Effects of Joint Attention and Tone on Adult Canadian English Speakers learning Mandarin Words

by

Elizabeth M Rousseau, B.A., Carleton University, 2011

Thesis
A thesis submitted to the Institute of Cognitive Science in partial fulfillment of the requirements for the degree of

Master of Cognitive Science

Carleton University, Ottawa
April 17
Acknowledgements

I would like to thank my supervisors, John Logan and Randall Gess, whose time and efforts made this work possible. I would also like to recognize James Howell, Peter Wozniak, and Lubna Safaa, whose contributions made it likely. Finally, I would like to show my appreciation to family, who made it happen.
Abstract

Effects of Joint Attention and Tone on Adult English Canadian Speakers learning Mandarin Words

Elizabeth M Rousseau, M. Cog. Sc.
Carleton University, 2014

Supervisors: Randall Gess & John Logan

Joint attention is regarded as a key factor in the development of communication skills, theories of the mind, and language. Joint attention is when our focus shifts, following the gaze of another person. This ability appears to play a large part in developing language skills. This study examines joint attention in the context of adult second language word learning to determine if it enhances attention to the perception of unfamiliar speech sounds. Naive participants (participants with English as their primary language) were instructed to match unfamiliar Mandarin syllables varying in tone to images, while a video partner was viewed examining the same sounds and images. Participants were tested on how many sounds they could learn to pair with images in several attention conditions over five blocks of trials. The overall goal was to determine the influence of the video partner on word learning. Results indicated that some attention conditions were more effective in promoting the learning of the sound-picture pairs than other attention conditions, and that the effect of attention condition interacted with type of tone. Performance also showed a modest improvement over the course of several
blocks, with the largest improvement occurring in the first three blocks. The results provide insight into how components of language are learned by adults, and especially how the influence of social context can facilitate the learning process.
Table of Contents

List of Figures ........................................................................................................ vi
1. Introduction ........................................................................................................... 7
2. Goals of the Present Study .................................................................................. 8
3. Joint Attention ..................................................................................................... 9
   3.1 Infants and Joint Attention .............................................................................. 9
   3.2 Eyes and Social Awareness .......................................................................... 10
   3.3 Cognitive Processes Related to Gaze in Adults ............................................. 12
   3.4 Cognitive Processes of Joint Attention ......................................................... 14
4. Attention and Memory .......................................................................................... 17
5. Mandarin ................................................................................................................ 20
   5.1 Mandarin as a Second Language ................................................................... 20
   5.1 Mandarin Tones .............................................................................................. 21
6. Method .................................................................................................................... 23
   6.1 Participants ...................................................................................................... 23
   6.2 Measures .......................................................................................................... 23
   6.3 Stimuli ............................................................................................................. 24
   6.4 Procedure ......................................................................................................... 27
7. Results ..................................................................................................................... 30
   7.1 Analysis ........................................................................................................... 30
   7.2 The Relation between Joint Attention and Autism ......................................... 35
   7.3 The Relation between Working Memory and Word Learning ....................... 36
8. Discussion ................................................................................................................ 38
   8.1 Joint Attention as a Learning Tool ................................................................. 38
   8.2 Memory and Attention .................................................................................. 40
   8.3 Bilingualism .................................................................................................... 41
   8.5 Limitations ...................................................................................................... 41
   8.6 Time Variability in the Separate Conditions ................................................. 42
9. Conclusion ............................................................................................................... 43
References .................................................................................................................. 44
Appendix ..................................................................................................................... 49
List of Figures

FIGURE 1. PITCH contours of MANDARIN TONES ................................................................................................................22
FIGURE 2. DIAGRAM of the different gaze conditions created by adjusting the timing of stimuli in a trial .................................................................26
FIGURE 3. DIAGRAM of the learning and testing phase sequence, hereafter referred to as blocks ........28
FIGURE 4. MEAN proportion correct across training blocks ..................................................................................................................30
FIGURE 5. PERCENT of correct responses for each joint attention condition ..........................................................31
FIGURE 6. PERCENT correct of each tone pair ..........................................................................................................................32
FIGURE 7. PROPORTION correct of both joint attention condition and tone ..........................................................33
FIGURE 8. TOP vs. BOTTOM performance across blocks ................................................................................34
FIGURE 9. INTERACTION of WM and word learning .................................................................................37
FIGURE 10 Top vs. Bottom Working Memory Score across Blocks ........................................................................38
1. Introduction

