Gradus 5.1 (2020) 71–98

Gradus
Revista Brasileira de Fonologia de Laboratório

The perception and comprehension of L2 English sentence types: cross-linguistic influence and task effects

Matthew Patience (matt.patience@utoronto.ca)
University of Toronto

Laura Colantoni (laura.colantoni@utoronto.ca)
University of Toronto

Gabrielle Klassen (gabrielle.klassen@mail.utoronto.ca)
University of Toronto

Malina Radu (malina.radu@mail.utoronto.ca)
University of Toronto

Olga Tararova (otararov@uwo.ca)
Western University

Abstract

L2 prosody is particularly difficult to acquire, because it requires an understanding of intonation, syntax, and pragmatics. For example, to acquire English sentence types, speakers must learn that statements (Ss) and absolute yes/no questions (AQs) are syntactically and prosodically marked, whereas the difference between Ss and declarative questions (DQs) is purely prosodic. Moreover, DQs can only occur in restricted contexts, such as to express surprise. In this paper, we examine the L2 perception and comprehension of English sentence types, by speakers of three typologically distinct L1s (Spanish, Mandarin, Inuktitut), with the goal of investigating the role of crosslinguistic influence (CLI). Spanish uses only intonation (a higher initial pitch accent and final rising boundary tone) to distinguish Ss from AQs and DQs, whereas in Mandarin, questions (AQs and DQs) can be syntactically identical to statements or marked by the lexical particle *–ma*. Mandarin also has a prosodic distinction between broad focus and echo questions (which are similar to English AQs and DQs). In contrast, Inuktitut has a very restricted use of pitch, and primarily marks questions morphologically. Learners of each L1 and English controls performed three tasks that varied in the amount of contextual and linguistic information available. Our results revealed evidence of both positive and negative CLI. Inuktitut and Mandarin speakers demonstrated some tendencies to focus more on syntax than intonation. Moreover, the Mandarin speakers were the most successful at acquiring the pragmatic distinction between AQs and DQs, which we argue is due to a similar contrast in Mandarin.

Article history
Received 2019-10-08
Revised 2020-05-22
Accepted 2020-05-26
Published 2020-08-10

Keywords:
crosslinguistic influence
L2 English
phonetics
prosody

Resumen

La prosodia es particularmente difícil de adquirir, porque requiere un entendimiento de la entonación, la sintaxis, y la pragmática. Para adquirir la distinción entre preguntas y declarativas en inglés, un hablante tiene que aprender que las declarativas (Ss) y preguntas absolutas sí/no (AQs) están sintáctica y prosódicamente marcadas, mientras que la diferencia entre Ss y preguntas declarativas (DQs) es únicamente prosódica. Además, las DQs solo ocurren en contextos limitados. En este estudio, examinamos la percepción y comprensión de las Ss, AQs, y DQs del inglés, por hablantes nativos de español, mandarín, e inuktitut, con el objetivo de investigar la influencia interlingüística (CLI). En español, la diferencia entre las Ss y las AQs/DQs es puramente prosódica, mientras que, en mandarín, las preguntas (AQs, DQs) pueden ser idénticas sintácticamente a las Ss, o estar marcadas por la partícula *-ma*. El mandarín tiene también una distinción prosódica entre preguntas de foco amplio y preguntas confirmatorias, similar a la del inglés. El inuktitut, en cambio, tiene un uso limitado de la prosodia; las preguntas están marcadas morfológicamente. Aprendices de cada una de estas lenguas y un grupo de control completaron tres tareas. Los resultados revelaron evidencia de CLI positiva y negativa. Los hablantes de mandarín e inuktitut se enfocaron más en la sintaxis que la entonación. Además, los hablantes de mandarín tuvieron más éxito en la adquisición de la distinción entre las AQs y DQs, lo cual, argüimos, fue debido al contraste pragmático existente en mandarín.

Palabras claves: transferencia interlingüística, L2 inglés, fonética, prosodia

Introduction

One of the principal goals in research on second language (L2) acquisition is to predict and explain what will be difficult for a learner acquiring a non-native language. L2 speech models, based on difficulty in perception[1] or markedness,[2] are focused on the acquisition of segments. Consequently, they are not easily applicable to L2 prosody and, as a result, we cannot effectively predict and explain difficulty in the acquisition of prosody. One of the challenges of learning L2 prosody is that it interacts with multiple linguistic domains, such as morpho-syntax and pragmatics. Therefore, an L2 learner needs to acquire not only the prosodic contours of a language, but also how these contours are associated with a given syntactic structure to be used in a specific context. A prime example of this difficulty is the acquisition of English sentence

[1] Flege, "Second language speech learning" (1995); Best and Tyler, "Nonnative and second-language speech perception" (2007).

[2] Eckman, "Typological markedness and second language phonology" (2008).

types, such as those exemplified in (1):

(1) a. Peter bought a piano.
b. Did Peter buy a piano?
c. Peter bought a piano?

English statements (S), as in (1a), are produced with a default falling final pitch contour. In contrast, absolute Y/N questions (AQs) are formed with subject-verb inversion or with *do* support (1b) and are produced with a final rising contour (and most likely with a higher initial pitch accent).[3] English also has declarative questions (DQs) (1c), which, like AQs, are produced with a final rising contour. However, in contrast to AQs, there is no subject-verb inversion or *do* support. Moreover, DQs are only used in restricted contexts.[4] To acquire English sentence types therefore requires knowledge of the prosody (falling contours for Ss, rising contours for AQs and DQs), syntax (subject-verb inversion and *do* support in AQs, but not Ss and DQs), and pragmatics (when to use DQs, AQs, or Ss). One of the objectives of the present paper is to examine whether L2 learners of English can distinguish between these sentence types using prosodic and syntactic information, and whether they can comprehend in which contexts each of these sentences are appropriate. To achieve this goal, it is ideal to analyze L1s that differ typologically (both in syntax and in prosody) from the L2, because this allows us to examine what role language experience plays. The present paper compares the perception and comprehension of English sentence types by L1 speakers of Mandarin, Spanish and Inuktitut. These three languages were chosen because they differ in terms of their prosodic typologies, and the ways in which questions are formed. Like English, Spanish is an intonation language. It marks questions with a high initial pitch accent and a rising as opposed to falling contour. However, syntactic differences are not used to mark question types in Spanish, and Ss and Y/N questions can have the same word order. Mandarin, in contrast, is a tonal language, in which questions are marked using syntactic and morphosyntactic cues, and differences in pitch. Finally, Inuktitut is an agglutinative language, with a very restricted use of prosody. It marks questions using a mood suffix (Indicative versus Interrogative). Due to prosodic and morpho-syntactic differences, we should expect varying levels of difficulty in the perception and comprehension of English sentence types.

[3] Saindon et al., "When is a Question a Question for Children and Adults?" (2017); Patience et al., "Initial Pitch Cues in English Sentence Types" (2018).

[4] Gunlogson, "True to Form" (2001); Truckenbrodt, "On rises and falls in interrogatives" (2009); Peterson, "Mirativity as surprise" (2016); Warren, *Uptalk* (2016).

Our second objective is to determine whether the type of task modulates the degree of crosslinguistic influence (CLI) in prosody, as it does with the acquisition of segments.[5] Previous work suggests that this is also the case with prosody.[6] However, the role of task in L2 prosody has not been tested extensively.

Our goals were addressed by examining the perception-com-

[5] Neufeld, "Phonological Asymmetry in Second-Language Learning and Performance" (1988).

[6] Grabe et al., "Perception of English Intonation by English, Spanish, and Chinese Listeners" (2003); Ortega-Llebaria and Colantoni, "L2 English intonation" (2014).

prehension of English sentence types across three tasks varying in the amount of linguistic and contextual information available to learners. In the intonation only (IO) task, learners discriminated between isolated statements and questions that had been low-pass filtered to remove segmental (and therefore lexical) information. This task was used to determine whether they could identify sentence types, based on their perception of rising versus falling contours. In the segments and intonation (SI) task, the learners heard the same sentences, unmodified, with both segmental and lexical information. If learners had difficulties distinguishing between statements and questions in the first task, and no difficulties distinguishing between AQs and DQs when the lexical information was present, it could indicate that they focused primarily on the syntactic cues. Finally, in the contextualized (C) task, learners were presented with a context, followed by three sentences (an S, an AQ, and a DQ), and had to identify which sentence was the most appropriate for the context. This task allowed us to determine whether learners had difficulty acquiring the pragmatic distribution of AQs and DQs.

Sentence types in English, Spanish, Mandarin, and Inuktitut

In this section, we present the prosodic, syntactic, and pragmatic characteristics of English sentences. This is followed by a summary of the relevant characteristics of Mandarin, Spanish, and Inuktitut sentence types, in order to highlight the expected role of CLI.

