grice 2015). Similarly, in the perception of Danish utterances differing in only f0, the most steeply falling intonation contours were identified as being declarative, and the least falling ones were identified as interrogative, which echoes the fact that questions have less pitch declination than declaratives in Danish (Thorsen 1980).

2.3 Moderating role of experience in perception

Besides the acoustic nature of questions, whether and how intonation is utilized in perceiving questions may also depend on the characteristics of the perceivers, for instance, their specific language exp-eriences. Compared with the salient semantic and syntactic cues in marking interrogatives, the more subtle prosodic cues have been regarded as harder to learn (Chun 1998; Taylor 1993) and consequently may require more language experience before integration into perception. For instance, in the assessment of knowledge of English intonation patterns by L2 speakers, Mok et al. (2016) found that L2 learners (Accuracy = 42.94%) performed significantly worse than first language speakers (Accuracy = 72.57%). Such perception differences in intonation may even hold for more advanced learners. For example, in the perception of English intonation patterns, although the advanced German learners performed similarly to English first language speakers for some sentence types (e.g., statements, yes/no questions), they performed significantly worse than the L1 comparison group in cases like open and closed tag questions, which was explained by the researchers as evidence for universal perception errors across L2 groups (Puga et al. 2017). In the perception of low-pass filtered English declarative and polar questions, Patience et al. (2020) also found that the overall accuracy rates were highest for the L1 English listeners (84%), followed by L1 Mandarin (80%), Spanish (76%), and Inuktitut (75%) speakers.

2.4 Interrogatives in Ktunaxa

Questions in Ktunaxa are marked morphologically, mainly by the morpheme *k* (see Section 1 for a comparison between polar questions and declaratives). The wh-question is further marked semantic-

ally by the presence of a sentence-initial question word (3), and such wh-words include: *qapsin(s)* (‘what/why’), *qałaˀ(s)* (‘who/whom’), *qaˀ(s)* (‘how/where/when/why’), *qaˀas* (‘where/when’), and *q̓aksa* (‘how many/how much’) (Dryer 1999; Sandoval and Zhou 2023).

(3) **Qas** **ki**ʔin ka kanuhus yuwat̓ʔ
Qas **ki**=ʔin ka kanuhus yuwat̓
where COMP=to.be 1.POSS red bee
‘Where is my red bee?’

Like tag questions in English (e.g., “You did that, right?”; Tottie and Hoffmann 2006), tag questions in Ktunaxa also consist of a declarative anchor followed by a question tag where the *k* morpheme appears (4). The question tag commonly takes the form of *kqaqa* in Ktunaxa, which can be translated to “Is that so?” in English (c.f. *qaqa*, ‘to be so’).

(4) Miksan ʔat pał ʔisqat̓ɬiʔitni ʔinnu niʔs naktasuʔks, **kqaqa**ʔ
Miksan ʔat pał ʔisqat̓-ɬiʔit-ni ʔinnu niʔ-s naktasuʔk-s **k=qa-qa**
but HAB EVID cold-place-IND M.DEM A.DEM-OBV January- OBV COMP=thus-exist
‘However, it is cold there in January, right?’

While the morphology and semantics of Ktunaxa questions have been studied by previous scholars, no study has been conducted to investigate the prosody of Ktunaxa questions except for an earlier study by the current researchers (Anderson et al. 2023). Through acoustic analysis, it was found that interrogatives in Ktunaxa lack a final rise but have an overall higher pitch than declaratives. It was also observed that different subtypes of questions vary in pitch patterns: wh-questions begin with the highest pitch of all types of questions before falling to the pitch level of declaratives, and tag questions have a major rise at the end of the phrase. However, with a small sample size, the observed patterns may not be robustly supported.

3 Study overview and hypotheses

Given the lack of prosodic studies in Ktunaxa interrogatives, this study serves as a first-pass investigation into the acoustic and perceptual aspects of intonation in Ktunaxa interrogatives. We aim to explore: (i) how intonation plays a role in marking different types of questions in Ktunaxa, including polar questions (yes/no questions), wh-questions, and tag questions; (ii) how intonation functions in the perception of Ktunaxa questions; (iii) whether language experience moderates the perception of Ktunaxa intonation. A series of hypotheses have been developed based on previous studies on interrogativity and Ktunaxa.

Since (i) the former study (Anderson et al. 2023) observed that wh-questions in Ktunaxa have the highest pitch sentence-initially (where wh-words always occur) and tag-questions in Ktunaxa have a significant pitch rising sentence-finally (where question tags always appear); (ii) the focus corresponds to a higher pitch among declaratives in Ktunaxa (McClay 2017); (iii) focus of wh-questions is on wh-words cross-linguistically (e.g., Japanese, Ishihara (2002); Mandarin, Li and Thompson (1979), it is hypothesized that:

H1. The focus of wh-questions is on the sentence-initial question word and the focus of tag questions

is on the sentence-final tag, and such question focus will correspond to a higher pitch on those words.

