Using Audio and Audio-Visual Feedback to Enhance Learners' Perception and Production of Tag Questions

Souheila Moussalli

A Thesis

In

The Department of

Education

Presented in Partial Fulfillment of the Requirements

for the Degree of Master of Arts (Applied Linguistics) at

Concordia University

Montreal, Quebec, Canada

March 2013

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: Souheila Moussalli

Entitled: Using Audio and Audio-Visual Feedback to Enhance Learners' Perception and Production of Tag Questions

and submitted in partial fulfillment of the requirements for the degree of

Master of Arts (Applied Linguistics)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

_____Marlise Horst_____Chair
_____Walcir Cardoso _____ Examiner
_____Pavel Trofimovich _____ Examiner
_____ Sara Kennedy _____Supervisor

Approved by _____

Chair of Department or Graduate Program Director

_____

Dean of Faculty

Date _____April 10, 2013_____

Abstract

Using Audio and Audio-Visual Feedback to Enhance Learners' Perception and Production of Tag Questions

Souheila Moussalli

Intonation is fundamental to human speech in conveying meaning. Teachers and researchers have raised the concern that the intonation of a second language is difficult both to teach and to learn in the language classroom. However, due to particular characteristics of technology-based instruction, there is an increasing tendency to teach and learn second language intonation through the use of technology and software. For the above reasons, an open-access tool which incorporates computer-based audio-visual feedback was designed and piloted to help Chinese learners of English improve their perception and production of rising and falling tag questions. Six international students (native speakers of Mandarin) participated in the study. The participants were pre- and post tested on their perception (in context and in isolation) and production (in context) of question tags. The participants also underwent training lasting approximately two weeks. Each week, the participants were asked to complete two blocks of the training materials, for about 15-20 minutes each time. Individual semi-structured interviews took place after the training with all participants. Based on observations, pre- and post-test scores, and the interviews conducted, the results revealed that the participants have become more aware of their perception of tag questions. Also, the participants showed some improvement in their production of tag questions. Participants also stated that the two applications used: WASP, a software that allows the recording, display and analysis of speech and Online Audio Recording, a plug-in for the e-learning platform Moodle were beneficial, interesting to use and helpful for their learning.

Acknowledgements

My sincerest thanks and appreciation go out to my supervisor, Dr. Sara Kennedy who was abundantly helpful and offered invaluable assistance, support and guidance during the course of writing this thesis. I would also like to thank my committee members, Dr. Walcir Cardoso and Dr. Pavel Trofimovich for their valuable insights and assistance. To my husband, thank you for the constant support, understanding, and endless love when it was most needed. I definitely could not have completed this project without your encouragement. Last, but definitely not least, I would like to express my gratitude to my family members – my grandfather, parents and siblings – for their support, blessings and wishes for the successful completion of this project.

Contribution of Authors

As first author for the included manuscript, Souheila Moussalli was the major contributor to the study's conception, design, data collection, and write-up. The organization of the manuscript was jointly developed by Souheila Moussalli and Sara Kennedy. Souheila Moussalli wrote substantive drafts of the manuscript, while Sara Kennedy provided extensive feedback on organization and content, re-worded some passages, and suggested possible interpretations of the data. The majority of the manuscript's conception and content was developed by Souheila Moussalli. This contribution is reflected in her status as first author.

Table of Contents

List of Figures

List of Tables

Chapter 1

Considering the Use of Technology for Learning

Over the past few decades, technological devices, including computers, have become an integral part of our daily lives. According to Gupta (2010), technology today is within the reach of more people because it has become cheaper. Moreover, compared to the past, many more students are very comfortable using technology in many domains. All these changes have led to the integration of technology in educational settings. Gupta (2010) distinguishes between two uses of technology: *technology of education* and *technology in education*. Technology *of* education is considered to be a medium which helps instructors optimize delivery of the prepared learning materials such as video lectures and digitized lectures. However, technology *in* education is used to improve learning and meet learners' needs. She gives an example of science students who use graphics and animations to visualize elements that static images cannot replicate. In terms of language learning, according to Kern (2006), the use of technology can be thought of in terms of the metaphors of tutor, tool, and medium:

> In the tutor role, computers can provide instruction, feedback, and testing in grammar, vocabulary, writing, pronunciation, and other dimensions of language and culture learning... In the tool role, computers provide ready access to written, audio, and visual materials relevant to the language and culture being studied... In the medium role, technology provides sites for interpersonal communication, multimedia publication, distance learning, community participation, and identity formation (p. 191, 192).

Based on Kern's (2006) view, this study will focus on the use of technology as a tutor and a tool for learning the perception and production of English tag questions.

Chapelle (2001) states that there are a number of themes researchers and teachers should consider when working with technology in language learning. These themes are:

1. Reliability and learner fit: e.g. Is the use of the technology too easy or difficult for learners? How do learners' individual differences affect their performance?

2. Authenticity and generalizeability: how does the use of the technology reflect technology use in the non-research environment? Can research results be generalized to other contexts?

3.-Construct validity and operationalization of learning conditions: what theoretical constructs underlie use of technology? How does use of technology reflect those constructs?

4. Language learning potential and operationalization of learning conditions: what potential does use of technology have for language learning? How is that potential realized in its use in research?

5. Interactiveness and meaning focus: How does the use of technology engage learners' meaningful use of communicative abilities?

6. Positive impact: how do learners benefit or suffer from use of technology?

7. Practicality: How easy is it to find, modify, and use the technology in a non-research context?

Clearly, using technology in language learning is not the same as using technology effectively in language learning. In this chapter, I will look at the benefits of using technology for language learning and the ways in which technology has been used

effectively to help language learners with a challenging aspect of English: perceiving and producing intonation patterns.

Benefits of Technology Use in Educational Settings

The advancement in technology is seen by many language educators as a great opportunity that fosters and adds great value to their classroom teaching in a variety of ways (Felix, 2001). Educators are aware that the use of technological activities in their classrooms can motivate their students. Students have shown positive perceptions toward the use of technology in their language classrooms (Weinberg, Peters, & Sarma, 2008). A study by Osuna and Meskill (1998) revealed that the students who worked on a series of web-based activities in their Spanish class rated the use of technology and the activities very favourably. Another study by Taylor and Gitsaki (2004) also showed that the students benefited from using the web in their English language class and thus they considered it a rewarding tool for language learning. Felix (2001) stated that the web provided the students with time, flexibility, privacy and access to a wide variety of information. Zhao (2003) stated that "application of technologies can be effective in almost all areas of language education. Modern technology can help enhance the quality of input, authenticity of communication, and provide more relevant and useful feedback" (p. 21).

These findings suggest that, properly used, technology can be a valuable add-on to face-to-face language teaching. In many language classrooms, there is limited time for learners to practise their oral production and communication. Technology can be used to provide additional opportunities for oral production. In a study by Holland, Kaplan and Sabol (1999), students' production of Arabic utterances improved as a result of an

interactive Microworld program with automatic speech recognition (ASR). Learners gave instructions and commands in Arabic which had to be recognized by the ASR system. Similarly, a study by Bradlow, Pisoni, Akahane-Yamada, and Tohkura (1997, 1999) showed that the production of /l/ and /r/ consonants by L1 Japanese speakers improved as a result of the computer-based training provided, incorporating perception and production of many different English tokens with /l/ and /r/.

Other studies have also shown that not only L2 production but also L2 perception improves with the help of technology. For example, Wang and Munro (2004) showed that the Cantonese participants' perception of three vowel contrasts in English (/i/-/I/, as in beat vs. bit, /u/-/ʊ/, as in Luke vs. look, and /e/-/æ/, as in bet vs. bat) on which they received computer-based perception training, improved significantly.

Using Visual Feedback Technology to Learn L2 Pronunciation

In this section, studies that used visual feedback for L2 learning purposes are introduced. A study by Hirata (2010) revealed that English learners of Japanese who used the CSL-Pitch Program to look at pitch and duration of speech, a computer speech lab that analyzes speech, learned to perceive and produce novel Japanese words correctly using appropriate pitch and durational contrasts. Anderson-Hsieh (1992) (as cited in Chun, 1998) showed that with the help of a program called Visi-Pitch, participants were able to more easily reproduce the targeted suprasegmentals. The program Visi-Pitch allowed the international teaching assistants to look at an accurate visual representation of the English suprasegmentals in real time paired with their own recording (auditory feedback) during speech. These studies, among others (Hardison, 2004; Thomson, 2011), have shown that using technology to deliver visual feedback about suprasegmental

production can lead to improved pronunciation and perception that also generalizes to new contexts.

The previous studies have shown how technological devices and software for the purpose of perceiving and producing human speech can be used to facilitate language learning and how feedback can be used in more effective ways. However, technology can also be ineffective for L2 learners' development of aural/oral abilities. One example is provided by Thomson (2011); he explains how a computer-assisted pronunciation training application using visual feedback to train learners' pronunciation is confusing and unhelpful. The application has a visual pronunciation meter which does not tell the learners what to modify in their pronunciation but rather tells them how native-like their pronunciation is. Thomson adds that the effectiveness of applications with visual feedback can vary, depending on the aspect of pronunciation. For example, visual feedback which relates to the phonetic properties of the learner's recorded speech and which allows comparison to the phonetic properties of a native speaker model tends not to be effective for teaching individual sounds because learners need to be able to interpret technical measures of speech phonetics to use the feedback. Visual feedback can be more effective for teaching suprasegmentals such as intonation because, as Thomson explains, particular kinds of visual feedback can be easily interpreted by learners.

Challenges in Learning English Intonation

According to Kang, Rubin and Pickering (2010), "intonation is typically defined as the linguistically meaningful use of vocal pitch level and pitch movement in phrases"(p.556). In general, intonation is considered a complex linguistic system. However, this complex system varies from one language to another. Non-native speakers

often find it difficult to acquire and produce the intended intonation pattern in comparison to native speakers. According to Verdugo and Trillo (2005), non-native speakers tend to overuse a very limited variety of patterns in the foreign language. In the case of L2 English learners, for example, Kang, Rubin and Pickering (2010) explain that it is common for English language learners from Korea, Greece and Indonesia to have a preference for falling tones in their speech. Moreover, a study by Wennerstrom (1994) showed that Japanese, Thai, and Chinese speakers of English displayed low, falling tones at boundaries between related propositions where native listeners would anticipate rising or mid-level tones. Another famous example is from Gumperz (1982) about native speakers misinterpreting the pragmatic meaning of non-native speakers' intonation. In a British cafeteria, customers interpreted the intonation and manner of non-native cafeteria staff to be rude. This was because the staff would use falling intonation rather than rising intonation when asking customers if they wanted gravy.

For native speakers of tonal languages, learning English intonation is especially challenging. For example, in Mandarin Chinese, intonation is influenced by lexical tone: every syllable is assigned a particular tone which carries lexical meaning. Utterances in Mandarin do have general intonation contours which overlay the lexical tones, but

> "compared to an 'intonation language' like English, there are far fewer possible intonation patterns […] When Chinese ELLs [English language learners] begin learning English, they start out with a much more limited set of contours than an English NS, and therefore they have a large number of contours to acquire in English (Park, 2012, p. 25).

The use of tag questions in English exemplifies the challenges of learning English intonation. According to Parkes (1989), "A tag question is a device used to turn a statement into a question […] When the statement is positive, the tag question is expressed in the negative; when the statement is negative, the tag question is positive"(as cited in Knudsen, 2011, p. 10). Tag questions are formed by declarative statements followed by a brief yes-no question (e.g., *It's snowing, isn't it?*).