English

Canadian English[7] is an intonation language. The Y/N question types we study here (AQs and DQs) differ prosodically from statements by the presence of a final rise on the nuclear contour.[8] Recent research on a small corpus suggests that questions also have a higher initial pitch accent than statements,[9] which is further supported by the fact that native English speakers can discriminate between questions and statements based on differences in initial pitch accents.[10] In addition to prosodic cues, English AQs are marked syntactically, by the pre-posing of the auxiliary verb to sentence initial position (e.g., Pete can go to school *Can Pete go to school?*) or by the insertion of the auxiliary verb *to do* (as in [1b], above). In contrast to AQs, DQs have the same word order as Ss. The difference between AQs and DQs, then, is that DQs are used

[7] Canadian English is the variety used in the present study, given that it is the one that the participants have had the greatest exposure to.

[8] Bartels, *The Intonation of English Statements and Questions* (2014); Haan, "Speaking of questions" (2002); Hedberg and Sosa, "A unified account of the meaning of English questions with non-canonical intonation" (2011).

[9] Patience et al., "Initial Pitch Cues in English Sentence Types" (2018).

[10] Saindon et al., "When is a Question a Question for Children and Adults?" (2017).

in restricted contexts,[11] such as to express surprise or incredulity,[12] or for echo questions.[13]

[11] Gunlogson, "True to Form" (2001).
[12] Truckenbrodt, "On rises and falls in interrogatives" (2009).
[13] Gunlogson, "True to Form" (2001); Warren, *Uptalk* (2016).

Spanish

Spanish is the language that most resembles English prosodically. Spanish, like English, is an intonation language, and has Y/N questions that are distinguished from statements by a final rise[14] and a higher initial pitch accent.[15] In contrast to English, Spanish has variable word order for both questions and Ss, with orders varying primarily according to verb type (transitive versus intransitive) and information structure,[16] as opposed to sentence type,[17] as illustrated in (2-3):

(2) a. Lola compró una casa.
b. ¿Lola compró una casa?
c. ¿Compró una casa Lola?

Lola bought a house. Lola bought a house? Bought a house Lola?

(3) a. Vienen los bomberos.
b. ¿Vienen los bomberos?
c. ¿Los bomberos vienen?

Come the firefighters. Come the firefighters? The firefighters come?

In (2), the sentences are constructed with a transitive verb. The two questions (b, c) have the same meaning, although frequency of use is likely different across dialects. In (3), the sentences are constructed with an intransitive verb; (b) has a broad focus interpretation, whereas (c) has a contrastive interpretation. Given that Spanish speakers use intonation to mark sentence type in the same way English speakers do, we should expect Spanish speakers to be capable of identifying English DQs and Ss perceptually. In contrast, the fact that word order is not the primary cue to sentence type could make these learners less sensitive to the use of syntactic cues for distinguishing between Ss and AQs.

[14] Estebas-Vilaplana and Prieto, "Castillian Spanish intonation" (2010).
[15] Sosa, *La entonación del español* (1999). Dialectal differences have been reported for Spanish nuclear contours in Y/N questions. For example, some Caribbean varieties have a rising-falling contour; Hualde and Prieto, "Intonational variation in Spanish" (2015). One of the present study's speakers is from a region (Venezuela) that has been reported to have falling contours as the unmarked Y/N question form.
[16] Zubizarreta, *Prosody, Focus, and Word Order* (1998).
[17] Hertel, "Lexical and discourse factors in the second language acquisition of Spanish word order" (2003).

Mandarin

Mandarin is a tonal language. Pitch is used at the lexical level to contrast lexemes, which can be distinguished by tone alone. Mandarin speakers are therefore highly sensitive to prosodic differences, especially at the lexical level. Like English, Mandarin

has Y/N questions that are marked (morpho)syntactically, or that are syntactically identical to statements and marked with only intonation. An example of a Mandarin statement and three Y/N questions types are displayed in (4-7).[23] In (5), questions are marked using a particle *-ma* that is placed at the end of the utterance. In (6), questions are marked using a verb-not-verb (V-not-V) sentence structure. In contrast, the question in (7) is identical syntactically to the S in (4).

[23] Lee, "The prosody of questions in Beijing Mandarin" (2005).

(4) Statement:
Tā qù yīyuàn. 他去醫院.
he go hospital
"He is going to the hospital."

(5) -ma particle questions:
Tā qù yīyuàn **ma**? 他去醫院嗎?
he go hospital [*question particle*]
"Is he going to the hospital?"

(6) V-not-V questions:
Tā qù bú qù yīyuàn? 他去不去醫院?
he go not go hospital
"Is he going to the hospital?"

(7) Declarative questions:
Tā qù yīyuàn? 他去醫院?
he go hospital
"He is going to the hospital?"

Regarding pragmatic use, the primary difference between (5) and (6) is that, while both can be used in neutral contexts where there is no assumption regarding the answer to the question, (5), but not (6), can also appear in non-neutral contexts.[18] That is, contexts in which the speaker had a previous assumption regarding the response, such as in echo questions or to mark surprise or incredulity, similar to DQs in English. The DQ in (7), like (5), can be used in either broad focus questions or echo questions. Therefore, while (7) is similar in structure to English DQs, it can be used in contexts in which AQs or DQs would be used, as can (5). With respect to prosodic differences, Lee observed a three-way distinction between statements, broad focus questions, and echo questions. The Ss consisted of a falling F_0 across the utterance, the echo questions consisted of a rising F_0 across the utterance, and the F_0 in broad focus questions remained stable across the utterance. In addition to the rising F_0, echo questions were produced with an expanded pitch range throughout the utterance. This was also the case for broad focus questions, but differences with Ss were minimal. In contrast to both echo questions and Ss, broad focus questions were realized primarily with local prosodic differences, occurring over the final NP, and in the final syllable. More specifically, a larger pitch range expansion, and higher max F_0 were observed. While some minor prosodic differences were observed in questions formed with *-ma* when compared to DQs (in the same pragmatic contexts), the differences in pitch primarily reflected the pragmatic uses (i.e., broad focus versus echo questions). The author did not examine the prosody of V-not-V structures.

[18] Lee, "The prosody of questions in Beijing Mandarin" (2005).

In sum, Mandarin uses a combination of syntactic, morpho-syntactic, and prosodic cues to mark sentence type. While not identical to English, there are many similarities. However, two of the three question types can be used for both broad focus questions (like the English AQs) and echo questions (a type of English DQ). This could result in some mapping problems for the learners, given that they may inadvertently use syntactically marked and unmarked utterances for the English AQ and DQ contexts. Nevertheless, they also have a crucial advantage over the Spanish (and, as we will see in the next section, Inuktitut) speakers, since Mandarin speakers are already experienced at perceiving differences in broad focus versus echo-type questions, even though the linguistic domain used to mark these sentence types is different (syntax in English, intonation in Mandarin). In other words, we expect L1 Mandarin learners to be sensitive to the differences between AQs and DQs and their respective interpretations when given enough context.

Inuktitut

Inuktitut is an Eskimo-Aleut language spoken in North Eastern Canada. Differences between statements and questions are marked morphologically by a mood marker,[19] as in (10).[20]

[19] Spalding, *SALLIQ* (1969); Smith, *Some grammatical aspects of Labrador Inuttut (Eskimo)* (1977).

[20] Smith, *Some grammatical aspects of Labrador Inuttut (Eskimo)*, p. 15, 65 (1977).

(10) taku-vaga.
see I it (Indicative, transitive, 1SG-SBV)

(11) taku-vagaa?
see I it (Interrogative, transitive, 1SG-SBV)

In terms of prosody, very few studies have been conducted, and thus, our understanding of its use in Inuktitut is limited. Research on other Eskimo-Aleut languages[21] suggests that Inuktitut may be classified as an edge language[22] with limited to no word-prosody. Acoustic studies on Inuktitut have found evidence supporting this, revealing very little pitch movement across the utterance.[23] Changes in pitch are primarily observed utterance-finally, where a final rise indicates that the speaker is not finished speaking (i.e., a continuation marker), whereas a final fall marks the end of a turn. Nevertheless, questions also can be marked with a final rise and lengthening of the final vowel.[24] This prosodic cue is not consistently present, however, and the extent to which it is used will have to be investigated in future work. Of the four languages included in the present study, Inuktitut relies the least on changes in prosody; thus, we expect the L1 Inuktitut speakers to be the least sensitive to prosodic cues, and to focus more on syntactic cues. Therefore, they should perform best when syntactic markers are present. In other words, they are expected to distinguish between AQs and Ss but will be less target-like than the other groups in perceiving DQs.

[21] Arnhold ("Prosodic Structure and Focus Realization in West Greenlandic").

[22] Jun, "Prosodic typology" (2014a).