Since (i) questions in Ktunaxa are marked by morphological cues (Sandoval and Zhou 2023); (ii) the role of intonation in marking interrogativity is secondary with the presence of syntactic and semantic markers (Haan 2001), it is hypothesized that:

H2. L1 and/or advanced Ktunaxa speakers will be able to perceive the intonation differences between declaratives and interrogatives in the absence of semantic cues, but their accuracy may only be slightly greater than chance.

H3. Manipulations of pitch height (e.g., changing the pitch differences between declaratives and polar questions) and contour (e.g., chopping out the relevant/focused part of questions such as the wh-word or the question tag) will change the perception accuracy.

Since previous studies have found that language experience positively correlates with the intonation knowledge of questions (Mok et al. 2016; Patience et al. 2020; Puga et al. 2017), we hypothesize that:

H4. Later Ktunaxa learners will perform worse than L1/advanced speakers in perception studies.

To test the above hypotheses, this study includes both an acoustic analysis of the tokens collected from first language Ktunaxa consultants and a series of perceptual tests utilizing the tokens collected.

4 Methodology

4.1 Data collection

For the purposes of acoustic analysis and perceptual study, samples of Ktunaxa speech were recorded from two L1 Ktunaxa consultants throughout the project. We recorded instances of both declarative and interrogative utterances. Within interrogatives, we solicited polar questions, wh-questions, and tag questions as three relevant subcategories from the consultants. A variety of methods were used to collect both spontaneous and non-spontaneous speech samples of the studied types, including direct translation tasks, storyboards, as well as a guessing game. To simplify the present study, we included audio tokens from only one consultant whom the researchers were able to elicit the majority of their tokens from.

4.2 Data processing

Each of the collected utterances was manually segmented by the researchers using Praat to remove any leading and/or trailing silences. During this manual tokenization phase, the researchers also excluded tokens that did not correspond to full sentences, contained disfluent pauses (pauses with a duration longer than 300ms), or with excessive background noise. After this exclusion process, we were left with 86 declarative tokens and 149 interrogative tokens. See Table 1 for a more detailed breakdown by utterance type.

The tokens were left unaltered for the acoustic analysis. However, for the perceptual study, we applied further data processing. This included applying a low-pass filter at 400 Hz as in Medrano-Miller (2022). A filter at this frequency effectively removes all phonetic information that robustly

Table 1: Number of tokens by utterance type

Utterance Type	Token Count
Declarative	86
Polar question	43
Wh- question	70
Tag question	36

cues segmental identity by removing segmental phonology and semantic (word-level) content from the utterances but preserving intonational information. These filtered tokens were then randomly shuffled with labels blinded.

4.3 Acoustic analysis methods

The acoustic analysis conducted in this project sought to quantitatively describe differences in prosody between different utterance types in Ktunaxa. We used VoiceSauce (Shue et al. 2011) to estimate and extract acoustic features from each of our utterances. Due to limitations related to consistency in the recording setup, we could not ensure a standard quality for measures related to loudness or intensity, so we excluded them from this analysis. For the purposes of this preliminary acoustic analysis, we focus solely on f0.

In order to be able to compare f0 contours across phrases of varying lengths, we normalized the f0 by time. To do this, we considered the mean f0 for each quantile of the audio duration (e.g., for some n within the interval [0, 100], we calculated the mean f0 at the n^{th} quantile time point of each audio token).

Afterward, to establish whether the difference between the f0 contours of different utterance types was statistically meaningful, we fitted a generalized additive model (GAM) to the normalized f0 contour data. This model used the f0 as the dependent variable while the independent variables are a term for utterance type (declarative, polar, wh-, or tag), a term quantile of time, and a smooth interaction term between quantile of time point and utterance type. We observed the covariates of the fitted GAM model with statistically significant p-values and effect sizes to be indicative of acoustic differences in f0 between utterance types in Ktunaxa.

4.4 Perceptual experiment paradigms

For the perceptual study, we conducted a series of perceptual forced-choice experiments where the participant was presented with low-pass filtered tokens and was made to predict the original utterance type of each token. In each of these individual experiments, the researchers randomly selected an even number of tokens of each relevant utterance type from the collected data. The tokens in each trial were then randomized and renamed using a Python script to blind the identity of the order and utterance type to both the researchers and participants.