A tag question can either have a rising intonation or a falling intonation depending on the speakers' expectation of the answer. If the speaker does not know the answer to the question, the speaker would use a rising contour. If the speaker knows the answer to the question and is only trying to confirm the assumption, then the speaker would use a falling contour. Therefore, in English the production and interpretation of the pragmatic meaning of tag questions largely depends on the intonation pattern produced. Not all tag questions are formed as they are in English, however.

In Mandarin, tag questions are formed by adding the particles *shi ba* and *shi ma* to the declarative statement. The particle *ma* in Chinese is usually matched with a rising contour to mean that the speaker does not know the answer to the question. Excerpts (1) and (2) below are from Zhu and Wu (2011).

(1) He goes to school, doesn't he?

Ta1 shang4xue2 shi4ma

On the other hand, the particle *ba* in Chinese is usually matched with a falling contour to mean that the speaker knows the answer to the question and is only trying to confirm.

(2) He goes to school, doesn't he?

Ta1 shang4xue2 shi4ba

Based on the above studies, it is clear that the inappropriate perception or production of intonation contours by language learners may lead to misunderstanding and misinterpretation. Thus, I aim to explore how technology can be used to help learners improve their perception and production of English tag questions with the help of audio and audio-visual feedback.

Following this chapter is a manuscript describing a study which examined the production and perception of tag questions using a set of open-access tools which incorporate computer-based audio and audio-visual feedback. The study explored the use of two tools: WASP, a software that allows the recording, display and analysis of speech, and Online Audio Recording, a plug-in for the e-learning platform Moodle which allowed students to record audio clips. Using these two tools to work on Chinese learners' perception and production of tag questions, the results showed that the participants became more aware of different intonation contours and pragmatic meanings of tag questions. Furthermore, the results revealed that the tools used were easy and interesting for participants to use and helpful for their learning.

Chapter 2

Manuscript

The role of intonation is fundamental to human speech because it is central to conveying meaning. Every native speaker of a language can understand and use intonation to communicate his or her message; however, non-native speakers are sometimes unable to use intonation in ways that suit the intended message. According to Verdugo and Trillo (2005), native speakers communicate effectively by automatically selecting the appropriate prosodic patterns, lexical and contextual information. They do so because of their prior knowledge and experience in a particular context. On the other hand, non-native speakers may have less experience, and thus may have insufficient knowledge of the second language (L2), which means they use inappropriate intonation patterns in certain contexts. Spaai and Hermes (1993) explain that language learners often superimpose the prosodic features (features of rhythm, stress, and intonation) of their first language (L1) on the sounds of the language learned. This is because the suprasegmental features and patterns of the speakers' L1 are deeply rooted. Consequently, language learners are often unaware of differences in intonation between their first language and a second language, and of how different intonation patterns can affect meaning in the second language.

In order for language learners to learn these features it is important for them to perceive and produce them. Studies on the learning of second language phonology have not conclusively shown whether L2 speakers learn to perceive before they learn to produce aspects of L2 phonology, or vice versa. For example, Archibald found that Hungarian and Polish native speakers, whose languages have predictable stress patterns,

had better scores in the perception of stress for English verbs and nouns than in the production of the stress (1993, as cited in Altmann, 2006). Altmann's 2006 study revealed results quite the opposite with participants who were native speakers of languages with predictable stress patterns. Altmann found participants with relatively lower performance in the perception of stress for English Novel words showed very good performance in the production task. Moreover, participants with high perception scores produced words with the least native-like stress patterns. The studies above suggest that the relative order of learning to perceive and produce aspects of L2 phonology is still unclear. Further investigation of the relationship between perception and production of suprasegmentals requires that learners, either cross-sectionally or longitudinally, engage in both perception and production tasks.

Teachers and researchers have raised the concern that the intonation of a second language is difficult both to teach and to learn in the language classroom (Park, 2011). However, due to particular characteristics of technology-based instruction, there is an increasing tendency to teach and learn second language intonation through the use of technology and software. In this article, the design and feasibility of a tool which uses open-access tools to improve L2 English learners' perception and production of English intonation is described.

With respect to English, research has shown that non-native speakers (especially speakers of Chinese languages) have difficulties using intonation to signal meaning and structure. One famous example was described in Gumperz (1982) with native speakers misinterpreting the pragmatic meaning of non-native speakers' intonation. In a British cafeteria, customers interpreted the intonation and manner of cafeteria staff who were

non-native English speakers to be rude. This was because the staff would use falling intonation rather than rising intonation when asking customers if they wanted gravy. Stibbard (1996) explained that Cantonese learners of English found it difficult to use English intonation properly both in "physiological/phonetic terms, that is, in the physical problem of how to produce authentic-sounding intonation, and in phonological/semantic terms, that is, in deciding how to use intonation meaningfully in their speech"(introduction, para. 4). Litzenberg (2003) states that the "pitch in Chinese is distinctive at the word level and belongs to the lexical representation of a word. In English, on the other hand, pitch distinctions are generally made at the sentence level and are not lexically distinctive" (p. 2).

These differences in intonation systems have implications for how non-native speakers produce intonation in English. Rui (2007) found that English major undergraduate students in China often demonstrated non-native like intonation patterns in their spontaneous and oral reading speech when compared to British speakers (as cited in Park, 2012, p. 25). Wennerstrom (1994) found that a group of L2 English speakers of Asian background spoke in a monotone in English because in tonal languages "pitch functions to distinguish lexical rather than discourse meaning" (p. 417). A study by Pickering (2001) revealed that in their instructional discourse, native Chinese teaching assistants used fewer rising tones than did the native speakers, which created misunderstanding for the students of the Chinese teaching assistants.

In addition, research has shown that the intonation contours of questions differ between native English speakers and non-native speakers. For example, a study by Wennerstrom (1994) showed that non-native speakers produced less contrast in pitch

(degree of highness or lowness of tone) in yes/no questions than native English speakers. Burleson (2007) showed that the pitch in tag questions differed between non-native speakers and native speakers. In the statement part of the tag question, non-native speakers tended to use higher pitch on more syllables in the statement, implicitly treating more words as new information. According to Wennerstrom, a high pitch accent is associated to 'individual words to indicate that these are new or contrastive in the discourse, to be added to the speaker and hearer's 'mutual belief space' (p. 405). This use of pitch accent by the non-native speakers resulted in an unusual pattern of cues for a listener trying to identify given and new information.

Nowadays, because of advances in technology and much more widespread access to computers and Internet connections, technological devices have become normalized in the sense that we have grown to rely on them as tools for communication, teaching and learning. In terms of teaching and learning, Thorne and Smith (2011) have identified at least three potential uses of technology in language learning:

> [T]o provide access to input and rehearsal (recordings, tutorials, and drills), to amplify possibilities for personal expression (text and media processing), to extend existing and enable new opportunities for interpersonal communication (synchronous and asynchronous messaging)…(p. 268).

Possible benefits of using technology in these three ways need to be explored with particular uses of particular technologies in particular contexts. Chapelle (2001) notes that when conducting research on the use of technology in language learning, researchers should consider several themes when evaluating particular technologies used in particular ways. The themes are the following:

1. Reliability and learner fit: e.g. Is the use of the technology too easy or difficult for learners? How do learners' individual differences affect their performance?
2. Authenticity and generalizeability: how does the use of the technology reflect technology use in the non-research environment? Can research results be generalized to other contexts?
3. Construct validity and operationalization of learning conditions: what theoretical constructs underlie use of technology? How does use of technology reflect those constructs?
4. Language learning potential and operationalization of learning conditions: what potential does use of technology have for language learning? How is that potential realized in its use in research?
5. Interactiveness and meaning focus: How does the use of technology engage learners' meaningful use of communicative abilities?
6. Positive impact: how do learners benefit or suffer from use of technology?
7. Practicality: How easy is it to find, modify, and use the technology in a non-research context?

The use of technology to learn the spoken form of a second language has increased exponentially. One advantage of using technology is its capability of providing "relevant and useful feedback" (Zhao, 2003, p. 21). According to Thomson (2011), there are a number of applications that provide visual feedback to language learners which is simple for them to interpret, helping learners to develop certain aspects of second language pronunciation. He adds that in general, visual displays or feedback which shows

contours of pitch, relative duration of segments and amplitude of segments, tend to be very effective for the teaching of suprasegmentals such as intonation, rhythm and stress.

Early on, de Bot (1983) showed that providing students with audio-visual feedback was effective in enhancing the students' production of intonation compared to those students provided with audio feedback alone. Moreover, a study by Hardison (2004) showed that learners receiving implicit training (no metalinguistic information) in intonation with real-time visual pitch display showed improved production at the suprasegmental level. Although the participants in the study did not receive explicit instruction as part of their training, they still improved significantly from the implicit training.

A study by Taniguchi and Abberton (1999) revealed that a group of Japanese EFL learners who were exposed to interactive visual feedback greatly improved in their production of English intonation and their naturalness in English intonation, in comparison to the group of learners who did not receive any visual feedback. In addition to that, several other studies have revealed that the use of visual intonation-display systems have had positive effects on the learning of intonation for foreign-language learners (Hengstenberg, 1980, Lane & Buiten, 1966, as cited in Spaai & Hermes, 1993).

According to de Bot (1983), audiovisual feedback allows participants to make decisions based on what they see in front of them, depending on how similar the imitated contours are to the originals. The participants' degree of success in imitation is visualized through a picture which motivates them to either repeat an imitation or try a new one. In contrast, audio feedback may not provide the participants with enough information about

their success in imitating a sound. According to de Bot, this might discourage the participants receiving audio feedback from repeating or practicing the sounds.

Typically, when audiovisual feedback was provided in these studies, it wascombined with explicit (metalinguistic) instruction in those particular suprasegmental aspects. Researchers usually did not explore whether participants who received solely audiovisual feedback could learn the intonation pattern. However, in Hardison (2004) the participants did not receive explicit instruction on the perception or production of the intonation patterns but still improved significantly in their production at the suprasegmental level. The participants received information via two channels, the aural and the visual, which may have bypassed a need for additional explicit metalinguistic information to help participants to notice the relevant aspects of native models and their own productions.,

To summarize the above: the production and perception of intonation in English plays a crucial role in interpreting utterances for meaning; non-native speakers have demonstrated consistent problems producing and interpreting English intonation appropriately; finally, technology has been successfully used to implicitly train non-native speakers to perceive and produce particular features of intonation. However, it is not clear to what extent learners of L2 intonation may benefit from feedback from two channels (audiovisual) as opposed to one channel (audio only). It is also unclear to what extent participants *need* both explicit metalinguistic training and audio-visual feedback to improve their production at the suprasegmental level. Moreover, the tools used in the above studies required specialized knowledge of intonation systems and instrumental phonetics and/or were expensive and not freely available to the public. Thus, in the

current study we explore the use of accessible and user-friendly technological tools to provide practice opportunities with audio or audio-visual feedback, without explicit instruction, to help learners improve their perception and production of English intonation, with a particular focus on tag questions.