[23] Shokeir, "Intonation in Inuktitut" (2009).

[24] Shokeir, "Intonation in Inuktitut" (2009).

L2 prosody

CLI and task effects in the acquisition of prosody

A growing body of research has documented the effects of CLI in the acquisition of L2 prosody. This literature has provided evidence of mostly negative transfer from the L1 into the L2 in a variety of structures, such as peak alignment patterns,[25] nuclear stress placement,[26] or realization of corrective focus.[27] In particular, results from production studies show consistent L1 influence in the phonetic realization of contours in the L2, even by very advanced speakers. Most studies, however, have focused on production data of read or scripted speech.

[25] Mennen, "Bi-directional interference in the intonation of Dutch speakers of Greek" (2004).
[26] Zubizarreta and Nava, "Encoding discourse-based meaning" (2011).
[27] Ortega-Llebaria and Colantoni, "L2 English intonation" (2014).

Perception studies are limited, but evidence of negative transfer from the L1 into the L2 emerged in the pioneer studies by Dupoux and colleagues.[28] Crucially for the present study, this research showed that L1 French-L2 Spanish speakers could auditorily discriminate words with different lexical stress patterns, but were not able to use stress for lexical retrieval. That is, whereas participants used auditory information to differentiate words, they were not using this information linguistically (i.e., to retrieve words in the lexicon), as native Spanish speakers do. Such differences between auditory and linguistic performance have also been reported in the perception of intonation. Grabe *et al.*[29] found that L1 Spanish, L1 Mandarin, and L1 English speakers discriminated equally well non-speech stimuli that differed in the type of rise or fall. However, the language groups differed in their discrimination of the same contours once segmental material (namely English noun phrases) was present. Ortega-Llebaria and Colantoni[30] also reported L1 based differences in the perception of English corrective stress by L1 Mandarin and L1 Spanish learners. Participants completed two different types of tasks that manipulated the degree of access to meaning. The Spanish group performed less accurately than the Mandarin group, which was expected, due to Mandarin more closely resembling English in the way that contrastive meaning is expressed. Moreover, results showed that learners had higher accuracy rates in the decontextualized task, demonstrating that additional access to meaning resulted in more difficulty.

[28] Dupoux et al., "Persistent stress 'deafness'" (2008).

[29] Grabe et al., "Perception of English Intonation by English, Spanish, and Chinese Listeners" (2003).

[30] Ortega-Llebaria and Colantoni, "L2 English intonation" (2014).

Findings from these relatively recent studies on prosody are consistent with results on CLI and task effects in the acquisition of segments, which also document higher rates of CLI in tasks that tap linguistic knowledge. For example, Neufeld,[31] who studied the perception and production of L1 English-L2 French speakers across various tasks, found that performance in production decreased when the complexity of the task increased. In particular,

[31] Neufeld, "Phonological Asymmetry in Second-Language Learning and Performance" (1988).

he reported a higher rate of segmental errors when comparing a sentence imitation to a semi-spontaneous narrative. In the present study, we not only compare a variety of L1s, but we also include three different tasks with the goal of furthering our understanding of how these factors interact in predicting L2 performance in perception.

L2 acquisition of English sentence types: syntax and prosody

Previous work on the L2 acquisition of English sentence types by adult learners has revealed that learners can successfully acquire the inversion of AQs. For example, Pozzan and Quirk[32] investigated the syntactic acquisition of L2 English AQs by intermediate to advanced L1 Chinese (Mandarin, Cantonese, and Shanghainese) and L1 Spanish speakers. The L1 Chinese group produced no inversion errors, whereas the Spanish group produced inversion errors only 2.7% of the time. The results demonstrate that even intermediate learners with different L1s are relatively successful when acquiring AQs. However, the acquisition of DQs was not examined, thus it remains unknown whether the learners might have had difficulty distinguishing between the two forms.

[32] Pozzan and Quirk, "Second language acquisition of English questions" (2014).

As concerns the perception of prosodic cues to sentence types, previous work has demonstrated that naïve learners with different language backgrounds identify words with higher intonation, and final rises, as questions, when presented with utterances from an unfamiliar language.[33] This suggests that learners may be predisposed to identifying a final rise as a question. Although the participants in this study were speakers of languages in which rising and falling contours are used to mark questions and statements respectively (see Rialland[34] for a different use of prosody to mark sentence type), this research suggests that at least two of the groups in our study (i.e., Spanish and Mandarin speakers) should not differ in their auditory discrimination of sentence type.

[33] Gussenhoven and Chen, "Universal and language-specific effects in the perception of question intonation" (2000).

[34] Rialland, "The African lax question prosody" (2009).

There are a handful of studies which have looked specifically at the perception of English sentence types. In an early study, Cruz-Ferreira[35] tested L1 Portuguese-L2 English and L1 English-L2 Portuguese learners in their discrimination and identification of S and DQ intonation. In both languages, a fall is associated with statements and a high rise with questions. However, in Portuguese, a low rise is also associated with questions, whereas in English, it is ambiguous between statements and questions. When the L1-English-L2 Portuguese learners listened to pairs of stimuli in which both utterances had a low rise, or in which one had a low rise and the other one a low fall, they were able to discriminate them. However, they could not reliably identify

[35] Cruz-Ferreira, "Non-native Comprehension of Intonation Patterns in Portuguese and in English" (1983).

the sentence type, suggesting transfer from their L1. Evidence of negative transfer was also found in Liu and Rodríguez,[36] who examined L1 Chinese[37]-L2 English speakers' perception of final intonational contours using a categorical perception paradigm. The stimuli used were short Ss in which the contour over the final word was manipulated, creating a continuum ranging from 200 to 300 Hz. Results of an identification task showed that both groups differed in their threshold of the perception of a question. For the L1 Chinese group, it was located at the height of a level tone; for the L1 English group, question responses required the presence of a low rise.

Taken together, these studies provide evidence of mostly negative, but also some positive CLI in the perception and production of sentence types. To our knowledge, no study has explicitly examined the acquisition of English DQs, AQs and Ss; thus, it remains to be determined whether and to what extent learners of English might experience difficulty with the AQ/DQ contrast. Nevertheless, given some of the similarities and differences that were identified between English, Mandarin, Spanish, and Inuktitut, we expect to see both positive and negative CLI.

[36] Liu and Rodriguez, "Categorical perception of intonation contrasts" (2012).

[37] This is the term used by the authors.

Current Study

Research Questions and Predictions

Our primary goal was to determine whether L2 learner groups differ among each other and from controls in the perception-comprehension of sentence types, in order to have a better understanding regarding how CLI interacts with task effects to predict L2 learners' performance. Our study aimed to answer the following four research questions, which are presented with their corresponding predictions.

RQ1: Can L2 learners distinguish between statements and questions using intonation only?

H1: Based on CLI, in the IO task, we expect the following accuracy hierarchy (lowest to highest): Inuktitut > Mandarin, Spanish. Inuktitut learners should experience the most difficulty, given that their L1 has a very restricted use of pitch, and questions are primarily marked by morphology. As a result, Inuktitut learners are expected to either identify Ss as questions, or questions as Ss, more frequently than the other groups. We predict that Mandarin and Spanish speakers will display similar results, given that both groups are sensitive to changes in intonation for marking sentence type.

RQ2: Can L2 learners distinguish between statements and questions when sentences are produced in isolation? Do they use syntax or intonation?

H2: Given the presence of redundant cues in AQs, we do not expect learners in the SI task to experience difficulty identifying AQs. Greater difficulty is expected with DQs, depending on the L1. We predict the following hierarchy: Inuktitut > Mandarin, Spanish. Mandarin and Spanish speakers should consistently identify DQs correctly, given that in both languages, DQ structures with a rising intonation are always questions. In contrast, Inuktitut speakers should be the least sensitive to prosodic cues, and are expected to identify DQs as Ss, or Ss as questions, more frequently than the other groups.

RQ3: Can L2 learners comprehend in which context each sentence type is appropriate? Are differences observed between AQs and DQs?

H3: Larger differences between groups are expected in the C task, given that it requires knowledge of prosody, syntax, and pragmatics. The AQ/DQ distinction is likely to be the most problematic. We expect the Mandarin speakers to have an advantage, due to the fact that Mandarin has a pragmatic distinction between broad focus and echo questions (marked prosodically), which has some parallels to the AQ/DQ distinction in English. In contrast, Inuktitut speakers should overgeneralize AQs in place of DQs (focusing on the syntax), whereas Spanish speakers are more likely to overgeneralize DQs in place of AQs. Moreover, consistent with H1 and H2, we predict that Inuktitut learners will focus less on intonation, thus should be more likely to identify Ss as DQs. Overall, we predict the following hierarchy: Inuktitut > Spanish > Mandarin.

RQ4: Does CLI play a greater role with increasing levels of contextual information?