Vocabulary learning can be one of the most daunting tasks for a second language (L2) learner. Speech sounds have to be accurately perceived and stored in long-term memory, and the sound patterns need to be mapped to the corresponding semantic and lexical information. Furthermore, vocabulary learning in a L2 can vary dramatically from person to person. Many factors contribute to individual differences in vocabulary learning performance. External factors that influence word learning success include the respective phonemic inventories of both the first language and target language (Flege, 1995; So & Best, 2010), the ease with which the concept can be linked to the given sound, and how the words are taught (Díaz et al., 2008; Hanulíková & Dedi, 2012; Yeung, Chen, & Werker, 2013). Individual differences also play a large role in word learning. These individual variables include cognitive ability, which also comprises attention and memory capacity differences, personality traits, motivation, and identification with the culture of the target language. All of these can influence the rate and accuracy of word learning (MacIntyre & Charos, 1996). Although word-learning is generally considered to be a process in which ultimate success is dependent on the characteristics of the individual, recent work has begun to examine how cognitive processes such as word learning are affected by social context. Social context is shaped by both the environment and the individuals participating in it. Language is a prime example of a social skill, and
exploring how social factors influence language learning could provide insights into
general cognitive processes as well as how humans use social contexts to learn.

2. Goals of the Present Study

In this thesis, I used a novel experimental paradigm to examine how much influence joint attention exerts on adult second language learners. Previous studies using live interaction paradigms have determined that joint attention promotes the allocation of additional resources to the adult word learning process. The goal of the study was to examine joint attention’s influence on word learning, specifically whether it can facilitate the learning of unfamiliar linguistic features. The present study employed a novel video interaction paradigm in which participants learned new words in four different social contexts. In the experiment, a video of a learner is presented to the participant on a screen, while the participant is learning non-word sound-image pairs. The minimal pairs were Mandarin monosyllables contrasting pairs of Mandarin tones. To create the joint attention contexts, a series of videos were made that depicted the face and head of a person, which was styled in appearance to resemble as closely as possible the avatar that was created in a joint attention study by Kim and Mundy (2012). The videos were manipulated to create four different interaction patterns that fall under the joint attention paradigm. The different interaction patterns were created by varying the direction and timing of the video learner’s eye gaze.

In addition to providing insight into how social contexts may influence word learning, research in this area may be constructively used to improve language programs where learners do not have the benefit of a social group to inform word learning. The
objective of the experiment was to replicate and extend findings obtained by Shimada Hirotani, Okazaki, Tanabe, and Sadato (2012) who found that specific joint attention patterns promoted word learning. While Shimada et al.’s study used a live interaction paradigm and English non-words with Japanese students, the present study is a video interaction paradigm that uses Mandarin words with English-speaking participants. The premise for the video-interaction paradigm was taken from Kim and Mundy (2009) who examined the interaction between image recognition and joint attention. Kim and Mundy successfully used an avatar that responded to eye movement to create the joint attention contexts.