Before the task, the participant was instructed about the potential types of filtered utterances that they would be judging. They were shown an example sentence in Ktunaxa for each utterance type being tested, as well as a sample original audio and filtered audio to familiarize the participant with the stimuli that they would be judging. They were then instructed to make a judgment to the best of their abilities as to which utterance type the token belongs to in response to each token that

was presented during the task. During the task, the low-pass filtered tokens were presented for the participant. The participant was then asked to make a judgment before being presented with the next token. However, the participant could request for the token to be replayed as many times as necessary before providing their response. The participant was not informed about the number of total tokens that they would be listening to during each session, nor were they informed about the proportion of each type of token in each task.

A summary of all of the perceptual experiments including the trial designations by which the rest of the report will refer to them hereafter is provided in Table 2 below.

Each trial varied in exact length but the researchers designed each trial to contain at least 16 of each type of utterance being tested, and included a five-minute break for the consultant after every 15 to 20 stimuli token to prevent fatigue.

Table 2: Summary of each trial

Trial Designation	**Judgement Task Items**	**Goal**
1.1	Two-choice judgment between polar questions and declaratives	Test H2. Whether Ktunaxa speakers can identify differences between interrogatives and declaratives in the absence of semantic information Repeated with Lx learners to test H4 and the role of experience in learning prosodic patterns
1.2	Three-choice judgment between polar questions, wh-questions, and tag questions	Test H2. Whether speakers can identify differences between interrogatives and declaratives in the absence of semantic information
2.1	Two-choice judgment between f0-manipulated polar questions and declaratives (increased f0 differences)	Test H3. Whether manipulating f0 height will affect the Ktunaxa speaker's recognition of prosodic differences between interrogatives and declaratives
2.2	Two-choice judgment between pitch-manipulated polar questions and declaratives (decreased pitch differences)	Test H3. Whether manipulating pitch height will affect the Ktunaxa speaker's recognition of prosodic differences between interrogatives and declaratives
3.1	Two-choice judgment between wh-questions and declaratives	Test H2. Whether Ktunaxa speakers can identify differences between interrogatives and declaratives in the absence of semantic information Repeated with Lx learners to test H4. and the role of experience in learning prosodic patterns

3.2	Two-choice judgment between tag questions and declaratives	Test H2. Whether Ktunaxa speakers can identify differences between interrogatives and declaratives in the absence of semantic information Repeated with Lx learners to test H4. and the role of experience in learning prosodic patterns
4.1	Two-choice judgment between focus-chopped wh- questions and declaratives	Test H3. Whether manipulating pitch contour will affect the Ktunaxa speaker's recognition of prosodic differences between interrogatives and declaratives
4.2	Two-choice judgment between focus-chopped tag questions and declaratives	Test H3. Whether manipulating pitch height will affect the Ktunaxa speaker's recognition of prosodic differences between interrogatives and declaratives
4.3	Two-choice judgment between pause removed tag questions and declaratives	Test H3. Whether manipulating pitch contour will affect the Ktunaxa speaker's recognition of prosodic differences between interrogatives and declaratives

In most experiments, the tokens used in the experiment are simply the low-pass filter version of the original token. However, for several experiments, the research-ers further perturbed the audio before applying a low-pass filter. For experiment 2.1 where the pitch was manipulated, the researchers manipulated the pitch of interrogatives up +0.5 ERB (Equivalent Rectangular Bandwidth, a psycho-acoustic measure; Nolan 2003) and that of declaratives down -0.5 ERB. This had the effect of increasing the f0 gap between the two utterance types. Similarly, for experiment 2.2, the f0 was manipulated in the opposite direction such that interrogatives were shifted down -0.5 ERB and declaratives up +0.5 ERB. The specific f0 manipulations were chosen in order to ensure that the altered audio remained within the f0 range of a typical adult human-speaking voice (i.e., 95—255 Hz; Baken and Orlikoff 2000).

In experiment 4.1, we manually removed the initial question word of wh-questions before filtering the audio whereas in 4.2, we similarly removed the final tag phrase of the tag question before applying the filter. For 4.3, we maintained the final tag phrase, but instead removed any prosodic pauses between the main phrase and the tag phrase before filtering. These modifications shift the prosodic contour of the utterances by omitting a prosodically focused element (the question word in wh- words, and the tag phrase in the tag questions) or a salient boundary in the case of the pause in tag questions.

After each experiment, the participant's responses were compared with the actual true utterance types of the tokens. Metrics for evaluating classification tasks such as accuracy, as well as class-specific tasks such as sensitivity, and precision are reported for each experiment to compare the participant's overall performance between experimental conditions. Accuracy is the proportion of the number of correctly classified tokens overall. In our case, using interrogatives as the reference class, precision (also known as true positive rate) is a proportion of the number of correctly identified questions divided by the number of tokens identified as questions (regardless of whether it is correct

or not). This approximately measures the quality of question identification as it penalizes false positives. Sensitivity (also known as recall) is a proportion of the number of correctly identified questions divided by the total number of actual questions. This approximately measures the quantity of questions being identified.