H4: For all learner groups, we expect CLI to increase, and be most prevalent in the contextualized task. This is based on previous work which has found that CLI plays a greater role in tasks involving more access to meaning.[38] Accordingly, given that Inuktitut speakers are expected to experience difficulty identifying sentence type based on intonation, we expect them to experience more difficulty correctly identifying Ss and DQs in the SI compared to the IO task, and the most difficulty in the C task. For Mandarin, we expect positive CLI to improve their performance as access to information increases. With respect to Spanish, CLI would predict that learners should experience difficulty in the C task, with the AQ-DQ contrast, given that this contrast is not marked syntactically in Spanish.

[38] Grabe et al., "Perception of English Intonation by English, Spanish, and Chinese Listeners" (2003); Ortega-Llebaria and Colantoni, "L2 English intonation" (2014).

Methodology

As part of a larger project, the participants performed a series of perception and production tasks, which were designed to answer questions related to the acquisition of a variety of structures. We only present the details of the perception tasks, given that they are the only ones relevant to the current study.

Participants

Fifteen L1 English controls, as well as 15 L1 Spanish, 15 L1 Mandarin, and 13 L1 Inuktitut adult learners of English, participated in the study. The controls were all native speakers of Canadian English. The L1 Spanish and L1 Mandarin groups were recruited from the Toronto area. The majority of the L1 Mandarin speakers were undergraduate students, having moved to Toronto as adults. They had, on average, been in Canada for two years. The L1 Spanish participants were either graduate students or members of the Hispanic community in Toronto. For this reason, the average age and length of exposure to English was higher in the Spanish compared to the Mandarin group (table 1).

Tabela 1: Participant profiles. Values reflect means; values in brackets indicate range. AoA=Age of onset of acquisition; LoR=Length of Residency; LoE=Length of exposure.

Group	N	Age	AoA (yrs)	LoR (yrs)	LoE (yrs)
English	15	25 (18-30)	n/a	n/a	n/a
Spanish	15	36 (19-71)	14 (6-26)	10 (2-26)	22 (3-59)
Mandarin	15	20 (18-23)	9 (4-14)	2 (1-4)	11 (5-15)
Inuktitut	13	35 (19-71)	5 5 (0-11)	n/a	30 (12-65)

However, the Mandarin group began learning English at an earlier age overall. This is because speakers typically begin learning English in school from an early age, although the amount of exposure to English during their early education is limited. In contrast to the L1 Spanish and L1 Mandarin groups, the L1 Inuktitut speakers were recruited in Ottawa, given the presence of a large Inuktitut-speaking community in the area. All Inuktitut participants were born in Northern Canada, from one of four regions (South Baffin (N=8), Nunatsiavut (N=3), North Baffin (N=1) and Churchill (N=1)), in which Inuktitut is widely spoken as the primary language (65% of the population of Nunavut has Inuktitut as the primary language, although Inuktitut-English bilingualism has increased since the 90s).[39] These speakers were first

[39] Lepage and Langlois, *Evolution of the language situation in Nunavut, 2001 to 2016* (2019); Dorais, *The Language of the Inuit* (2010).

exposed to Inuktitut at home, as most speakers (N=10) had two parents who spoke Inuktitut, whereas the remaining three had one parent who spoke Inuktitut. The Inuktitut speakers were educated in English, and thus grew up as sequential bilinguals. The participants reported that they continue to speak Inuktitut, with daily interactions in Inuktitut ranging from 25-90%. Given the learners' backgrounds, we must consider the possibility that any differences observed in the results could be partly due to the heterogeneity of their profiles. Nevertheless, we should still be able to identify which non-target responses are due to differences in the L1. Table 1 summarizes the details of the learner profiles.

Tasks and Stimuli

The participants performed three tasks that varied in the amount of linguistic and contextual information available. The IO task involved listening to isolated, low-pass filtered Ss, AQs, and DQs. The participants had to identify whether they heard a question, statement, or exclamation. The SI task consisted of the same isolated sentences, without the low-pass filtering. The C task comprised listening to a short scenario. Participants were then presented with three options (aurally): an S, DQ, or AQ. They had to select the utterance that was most appropriate to the context.

The stimuli for the IO and SI tasks consisted of 10 target sentences per context (S, AQ, DQ), in addition to 25 distracters (exclamations, *wh*-questions). The stimuli were produced by a 24-year-old female, native speaker of Canadian English (Ontario). The recordings were created in a quiet room using a Marantz professional solid-state recorder PMD661. The stimuli were read four times each, and subsequently listened to by three trained phoneticians, who selected the most natural sounding recordings of each stimulus. These stimuli were then low-pass filtered to be used for the IO task. This was achieved by identifying the minimum and maximum pitch in PRAAT,[40] and subsequently a pass Hann band filter was applied to remove acoustic data with an F_0 above the maximum, effectively removing any segmental information. An example of the pitch profile of each sentence type is displayed in fig. 1.

[40] https://www.fon.hum.uva.nl/praat/.

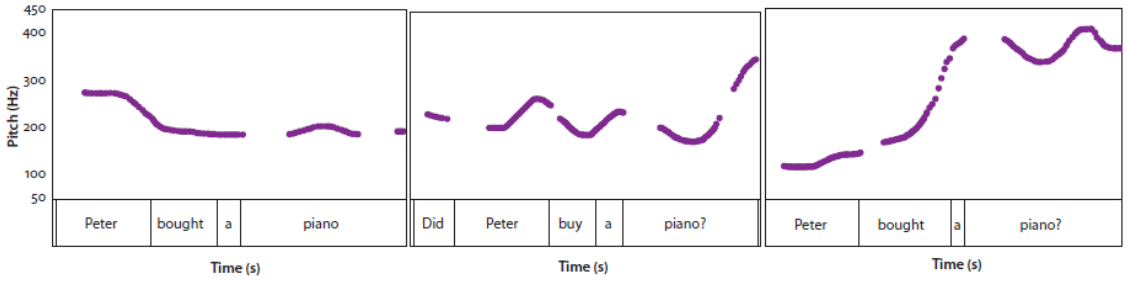

Figure 1: Example of pitch contours of the stimuli used in the IO and SI task. Note: S (left), AQ (middle), and DQ (right).

The stimuli for the C task consisted of 6 target scenarios per sentence type (S, AQ, DQ). No distracters were included in this task, in order to prevent it from becoming too long and cognitively demanding. The contexts were created such that only one of the three possible responses was felicitous. For example, the following context was designed to elicit a DQ:

> Before coming to Toronto from Australia, Mary heard about raccoons, looked at some pictures, and thought they were cute little things. One evening, she was eating outside with friends and saw a mid-sized animal crossing the street and thought it was a dog. Her friends commented that it was a raccoon, and she asked...
>
> a. This is a racoon.
> b. This is a racoon?
> c. Is this a racoon?

The contexts and stimuli were recorded by the same Canadian English speaker who recorded the stimuli for the other two tasks, using the same equipment. In order to test the appropriateness of the contexts for eliciting the target responses, the task was piloted with native and non-native speakers. During pilot testing, we eliminated and/or modified the context and responses, because some of the vocabulary items were too difficult for the L2 speakers. We then selected the contexts that consistently elicited the target response.

Procedure

Participants performed the tasks using SuperLab Pro, version 5. The task order was fixed, beginning with the IO task, followed by the SI task, and finishing with the C task. During the IO and SI tasks, participants heard the stimulus a single time, and could not repeat it. They subsequently had to respond using the SuperLab Pro response pad, which had three different colored buttons corresponding to the responses: one for a statement, a question, and an exclamation. Upon pressing the button to indicate a response, the next stimulus was presented immediately. The exclamation option was included for two reasons. First, to have a third option, so that pure guessing would only give a 33% chance of a correct response; and second, because previous work has demonstrated that DQs can be interpreted as exclamations.[41] It therefore further increased the difficulty of the task. During the C task, the participants listened to the context while also following it orthographically on a computer screen. As soon as the context was finished, speakers were presented aurally with three possible options: S, AQ, DQ. They indicated their selection on the same response pad used for the other two tasks. The participants

[41] Warren, *Uptalk* (2016).

were only able to select a response after hearing all three possible options. A brief practice session was included before each task, to ensure they understood how to perform the task. In between tasks, participants had a short break.