3. Joint Attention

3.1 Infants and Joint Attention

Insight into how infants learn words and develop social knowledge may be foundational to understanding how social contexts apply to word learning in adults. Infants are believed to learn words through basic memory and attentional skills (Samuelson & Smith, 1998). Infants follow social cues such as head motions, pointing, and gaze to determine what objects are being referenced (Mundy & Jarrold, 2010). This creates an environment where both participants are attending to the same object, with the awareness that their focus is shared. This social context is referred to as joint attention (Kim & Mundy, 2012). Developmental research has found that joint attention plays a key role in word learning among infants. Baldwin suggested that infants’ comprehension of social contexts and non-verbal cues are necessary for successfully mapping labels to objects, skills which are key to vocabulary learning (Baldwin et al., 1996; Baldwin, 1991). Infants at nine months are already attending to the movements of the head for social cues, and subsequent development is swift (Carpenter & Nagell, 1998). Following
adults’ gaze and engaging in joint attention with parents predicts not only subsequent receptive vocabulary but also demonstrates the progression of the infants’ developing social awareness (Tomasello & Farrar, 1986). Laboratory studies have also replicated the effects of joint attention. In a electroencephalogram (EEG) study, Hirotani, Stets, Striano, and Friederici (2009) found that even within the span of one examination session, 20 month-old infants demonstrated greater word learning of novel words in conditions where attention was mutually engaged by the infant and the experimenter than in non-joint attention conditions. An experimenter verbally labeled both known and unknown objects in two different conditions. In the joint attention condition, the experimenter initiated eye contact with the infant before labeling the object. In the non-joint attention condition, the experimenter looked at the infant but did not meet their eyes while labeling the object. In the testing phase, congruent and incongruent image-label pairs were presented. The event related potentials (ERP) evoked by the study were found to be more negative in the incongruent condition than the congruent, and the negativity was significantly more pronounced for the images introduced in the joint attention conditions. Therefore, the infant appeared to be more aware of incorrectly matched stimuli that had been initially introduced in the joint attention condition. Bayliss, Bartlett, Naughtin, and Kritikos (2011) attribute the improved word-learning abilities of the infants to the eye-gaze conditions aiding the infant’s focus and guiding it to the proper target, as well as promoting the social motivation to communicate.

3.2 Eyes and Social Awareness

Eyes play a significant role in developing social understanding in joint attention contexts. The eyes cue observers to the individual’s focus of attention, which is then used
to infer the intentions of the observed individual (Mundy & Newell, 2007). Eye gaze is so central to the human experience that cognitive processes related specifically to gaze following and social development start almost immediately after birth (Tomasello & Farrar, 1986). Aside from language development, eye gaze plays a key role in infant mental development. It has been theorized that through observing their own actions in conjunction with the behavior of others infants create concepts of both the self and others. (Mundy & Jarrold, 2010). Furthermore, the ability to share a conceptual space through mutual attention to an object facilitates the shared intentionality that makes human collaboration possible (Baldwin & Markman, 1989; Baldwin, 1991; Tomasello, Carpenter, Call, Behne, & Moll, 2005).

A notable exception to typical joint attention development is the case of people with autism and Asperger’s Syndrome. Limited or nonexistent joint attention behavior is a predictor of autism. People with autism are noted for their reduced social abilities and impaired social development, and can also have restricted language development, such as a limited vocabulary (Mundy & Newell, 2007; Tanabe et al., 2012; Williams, Waiter, Perra, Perrett, & Whiten, 2005). Furthermore, individuals with a few of the characteristics of autism are also observed to have reduced sensitivity to the effects of gaze cues (Frischen, Bayliss, & Tipper, 2007). Conversely, these same types of individuals appear to have enhanced perceptual processing, with a greater sensitivity to pure tone cues as well as segmental information (Bouvet, Mottron, Valdois, & Donnadieu, 2013; Yu, 2010).
3.3 Cognitive Processes Related to Gaze in Adults

Biologically, human eyes are ideal for indicating gaze direction because the large amount of white sclera on either side of the relatively small iris makes the task of discriminating gaze direction much easier than it would be in other primates. Humans also have some predispositions when it comes to perceiving gaze. People are extremely sensitive to direct gaze, which captures the viewers’ attention. Sato and Itakura (2013) examined the relationship between gaze and attention in a study that also explored learning. Participants looked at a picture of an individual that was depicted either with eyes directed forward, toward the viewer, or with gaze averted to the side. While attending to the picture, the participants were told to press two buttons at will, provided they pressed both an equal number of times. While the participants were not told of any specific associations between the buttons, pressing the buttons changed the picture to depict the tongue sticking out or ballooning the cheeks. After performing this task, participants switched to a new task where the participants were told to deliberately change the image to either sticking out the tongue, or puffing the cheeks. Quick adaptation to this task would indicate that the participants had made the association between the buttons and the images presented in the previous task. Participants that had the gaze forward condition, with the face that appeared to be looking at them, were quicker to map the association between their actions and the different images than participants who were looking at the averted gaze version.

A degree of automatic cueing occurs when perceiving shifts in gaze that may automatically engage joint attention. Driver, Davis, Ricciardelli, Kidd, Maxwell and Baron-Cohen (1999) found that spatially orienting objects was facilitated by viewing a
face that was attending the same direction, even when, in the course of the study, gaze direction had no relevance to the location of the objects. They found that reaction times (RTs) were smaller when the gaze was congruent with the location of the object. This effect persisted even when it was more likely that the object would appear in the opposite direction than the depicted gaze, thus demonstrating an automatic tendency to follow gaze cues.