As Bolinger (1989) describes them "tag questions are hybrids, part statement (as a rule) and part question. The term is generally restricted to utterances that contain a statement which is then immediately questioned by repeating the auxiliary verb" (p. 115). An example is: *You're late, aren't you?* English tag questions, in general, can end either in a rising-falling or a rising intonation depending on what the speaker is expecting as an answer. Huang (1980) explains these two intonation patterns in the following manner: "[T]he speaker's intonation indicates how strong his or her presupposition is that the assumption - positive or negative - will be confirmed by the listener. If the speaker uses rising intonation, the expectation is weak. If he or she uses rising-falling intonation, the presupposition of confirmation is strong" (as cited in Celce-Murcia & Freeman, 1998, p. 262). Take, for example, the tag question: *You've lost your keys again, haven't you?* If the speaker uses a rising intonation at the end, it means that the speaker's expectation is weak and the speaker does not know the answer. The speaker is asking for one. When the speaker uses a falling intonation at the end, it means that the speaker's supposition of the hearer losing the key again is strong. The speaker is sure of that and is thus seeking a confirmation.

Using tag questions is often problematic for learners of English. Parmelee (n.d) stated that asking and answering tag questions is a consistent problem for many foreign learners of English, especially Asians. Parmelee observed that Thai and other Asian

learners usually answer a tag question solely with regard to the truth or falsity of the statement contained in the question, whereas native speakers respond to the truth or falsity of the tag following the statement. A study by Verdugo and Trillo (2005) on Spanish speakers revealed that Spanish speakers tended to over-generalize the use of rising tone instead of the falling tone in interpreting and expressing English tag questions. The Spanish speakers tended to use the rising tone in a tag question even when a falling tone is required to achieve the intended message (expressing certainty about the proposition being stated).

In a similar way, tag questions in Mandarin can either end in a final rising-falling or final rising intonation depending on what the speaker is expecting as an answer. The tag questions are usually formed by adding the particles *shi ba* and *shi ma* to the declarative. The particle *ma* in Chinese tags is matched with a rising tone while the particle *ba* in Chinese tags is matched with a falling tone. For example, the tag question *He goes to school, doesn't he?* when used with a rising intonation would be *Ta1 shang4xue2 shi4ma*. In this situation, the speaker does not know the answer to the question and is asking for one. When the same tag: *He goes to school, doesn't he?* is used with a falling intonation, it would be *Ta1 shang4xue2 shi4ba* In this situation, the speaker knows the answer to the question and is trying to confirm it by seeking an agreement (Zhu & Wu, 2011).

The difference between the English and the Chinese tags is that in Chinese, the intonation of the tag question is attached to a certain particle (*ma* or *ba*). However, in English, the intonation of the tag is not attached to a certain particle but can be applied to any tag depending on the speaker's expectation. Thus, Chinese participants might have

difficulties perceiving and producing tag question intonation which is not part of a particular lexical phrase.

To sum up, non-native speakers can have difficulty learning and using tag questions. Interventions with computer technology have shown some success in helping non-native learners perceive and produce English intonation patterns. These interventions, though, have typically been with expensive software requiring specialized designer knowledge. To address this challenge, we decided to provide implicit, repetitive and restricted practice with the use of freeware. To our knowledge, there are no published studies on the use of technology to help non-native learners perceive and produce tag questions. Therefore, we undertook to design and pilot an open-access tool which incorporates computer-based audio feedback (through the activity module for Moodle: Online Audio Recording) or audio-visual feedback (through the application WASP) to help Chinese learners of English improve their perception and production of rising and falling tag questions. We choose to work with these two types of feedback (audio- visual and audio) in order to further explore the results of feedback from two channels and from one channel (audiovisual and audio, respectively). Our research questions for this feasibility study were the following:

Research Questions

1. How easy are the tools (WASP and Online audio recorder) for participants to use?
2. Do the tools meet the learners' needs in terms of learning?
3. What effects does implicit training with the tools (WASP and Online audio recorder) have on participants' perception, production and pragmatic interpretation of tag questions?

Hypothesis – Research Question 3

We predict that that learners exposed to the audio visual feedback might perform better in their perception and production of the tag question intonation than the learners exposed to the audio-only feedback.

Participants

Two groups of participants participated in the study: an audio-visual feedback group (AV) and an audio-only feedback group (AO). The AV group and the AO group consisted of three students each (n=6), with the sole male in the AO group. The participants in both groups were international students recruited from an English-medium university in Montreal (M_{age}= 24). Participants were native speakers of Mandarin, and had completed elementary and high school in China. They were in their second year of residence in Canada (M=27.8 months) and had studied at the university for a year or less (M=11.3 months). They rated their own proficiency in English on a 9-point Likert-type scale (1= Extremely Poor, 9= Extremely Fluent), with a mean rating of 5.9. All participants were given pseudonyms. Further participant information is seen in Table 1.

Table 1 Participant information

Participant	Gender	Age	Period of full-time studies to date (months)	Length of residency in Canada months	Proficiency in English
Audio-visual Group					
Hao	F	22	32	22	5.5
Yin	F	23	6	19	6.25
Kew	F	21	6	18	5.25
Mean		22 (1)	14.6 (15.0)	19.6 (2.08)	5.6 (0.52)
Audio Group					
Xin	F	23	10	11	6.75
Ji	F	24	6	7	5.5
Fu	M	36	10	90	6.5
Mean		27.6 (7.23)	8.6 (2.30)	36 (46.8)	6.25 (0.66)

Note. Likert-type scale 1-9 (1 = extremely poor, 9 = extremely fluent)
* Numbers in parentheses represent standard deviations

Method

This study had three phases: a familiarization and pre-test phase, a training phase, and a post-test phase. Table 2 shows the tasks and time periods for each phase.

Table 2

Summary of Design Study

Phase	**Familiarization/ Pre-test**	**Training**	**Post-test**
Time period	5 minutes/30 minutes	Twice a week over 2 weeks	2 weeks after end of training/30 minutes
Tasks	Listen and select final intonation contours: 5 declarative sentences with up or down final intonation; Read aloud and record: 2-3 line text/ Listen and select final intonation contours: 20 tag questions, 1 video with 4-5 tag questions; Read, select final intonation contours, and read aloud and record: 1 written paragraph with 4-5 tag questions	**Audio-visual group** Listen and record own version, comparing own contours of final intonation contour to original final intonation contour: 15 tag questions; Read, select final intonation contours, and read aloud and record: 1 written text with 4-5 tag questions **Audio-only group** Listen and record own version: 15 tag questions; Read, select final intonation contours, and read aloud and record: 1 written text with 4-5 tag questions	Same tasks as pre-test; similar materials

Instruments/Materials

All instruments and materials were piloted before data collection (described below). In pre- and post-tests delivered through Moodle (2012), an online course-management system, two perception tasks and one production task were used to assess participants' perception and production of tag questions. Before pre-testing, participants completed a written questionnaire on their language learning history and became familiar with Moodle and the types of tasks used in the pre-test. Two types of perception and production tasks were used for the training phase, delivered through Moodle and, additionally for the audio-visual group, through WASP (Huckvale, 2012), a simple application for the recording, display and analysis of speech. Participants also underwent a semi-structured interview after post-testing.

Preparing the Material

Questionnaire.

The participants were asked to complete a written questionnaire at the beginning of the study to gather some information about them. The questionnaire consisted of 30 questions in English about participants' language background and language learning history, drawn from a questionnaire used in the second author's previous research (see Appendix A).

Semi-structured interview.

A semi-structured interview took place with all participants after completing the post-test. The interview consisted of questions regarding the participants' performance

and opinions on the material in general. The researcher created 19 questions that were piloted with a pilot group. The pilot testing results are described below in the Piloting section.

Declarative sentences.

Five declarative sentences were created by the second author, a native speaker of English. The declarative sentences were used to familiarize the participants with the task and instructions using the computer. Three of the five sentences were spoken with a rising intonation. The sentences were an average of 4.2 words in length and used the simple present. 90.48 % of the words fell within the 2000- word frequency band (Cobb, 1994). The sentences were recorded in a quiet room and were pilot tested.

Pre- and post-test.

The tests included three types of materials: a set of audio tag questions with corresponding text, a brief video, and a brief written paragraph.

Tag questions. 100 tag questions (50 rising and 50 falling) were created by the first author, a near-native speaker of English and verified for authenticity by the second author, a native speaker of English. The sentences were an average of 6.2 words in length and used the simple present, the past tense and the future construction will. The tags were formed using positive sentences with a negative tag and vice versa. 96% of the words fell within the 2000-word frequency band (Cobb, 1994). The questions were recorded by three native speakers of English (two females and one male). The first version of the tag questions was recorded in a quiet room by one of the female native speakers and the male native speaker onto a Macintosh laptop using Garage band and Audacity software and a Plantronics headmounted stereo microphone. Each speaker recorded 50 of the 100 tag

questions (25 rising and 25 falling) by reading them aloud one after the other at least twice. Then the recorded tag questions were edited using Audacity (2013), a sound-editing software. A period of 500 milliseconds of silence was inserted before and after each tag question. The edited tag questions were then viewed with WASP. Of the 100 tag questions, 40 (20 rising and 20 falling) did not display the targeted intonation contour for the tag. Therefore, a second female native speaker recorded those 40 tag questions. The tag questions were read aloud one after the other at least twice. Then, the re-recorded questions were edited and viewed with WASP. The final set of 100 tag questions chosen from the pool of questions recorded by the three speakers were selected because those questions had the best sound quality and most clearly displayed the targeted final intonation contour for the tag. The 100 questions were normalized for peak volume, then they were randomly divided into 5 groups of 20 questions. These 5 groups of tag questions were pilot tested. The results are explained in the Piloting section below.

Videos. Scripts for three short video clips of an average 238 words were created by the first author and verified for authenticity by the second author. They were two-sided phone conversations of 3-4 minutes in length. Each conversation involved a common service transaction between a customer and a company representative (i.e., fast food order, hotel reservation and floral order). The conversations included 4-5 rising and falling tag questions, in equal or almost equal numbers, asked by the customer. These questions were preceded by language from the customer which demonstrated his or her (lack of) certainty about the answer to the question. The customer's turns lasted an average of 13.3 words while the other speaker's turns lasted an average of 11.9 words. 86.8% of the vocabulary used in the videos fell within the 2000-word frequency band

(Cobb, 1994). The videos were recorded in a quiet room by the same male and female native speaker of English as for the tag questions, directly onto a Macintosh laptop using a digital video camera and two lapel-mounted wireless microphones, a wireless transmitter and receiver, and an audio adapter which combined the two audio tracks onto one sound file. The customer in each video was always shown on-screen, with the customer representative off-screen but fully audible. Four takes of each video were recorded, with the actors switching roles each time. The videos were edited using VideoPad video editing software (NCH software, 2008). The final versions of each video were created from combining sections from different takes which had the best audio and picture quality. The final versions of videos for piloting were edited into sections divided by tag questions; each section ended after a tag question in order to allow the participants time to decide whether the tag rose or fell. After making their choice, the participants had to click on the next video section to view the video. The videos were then piloted.

Texts. Nine texts of an average two hundred and twenty-seven words were created by the first author and reviewed for authenticity by the second author. The texts were dialogues on various topics which participants might encounter in their daily lives (e.g., classmates arguing, two people gossiping). 90.6%of the vocabulary used in the texts fell within the 2000-word frequency band. Each text included 4-5 rising and falling tag questions in equal or almost equal numbers. These questions were preceded by language which demonstrated the speaker's certainty (or lack of) about the answer to the question. The texts were then pilot tested.