For this study, each of the experiments was completed with the consultant who is an L1 Ktunaxa speaker. Some were repeated with three Ktunaxa learners (English L1) recruited from Ktunaxa language classes or UBC Field Methods classes to obtain a non-L1 speaker comparison.

5 Results

5.1 Acoustic analysis results

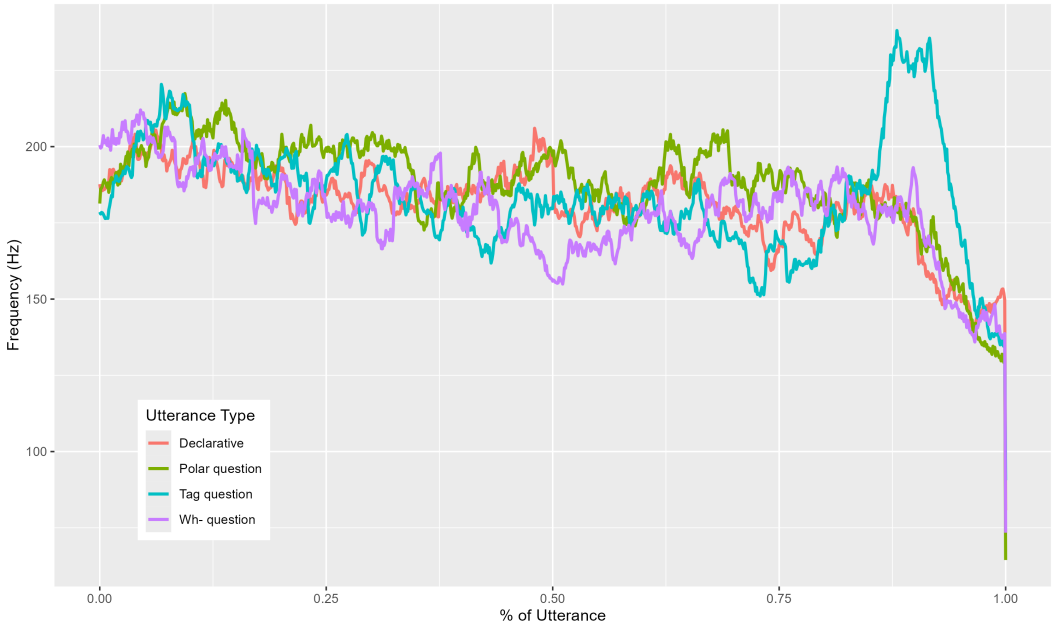

Figure 1: Mean time-normalized f0 contours across Ktunaxa utterance types

When plotting the f0 contours of the collected data (see Figure 1), we see several clear trends in f0 across utterance types. We observe that polar questions have a higher f0 contour than declaratives across the entire utterance. Furthermore, we see that they also tend to have higher mean f0 than other question types for the majority of the utterance. However, two characteristic exceptions are observable. Wh-questions tend to begin with the highest f0 of all utterance types before falling to the lowest overall f0 of all utterance types. Tag questions have a relatively average f0 contour compared to the other types but have a pre-final rise and fall pattern that even exceeds the f0 values of polar questions. In general, we also see that utterances in Ktunaxa generally follow a falling intonation pattern overall with an especially pronounced drop off towards the last quartile of each utterance.

When observing the fitted GAM model (see Figure 2), we find the same patterns within this model as the actual data, including that polar questions have the highest overall pitch contour, wh-questions have the highest initial f0, and the pre-final rise in tag questions. When observing the