Bayliss et al. (2011) observed that our response to shifts in eye gaze is an automatic process, in a study in which eye gaze cues were changed to other directional cues, such as arrows and pointing. These different directional cues did not have the autonomous processing that occurred in response to the eyes. Frischen, Bayliss, and Tipper (2007) have proposed that several cognitive processes work in concert to coordinate joint attention. The superior temporal suculus (STS) is active in social perception, which means it contributes to joint attention, since the STS is responsive to all kinds of body movements, in particular the eyes and mouth. It has been found that while cells in the STS can be tuned to specific regions of the body, such as just the head or a hand, these cells can also be influenced (as measured by response magnitude) by other parts of the body, such as gaze direction. So we react more quickly to someone moving left if they are already looking in that direction. This would indicate that the STS is part of a distributed network that collaborates to determine another person’s intentions. Brain regions also perform specific roles, such as determining if a gaze shift is to the left or right, not just forward or averted, indicating particular sensitivity to directional information. Gaze processing would even appear to be influenced by the perceived goal of the eye movement, suggesting that not all eye shifts are assessed equally. The authors suggested that individuals seem to be able to tell the difference between random eye
shifts and purposeful ones. Importantly, the STS is more responsive to eye movements that appear to provide meaningful information.

3.4 Cognitive Processes of Joint Attention

Having learned from infancy that people are thinking about what they are looking at, adults are quite sophisticated at using mental models to infer the goals of others based on direction of gaze. Some research suggests that our mental models of others based on gaze provide adults a cognitive boost in a variety of tasks. Böckler, Knoblich, and Sebanz (2011) created a task where participants looked at pictures of hands, and had to identify whether or not the hand was the right or left hand. The images were rotated to varying degrees to add complexity to the task. Participants responded faster when observing the images in a joint attention condition than in the control condition. Böckler et al. suggested that the joint attention condition somehow improves the saliency of the stimuli, and the presence of a fellow participant may also improve motivation. Kim and Mundy (2012) tested stimulus saliency, attention and memory with an avatar replacing one of the participants. The individual participants were monitored with a head-mounted visual display system which was integrated with an eye-tracker. The participants saw a digital avatar on the display. In a trial, images were presented to the right and left of the avatar, which either initiated looking at the images, known as reactionary joint attention from the perspective of the participant, or appeared to follow the participant’s gaze, meaning that the participant initiated the joint attention condition, known as initiating joint attention. Afterwards, the participants were tested by asking them to identify what image they saw. There were three different types of images: faces, buildings, and abstract shapes. An interaction was found between condition and image type. House and abstract
images observed in the initiating joint attention condition were correctly identified more frequently than in the reactionary condition. Overall, Kim and Mundy found that interacting with the gaze of an avatar improved the participants’ ability to recall the images later, suggesting that the simulated joint attention condition resulted in an improvement in memory and attention.

So far, these studies have been attempting to use observed behavior as an indicator of cognitive processes. Williams et al. developed video stimuli in order to examine joint attention in an fMRI study. They created videos of the head and shoulders of a model in the top two-thirds of a screen, while the bottom third was a flat colour. The bottom third of the screen was digitally manipulated to make it appear that a red dot was moving around on it. In one condition, the head and eye movement of the model closely follows that of the dot, as if the model is watching it. In the second condition, the model’s movements do not match those of the dot. Participants in the fMRI watched both types of videos, simply being told to follow the dot with their eyes. What they found was that in the condition where the participant and the model in the video appeared to be tracking the red dot together, brain regions were being activated beyond what the basic dot-tracking required. The additional regions included the right ventral medial frontal cortex, the left anterior frontal cortex, and the right anterior frontal lobe. The ventromedial frontal cortex has been implicated in “theory of mind” processes. This brain region had not been observed to respond to eye movement, but Williams et al. believe that their results support the view that mutually directed attention is a type of joint attention, and its overlap with theory of mind processes support their linked development. The activation of the left superior frontal gyrus, suggesting executive functions, was the other marked difference in the joint attention condition. Its activation would seem unusual, as both tasks were identical. However, if the brain was abstracting the model’s perceived
processes with their own behavior, then such additional activations would not be surprising, and would provide further evidence of joint attention functions. They also found that regions of the brain that control attention had greater activation in the joint attention condition, adding evidence to the earlier assertion that observing gaze facilitated one’s own attentional shifts. A concern in the study was the lack of intentionality between the participant and the model – there was no link between the two except that they were performing the same behavior (Williams, Waiter, Perra, Perrett, & Whiten, 2005).