Piloting

One group participated in the pilot study. The group consisted of 3 male and 2 female participants. The participants were international students recruited from an English-medium university in Montreal, who were native speakers of Mandarin and had completed elementary and high school in China. They rated their own proficiency in English on a 9-point Likert-type scale at a mean of 5.9 (1= extremely poor, 9= extremely fluent). Each participant was given a set of material to work on: tag questions, videos and texts using WASP and Moodle. Five groups of twenty tag questions were randomly created, with each participant piloting one group of questions. All videos and declarative sentences were piloted by all participants. Each participant piloted 2-3 of the 9 texts, working in a Macintosh computer laboratory in the presence of the first author, who gave guidance in navigating Moodle and WASP and noted any problems which were mentioned or observed. The participants were then interviewed after completing the tasks. They were asked about the quality of the tag questions, videos and texts (sound quality and level of difficulty). They were also asked about using WASP, a simple application for the recording, display and analysis of speech, and Online Audio Recording, a flash-based audio recorder that can be installed into Moodle as an activity module, in terms of how easy it was to understand instructions and to use these applications. Participants took one hour to complete the tasks and the interview. The piloted materials were not modified except for one of the videos, which was discarded because of its poor sound quality and some minor problems in audio. The instructions created for WASP users were edited to increase their clarity and ease of use.

Final Materials

From the 100 piloted tag questions, 2 groups of 20 -were randomly chosen for the pre- and post-tests. From the remaining 60 tag questions, 4 groups of 15 tag questions were randomly selected to be used for the purposes of training. The two remaining videos were randomly assigned to the pre- and post test. Seven of the nine piloted written texts were randomly chosen for the study: one of the seven was used for familiarization training, two were assigned to the pre- and post-test, with the remaining four used for the perception and production training. The piloted semi-structured interview used at the end of the study consisted of 19 questions. The questions were about the participants' performance and opinions on the material in general. The testing and training tasks themselves are described in the Procedure section.

Pre- and Post-Test Materials

The testing materials for the pre- and the post- tests were similar but not equivalent. The tests were delivered through Moodle (described below). The tests included three sections: a set of 20 audio tag questions (10 falling and 10 rising, ordered randomly), with corresponding text, a three-minute video clip, and a written paragraph of 15-25 lines of an average 225 words. The set of tag questions, the video clip, and the paragraph used vocabulary which fell primarily (90.5%) in the 2000-band of word frequency (Cobb, 1994).

Training Materials

Moodle familiarization. Before the pre-test, students were made familiar with Moodle and the type of tasks to be presented. The Moodle familiarization tasks included

a perception task with recordings of five declarative sentences (with final intonation contours going up in three sentences and down in two sentences) and a recording task with two to three lines of written text randomly chosen from one of the three texts not used for the testing or training, to be recorded using the Online Audio Recording activity module for Moodle.

Perception and production training: the training materials were delivered through Moodle and WASP, depending on the group of participants. Two blocks of training materials were made available each week, for a total of four training blocks over two weeks. Each training block contained a set of 15 recorded tag questions (with similar numbers of rising and falling questions) and a brief written paragraph of 227 words, on average, containing a minimum of 4 tag questions (similar numbers of rising and falling).

Testing and Training Applications

Moodle. Moodle is an Open Source Course Management System (CMS). In the familiarization phase, the first author helped participants to create an account and then log in using their user-names and passwords to access the activities. The format of testing and training materials was designed by the first author using the Moodle tools Assignments and Quizzes. The materials (short paragraphs) in Assignments were used with the Online Audio Recording activity module so participants could record texts. The materials in Quizzes contained the tag questions and the videos. All of participants' responses to the training and testing tasks were saved on a directory on the server hosting Moodle. The Assignments and Quizzes tools allowed the first author to show or hide items for training or testing. The amount of control participants had over the content of

the Moodle items was quite limited. The sound and video files and texts were uploaded, displayed and hidden by the first author, using a fixed schedule. Items in a task had to be completed in a fixed order. During a pre- or post-test, participants were able to go back and re-do an item in the task if they had enough time left and if they haven't submitted their work yet. Nevertheless, the participants had control over the recordings they made and the file names they gave to the recordings.

WASP. WASP was used by the audio-visual group during the training phase. It is a simple application for the recording, display and analysis of speech. The group using WASP could record and replay speech utterances, save them and open them from the folder in laptop where they were saved. The group could also view waveforms and pitch display tracks for each sound file (see Figure 2).

Online Audio Recording. Online Audio Recording was used by all participants in testing tasks. It is a Flash-based audio recorder that can be installed into Moodle as an activity module, not as a standalone plug-in. Files can be recorded and uploaded using the Online Audio Recording application. However, the uploaded files created by the recorder cannot be played by a playback device within Moodle; they must be downloaded and listened to using an external media player. Online Audio Recording also allows for manual uploads of audio files produced using other recorders, as was done for the training of the audio-only group.

Procedure

Throughout the study, participants were never informed by the authors about how final intonation patterns of tag questions in English could shape the pragmatic interpretation of a question. This is because we were interested in how practice solely in

perception and production of final intonation contours might be related to participants' awareness and use of pragmatic meaning. Participants started by filling out a questionnaire and signing a consent form (see Appendix B). The questionnaire consisted of 30 questions about participants' language background and language learning history (see Appendix A).

The familiarization and pre-testing was conducted in a Macintosh computer laboratory. With the assistance of the first author, the participants logged in to Moodle, created an account and logged into the study site. There they completed five tasks: two familiarization tasks and the three pre-test tasks.

Familiarization phase. For the perception task, participants read written instructions to listen to a set of five declarative sentences while reading the sentence texts, then to select one of the two options for final intonation, up or down, by clicking on a radio button. The first author checked to ensure that participants understood how to do the task. For the production task, participants read written instructions to use the Online Audio Recording activity module to record two to three lines from one of the texts which were not included in test or training materials. The participants received written and oral instructions on how to record and save on their screens before beginning the task. They had to record, upload and save their work. They were not given a time limit or a limited number of attempts. For both tasks, the first author observed participants, but no participants reported difficulties with doing the task. The familiarization phase lasted, on average, five minutes.

Pre-testing phase. This phase took place immediately after the familiarization phase in the same computer laboratory, in the presence of the first author. It took

participants on average 20 minutes to complete the test. Participants were already logged into Moodle, and were instructed orally to select the first exercise in the Pre-test activity. Similarly to the familiarization phase, participants then read instructions to listen to the tag questions, read the transcripts and click one of two options which described the final intonation: up or down (see Figure 1). They had 5 minutes to complete the exercise and were allowed only one attempt to complete the entire exercise. Before submitting the exercise, participants could re-listen to any tag question and change their answer. However, once the participant chose to submit the exercise, no other changes could be made. All participants submitted the exercise in less than five minutes.

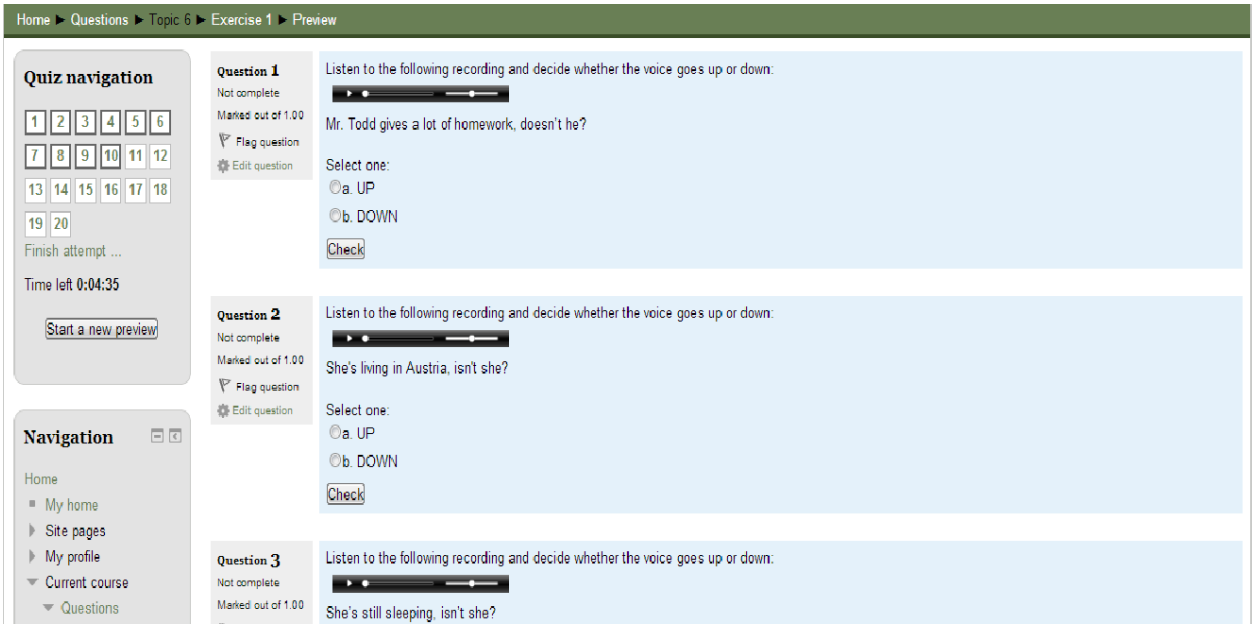

Figure 1. Screen shot of the first exercise in Moodle

Participants were then instructed to select the next pre-test exercise. They received written instructions to watch the video and to click the up or down option after hearing each tag question. The entire video was made up of different segments. Each

section ended after a tag question, allowing participants to make their selection. Participants then clicked on the next video segment. Before submitting the exercise, participants could re-listen to any video segment and change their answer, but could not change any answers after submitting. Participants had four minutes to complete the exercise, and all finished in the allotted time.

Participants were then instructed to select the final pre-test exercise. They received oral and written instructions to read the paragraph, and for each tag question in the paragraph, to select the final intonation they expected for that question. Different from the first two exercises, participants did this selection by writing on a paper with the paragraph, which included the bolded and italicized tag questions followed by the two options of up or down (see Appendix C). Participants circled one of the options for each question. Participants were then instructed to read aloud and record the paragraph, including the tag questions with their expected final intonation. Participants were not given a time limit or a limited number of attempts to complete their task. The paragraph with the intonation selections was then collected by the first author.

Training phase. A few days after the pre-test, the participants were randomly divided into two groups: the audio-visual group and the audio-only group, and began training. After the pre-test, the training materials were delivered through Moodle or WASP, depending on the group of participants. The training lasted approximately two weeks.

Audio-visual group. Participants in the audio-visual group had to attend training sessions with the first author twice a week for a period of two weeks. For each training session, each participant worked individually in a quiet room for around 30 minutes. The

participants were provided with written instructions on how to use WASP with Macintosh computers. The first author briefly checked participants' understanding of how to work with WASP. Participants' training involved reading texts, recording their own speech and analyzing their speech by looking at the pitch track of their recordings on WASP. By looking at the pitch track (a line connecting the points of a person's basic rate of vocal cord vibration at each sampled point in time), participants could see a wave showing the pattern of rises and falls in pitch (the intonation pattern) for a given utterance (see Figure 2).