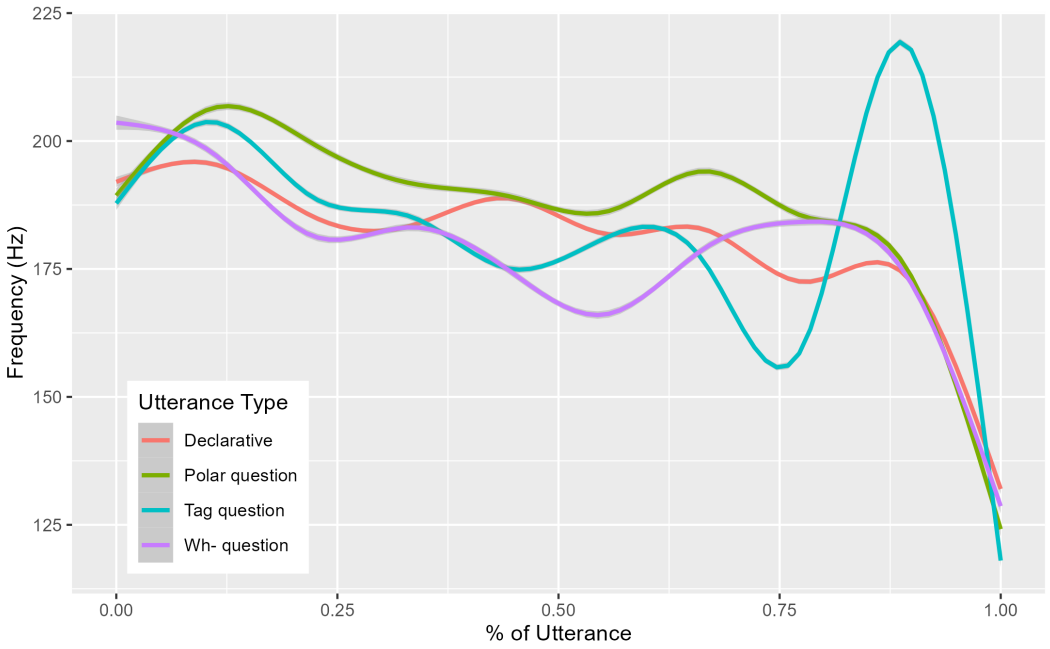

Figure 2: GAM model of time-normalized f0 contours across Ktunaxa utterance types

model summaries (see Table 3 and 4), all covariates including the interaction terms are found to be statistically significant from delcaratives, the reference level in the model.

Table 3: GAM parametric coefficients model summary

Covariate	**Beta estimate**	***p*-value**
(Intercept)	181.36	<2e-16
Polar question	6.83	<2e-16
Wh- question	-2.17	<2e-16
Tag question	2.93	<2e-16

Table 4: GAM smooth term summary

Interaction Term	**F statistic**	***p*-value**
normalized duration: declarative	2297	<2e-16
normalized duration: polar question	1943	<2e-16
normalized duration: wh- question	1346	<2e-16
normalized duration: tag question	1906	<2e-16

5.2 Perception study results

For the perceptual study, we present the results for each individual experiment in terms of the participant's overall accuracy as well as the sensitivity and precision regarding the interrogative

class. Results of the L1 Ktunaxa-speaker consultant can be found in Table 5 whereas the results for each of the Ktunaxa language learners are in Table 6, 7, and 8, respectively.

Table 5: Consultant perceptual experiment results summary

Trial Designation	Accuracy (%)	Sensitivity (%)	Precision (%)
1.1	60.53	26.32	83.33
1.2	46.67		
2.1	76	55	100
2.2	46	15	42.86
3.1	60	43.75	70
3.2	83.33	75	92.3
4.1	76.7	50	88.9
4.2	43.75	25	40
4.3	68.75	56.25	75

Table 6: Learner 1 perceptual experiment results summary

Trial Designation	Accuracy (%)	Sensitivity (%)	Precision (%)
1.1	57.89	47.37	60
3.1	50	43.75	53.85
3.2	53.33	50	57.14

Table 7: Learner 2 perceptual experiment results summary

Trial Designation	Accuracy (%)	Sensitivity (%)	Precision (%)
1.1	60.53	57.89	61.11
3.1	56.67	62.5	58.82
3.2	76.67	75	80

In trial 1.1 (polar questions vs. declaratives) we find that the consultant has a better-than-chance level classification accuracy. However, there is a tendency to treat both types of utterances as declarative. However, they do perform the best out of all participants (although only marginally better than most other participants). In trial 1.2 (polar questions vs. wh-questions vs. tag questions), we find that the consultant also achieves greater-than-chance accuracy at identifying all the three types of questions.

In trial 2.1 (f0 gap increased polar questions vs. declaratives), we find that augmenting the f0 difference between declaratives and interrogatives by -0.5 and +0.5 ERB respectively, increases the consultant's classification accuracy. Notably, it also results in perfect precision at identifying polar questions. In the similar trial 2.2 (f0 gap decreased polar questions vs. declaratives), we find that the classification of polar questions remains relatively low, albeit higher than the classification of the non-f0 altered tokens.

In trial 3.1 and 3.2 (wh-questions vs. declaratives, tag questions vs. tag questions), we find that the consultant is able to identify wh- and tag questions at a greater-than-chance rate, and with much higher precision and sensitivity than polar questions. This is especially the case for tag questions.