So how does joint attention influence adult language learning? Shimada, et al. (2012) used a live interaction paradigm to determine if joint attention affected word learning. The participants were Japanese speakers learning English non-words. Eighty English non-words were paired with eighty unfamiliar images, and these sound-image pairings were taught to the participants by displaying the image on a monitor and playing the sound over headphones. In the experiment, pairs of participants were seated across from each other at a table, with monitors set up at either end of the table, to the right and left of the participants. Participants were told to listen for an audio cue in one ear of their headphones, and to look towards the monitor in the direction of the cue where they would see an unfamiliar image and hear the word associated with it. At the beginning and end of every trial, they were to look forward and make eye contact with their fellow learner. By adjusting the audio cues to the participants’ headphones, the researchers were able to construct four different learning conditions. The first condition was the non-joint attention condition, wherein the participants looked toward different monitors. The second condition was the simultaneous joint attention condition, which was coordinated so that the participants looked toward the same monitor simultaneously. The third and fourth conditions were the initiating joint attention condition, where one participant
looked first, and the other followed. The following participant experienced the reactionary joint attention condition, as they were following the attention of the other participant. The head-turning behavior was monitored by cameras in the room to account for head motion, while the eyes were monitored with a head-mounted eye-tracker on each participant. After all the sound-image pairings were presented, the learning phase was complete and the participants moved on to the testing phase. Participants were tested by asking them to match the words with the image they were associated with. All the words were tested in each testing phase. The word-learning phase alternated with the testing phase five times. The simultaneous joint attention condition showed the best word-learning performance, though it was not markedly better than any of the other conditions. That simultaneous joint attention was more effective than initiating joint attention was not predicted by the results of earlier literature such as earlier studies on infants, or Kim and Mundy’s (2012) study on adults using an avatar. Shimada et al. did not speculate about the results, saying further research was needed before any conclusions could be drawn about the differences between adult and infant word learning in social contexts.

So far, the studies reviewed to this point have established that people are sensitive to gaze cues, and the information that can be imbedded within said cues. That said, there are additional factors that can influence an individual’s receptiveness to eye gaze cues, including attention and memory.

4. ATTENTION AND MEMORY

The increased attention provided by automatic cueing through shifts in gaze may affect the working memory of the participant. Baddeley’s model of working memory (Baddeley & Hitch, 1974; Baddeley, 2002) consists of a central executive that operates
two “slave systems”, the phonological loop and the visual-spatial sketchpad, and in later forms of the model, the addition of an episodic buffer. It was proposed that working memory handles immediate sensory and long-term memory inputs to enable complex cognitive activities. Performance on tasks that test working memory, such as operation span or reading span tasks, have been found to correlate with a variety of other demanding cognitive tasks. Baddeley argues that the main role of the phonological loop is facilitating language learning. Baddeley, Gathercole, and Papagno (1998) found that Finnish children who received high scores in English non-word repetition (which is thought to rely exclusively on the phonological loop) had both a larger Finnish vocabulary as well as better foreign language abilities two years later. They propose that the phonological loop is responsible for bridging the gap between speech input and the creation of phonological representations that allow for long-term storage. Once a structure is in the mental lexicon, it is easy to retrieve and produce. The phonological loop mediates between the input and the eventual structural representation. However, as noted above, the performance of the phonological loop can vary from individual to individual, contributing to differences in second language performance among typical populations. Baddeley et al. (1998) observe that polyglots perform significantly better than average on phonological tests, suggesting that a high-performing phonological loop is required for advanced second language learning.