Data Preparation

The responses were extracted from SuperLab Pro into Excel, and the accuracy rates for each task were calculated (number of correct responses / number of total responses). We used mixed effects binomial logistic regressions to examine the role of the predictor variables on the accuracy outcome variable. In all cases unless otherwise specified, our random effects structure consisted of an intercept for participant, and a by-participant random slope for the effect of sentence type. For each task, a series of mixed effects binomial logistic regressions were run to determine the best model fit, comparing the base model with our random effects (M0), to a model including language (M1), language and sentence type (M2), and the interaction of language*sentence type (M3). The best model fit was determined according to differences in AIC and AIC weights, following the recommendations discussed in Wagenmakers and Farrell.[42] We then treated the best supported model as our main effects model. For all models, English was the reference level for language, whereas Ss were the reference level for sentence type. Post-hoc pairwise comparisons were performed (with a Tukey adjustment) when an interaction was present, to examine in greater detail the differences between sentence types for each language group. Statistics were performed using R v.3.6.1.[43]

[42] Wagenmakers and Farrell, "AIC model selection using Akaike weights" (2004).

[43] https://www.r-project.org/.

Results

In this section, we present the results by task and we conclude with a task comparison, so as to follow the order of our four research questions.

Intonation Only Task (IO)

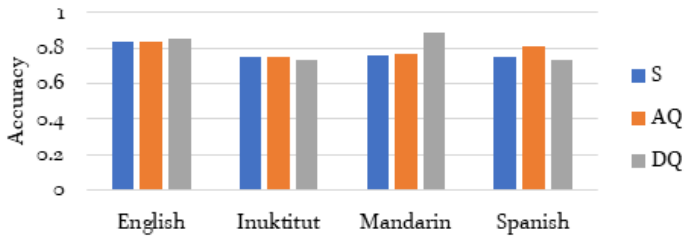

Figure 2: Bar chart displaying accuracy results in the IO task, by language.

Figure 2 displays the results of the IO task, grouped by language and sentence type. Overall accuracy rates were highest for the controls (84%), followed by the Mandarin (80%), Spanish (76%), and Inuktitut (75%) speakers.

The model comparison of the mixed effects binomial logistic regressions revealed that the base model was the best fit (M0: $AIC = 1693$, $AIC\omega = .41$; M1: $AIC = 1694$, $AIC\omega = .22$; M2: $AIC = 1694$, $AIC\omega = .15$; M3: $AIC = 1694$; $AIC\omega = .22$). Adding the predictors did not improve the models' explanatory power, indicating that neither language nor sentence type had an effect on accuracy. Therefore, while some variability was observed between sentence types, the learner groups were no different from the controls. Nevertheless, all groups displayed some errors, and we were interested in the types of errors produced, in order to identify the source of difficulty. Error counts by context are presented in Table 2. In order to determine whether group differences existed with respect to error types, we ran a conditional inference tree analysis, with sentence identification (S, Q, exclamation) as the response variable, and sentence type and language as the independent variables. Sentence type, not surprisingly, had the strongest association with the response variable, and error types were split first into questions (AQ/DQ) and Ss ($p < .001$). Language also had an effect. In the S condition, Inuktitut was grouped apart from the other three languages ($p = .008$). This was primarily due to the fact that the Inuktitut group misidentified Ss with similar frequencies as questions or exclamations, whereas the other languages misidentified Ss primarily as exclamations.

Tabela 2: Error counts by context and language in the IO task.

	Context					
	AQ		**DQ**		**S**	
Identified as:	**Excl.**	**Ss**	**Excl.**	**Ss**	**Excl.**	**Q**
English	17	8	14	7	21	4
Inuktitut	23	9	26	9	17	15
Mandarin	24	11	14	3	31	5
Spanish	19	10	35	5	29	8

Segments and Intonation Task (SI)

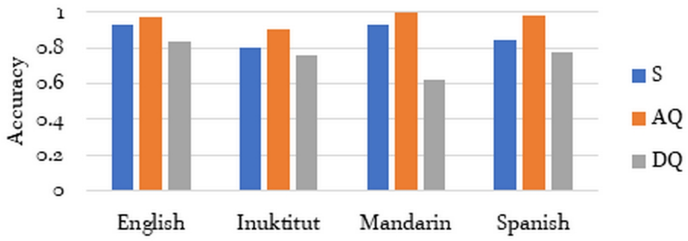

Figure 3: Bar chart displaying accuracy results in the SI task, by language.

The accuracy results of the SI task are displayed in Figure 3. They were higher overall than in the IO task, although some difficulty was experienced with the DQ context. Controls had the highest overall accuracy rate (91%), followed by the Spanish (87%), Mandarin (84%), and Inuktitut speakers (82%).

Our model comparison revealed that Model 3 was clearly the best fit, indicating an interaction between language and sentence type (M0: $AIC = 1098$; $AIC\omega = 0$; M1: $AIC = 1096$; $AIC\omega = 0$; M2: $AIC = 1053$; $AIC\omega = .03$; M3: $AIC = 1046$; $AIC\omega = .97$). We then performed a post-hoc analysis to examine the extent to which different sentence types influenced accuracy, for each language group (Table 3). The results revealed a higher accuracy for AQs as opposed to DQs for English, Inuktitut, and Spanish speakers. The Spanish group also had greater difficulty with the Ss than the AQs. Regarding the Mandarin group, Ss were more difficult than DQs. Moreover, they achieved 100% accuracy in the AQ context. Consequently, the Mandarin results comparing AQs with Ss and DQs are unreliable, due to quasi-complete separation of the data, which lead to inflated estimates, standard errors, and incorrect p-values.[44] Nevertheless, we can see in Figure 3 that the DQ context was clearly more difficult than the AQ context, with an accuracy of only 62%, compared to the 100% for AQs, and 93% for Ss.

[44] The quasi-complete separation is caused by 100% of the AQs being answered correctly. Consequently, the AQ context perfectly predicts a target response, which is problematic in a binomial logistic regression. While the quasi-complete separation prevents a reliable analysis of the S-AQ and AQ-DQ contrasts, it does not affect the results of the overall model. A variety of solutions have been proposed to address quasi-complete separation, with a penalized maximum likelihood estimation or a Bayesian estimation being the most common. Nevertheless, neither of these were adequate for our data. However, there can be no doubt in the data that the DQs were more difficult than the AQs for the Mandarin speakers, and the results strongly suggest that Ss were as well. This is further supported if we analyze data in which a single instance of a non-target response for the AQs were observed, eliminating the quasi-complete separation, and demonstrating that Ss and DQs were indeed more difficult (Ss: $\beta = -2.44$; $SE = 0.39$; $z = -6.24$; $p < 0.001$; DQs: $\beta = 4.86$; $SE = 1.01$; $z = 4.80$; $p < 0.001$). Levshina, *How to do Linguistics with R* (2015); Allison, "Convergence Failures in Logistic Regression" (2008).

Tabela 3: Post hoc pairwise comparisons of sentence types, grouped by language, for the SI task.

Language	Contrast	Est.	SE	z	p	
English	S-AQ	3.09	1.40	2.21	0.070	
	S-DQ	−0.18	0.82	−0.22	0.974	
	AQ-DQ	3.27	1.23	2.66	0.022	*
Spanish	S-AQ	3.99	1.29	3.09	0.006	*
	S-DQ	0.07	0.71	0.09	0.995	
	AQ-DQ	3.92	1.14	3.44	0.002	*
Mandarin	S-AQ	19.09	43.16	0.44	0.898	
	S-DQ	−2.17	0.73	−2.98	0.008	*
	AQ-DQ	21.27	43.17	0.49	0.875	
Inuktitut	S-AQ	2.92	1.33	2.20	0.072	
	S-DQ	0.17	0.76	0.22	0.973	
	AQ-DQ	2.75	1.16	2.37	0.047	*

Recall that, in the SI task, segmental (and therefore lexical) information was present. If learners focus more on the syntax as opposed to intonation, we might expect them to misidentify DQs as Ss, while correctly identifying AQs as questions. The latter was supported by the results, as all groups were at or near ceiling for the AQ context. We also examined the types of errors produced in the DQ and S contexts (Table 4), to better understand the sources of difficulty. As previously mentioned, DQs are sometimes identified as exclamations;[45] thus, if learners identified the utterances as exclamations, it would be difficult to determine whether it was because they did not focus on the intonation, or because they misinterpreted question intonation for exclamation intonation. However, if the DQs were identified as Ss, it would suggest that the intonation was secondary to the syntax. We performed a conditional inference tree analysis to examine the effect of sentence type and language on the error counts, and, as with the results in the IO task, sentence type had the strongest effect, splitting the data into AQs/DQs and Ss ($p < 0.001$). Within the AQs/DQs, Mandarin was grouped apart from the other three languages ($p < 0.001$). Table 4 indicates that this was due to the fact that, while all learner groups identified DQs to some degree as Ss, the Mandarin group did so more frequently (more than 50% of the time). Given that the Mandarin group had no difficulty using prosody to discriminate questions from Ss in the IO task, the results of the SI task suggest that, for the Mandarin speakers, syntax is a primary/stronger cue to sentence type.

[45] Warren, *Uptalk* (2016).

Tabela 4: Error counts by context and language in the SI task.