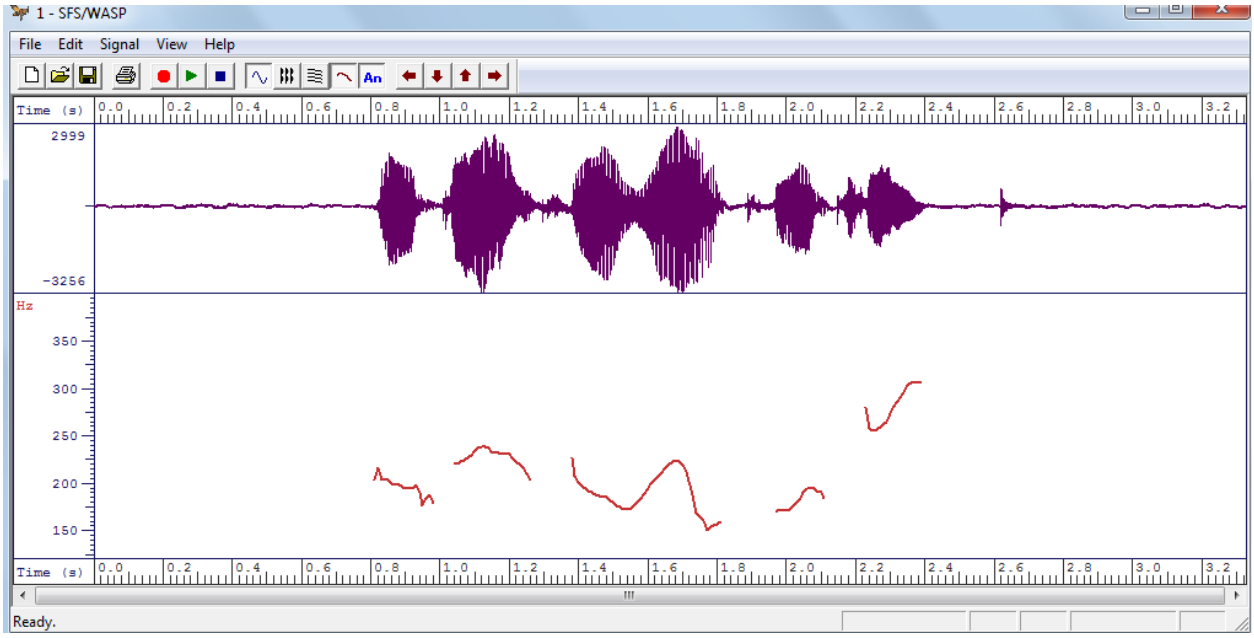

Figure 2. Visual display of the utterance *you're going for a walk, aren't you?* in WASP

Training materials consisted of, first, recorded tag questions (with no transcript) and second, a short written text. Participants were instructed to listen to the tag questions, repeat them and record their own versions as much as they pleased. As the participants

recorded their versions, visual feedback appeared in the form of a pitch track. The participants were asked to look at their pitch tracks and compare them to the originals, which could also be seen on WASP. The short written text was provided on paper, and participants were instructed to read the text, and for each tag question in the text, to select the final intonation they expected for that question by circling either of the two choices: up and down. Participants then recorded their own version of the text and had the possibility of looking at the pitch track for their own recorded versions. No model recordings of the written text were available to the participants.

Audio-only. The audio-only group accessed Moodle remotely, and was asked to do the training for a period of 15-20 minutes twice a week by completing the two blocks of training material each week. Similarly to the audio-visual group, the training required that the audio-only group listen to the tag questions and repeat and record their utterances. They were also asked to read a text containing tag questions and select the appropriate final contour (up or down) for each question, then make their own recordings of the tag questions. However, unlike the audio-visual group, this group did not receive any visual feedback. The participants in this group were able to listen to their recordings and aurally compare their individual tag questions (not those from the text) to the original recordings. It was possible to confirm that all participants in the audio-only group had accessed all the training exercises during the training period.

Post-test. After their training, all participants were then tested again in the computer lab on their perception and production of tag questions, using the same tasks and the same procedure as in the pre-test, with different content. All participants submitted the first two tasks in the allotted time except for one participant, who did not

complete the first task in the allotted time. All participants completed the third untimed task, with the written paragraph collected by the first author. The researcher conducted the interviews face-to-face in English after the post-test. They were recorded using Audacity software and then transcribed by the first author. The interviews lasted about 15 minutes.

Data Processing and Analysis

The first exercise for the pre-and post-tests included 20 tag questions and was scored out of 100 points. Each tag question was given a value of 5 out of 100. Based on the prepared answer key, the participants would be awarded 5 points for every correct choice and 0 points for every incorrect choice they made.

The second exercise for the pre- and post-test was given a total score of 25 points. Each of the 5 tag questions was given a score of 5 points. Based on the prepared answer key, the participants would be awarded 5 points for every correct choice and 0 points for every incorrect choice they made.

The third exercise was also given a score calculated out of 40 points. However, the scoring for this exercise was divided into two parts: selection and production of tag questions. For the selection part, the participants were scored out of a total 20 points. The participants were awarded 5 points for every correct selection they made and 0 points for every wrong selection they made, according to the answer key. For the production part, participants were scored out of a total 20 points. The participants' productions of tag questions were first analyzed using the WASP application to check if the contour went up or down. Then, participants were awarded 5 points for every accurately produced tag question and 0 points for every inaccurately produced tag question, according to the

answer key. Mean scores and standard deviations of the pre- and post- tests were calculated for all individual, group and total scores using the application Excel.

Interview

After the interviews were transcribed, data were analyzed through topic-oriented cyclical data analysis (Watson-Gegeo, 1988). Interview questions were categorized as relating to one of the three research questions. Then, through repeated readings of the transcripts, participants' responses were classified according to the research question categories, although a response could be relevant to more than one research question at a time.

Results

The first research question was: *How easy are the tools (WASP and Online audio recorder) for participants to use?*

Based on the interviews conducted and the researcher's observations, the participants found that the two applications WASP and Online audio recorder were user-friendly and practical for educational practices. As Yin put it, "interesting attached to education"[it was interesting for her to make use of the applications for learning purposes]. The participants thought that using WASP was helpful and interesting "yes, [using the application] was fine, it was interesting" (Kew). Yin said, "I enjoyed using the application", and this was true of all participants. When the audio-visual group was asked about the application they preferred the most, one out of the three said that she thought using the Online Audio Recorder was much easier because fewer functions were needed to listen, record and save: "Yeah, cause this one I have to record it, I have to save

it, I have to open another file and do it again and I have to wait… but for this one, it is easier to cooperate…" (Hao)

The second research question was: *Do the applications meet the learners' needs in terms of learning?*

On the pre-test, all the participants performed well on all three tasks, especially on the first and second exercises (see Table 3). It appears that participants were already able to accurately perceive final intonation contours at the beginning of the study. However, all participants stated that the applications were helpful for their knowledge about the nature of tag questions. For example, at the beginning the participants were not aware of the various intonation patterns for producing tag questions. Yin said that she had become more aware:"It just feel clearer than before… because the first time I did the testing in the lab I really didn't know about it, but now I know ..ahh.. it's going to be like this" [now the intonation is clearer … at the beginning I did not know about the intonation of the tag question, but now I know when I hear it that it is either going up or down]. The following sections discuss the results of the participants' perception and production along with their interpretations of the tag questions.

Table 3

Pre-Test Results

Participants	Exercise 1 /100	Exercise 2 /25	Exercise 3			Individual Total scores /165
			Accurately selected /20	Accurately produced /20	Total score /40	
Audio-visual						
Hao	100	25	10	5	15	140
Yin	100	25	15	-	-	-
Kew	100	25	15	15	30	155
Mean	100 (0)*	25 (0)	13.3 (2.88)	10 (7.07)	22.5 (10.6)	147.5 (10.6)
Audio-only						
Xin	90	15	15	10	25	130
Ji	100	20	15	15	30	150
Fu	100	20	5	5	10	130
Mean	96.6 (5.77)	18.3 (2.88)	11.6 (5.77)	10 (5)	21.6 (10.4)	136.6 (11.54)
Total Mean	98.3 (4.08)	21.6 (4.08)	12.5 (4.18)	10 (5)	22 (9.08)	141 (11.4)

Note. Numbers following forward slashes (/) represent total number of points
*Numbers in parentheses represent standard deviation.

The third research question was: *What effects did the training have on participants' perception, production, and interpretation of tag questions?*

Perception

No statistical tests were run to explore group differences because of the small sample size. Because this study was primarily a feasibility study, we were not greatly concerned with within-group or between-group effects. However, in general, the participants' performance in the pre-test and the post-test was similar (see Table 4 for post-test results). The participants' scores were high on both the first and second exercises of the pre- and post- test. In the pre-test, almost all participants excelled on the first exercise, with a mean score of 98.33 out of a possible 100 points. The Audio–visual

group had a mean score of 100 and the Audio-only group had a mean score of 96.6. Similarly, in the post-test, participants excelled in the first task with a mean score of 93.33 out of 100 points. The Audio–visual group again had a mean score of 100 and the Audio-only group had a mean score of 86.6. The reason the mean scores for the audio-only group were much lower than the other group had to do with one of the Audio-only participants who could not finish the task on time and thus had a lower score than the others in the same group.

As for the second exercise, the participants also excelled in the pre-test with a mean score of 21.6 out of 25. The Audio–visual group performed better than the Audio-only group with a mean score of 25 compared to a mean of 18.3 for the Audio-only group. However, in the post-test all the participants received perfect scores of 25.

Table 4

Post-test Results

Participants	Exercise 1 /100	Exercise 2 /25	Exercise 3			Individual Total scores /165
			Accurately selected /20	Accurately produced /20	Total score /40	
Audio- visual						
Hao	100	25	20	10	30	155
Yin	100	25	15	10	25	150
Kew	100	25	15	15	30	155
Mean	100 (0)*	25 (0)	16.66 (2.88)	11.66 (2.88)	28.33 (2.88)	153.33 (2.88)
Audio-Only						
Xin	95	25	15	5	20	140
Ji	65	25	20	20	40	130
Fu	100	25	20	15	35	160
Mean	86.66 (18.92)	25 (0)	18.33 (2.88)	13.33 (7.63)	31.66 (10.40)	143.33 (15.27)
Total Mean	93.3 (14.02)	25 (0)	17.5 (2.73)	12.5 (5.24)	30(7.07)	148.33 (11.25)

Note. Numbers following forward slashes (/) represent total number of points
*Numbers in parentheses represent standard deviation.

Perception and Awareness in the Real World

Although most participants performed very well in the perception tasks even at the beginning of the study, most of the participants stated that the training made them more aware about the existence of tag questions in the real world: "I had no idea about this tag questions" (Yin). They also explained that the training helped them to recognize and perceive the final intonation contours of tag questions used in real communication. "I'm able to tell… I can recognize [if it is going up or down]" (Kew).

Production

Detailed analyses of the results for the production of tag questions in Exercise Three (tag questions produced in context) were based on two criteria:

1. If the participants chose the right tags (5 points for each accurate selection)

2. If the participants produced the right tags (5 points for each accurate production)

The individual scores for selection and production in pre- and post-tests are in Table 5.

Pre-test. In the pre-test, participants had to read the paragraph, and for each tag question in the paragraph, to select the final intonation they expected for that question. The participants had to select by circling one of the options for each question. Participants were then instructed to read aloud and record the paragraph, including the tag questions with their expected final intonation.