But what underlies high levels of performance in the phonological loop, or even in working memory in general? Executive attention may be a defining factor for the overall effectiveness of the working memory system. Performance on working memory tasks, such as reading span or operation span tasks, is correlated with fluid intelligence and demanding cognitive tasks such as reading comprehension and importantly, language learning (Conway, Kane, & Engle, 2003; Turner & Engle, 1989). Individuals with equal
skills in the processing tasks (the reading or operation span tasks) still displayed variability in the number of words recalled. This difference has been ascribed to differences in the attentional system. Attention is defined as the ability to maintain items until they are ready to be retrieved. Engle found that high working memory individuals had significantly greater attention abilities than low-working memory individuals in a number of attentional tasks (Engle, 2002), a finding consistent with other researchers (Lépine, Barrouillet, & Camos, 2005; Lépine, Bernardin, & Barrouillet, 2005). Furthermore, other researchers theorize that the mechanisms of attention and working memory are the same as those that handle perception and attention in the real world (Jonides, Lacey, & Nee, 2005). Attending to the actions of others may be viewed as a form of Hebbian learning, (colloquially summarized as “cells that fire together, wire together”) which may prove fruitful for strengthening neural pathways, and reinforcing learning (Page & Norris, 2009; Postovsky, 1974; Tomasello et al., 2005).

Learning unfamiliar words may fall under the theoretical domain of the Auditory Short Term Memory, which is the ability of the brain to retain sounds, both verbal and non-verbal, such as music. The Auditory Short Term Memory holds items for integration into larger elements over time. In the case of language, this would be sounds on the phonetic level, before they are categorized. Using this model, Nolden, et al. examined what regions of the brain were activated when remembering pure tones. Using magnetoencephalography, they attempted to investigate the effect of interfering tones, They found that the superior temporal gyrus in both hemispheres was involved in the memory task, as was the inferior frontal gyrus, and the parietal structures. Nolden et al. (2013) also suggested that further regions of the brain may become active if the tones carried linguistic information. Furthermore, these nonverbal short-term memory items are susceptible to decay over delays as brief as 30 seconds (Mercer & McKeown, 2013).
These studies deliberately created pure tonal stimuli in order to determine if there were specific regions of the brain dedicated to processing tone. More work has to be done with tone in a linguistic framework to ascertain how much processing is shared between pure tones and tone in language.

5. Mandarin

5.1 Mandarin as a Second Language

As noted earlier, the relationship between the characteristics of the first (L1) and second language (L2) affects L2 word learning. The focus of the present study is the role of joint attention in English speakers learning Mandarin words. The contrast between English and Mandarin is a prime example of the phonological disparity that can make language acquisition challenging. English uses pitch to provide information about what is being said, communicating the whether the speaker is asking a question, making a statement or command, while also making clear what information in a sentence may be new. In Mandarin, four different pitch patterns are also used to contrast words lexically. These tonal contrasts are observed to be challenging for English speakers to discriminate (Wayland & Guion, 2004).

Several studies have examined the effects of acoustic training to improve English listeners’ abilities to perceive Mandarin tone. Most found that with appropriate training, English speakers show improvement in perceiving Mandarin tones (X. Wang, 2013; Y. Wang et al., 1999; Wayland & Guion, 2004; Liu et al., 2011). All the studies were concerned with improving the perception of the English learners in order that they might build new phonological categories for the tones. Liu et al. (2011) used multimodal teaching systems, pairing acoustic stimuli with visual cues to promote perception. The
visual cues included a romanized spelling of the word, and a system for expressing the
tonal pattern, either as a numerical, or a visual depiction. The multimodal system yielded
varying degrees of success, with the more salient visuals – the depictions of the contours –
aiding the word-learning process more than the symbolic aids, the numerical labels.
Instead of using multimodal systems, Wayland and Guion (2004) focused on contrasting
the tones with each other for the participants to identify so that, as in natural learning
conditions, the learners would learn the differences on their own. Their timeline was only
five training sessions over five days. Although most of the studies found an overall
improvement on all participants’ ability to perceive the tones, those studies that examined
individual results noted substantial variability in terms of how successfully each
participant identified tones (X. Wang, 2013).