	Context					
	AQ		DQ		S	
	Excl.	Ss	Excl.	Ss	Excl.	Q
English	3	1	22	3	11	0
Inuktitut	8	4	23	8	23	3
Mandarin	0	0	24	33	10	0
Spanish	1	2	25	9	22	1

Contextualized Task (C)

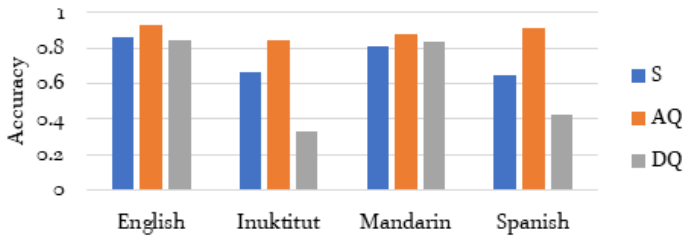

Figure 4: Bar chart displaying accuracy results in the C task, by language.

The C task had the lowest overall accuracy rates and, as predicted, was particularly difficult for the Inuktitut (62% accuracy) and Spanish groups (66%). This contrasts with the much higher accuracy rates achieved by the Mandarin (84%) and control groups (88%). In Figure 4, we can see that the DQ context in particular was highly variable.

The interaction model (Model 3) was again the best fit (M0: $AIC = 1087$; $AIC\omega = 0$; M1: $AIC = 1087$; $AIC\omega = 0$; M2: $AIC = 1005$; $AIC\omega = 0.01$; M3: $AIC = 990$; $AIC\omega = 0.99$), demonstrating that accuracy rates for each language group differed across sentence type. Due to convergence errors, only a random intercept for participant (and not a random slope) was included. A post-hoc analysis was performed to compare differences in sentence types by language (Table 5). The results only revealed differences for the Spanish and Inuktitut groups, who demonstrated identical trends. They had more difficulty correctly identifying AQs compared to Ss, AQs compared to DQs, and Ss compared to DQs.

Tabela 5: Post hoc pairwise comparisons of sentence types, grouped by language, for the C task.

Language	Contrast	Est.	SE	z	p	
English	S-AQ	0.81	0.53	1.54	0.273	
	S-DQ	−0.19	0.43	−0.43	0.904	
	AQ-DQ	0.99	0.51	1.93	0.130	
Spanish	S-AQ	1.84	0.44	4.16	< 0.001	*
	S-DQ	−1.00	0.32	−3.09	0.006	*
	AQ-DQ	2.83	0.44	6.39	< 0.001	*
Mandarin	S-AQ	0.55	0.43	1.28	0.406	
	S-DQ	0.16	0.40	0.41	0.912	
	AQ-DQ	0.39	0.44	0.88	0.652	
Inuktitut	S-AQ	1.10	0.41	2.68	0.020	*
	S-DQ	−1.55	0.36	−4.27	0.000	*
	AQ-DQ	2.66	0.42	6.28	< 0.001	*

Given the low accuracy rates in the DQ context, we were interested in analyzing the source of the errors. To do so, we performed a conditional inference tree analysis that examined the effect of sentence type and language on the error counts (Table 6). The analysis revealed that sentence type again had the strongest effect;

however, contrary to the results of the IO and SI tasks, the results were grouped into AQs/Ss and DQs ($p < 0.001$). Language had no effect; nevertheless, the data revealed some clear trends. The Inuktitut speakers tended to misidentify DQs as AQs 59% of the time, and only correctly identified them as DQs 33% of the time. Similarly, Spanish speakers incorrectly identified DQs as AQs 49% of the time, while correctly identifying them as DQs 42% of the time. Note that for both groups of speakers, Ss were rarely selected, indicating that the learners knew a question was required for the context; they just did not select the correct question type. This suggests that these learners had not learned the pragmatic distinction between the two question types. The results also revealed that Inuktitut learners identified Ss as DQs more frequently than the other language groups (18% of the time).

Tabela 6: Error counts by context and language in the C task.

	Context					
	AQ		DQ		S	
	DQs	Ss	AQs	Ss	AQ	DQ
English	6	0	13	1	8	4
Inuktitut	7	5	46	6	12	14
Mandarin	10	1	13	2	9	8
Spanish	7	1	44	8	21	11

Task Comparison

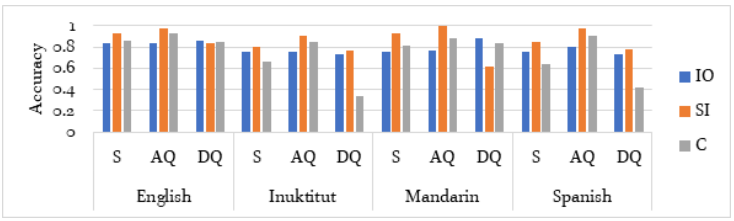

Figure 5: Result across tasks for each language group.

We finish this section by comparing the accuracy rates obtained across tasks, by language group (see Figure 5). The highest overall accuracy rate was obtained by the control group at 88% (IO task=84%; SI task=91%; C=88%), followed by the Mandarin group at 83% (IO task=80%; SI task=85%; C=84%), the Spanish group at 76% (IO task=76%; SI task=87%; C=66%), and finally, the Inuktitut at 73% (IO task=75%; SI task=82%; C=62%).

We ran another set of mixed effects binomial logistic regressions in order to examine whether differences were observed across tasks, for each language group. The same four models as in the previous tasks were run, except we replaced sentence type with task. For this analysis as well, only a random intercept of participant was included, due to convergence areas when including a random slope. Model 3 (interaction between language and task) was the best model (M0: $AIC = 4181$; $AIC\omega = 0$; M1: $AIC = 4172$; $AIC\omega = 0$; M2: $AIC = 4111$; $AIC\omega = 0$; M3: $AIC = 4092$; $AIC\omega = 1$), indicating that task differences varied by language. A post-hoc analysis examining these differences (Table 7) revealed that the control group achieved a higher accuracy in the SI as opposed to the IO task. No differences across tasks for the Mandarin group were observed. For both the Inuktitut and Spanish groups, the IO task was more difficult than the SI task, and the C task was the most difficult overall.

Tabela 7: Post hoc pairwise comparisons of tasks, grouped by language.

Language	Contrast	Est.	SE	z	p	
English	IO-SI	−0.70	0.21	−3.30	0.003	*
	IO-C	0.36	0.23	1.56	0.262	
	SI-C	−0.34	0.25	−1.34	0.373	
Spanish	IO-SI	−0.73	0.18	−4.03	< 0.001	*
	IO-C	−0.55	0.18	−3.16	0.005	*
	SI-C	−1.28	0.19	−6.61	< 0.001	*
Mandarin	IO-SI	−0.35	0.18	−1.92	0.135	
	IO-C	0.26	0.21	1.27	0.414	
	SI-C	−0.08	0.22	−0.39	0.921	
Inuktitut	IO-SI	−0.51	0.19	−2.75	0.016	*
	IO-C	−0.68	0.19	−3.61	0.001	*
	SI-C	−1.19	0.20	−5.95	< 0.001	*

Discussion

Summary of results

RQ1: Can L2 learners distinguish between statements and questions using intonation only?

The results revealed that learners were overall relatively successful when identifying sentence types using only intonation contours, and were no better with one sentence type over another. Crucially, and against our initial prediction, these findings demonstrate that all learner groups, including the Inuktitut speakers, could successfully distinguish between statements and questions, when only presented with the intonation contour, regardless of the use of intonation in their native language. These findings are consistent with Grabe *et al.*'s[46] and Gussenhoven and Chen's[47] claims that there is a universal bias toward the perception of rising contours as questions. Nevertheless, our error analysis revealed that the Inuktitut speakers identified Ss as questions 12% of the time, compared to only 3-5% observed in the other groups. This suggests that the Inuktitut group was less sensitive to pitch to some degree, and was consistent with our CLI-based prediction that Inuktitut learners would be more likely to identify Ss as questions.

[46] Grabe et al., "Perception of English Intonation by English, Spanish, and Chinese Listeners" (2003).

[47] Gussenhoven and Chen, "Universal and language-specific effects in the perception of question intonation" (2000).

RQ2: Can L2 learners distinguish between statements and questions when sentences are produced in isolation? Do they use syntax or intonation?

The accuracy rates for the SI task were the highest overall. This is unsurprising in the AQ and S contexts. Nevertheless, we found that, despite the higher overall accuracy rates, the DQ context was still challenging for all speakers, with an overall mean error rate of 25%. The results, though, were not consistent with the CLI-based predictions. The DQ context was especially difficult for the Mandarin speakers, who identified DQs as Ss 22% of the time (58% of DQ errors), compared to 6% of the time by the other two learner groups, and 3% by the controls. In Mandarin, Y/N questions can have a DQ structure, with a rising intonation. Thus, a transfer of this structure would predict little difficulty with English DQs. The results from the IO task revealed that Mandarin speakers were able to identify questions using only prosodic information. Therefore, the difficulty experienced in the DQ context in the SI task was most likely due to giving greater weight to the syntactic cues. While the Spanish and Inuktitut groups also had relatively high error rates in the DQ context (23% and 24%, respectively), they tended to identify the DQs as exclamations (74% of all errors, for both groups). This was counter our prediction for Inuktitut speak-

ers, who were expected to identify DQs as Ss.