Table 5

Results of Exercise Three

	Pre-test												**Post-test**											
	Tag1			Tag2			Tag3			Tag4			Tag1			Tag2			Tag3			Tag4		
	S	**P**	**M**	**S**	**P**	**M**	**S**	**P**	**M**	**S**	**P**	**M**	**S**	**P**	**M**	**S**	**P**	**M**	**S**	**P**	**M**	**S**	**P**	**M**
Hao	U	U	A	U	U	A	A	A	A	A	U	U	A	U	U	A	A	A	A	A	A	A	U	U
Yin	U	-	-	A	-	-	A	-	-	A	-	-	A	A	A	A	A	A	A	U	U	U	U	A
Kew	U	U	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	U	A	U	A	U	U
Xin	U	U	A	A	A	A	A	A	A	A	U	U	A	U	U	A	A	A	A	U	U	U	U	A
Ji	U	U	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
Fu	U	U	A	U	U	A	A	A	A	U	U	A	A	A	A	A	U	U	A	A	A	A	A	A

Note: A: accurate/ U: unaccurate / Audio-only group : Shaded Area
S: selected, P: produced, M: matched , - : no data available

Participants' accuracy sometimes varied for different tag questions. All participants were completely accurate in their selection and production of the third tag question. However, for the first tag question, the selection and production was inaccurate for all participants. The fourth tag elicited the most variable performance. All participants except one selected the correct tag, but three of those participants produced a tag which was different from their initial selection. The differences across different questions will be considered in the Discussion section.

In terms of performance by group, in the audio-only group, one participant (Fu) made inaccurate selections and productions for three tags, while the other participants

were inaccurate in their selection and production for one tag. In the audio-visual group, one participant (Hao) made inaccurate selections and productions in two tags, while another participant was inaccurate in the selection and production for one tag. However, because of equipment failure, the productions of one participant in the audiovisual group (Yin) could not be analyzed. In general, both groups, the audio-visual (AV) and the audio-only (AO) group performed similarly in the pre-test with mean scores of 22.5 for the AV group and a mean score of 21.6 for the AO group.

Post-test. In the post-test, the task had similar parameters but different content. Again, participants' accuracy was different across different questions. The third tag was accurately selected by all participants except for one (Kew). However, only two participants inaccurately *produced* the third tag. As for tag four, it elicited the most variable performance. Most of the participants were accurate in their *selection* and then inaccurate in their *production* of the tag question. Only two of the participants were completely accurate in their selection and production of the fourth tag question.

In terms of performance by group, in the audio-only group, only one participant (Ji) was completely accurate in the selection, production, and matching of all the tags. One of the participants was inaccurate only in the production of one tag (Fu), while the last participant in this group was accurate in the selection of three tags out of four but inaccurate in the production of three tags out of four (Xin). In the audio-visual group, all three participants were accurate in their selection and production of two tags, while they had variable performance on the other two tags: selecting and producing inaccurately. Clearly, a participant's accurate selection of a final contour did not mean the production would also be accurate.

Pragmatic Meaning

During the interview, all participants were able to produce both up and down contours accurately when prompted to. However, half of the participants said it was not easy for them to accurately produce a rising or a falling tag question based on a given context. One of the participants explains: "Not very hard, but sometimes I got confused about whether it goes up or down" (Hao). Another participant explained that she still had a little confusion in choosing up or down but now not: "I still had some little confusion I have to choose to go up or down" (Yin)

All participants stated that before the training, they were not aware of the different interpretations of rising or falling tag questions (rising: seeking information or falling: confirming information). This is very evident from the comments of one of the participants, who reported always using a rising tag before and during the pre-test: " I think is always use up pronunciation, but I heard some experimental is use some down pronunciation [I have always used up but with the help of the study I realized that a falling intonation exists]… but always there is some confused because I don't know what time I need to use the up, what time I need to use the down because before I always use the up pronunciation [I was taught to always use the rising intonation but now that I realize there is a rising and a falling intonation I am a bit confused when to use either one]"(Fu). In all four out of the six participants were able to explain the pragmatic meaning of the English tag questions.

Some participants were still in the process of learning the pragmatic meaning of rising and falling tag questions, such as Fu, who started to use more of the falling intonation in the post-test, but all participants had at least become aware that there were

different pragmatic meanings. This awareness developed without explicit information being provided about the pragmatic meaning of final intonation contours.

Thus, the results show that the use of the two tools: WASP and Online Audio Recording were beneficial, interesting to use and helpful for the participants' learning. Not only did the participants become more aware of their perception of tag questions, but also, they have improved in their production of tag questions.

Discussion

The current study used the tools Moodle, WASP and Online Audio Recording to help Chinese learners of English with their perception and production of rising and falling tag questions. Six international students (native speakers of Mandarin) participated in the study. The participants were pre- and post tested on their perception and production of question tags. Participants were tested on their perception in context and in isolation. They were also tested on their production in context. The participants underwent training which lasted approximately two weeks. Each week, the participants were asked to complete two blocks of the training material (15-20 minutes). A semi-structured interview took place with all participants. What follows is a discussion of the results presented earlier.

1. How easy are the two tools (WASP and Online audio recorder) for participants to use?

The results revealed that the participants enjoyed using the tools and found them very user-friendly. These results have implications for researchers' and teachers' use of freeware to design tools for speech perception and production. The first author, who used the tools and Moodle, a delivery system, to design the pre-test, post-test, and training tasks, has average computer literacy and technological skills. However, she found and

made use of these applications quite easily. The applications do not require designers with advanced technological skills. Users with minimal technological knowledge can easily use these user-friendly tools to create the kind of materials needed for pronunciation, listening or speaking purposes.

2. Do the applications meet the learners' needs in terms of learning? And 3. What effects did the training have on participants' perception, production, and interpretation of tag questions?

Probably because of the small group of participants and the training that lasted over two weeks, the lack of difference between the two groups at the pre- and post-test was not surprising. Both the Audio-visual and the Audio-only group performed similarly. The results for the first and second perception tasks (in isolation) in the pre- and post-test revealed that the participants performed at ceiling on these tasks. This shows that the participants in general did not have any problems with the perception of the intonation contours in terms of acoustics (i.e., in terms of perceiving high or low pitch). However, the participants' performance on the third task (production in context) elicited variable performance. For example, in the pre-test, all the participants inaccurately selected and produced the first tag question, while the fourth tag question elicited the most variable performance. For the fourth tag question, all participants except one selected the correct tag, but three of those participants produced a tag which was different from their initial selection. Similarly in the post-test, the third tag was accurately selected by all participants except for one. However, two participants inaccurately *produced* the third tag. As for tag four, it elicited the most variable performance. Most of the participants were accurate in their selection and then inaccurate in their production of the tag

question. Only two of the participants were completely accurate in their selection and, production of the fourth tag question.

One explanation for why some participants selected accurately and produced inaccurately or vice versa could be due to the nature of the particular tag questions in context. In the pre-test, for example, the reason why the first tag question was inaccurately selected and produced may have to do with the way the tag question was presented in the text. In general, the texts included key words that would allow the participants make their choices based on them. These key words were mostly placed before the tag question in the text to provide enough information for the participants to make an accurate selection and production. However, the key words for the first tag in the pre-test, shown here in Excerpt (1) were placed after the tag question was presented in the text, which might have confused the participants into selecting and producing inaccurately. The underlined words are the keywords and the italicized words are the tag questions.

(1) Tag 1: Pre-test

Then, a policeman went to Mr. Dominic and asked him, "What time did you hear the gunshots?" Mr. Dominic said, "*Wait, wait! You're supposed to tell me that I have the right to remain silent, aren't you? (UP / DOWN)* <u>That's what happens on all the TV shows I watch</u>."

In the post-test, however, the key words for the third and fourth tag questions in the post-tests were placed before the tag questions and still there was variable

performance. Excerpt (2) shows the fourth tag question. So, one might ask, then: why this variable performance?

(2) Tag 4: post-test

Please, can't I go to bed! It's been a very difficult night!" The policeman said, "I understand. *It's very upsetting to think about this violent crime, isn't it?* (UP / DOWN) Go to bed and get some sleep.

Another explanation could be that the participants were starting to pay attention to the context of the tag questions which could help them to make their choice. In other words, the participants have used or adopted some context-based selection process when forced to make a choice of up or down. The production task in the pre-test, training and post-test forced the participants to make their choices (up or down) in context. The participants were not given any hints or feedback, nor were they provided with any audio models to show which direction the intonation contour should go. In the post-test, almost all the *selections* of question tags were accurate. According to Schmidt's Noticing hypothesis (1990), learners have to notice the input for learning to happen. As some participants stated in their interviews, at the beginning of the study, they were not aware of the different pragmatic meanings behind the two intonation patterns for producing tag questions. But at the pre-test and throughout the training, learners were asked to choose between the up and down intonation patterns in the contextualized production task. The learners might have noticed a gap in their knowledge of how to select or even produce the accurate intonation contour (up or down) in context. During the training, participants were exposed to models of question tags only in isolation. There were no contextualized

models of question tags which could help learners notice the links between intonation contours and pragmatic meaning. In the post-training interviews, participants explained that the training helped them perceive and recognize the final intonation contours of tag questions when participants were involved in or heard authentic communication.

Therefore, participants may have drawn on authentic input outside the training context in order to develop 'rules of thumb' for how up and down intonation contours match up with pragmatic meaning. This may explain why the participants' accuracy in selecting intonation contours in context improved noticeably from pre-test to post-test.

However, participants were more accurate in selecting intonation contours in context than they were in producing them in context. A complementary explanation for this variable performance could be related to Skill Acquisition theory, proposed by Robert DeKeyser (2007). DeKeyser (2007) explains that there are three stages of development (first declarative, then procedural and finally automatic) which are characterized by differences in the nature of knowledge and its use. At the beginning learners acquire some knowledge through explicit instruction or modeling, careful observation, and analysis of others using that knowledge. Declarative knowledge involves the acquisition of rules and facts, but learners do not try to use this knowledge. In the second stage, this knowledge is acted upon, turning into behavior or, as DeKeyser puts it, procedural knowledge. This kind of knowledge is now available for the learner once the conditions for the behavior are met. However, this acquired knowledge is far from being spontaneous and fluent, and there are consistent inaccuracies in the application of knowledge. In other words, a lot of practice is needed for application of the knowledge to become more fluent. Through repeated practice and use, the use of

knowledge gradually becomes less effortful and more fluent, which DeKeyser refers to as automatic.

The learners or participants may have acquired some declarative knowledge in the form of rules of thumb about the match between different intonation patterns and different pragmatic meanings. This would explain their improved accuracy of selection of contours in the post-test. This declarative knowledge may have been turning into procedural knowledge, although this was not yet robust or fine-tuned (DeKeyser, 2007, p.98). The results from the third and fourth tag question in the post-test production task may be evidence of this intermediate state. Most of the participants accurately *selected* the appropriate contour in context but could not accurately *produce* the appropriate contour. Further evidence comes from the semi-structured interviews: the participants were asked if they could produce the tag questions and most of them said yes, but it was not easy for them. It may be that if the training had included models of tag questions in context, with intonation contours matched to pragmatic meaning in context, participants' procedural knowledge of how to produce contours appropriate to the context may have become more accurate and more automatic.