Table 8: Learner 3 perceptual experiment results summary

Trial Designation	Accuracy (%)	Sensitivity (%)	Precision (%)
1.1	55.26	52.63	55.56
3.1	66.67	68.75	68.75
3.2	83.33	93.75	78.95

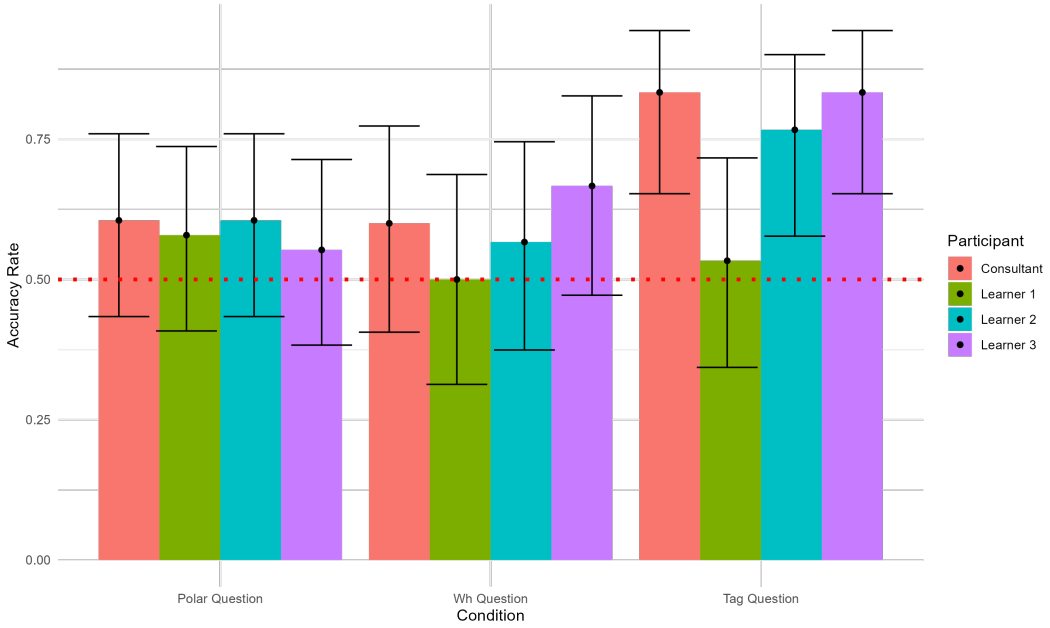

Figure 3: Participant Classification Accuracy for All Question Types vs. Declaratives

In the aforementioned trials which were both repeated by the researchers, the consultant performed at the same level or better than most learners (except for one learner in trial 3.1).

In trial 4.1 (wh-word removed questions vs. declaratives), we find that the consultant performs better than chance and is in fact above their classification baseline set in trial 3.1 (plain wh-questions vs. declaratives). In trials 4.2 and 4.3 (tag fragment removed questions vs. declaratives, tag pause removed questions vs. declaratives) we find that the consultant performs worse in both trials than the baseline set in trial 3.2 (plain tag questions vs. declaratives).

6 Discussion

We interpret these results as providing a variety of positive evidence supporting our first hypothesis that the focus of wh-questions and tag questions (i.e., the initial wh-word in wh-questions and the tag fragment at the end of a tag question) will correspond to higher f0. In our acoustic analysis, we find that there is a statistically significant difference between the f0 contours of different types. In our GAM model of the pitch contour (see Figure 2), we see that the wh-questions have the highest initial f0, and tag questions have a prominent rise in final f0, each corresponding to the position of our hypothesized focused element.

We also have evidence in favor of our second hypothesis that Ktunaxa speakers should be

able to perceive the differences between declaratives and interrogatives from intonation only. In all of the trials where participants were asked to differentiate between a type of interrogative and declarative, the consultant performed better than chance. However, tag questions have a markedly higher classification accuracy compared to the other two question types.

Furthermore, in trial 1.2 which involves the identification of specific interrogative subtypes in a three-way task, the consultant again performed better than chance. We propose that the gap in classification accuracy among question types is due to the prominent contour differences in tag questions that differentiate them from declaratives and other questions. Polar questions and wh-questions have a relatively flatter f0 contour over the utterance similar to declaratives, therefore it is sensible that Ktunaxa speakers perform worse at predicting these types of questions as opposed to tag questions with their prominent pre-final rise. Some evidence in favour of this proposal is found in the results of trials 2.1 and 2.2 where the polar questions have their f0 manipulated to be more different or more similar to declaratives. In the case of increasing the gap by shifting polar questions up and declaratives down, we increased the consultant's overall accuracy significantly, and maximized their precision for identifying questions. The opposite result was found when polar questions were shifted down and declaratives were shifted up. In that trial, we found that the consultant's performance was much worse than in the increased f0 gap trial albeit not worse than the baseline. These results suggest that Ktunaxa speakers are implicitly aware of how pitch differs between polar questions and declaratives. Moreover, given the overall lower-than-ceiling perception accuracy scores across subtypes of questions, we argue that the role of intonation in marking interrogativity in Ktunaxa may be secondary given the evident semantic and morphological markers, which also echoes the cross-linguistic Functionalist Hypothesis (Haan 2001).