RQ3: Can L2 learners comprehend in which context each sentence type is appropriate? Are differences observed between AQs and DQs?

Larger differences between groups were expected in the contextualized task, given that it requires knowledge and use of prosody, syntax, and pragmatics. This prediction was partially confirmed for the Inuktitut and Spanish groups; the contextualized task was particularly difficult for these groups (38% and 34% error rates, respectively), as they tended to overgeneralize AQs to DQ contexts. While this pattern was expected for the Inuktitut speakers, the Spanish speakers were expected to overgeneralize DQs to the AQ context, which was not the case. In contrast to the Inuktitut and Spanish learners, the Mandarin learners had a much lower overall error rate (16%). This was primarily due to a higher accuracy rate in the DQ context (83% versus 58% for Spanish and 67% for Inuktitut), which is consistent with our prediction that Mandarin learners would transfer their pragmatic knowledge of Mandarin questions. These results strongly suggest that the similar pragmatic distinction in Mandarin Y/N questions facilitated the acquisition of the English AQ-DQ contrast. The C task also revealed that the Inuktitut speakers identified Ss as DQs more frequently than the other learner groups. This finding is consistent with the result from the IO task, and suggests that Inuktitut learners pay less attention to intonational contours.

RQ4: Does CLI play a greater role with increasing levels of contextual information?

We predicted that CLI would increase with increasing access to context, based on previous work.[48] Our results arguably support this, at least when comparing the SI to the C task. Mandarin speakers improved with increasing context, and this improvement was likely the result of positive CLI from their L1. Spanish and Inuktitut speakers performed worse in the C task, mainly due to their difficulty with the AQ-DQ distinction, which is not marked syntactically in their L1. Moreover, the Inuktitut speakers showed a tendency to identify Ss as DQs, which was expected based on their L1. Given that they did not have difficulty with Ss in the SI task, it suggests that the additional task complexity resulted in more L1-based transfer.

[48] Grabe et al., "Perception of English Intonation by English, Spanish, and Chinese Listeners" (2003); Ortega-Llebaria and Colantoni, "L2 English intonation" (2014).

Crosslinguistic influence

The primary goal of the present paper was to determine the role of CLI in the L2 acquisition of English sentence types, with a focus on prosody, and to understand whether CLI increases with the

complexity of the task. Speakers from three typologically different L1s participated in our study, and thus, CLI was expected to play an important role. In particular, the Inuktitut speakers were expected to be the least sensitive to prosody, in contrast to the Spanish speakers, who were expected to focus primarily on prosody instead of syntax. Indeed, CLI did account for some of the patterns observed in the three groups. Overall, the L1 Mandarin group was the most successful in both the perception and comprehension of sentence types. This could be attributed to both a positive transfer of their ability to perceive tonal differences (IO task) and of their syntactic/pragmatic distinction of sentence types, which gave them an advantage in the C task. Recall from our discussion on Mandarin sentence types that there is a prosodically marked pragmatic distinction between broad focus questions, and echo questions or questions expressing surprise/incredulity. This pragmatic distinction is similar to the AQ/DQ distinction in English. The fact that the Mandarin speakers did not experience difficulty with the AQ/DQ contrast in the C task, whereas the Inuktitut and Spanish group did, suggests that the Mandarin learners transferred their pragmatic knowledge to English, which would have facilitated their acquisition of the AQ/DQ distinction. Instances of negative CLI may also have been observed in the Mandarin group, particularly in the SI task in which the Mandarin speakers tended to confuse Ss and DQs. The Mandarin speakers appeared to focus on the syntactic cues (when no context was present) which, at least in the case of AQs, are highly salient. This was potentially due to the fact that, in Mandarin, questions are not always marked by a final rise[49] and the intonation contour is partially dependent on the utterance's final tone. The results suggest that for Mandarin speakers, the syntax is a stronger cue than the prosody; nevertheless, the context plays an important role for them. Why might this be? The fact that context is also important in Mandarin for determining question type and structure may make them more likely to focus on the context when learning an L2 than speakers of other languages, such as Spanish and Inuktitut, whose Y/N questions do not vary to the same extent with context. Another possible explanation could be that the L2 Mandarin speakers were more advanced than the other groups, and only for this reason had acquired the contextual distinction. However, our participant profiles reveal that they had less experience with English than the other two groups of speakers, and less time living in an English dominant context. Proficiency is therefore an unlikely explanation.

[49] Liu and Rodriguez, "Categorical perception of intonation contrasts" (2012).

The L1 Spanish group presented an unexpected behavioural pattern, particularly in the C task. A positive CLI effect would have predicted that Spanish speakers would have a bias for DQs rather than for AQs, since Y/N questions in Spanish do not require inversion. However, the opposite was observed, which could be an

instance of overgeneralization.[50] We can speculate that L1 Spanish speakers initially learn that Y/N questions in English require inversion. They then generalize this pattern to all contexts, which is likely influenced by the more frequent use in the input of the default AQs, as opposed to the DQs. An alternative explanation is that participants had not had enough input to uncover in which contexts non-inverted questions occur in English. However, this seems less likely given that the Spanish learners had been residing in Canada for 6-14 years. This is an interesting scenario from a learning perspective, since we have word-order variation in the L1, which does not systematically associate with a given pragmatic context that maps onto two different pragmatic contexts in the L2.

[50] Odlin, "Cross-linguistic influence" (2003).

Inuktitut speakers behaved similar to Spanish speakers; thus, against our predictions, they were not always outperformed by the other learner groups, particularly in the IO and SI tasks. One possible explanation for their relative success in these two tasks is that they had had significant exposure to English. Although the main language spoken while these participants lived in Northern Canada was Inuktitut, they were educated in English, and English was prevalent in the area. The additional exposure of these speakers to English, from a young age, may have influenced their perceptual acuity with respect to prosody. Nevertheless, the fact that Inuktitut speakers more frequently identified Ss as questions, when compared to the other groups (in the IO and C tasks), suggests that they were somewhat less sensitive to prosodic contours and their use for identifying sentence types. Indeed, these errors were largely predicted based on CLI effects. The Inuktitut group also had a high error rate in the C task. Like the Spanish participants, they tended to select AQs more frequently than DQs in DQ prompting contexts. This was expected based on our prediction that Inuktitut speakers would focus on syntactic cues as opposed to prosodic cues. Overall, when comparing the results of the Inuktitut speakers to those of the Spanish and Mandarin speakers, they suggest that speakers of a language with limited use of intonation are relatively less sensitive to the intonation of an L2 than speakers of languages that use intonation extensively. Future work should investigate this observation in more detail, by examining if it is also observed for different prosodic structures.

Limitations and future avenues

The primary limitation of the present study is the heterogeneity of our learner groups. Ideally, our learner groups would have very similar characteristics and would therefore be more comparable. Nevertheless, given our interest in exploring a variety of L1s, it is challenging to recruit learners who have similar profiles. Despite the different characteristics, we expect to be able to more

effectively determine the role of CLI in the acquisition of prosody when we compare these results to production data obtained from the same learners. Moreover, an analysis of the speaker variables against the perception and production results should allow us to uncover to what extent linguistic experience plays a role in the perception and production of L2 intonation.

References

Allison, Paul D. (2008). "Convergence Failures in Logistic Regression". *SAS Conference Proceedings* (16th March 2008–19th March 2008). San Antonio (TX).

Arnhold, Anja (2014). "Prosodic Structure and Focus Realization in West Greenlandic". In: *Prosodic Typology II. The Phonology of Intonation and Phrasing*. Edited by Sun-Ah Jun. Oxford: Oxford University Press, pp. 216–251. ISBN: 9780199567300.

Bartels, Christine (2014). *The Intonation of English Statements and Questions. A Compositional Interpretation*. New York: Routledge. ISBN: 9781135685027.

Best, Catherine T. and Michael D. Tyler (2007). "Nonnative and second-language speech perception. Commonalities and complementarities". In: *Language Experience in Second Language Speech Learning. In Honor of James Emil Flege*. Edited by Ocke-Schwen Bohn and Murray J. Munro. Paper 360. Amsterdam: John Benjamins, pp. 13–34. ISBN: 9789027219732.

Cruz-Ferreira, Madalena (1983). "Non-native Comprehension of Intonation Patterns in Portuguese and in English". PhD thesis. University of Manchester.