Conclusion

Although a number of studies have explored the use of technology for learners' development of prosody (the rhythm, stress, and intonation of speech), very few studies have targeted the development of intonation in English tag questions, particularly for Chinese learners. The aim of the study was to investigate the feasibility of using computer-based training and feedback to help Chinese ESL learners improve their perception and production of tag questions. Two open-access tools which incorporate

computer-based audio (Online Audio Recording) and audio-visual feedback (WASP) were used to explore the Chinese learners' perception and production of English tag questions. In the framework of Chapelle's (2001) themes, these two tools were practical, showed learner fit and provided some evidence of construct validity, had a positive impact, and had language learning potential and potential for meaning focus. The practicality was shown by the ease with which the first author used these tools to implement training. The results demonstrated learner fit because participants quickly learned how to use these tools. The evidence of construct validity was seen in the way both the pre- and post-test results could differentiate between participants' ability to perceive intonation contours and their ability to match intonation contours to pragmatic meaning in tag questions. Perception and production of L2 sounds have been shown to be separate, though related, constructs. For example, a study by Lane and Schneider (1963), revealed that training the participants to discriminate tones did not improve their production of Thai lexical tones (as cited by Chun, 2002, p. 93). The tools show the potential for enhancing language learning, particularly for learners who struggle to perceive intonation contours in English and for learners who are not aware of the acoustic or pragmatic characteristics of intonation contours for tag questions. Although these tools were not used interactively, there is potential for meaning focus because the productive task required participants to select contours in context. As supported by the interviews, this task can encourage learners to focus on the pragmatic meaning of intonation contours. Overall, the use of these tools for training had a positive impact on participants because through using the tools, participants reported becoming more aware of the characteristics of tag question intonation in context.

Future Directions

The study was conducted with a relatively small number of participants because we were interested in looking at the feasibility of the tools to help second language learners improve their perception and production of tag questions. Future studies with larger numbers of participants, different groups and different combinations of tools are needed in order to generate findings which may be generalizeable to different participants and different contexts. In addition, the training was conducted over only two weeks; it is therefore difficult to make firm links between the training and participants' learning. Although one participant considered the training period to be somewhat long, future research could explore the effects of longer training periods and less concentrated blocks of training. Another area to investigate is the use of interactive technology in order to reflect technology use in the non-research environment, which may heighten the generalizeability of results to other contexts. Results from the current study show that early on, these learners could accurately perceive tag question intonations in isolation, but they were still developing control over their pragmatic understanding and productive use of tag question intonation. Future research on the use of technology for learning intonation should draw attention to the development of pragmatic knowledge and use.

Chapter 3

The goal of the study was explore the perception and production of English tag questions by Chinese learners of English using technological devices. The tools (WASP and Online Audio Recording) used in the study provided participants with audio-visual feedback and audio-only feedback in order to develop their production of tag questions as well as their understanding of the pragmatic use of tag questions. The WASP tool allowed the participants to look at the pitch track of their performance and compare it to the original model provided, unlike the Online Audio Recording which allowed participants to simply listen to their own recording and compare it to the original model. According to Chun, Hardison, and Pennington (2008), in the past, learners did not have easy access to computers and computer software that allowed them to look at intonation in terms of visual pitch changes in their speech because of financial and technical inaccessibility of the hardware and software. For the future, Chun et al. proposed that computers and computer software should be used to:

1. Provide learners with visualization of their intonation patterns and with immediate feedback to help improve their speech perception and production;

2. Provide learners with models of authentic speech;

3. Facilitate, record, and analyze speech including interactions of two or more speakers;

4. Offer tools for research, e.g., data collection tools to record students' performance and their steps toward self correction. (p. 336)

For example, Hardison (2005) used two computer-based tools for different types of contextualized training using participants' own oral production (authentic speech) to investigate the effects of production on discourse-level English prosody. Using tools like web-based Anvil and Kay Elemetrics Real-Time Pitch (RTP) program, Hardison was able to provide the participants with a screen display integrating the audio and video components recorded speech with the associated time-aligned pitch contour created in Praat. The participants could also see "a pitch contour in real-time that allowed on-screen comparison of a learner's utterance with that of a native speaker (NS) for feedback including overlay of one contour on the other" (p.170).

The tools used in the study could also be seen in terms of Chun et al.'s suggestions. Starting with the first suggestion, one tool chosen for this study (WASP) does provide learners with immediate feedback that shows the participants' intonation patterns (pitch track). With WASP or Online Audio Recording, participants could holistically compare aural and/or visual displays of their speech with a mode. Nevertheless, like many other tools currently used to provide feedback on intonation(for example, Hardison, 2004, 2005; Spaai & Hermes, 1993), the tools used in the current study do not provide easily interpretable *diagnostic* information about learners' strengths and weaknesses, especially regarding production of intonation. As for Chun et al.'s second and third proposal which are related to modeling authentic speech to learners and facilitating, recording and analyzing speech including interactions between speakers, the tools WASP and Online Audio Recording on Moodle did provide the participants with models of authentic speech and allowed the audio-visual participant group to easily record and analyze their recordings by looking at their pitch track (using the WASP

application). For example, during the training, the participants were able to listen to the previously recorded tag questions and compare them to their own recordings. In this study, participants recorded their individual speech. The audio files could potentially be opened in another tool (e.g., Praat) in order to allow for phonetic analysis by trained phoneticians. It would also be possible to use these tools to record speakers interacting. However, it would be difficult to analyze overlapping speech because both speakers would be recorded onto one track.

As is clear from the study, WASP and Online Audio Recording on Moodle can be productively used for research to record participants' performance and development in L2 speech. Using Moodle, it was possible to track participants' performance on the perception and production tasks; the length of the training period was such that identifying steps in participants' development was challenging. Nevertheless, the semi-structured interview offered some insights into the steps participants took towards self-correction. As one participants explains: 'Yes, because sometimes I look the line [I look at the pitch track] it's going up, but my voice is going down [in comparison to my own production which is going down], so I can do, do it again [I try to produce it again until it is similar to the original recording]' (Kew).

Chun et al.'s (2008) proposals for the use of technology in learning and practising L2 prosody are in line with Kern's (2006) suggestion that the use of technology could be thought of in terms of the metaphors of tutor, tool, and medium. Kern explains that in the tutor role, computers can provide instruction, feedback, and testing in all the different dimension of language and culture learning. Moreover, computers can be used as tools to gain access to written, audio, and visual materials relevant to the language and culture

being studied. And finally, in the medium role, technology is able to provide sites for interpersonal communication, multimedia publication, distance learning, community participation, and identity formation. Taking into consideration Kern's framework, WASP, Online Audio Recording, and Moodle were used in the role of tools, providing accessible English materials in the form of text, audio, and visuals which targeted the feature presented (intonation of tag questions). WASP and Online Audio Recording also worked as tutors, providing the participants with visual feedback, allowing self-assessment of intonation, and raising awareness of connections between language and culture, such as the pragmatically appropriate interpretation and use of tag questions. The connection between pronunciation and pragmatic competence is crucial for language learners who are developing their (receptive and productive) communicative abilities. For example, one of our participants explained that it was hard to communicate with people in English: "… when I wanted to ask questions in English, people didn't know is a question or sometimes I just said something generally but they felt like are you asking me?" (Yin). That is why learners wanting to improve their ability to communicate appropriately in the target language should develop their understanding of the link between pronunciation and pragmatics. According to Chun (2002), intonation should be viewed and interpreted based on the context it is produced in. Chun explains that it involves an ongoing process of interaction where the speaker reacts to the way the interlocutor is using intonation in such a way to either 'conform' to it or 'break away' from it when necessary.

According to Zhu (2012), one possible way of doing that is with the help of technological devices and tools that allow us to develop more communicative activities.

These communicative activities can help make the teaching and learning aspect of pragmatics more accessible and appropriate. Also, using more interactive and communicative activities and tools might help learners 'access and integrate the sociopragmatic and pragmalinguistic knowledge more quickly and efficiently' (Rose, 1994, as cited by Zhu, 2012 p.233). For example, Cohen (2008) conducted a small-scale research in February 2007 with advanced learners of Spanish to determine the impact of an interactive website 'Dancing with Words' on learners' pragmatic skills. The focus of the website was on strategies for learning and performing pragmatics through speech acts. The results revealed that the students improved their pragmatic performance and they have become more aware of the pragmatic meaning. Cohen explains that the results could be seen in the light of the role that technology plays in making pragmatics more accessible to learners.

This study investigated the feasibility of using computer-based training and feedback to help second language learners improve their perception and production of tag questions. The results revealed that the use of the two tools, WASP and Online Audio Recording, were beneficial, interesting to use and helpful for the participants' learning. Not only did the participants become more aware of their perception of tag questions, but they also developed in their awareness of the pragmatic meanings of tag questions. However, this study motivates further research to address issues related to computer-based training and feedback to help learners improve perception and production.

References

Altmann, H (2006). The perception and production of second language stress: A corss-linguistic experimental study. (Doctoral dissertation/Master's thesis). Retrieved from ProQuest Dissertations and Theses Database. (10,April 2013).

Audacity (2.0.3) [Software]. (2013) Available from http://audacity.sourceforge.net/

Bolinger, D. (1989). *Intonation and its uses: Melody in grammar and discourse*. Stanford, Calif: Stanford University Press.

Bradlow, A. R., Akahane-Yamada , R., Pisoni, D. B., & Tohkura, Y. (1999). Training Japanese listeners to identify English /r/and /l/: Long-term retention of learning in perception and production. *Perception & Psychophysics*, *61*(5), 977-985.

Burleson, D. F. (2007). Intonation patterns in English tag questions of Japanese speakers of English as a second language. *IULC Working Papers Online*, *7*.

Celce-Murcia, M., & Larsen-Freeman, D. (1998). *The grammar book: An ESL/EFL teacher's course*. Boston, Mass: Heinle & Heinle.

Chapelle, C. (2001). *Computer applications in second language acquisition: Foundations for teaching, testing and research*. Cambridge: Cambridge University Press.

Chun, D. M., Hardison, D. M., & Pennington, M. C. (2008). Technologies for prosody in context: Past and future of L2 research and practice. In J. G. Edwards, & M. L. Zampini (Eds.). *Phonology and second language acquisition* (pp. 323-346). Amsterdam: J. Benjamins.

Chun, D. M. (2002). *Discourse intonation in L2: From theory and research to practice*. Amsterdam: J. Benjamins.

Chun, D. M. (1998). Signal analysis software for teaching discourse intonation. *Language Learning & Technology*, *2*(1), 74-93.

Cobb,T. Web Vocabprofile [accessed 01 September 2012 from http://www.lextutor.ca/vp/], an adaptation of Heatley & Nation's (1994) Range.

Cohen, A. D. (2008). Teaching and assessing L2 pragmatics: What can we expect from learners? *Language Teaching*, *41*(2), 213-235.

de Bot, K. (1983). Visual feedback of intonation 1: Effectiveness and induced practice behavior. *Language and Speech*, *26*(4), 331-350.

DeKeyser, R. (2007). Skill Acquisition Theory. In B. VanPatten, & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 97-113). Mahwah, N.J: Lawrence Erlbaum Associates.

Felix, U. (2001). The web's potential for language learning: the student's perspective. *ReCall*, *13*(1), 47-58.

Gumperz, J. J. (1982). *Discourse strategies*. Cambridge [Cambridgeshire: Cambridge University Press.

Gupta, R. (2010). Language in India. *Using technology in the English language classroom*, *10*, 61-77.

Hardison, D. M. (2005). Contextualized computer-based L2 prosody training: Evaluating the effects of discourse context and video input. *CALICO Journal*, *22*(2), 175-190.

Hardison, D. M. (2004). Generalization of computer-assisted prosody training: Quantitative and qualitative findings. *Language Learning & Technology*, *8*(1), 34-52.