With regards to our third hypothesis that manipulating pitch height and contour can change perception accuracy, we find positive evidence for this hypothesis in the case of polar questions and tag questions. In the previously mentioned trials 2.1 and 2.2 where the f0 gap was artificially increased or decreased between polar questions and declaratives, we found that when the gap is increased, classification accuracy increases, whereas when the gap is decreased, classification accuracy decreases. This is consistent with our hypothesis, since we successfully shifted the participant's classification by exaggerating the characteristics of the utterance (increasing the f0 of polar questions and lowering the f0 of declaratives). Similarly, in trial 4.2 where we removed the tag fragment from the tag question, the consultant performed significantly worse than their baseline identification of unperturbed tag questions. The effect was also seen in trial 4.3 where we removed the pause between the main phrase and the tag fragment. These manipulations had the effect of changing the overall contour of the tag phrase. These results show that removing the characteristic contour of a tag phrase does reduce classification accuracy as expected. It also suggests that speakers are aware of and use both the intonational differences in the final tag fragment as well as the pause boundary to identify tag questions.

The only exception to the predictions we make from our hypotheses is in the case of wh-questions and removing the initial question word as in trial 4.1. In this trial, the consultant performed better in their classification of wh-questions once the utterance-initial question word was removed than the baseline. We offer several possible explanations for the seemingly mysterious higher perception accuracy in the chopped wh-questions compared with the original tokens. First, the sentence-initial wh-words may not be the (sole) focus of wh-questions in Ktunaxa, despite their correspondence to a higher f0 found by the acoustic analysis. As argued by Ladd (2008), wh-words are not always the focus of questions typologically. Although in languages such as Romanian or Greek, the neutral

location for the main accent in wh-questions is on the wh-word, in other languages such as English, wh-questions share the same sentence stress principles as other sentence types, i.e., a neutral location of stress on the verb, with exceptions of echo wh-questions without wh-movement (e.g., "You did WHAT?" "They went WHERE?"). In this sense, the intonation of wh-questions in Ktunaxa may be more comparable to English. Moreover, the focus of wh-questions can be influenced by pragmatic factors. In an investigation of Mandarin intonation, Liu and Xu (2005) found that the intonation of focus and interrogativity can be encoded separately. Depending on which part of the sentence the speaker aims to emphasize, the focus (i.e., a higher-pitched part) could lie in different positions within a wh-question, including sentence-initial, sentence-medial, or sentence-final. Therefore, for those Ktunaxa wh-questions in which sentence-initial wh-words are not the (sole) focus, the chopping of wh-questions may not cause a significantly adverse impact on the perception accuracy. Secondly, the Ktunaxa speakers may rely on the overall more prominent pitch contour in wh-questions, rather than any specific focused elements, for discrimination. In a post-test interview with our consultant VB, the consultant mentioned that, if two fluent Ktunaxa speakers are talking with each other in daily conversation, you can feel the ups and downs in their speech if they are asking a question. This is also consistent with the findings that the perception accuracy for wh- and tag questions (with more dramatic f0 contours) were higher than for polar questions (with flatter f0 contours).

As for our final hypothesis pertaining to the performance of Ktunaxa learners compared to L1 speakers at intonation perception, we find mixed results. In general, the consultant performs better or as well as any of the learners. Our original expectation was that experience in Ktunaxa would correlate positively with classification accuracy due to greater knowledge of the intonational patterns. However, the scenario could be more complicated, as languages exhibit mutual influence (for an overview see Chang 2019). In the context of pitch, studies found that L1 tonal language speakers may have a more accurate perception of L2 intonation than L1 non-tonal speakers (e.g., Patience et al. 2020). Moreover, the lack of an experience effect could be related to some limitations in our experimental setup. In our experiments, the participants of the study had varying degrees of previous exposure to Ktunaxa, as well as exposure including specifically to the Ktunaxa of the consultant. This may make it difficult to draw general conclusions from their performance as a result of the bias caused by a familiarization effect.

7 Conclusion

This study is, to the authors' knowledge, the first acoustic and perceptual study into the nature of Ktunaxa interrogatives vs. declaratives. Through the acoustic analysis of tokens collected from Ktunaxa speakers, we find the intonational differences between declaratives and interrogatives, with further differentiation between different subtypes of interrogatives such as polar, wh-, and tag. Through a series of perceptual studies, we also find strong evidence that Ktunaxa speakers are aware of the intonational differences between utterance types and are able to use this information to classify utterances, even in the absence of semantic information. This study provides an overview of the prosody of interrogatives and declaratives in Ktunaxa and raises potential avenues for future studies. These include a more rigorous analysis that considers the interaction between prosody and other semantically important elements apart from question words. Future studies could also extend on the work on wh-questions by considering types of wh-questions in Ktunaxa that lack an explicit question word.