Dorais, Louis-Jacques (2010). *The Language of the Inuit. Syntax, Semantics, and Society in the Arctic*. Montreal: McGill-Queen's University Press. ISBN: 9780773536463.

Dupoux, Emmanuel, Núria Sebastián-Gallés, Eduardo Navarrete and Sharon Peperkamp (2008). "Persistent stress 'deafness': The case of French learners of Spanish". *Cognition* 106.2, pp. 682–706. DOI: 10.1016/j.cognition.2007.04.001.

Eckman, Fred R. (2008). "Typological markedness and second language phonology". In: *Phonology and Second Language Acquisition*. Edited by Jette G. Hansen Edwards and Mary L. Zampini. Amsterdam: John Benjamins, pp. 95–115. ISBN: 9789027291394.

Estebas-Vilaplana, Eva and Pilar Prieto (2010). "Castillian Spanish intonation". In: *Transcription of Intonation of the Spanish Language*. Edited by Pilar Prieto and Paolo Roseano. München: LINCOM, pp. 17–48. ISBN: 9783862901845.

Flege, James Emil (1995). "Second language speech learning. Theory, findings and problems". In: *Speech perception and linguistic experience. Issues in cross-language research*. Edited by Winifred Strange. Timonium: York Press, pp. 233–277. ISBN: 9780912752365.

Grabe, Esther, Burton S. Rosner, José E. García-Albea and Xiaolin Zhou (2003). "Perception of English Intonation by English, Spanish, and Chinese Listeners". *Language and Speech* 46.4. PMID: 15198113, pp. 375–401. DOI: 10.1177/00238309030460040201.

Gunlogson, Christine (2001). "True to Form. Rising and Falling Declaratives as Questions in English". PhD thesis. Santa Cruz: University of California.

Gussenhoven, Carlos and Aoju Chen (2000). "Universal and language-specific effects in the perception of question intonation". *Proceedings ICSLP 2000*. Sixth International Conference on Spoken Language Processing (16th October 2000–20th October 2000). Edited by Zhongguo sheng xue xue hui and Zhongguo ke xue yuan. Volume 2. International Speech Communication Association. Beijing, pp. 91–94. ISBN: 9787801501141.

Haan, Judith (2002). "Speaking of questions. An exploration of Dutch question intonation". PhD thesis. Utrecht: LOT Graduate School of Linguistics, University of Utrecht.
Hedberg, Nancy and Juan Manuel Sosa (2011). "A unified account of the meaning of English questions with non-canonical intonation". International Seminar on Prosodic Interfaces. Jawaharlal Nehru University. URL: http://www.sfu.ca/~hedberg/Hedberg_Sosa_ISPI_paper_CORRECTED.pdf.
Hertel, Tammy Jandrey (2003). "Lexical and discourse factors in the second language acquisition of Spanish word order". *Second Language Research* 19.4, pp. 273–304. DOI: 10.1191/0267658303sr224oa.
Hualde, José Ignácio and Pilar Prieto (2015). "Intonational variation in Spanish. European and American varieties". In: *Intonation in Romance*. Edited by Sónia Frota and Pilar Prieto. Oxford: Oxford University Press, pp. 350–391. ISBN: 9780199685332.
Jun, Sun-Ah (2014). "Prosodic typology. By prominence type, word prosody, and macro-rhythm". In: *Prosodic Typology II. The Phonology of Intonation and Phrasing*. Edited by Sun-Ah Jun. Oxford: Oxford University Press, pp. 520–540. ISBN: 9780199567300.
Jun, Sun-Ah, editor (2014). *Prosodic Typology II. The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press. ISBN: 9780199567300.
Lee, Ok Jo (2005). "The prosody of questions in Beijing Mandarin". PhD thesis. Ohio State University.
Lepage, Jean-François and Stéphanie Langlois (2019). *Evolution of the language situation in Nunavut, 2001 to 2016*. In collaboration with Martin Turcotte. Ontario: Statistics Canada. ISBN: 9780660310725. URL: https://www150.statcan.gc.ca/n1/en/pub/89-657-x/89-657-x2019010-eng.pdf.
Levshina, Natalia (2015). *How to do Linguistics with R. Data exploration and statistical analysis*. Amsterdam: John Benjamins. ISBN: 9789027268457.
Liu, Chang and Amanda Rodriguez (2012). "Categorical perception of intonation contrasts. Effects of listeners' language background". *The Journal of the Acoustical Society of America* 131.6, EL427–EL433. DOI: 10.1121/1.4710836.
Mennen, Ineke (2004). "Bi-directional interference in the intonation of Dutch speakers of Greek". *Journal of Phonetics* 32.4, pp. 543–563. DOI: 10.1016/j.wocn.2004.02.002.
Neufeld, Gerald G. (1988). "Phonological Asymmetry in Second-Language Learning and Performance". *Language Learning* 38.4, pp. 531–559. DOI: 10.1111/j.1467-1770.1988.tb00166.x.
Odlin, Terence (2003). "Cross-linguistic influence". In: *The Handbook of Second Language Acquisition*. Edited by Catherine J. Doughty and Michael H. Long. Oxford: Blackwell, pp. 436–486. ISBN: 9781405151887.
Ortega-Llebaria, Marta and Laura Colantoni (2014). "L2 English intonation. Relations between Form-Meaning Associations, Access to Meaning, and L1 Transfer". *Studies in Second Language Acquisition* 36.2, pp. 331–353. DOI: 10.2307/26328943.
Patience, Matthew et al. (2018). "Initial Pitch Cues in English Sentence Types". *Proceedings of Speech Prosody 2018*. 9th International Conference on Speech Prosody (Faculty of Modern Languages and Literatures of Adam Mickiewicz University). Edited by Grażyna Demenko, Jolanta Bachan and Daniel Śledziński. International Speech Communication Association. Poznań, pp. 463–467. ISBN: 9788394760960. DOI: 10.21437/SpeechProsody.2018-94.
Peterson, Tyler (2016). "Mirativity as surprise. Evidentiality, information, and deixis". *Journal of Psycholinguistic Research* 45.6, pp. 1327–1357. DOI: 10.1007/s10936-015-9408-9.
Pozzan, Lucia and Erin Quirk (2014). "Second language acquisition of English questions. An elicited production study". *Applied Psycholinguistics* 35.6, pp. 1055–1086. DOI: 10.1017/S0142716412000690.
Rialland, Annie (2009). "The African lax question prosody. Its realisation and geographical distribution". *Lingua* 119.6, pp. 928–949. DOI: 10.1016/j.lingua.2007.09.014.
Saindon, Mathieu R., Sandra E. Trehub, E. Glenn Schellenberg and Pascal H. H. M. van Lieshout (2017). "When is a Question a Question for Children and Adults?" *Language Learning and Development* 13.3, pp. 274–285. DOI: 10.1080/15475441.2016.1252681.
Shokeir, Vanessa (2009). "Intonation in Inuktitut". Master's thesis. University of Toronto.

Smith, Lawrence R. (1977). *Some grammatical aspects of Labrador Inuttut* (*Eskimo*). *A survey of the inflectional paradigms of nouns and verbs*. University of Ottawa Press. ISBN: 9781772822007.

Sosa, Juan Manuel (1999). *La entonación del español. Su estructura fónica, variabilidad y dialectología*. Madrid: Cátedra. ISBN: 9788437617435.

Spalding, A. E. (1969). *SALLIQ. An Eskimo Grammar*. Government of Canada: Department of Indian Affairs and Northern Development.

Truckenbrodt, Hubert (2009). "On rises and falls in interrogatives". *Actes de la conférence Interface Discours & Prosodie* (*IDP*) *2009*. Interfaces Discours & Prosodie (9th September 2009–11th September 2009). Edited by Hi-Yon Yoo and Elisabeth Delais-Roussarie. Paris, pp. 33–46. URL: http://makino.linguist.jussieu.fr/idp09/actes_fr.html.

Wagenmakers, Eric-Jan and Simon Farrell (2004). "AIC model selection using Akaike weights". *Psychonomic Bulletin & Review* 11, pp. 192–196. DOI: 10.3758/BF03206482.

Warren, Paul (2016). *Uptalk. The Phenomenon of Rising Intonation*. Cambridge: Cambridge University Press. ISBN: 9781107123854.

Zubizarreta, Maria Luisa (1998). *Prosody, Focus, and Word Order*. Cambridge (MA): MIT Press. ISBN: 9780262240413.

Zubizarreta, Maria Luisa and Emily Nava (2011). "Encoding discourse-based meaning. Prosody vs. syntax. Implications for second language acquisition". *Lingua* 121.4. Acquisition at the Linguistic Interfaces, pp. 652–669. DOI: 10.1016/j.lingua.2010.06.013.