Heatley, A., Nation, P. (1994). Range. Victoria University of Wellington, NZ. [Computer program, available at http://www.vuw.ac.nz/lals/.]

Holland, V. M., Kaplan, J. D., & Sabol, M. A. (1999). Preliminary tests of language learning in a speech-interactive Graphics Microworld. *CALICO Journal*, *16*(3), 339-359.

Huckvale , M., (2012). WASP (1.53) [Software]. Available from Speech, Hearing and Phonetic Sciences UCL division of Psychology & Language Sciences: http://www.phon.ucl.ac.uk/resource/sfs/wasp.htm

Kang , O., Rubin, D., & Pickering, L. (2010). Suprasegmental measures of accentedness and judgments of language learner proficiency in oral English. *The Modem Language Journal*, *94*(4), 554-566.

Kern, R. (2006). Perspectives on technology in learning and teaching languages. *TESOL Quarterly*, *40*(1), 183-210.

Knudsen, A. (2011). Cross-linguistic interference of intonation in regard to tag questions in native, second and third language use (Doctoral dissertation/Master's thesis). Retrieved from ProQuest Dissertations and Theses Database. (29, Novermber 2012).

Litzenberg, J.J. (2003) Yes/No-Question Intonation of native speakers and international teaching assistants in academic discourse (Doctoral dissertation/Master's thesis). Retrieved from ProQuest Dissertations and Theses Database.(24, May, 2011).

Moodle, version 2.2 (2012) http://moodle.org (accessed on November, 2012).

NCH software (2008).Videopad (3.01)) [Software]. Available from NCH software: http://www.nchsoftware.com/videopad/index.html#102

Osuna, M. M., & Meskill, C. (1998). Using the World Wide Web to integrate Spanish language and culture: A pilot study. *Language Learning & Technology*, *1*(2), 71-92.

Park, M. W. (2011). Teaching intonation patterns through reading aloud (Doctoral dissertation/Master's thesis). Retrieved from ProQuest Dissertations and Theses Database. (03,January,2013).

Parmelee, J. (n.d.). Tag Questions Made Easy for Asian EFL/ESL Learners. *Employer Preferred TEFL Courses for Teaching English Abroad*. Retrieved from http://www.teflteachthai.com/Tag_Questions_Made_Easy.html

Pickering, L. (2001). The role of tone choice in improving ITA communication in the classroom. *TESOL Quarterly*, *35*(2), 233-255.

Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*(2), 129-158.

Spaai, G. W., & Hermes, D. J. (1993). A visual display for the teachin g of intonation. *CALICO Journal*, *10*(3), 19-30.

Stibbard, R. (1996). Teaching English intonation with a visual display of fundamental frequency. *The Internet TESL Journal*, *2*(8). Retrieved from http://www.aitech.ac.jp/~iteslj/Articles/Stibbard-Intonation/

Taniguchi , M., & Abberton, E. (1999). Effect of interactive visual feedback on the improvement of English intonation of Japanese EFL learners. *Speech, Hearing and Language: work in progress*, *11*, 76-89.

Taylor, R., & Gitsaki, C. (2004). Teaching WELL and loving IT. In S. Fotos, & C. Browne (Eds.), *New perspectives on CALL for second language classrooms* (pp. 131-147). Mahwah, New Jersey, U.S.A.: Lawrence Erlbaum Associates.

Thomson, R. I. (2011). Computer assisted pronunciation training: Targeting second language vowel perception improves pronunciation. *CALICO Journal*, *28*(3), 744-765.

Thorne, S. L., & Smith, B. (2011). Second language development theories and technology-mediated language learning. *CALICO Journal*, *28*(2), 268-277.

Verdugo, D. R., & Trillo, J. R. (2005). The pragmatic function of intonation in L2 discourse: English tag questions used by Spanish speakers. *Intercultural Pragmatics*, *2*(2), 151?168.

Wang, X., & Munro, M. J. (2004). Computer-based training for learning English vowel contrasts. *System*, *32*(4), 539?552.

Watson-Gegeo, K. A. (1988). Ethnography in ESL: Defining the essentials. *TESOL Quarterly*, *22*(4), 575-592.

Weinberg, A., Peters, M., & Sarma, N. (2005). Préférences des apprenants face à l'utilisation de la technologie dans l'apprentissage des langues. *RCLA/CJAL*, *8*(2), 211-231.

Wennerstorm, A. (1994). Intonational Meaning in English Discourse: A Study of Non-Native Speakers. *Applied Linguistics*, *15*(4), 399-420.

Zhao, Y. (2003). Recent Developments in Technology and Language Learning: A Literature Review and Meta-analysis*. *CALICO Journal*, *21*(1), 7-27.

Zhu, C., & Wu, X. (2011). A study of yes/no questions in English and Chinese: With special reference to Chinese EFL learners? understanding of their forms and functions. *Journal of Pragmatics*, *43*(2), 632?647.

Zhu, W. (2012). Polite requestive strategies in emails: An investigation of pragmatic competence of Chinese EFL learners. *RELC Journal*, *43*(2), 217 ?238.

Appendix A

Date of testing: _____ Participant Code_____

Language Background Questionnaire

Name: _____ Gender: Male____ Female _____

Phone number: _____ Email address: _____

Is your hearing normal as far as you know? Yes: _____ No: _____

Program at this university : _____

Date of beginning of full-time studies: _____

Expected graduation date: _____

Date of birth: _____ Birthplace (City, Country) : _____

What do you consider to be your native language?

Have you used this language since birth? Yes: _____ No: _____

What do you consider to be your second language? English:_____ Other:_____

At what age did you start learning your second language? _____

What language do you speak at home now? _____

What is the native language of your mother? _____ Your father? _____

In what language did you attend school in your country? Please circle the appropriate one

-elementary school: native language only English only Other: _____

-high school: native language only English only Other _____

-university: native language only English only Other _____

What ESL courses have you completed at this university? _____

What ESL courses are you taking now at this university? _____

Please rate your ability to speak, listen to, read and write **your native language** by using the scales in the box below. Please note that 1= extremely poor and 9= extremely fluent

Your native language **1=Extremely Poor** **9= Extremely Fluent**

speaking	listening	Reading	Writing
1 2 3 4 5 6 7 8 9	1 2 3 4 5 6 7 8 9	1 2 3 4 5 6 7 8 9	1 2 3 4 5 6 7 8 9

Please rate your ability to speak, listen to, read and write **English** by using the scales in the box below. Please note that 1= extremely poor and 9= extremely fluent

English **1=Extremely Poor** **9= Extremely Fluent**

speaking	listening	Reading	Writing
1 2 3 4 5 6 7 8 9	1 2 3 4 5 6 7 8 9	1 2 3 4 5 6 7 8 9	1 2 3 4 5 6 7 8 9

What other languages do you speak? How well do you speak them?

At what age did you start learning English? _____

Please list amount of time (per week) that you learned English in:

Primary school (time) _____ for (number of years) _____

Middle school (time) _____ for (number of years) _____

High school (time) _____ for (number of years) _____

University (time) _____ for (number of years) _____

In your country, did you learn any English outside of school? How?

_____

Have you ever lived in any country other than your native country and Canada? For how long?

For what purpose? _____

When did you arrive in Canada ? _____

What are your strengths in speaking English?

_____

What are your weaknesses?

_____

Appendix B

CONSENT FORM – Treatment groups

This is to state that I agree to participate in a program of student research being supervised by Dr. Sara Kennedy (LB 529-6, 1455 de Maisonneuve Blvd. West, Montreal, H3G 1 M8; phone: 514-848-2424 ext. 8935; email:sara.kennedy@education.concordia.ca) and conducted by Souheila Moussalli (phone: 514-578-8155; email: souheilamoussally@hotmail.com) of the Department of Education at Concordia University.

A. PURPOSE

I have been informed that the purpose of this research is to study how non-native learners of English develop in the perception and production of rising and falling tag questions using computer-based activities.

B. PROCEDURES

I have been informed that (1) this study will take place in computer Lab and in classrooms on the Sir George Williams campus at Concordia; (2) the tasks I will be asked to complete will take place over the Summer semester of 2011-2012. The tasks will consist of:

(a) filling out a questionnaire on-line

(b) engaging in three computer-based listening and speaking tasks twice a week for a period of 2 weeks

(c) participating in an interview

I have also been informed that (1) my performance in the tasks and interviews may be recorded; (2) some or all of these tasks may be carried out by Sara Kennedy, Souheila Moussalli, or another research assistant; and (4) my participation will extend over the Summer semester, and the total time required to complete all tasks will be approximately 1 hour per week for 2 weeks.

C. CONDITIONS OF PARTICIPATION

- I understand that I may decide not to participate in the experiment without any negative consequences.
- I understand that I am free to withdraw my consent and stop my participation at any time without negative consequences by phone or in writing.

- I understand that my participation in this study is confidential (i.e., the researcher will know but will not disclose my identity unless I give the researcher written permission to do so).
- I understand that there is a potential risk that I might find the tasks repetitive.
- I understand that there is a potential benefit of learning to better perceive and produce tag questions in English.
- I understand that the data from this study may be published or presented at a scientific conference; data will be reported in a way that protects each participant's identity.
- I will be participating in a draw at the end of the 2 weeks. After completing the tasks required, my name will be entered into a draw for a prize of a gift card or a meal valued between $30 and $50. Besides being entered in the lottery for the gift card, I will be paid $20 for completing the study: $5 at the beginning and $15 upon completion of the two weeks.
- I will receive at some later time a digital recording of my interview.
- I understand that if I request a copy of the final research report that one will be sent to me. I can make this request to a researcher while completing the consent form or later in writing.

☐ I wish to receive a copy of the final research report about this study.

I HAVE CAREFULLY STUDIED THE ABOVE AND UNDERSTAND THIS AGREEMENT.

I FREELY CONSENT AND AGREE TO PARTICIPATE IN THIS STUDY.

NAME (please print):

_____

SIGNATURE: _____

RESEARCHER SIGNATURE:

_____

DATE: _____, 2012.

If at any time you have questions about your rights as a research participant, please contact Kyla Wiscombe, Interim Research Ethics and Compliance Officer, Concordia University, at (514) 848-2424 ext. 7481 or by e-mail at: kwiscomb@alcor.concordia.ca

Appendix C

Sample Tag Sentences in Context

Ella and Dillan were at home getting ready to go to work. Ella was having breakfast when Dillan came rushing into the kitchen and started moving everything around. Ella looked at him and said "It happens every morning. ***You've lost your keys again, haven't you?'*** (UP / DOWN)

Dillan didn't answer and kept moving objects around until the entire room was a big mess. Dillan turned to Ella and said, "I have no idea where I put them! Maybe you can help. ***You don't know where my keys are, do you?"*** (UP / DOWN)

Ella thought carefully. "Let me think… I know your habits well. ***You always put your keys on this table, don't you?*** (UP / DOWN) But they aren't there now. Why not? " Dillan said, "Well, when I came home last night, my arms were full of groceries. I couldn't also put down my keys. It would have been impossible." Ella asked, "What did you do after you came in with the groceries?" Dillan answered, "Why do you want to know that? I walked into the kitchen and put the groceries away. ***I didn't put the food in the wrong place, did I?"*** (UP / DOWN) "Aha!!" said Ella. "Wait here." Ella walked to the refrigerator and opened the door. The keys were lying on the shelf next to the orange juice!