References

Anderson, L., B. Diep, and C. X. Xu. 2023. The intonation of declaratives and interrogatives in Ktunaxa. Poster Presentation.

Baken, R. J., and R. F. Orlikoff. 2000. *Clinical measurement of speech and voice*. San Diego: Singular Thomson Learning.

Caldecott, M. 2016. St'át'imcets intonation contours: A preliminary study. *Canadian Journal of Linguistics/Revue canadienne de linguistique* 61:119–155.

Chang, Charles B. 2019. The phonetics of second language learning and bilingualism. In *The Routledge Handbook of Phonetics*, 427–447. Routledge.

Chun, D. 1998. Signal analysis software for teaching discourse intonation. *Language Learning & Technology* 2:61.

Dryer, M. S. 1999. A comparison of preverbs in Kutenai and Algonquian. *Algonquian Papers-Archive* 30.

First Peoples' Language Heritage Language and Cultural Council. 2016. First peoples' language map of British Columbia: Ktunaxa.

Haan, J. 2001. *Speaking of questions: An exploration of Dutch question intonation*. Utrecht: LOT.

Hedberg, N., J. M. Sosa, and L. Fadden. 2004. Meanings and configurations of questions in English. In *Speech Prosody 2004, International Conference*.

Hirschberg, J. 2000. A corpus-based approach to the study of speaking style. In *Prosody: Theory and experiment: Studies presented to Gösta Bruce*, ed. M. Horne, 335–350. Dordrecht: Springer Netherlands.

Ishihara, S. 2002. Syntax-phonology interface of wh-constructions in Japanese. In *Proceedings of TCP*, 165–189.

Kootenai Culture Committee. 1999. *Ksanka ʔA·kłukaqwum - Kootenai Dictionary*. Elmo, Montana: Confederated Salish and Kootenai Tribes.

Ladd, D. R. 2008. *Intonational phonology*. Cambridge: Cambridge University Press.

Li, C. N., and S. A. Thompson. 1979. The pragmatics of two types of yes-no questions in Mandarin and its universal implications. In *Papers from the Fifteenth Regional Meeting Chicago Linguistic Society*, 197–206. Chicago, Ill: Chicago Ling. Soc.

Liu, Fang, and Yi Xu. 2005. Parallel encoding of focus and interrogative meaning in mandarin intonation. *Phonetica* 62:70–87.

McClay, Elise K. 2017. Focus in Ktunaxa: Word order and prosody. M.A. Thesis, University of British Columbia.

Medrano-Miller, C. V. 2022. The intonation of yes/no questions and declaratives in Secwepemctsín. BA Thesis, University of British Columbia.

Mok, P., Y. Yin, J. Setter, and N. M. Nayan. 2016. Assessing knowledge of English intonation patterns by L2 speakers. In *Speech Prosody*, 543–547.

Morgan, Lawrence. 1991. A description of the Kutenai language. Doctoral Dissertation, University

of California, Berkeley.

Nolan, F. 2003. Intonational equivalence: An experimental evaluation of pitch scales. In *Proceedings of the 15th International Congress of Phonetic Sciences*, volume 771.

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

Patience, M., L. Colantoni, G. Klassen, M. Radu, and O. Tararova. 2020. The perception and comprehension of L2 English sentence types: Cross-linguistic influence and task effects. *Gradus-Revista Brasileira de Fonologia de Laboratório* 5:71–98.

Puga, K., R. Fuchs, J. Setter, and P. Mok. 2017. The perception of English intonation patterns by German L2 speakers of English. In *Proc. Interspeech 2017*, 3241–3245.

Roettger, T. B., and M. Grice. 2015. The role of high pitch in Tashlhiyt Tamazight (Berber): Evidence from production and perception. *Journal of Phonetics* 51:36–49.

Sandoval, S., and K. Y. Zhou. 2023. Wh-questions in Ktunaxa. Lecture Handout.

Shue, Yen-Liang, Patricia Keating, Chad Vicenik, and Kristine Yu. 2011. VoiceSauce: A program for voice analysis. In *Proceedings of the 17th International Congress of Phonetic Sciences*, volume 3, 1846–1849. Hong Kong. URL https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2011.

Taylor, D. S. 1993. Intonation and accent in English: What teachers need to know. *International Review of Applied Linguistics in Language Teaching* 31:1–21.

Thorsen, N. G. 1980. A study of the perception of sentence intonation—evidence from Danish. *The Journal of the Acoustical Society of America* 67:1014–1030.

Tottie, G., and S. Hoffmann. 2006. Tag questions in British and American English. *Journal of English Linguistics* 34:283–311.

Xu, B. R., and P. Mok. 2012. Cross-linguistic perception of intonation by Mandarin and Cantonese listeners. In *Proceedings of the Sixth International Conference on Speech Prosody 2012*, 99–102.