5.1 Mandarin Tones

Mandarin is a language that uses level and contour tones. There are several
acoustic cues for Mandarin tones. F0 height, F0 shape, and (particularly to assist in
contrasting tones 2 and 3) F0 contrasts in the onset and turning point of the tone (Lei &
Zhang, 2008). Figure 1 below is a visual approximation of the pitch for each of the
Mandarin tones.
Currently, the results are mixed on whether tonal information is handled by linguistic processing or some other nonlinguistic brain process (Benassi-Werke, Queiroz, Araújo, Bueno, & Oliveira, 2012). Therefore, we can’t be sure whether general linguistic experience would predict abilities to learn tone.

When learning language, the first thing infants acquire is familiarity with the prosodic contours of their own language. Infants show preferences for prosodic contours of their own language before any other language feature. They prefer the pitch and stress patterns of their own language at 4-6 months, while preferences for vowels and consonants appear at 6-8 and 8-12 months respectively (Yeung, Chen & Werker, 2013). Examination of second language speakers of English found that pragmatic cues imparted through intonation were not used (Chen, 2009). While intonational information is not commonly taught, this shows that the second language learners were not able to reproduce these patterns on their own. Invariably, their own production lacked the intonational features that carried pragmatic information. In contrast, it is not uncommon for teachers of Mandarin to use comparisons with English intonational contours to assist learners with perceiving the different types of Mandarin tones. Current explorations of
teaching materials would suggest that the intonational contours of Mandarin sentences are not made explicit. What all this suggests is that suprasegmental features may be uniquely challenging to second language learners.

6. Method

6.1 PARTICIPANTS

Participants were recruited using the Carleton’s psychology recruitment system, (SONA). Fifty-one students (36 female, 15 male) participated in the study. All had normal or corrected-to-normal vision, with no hearing problems or reported neurological disorders. The age range was 19-32 with a median age of 20.

6.2 MEASURES

The short form Autism Quotient test from the Autism Research Centre in Cambridge was administered to all participants. The Autism Quotient test is comprised of 50 questions (e.g., “I am fascinated by dates”) that are scored on a four-point scale from ‘definitely disagree’ to ‘definitely agree’. A language experience questionnaire (See Appendix F) was also administered to determine what languages were spoken at home or at school and when the participant was growing up. Previous language experience may suggest higher working memory abilities and other better-than-average language skills, and this information would identify these additional variables among participants.
6.3 Stimuli

Three types of stimuli were used in the main experiment. First, the same non-images from the Shimada et al. (2012) study were used. These non-images were originally developed by Kroll and Potter (1986). The images are black and white depictions of unfamiliar objects (See Appendix A for examples). Kroll and Potter had the images rated for their unfamiliarity, and these ratings were used to ensure that the four groups of images have equivalent average scores for familiarity. The 88 images were blocked into four groups, all with an average familiarity rating of 4.37.

Second, a list of Mandarin monosyllabic words was selected. The recorded Mandarin words were acquired with permission from Yabla.com, an online language teaching program. The words were recorded by a speaker from Shanghai. The words were reviewed by a Mandarin teacher who is a native speaker of Mandarin to ensure 100% identification of tone and standard pronunciation. The Mandarin words were chosen based on whether their consonants would appear to be familiar to English consonants (as judged by the author). The 88 words were composed of 44 minimal pairs. The minimal pairs contrasted different tones. There were four tone pair contrasts, with 11 contrasts in each group. The tone pairs contrasted tones 2 and 3 (I), tones 1 and 4 (II), tones 3 and 4 (III), and tones 1 and 2 (IV) (See Appendix B for Mandarin stimuli).

Every word was matched with two 188-pixel non-images. One was the target image, the image that the participants were required to associate with the sound, while the other was a distractor image, which they were to ignore. The target image was always highlighted with a 2 pixel wide orange border.

The third stimulus was the video for the social contexts. There were seven videos of the head and shoulders of a female individual, on a black background, looking directly forward at the viewer. At specifically timed intervals the eyes shifted to the left or to the
right. The videos were recorded against a dark background with even lighting on the face. The size of the final video image was 378x160 pixels. There were four different learning conditions that each participant experienced: gaze forward condition, which was simply a baseline condition, and three joint attention conditions: initiating joint attention, responsive joint attention, and simultaneous joint attention. These different gaze conditions were created by varying the times that the image and audio stimuli appeared, and when the video participant’s gaze shifted. Because of the nature of the different conditions, not every element could be made identical in all the conditions. While all the conditions matched in total length, and the duration of the condition after the sound had been presented, the rest of the elements all varied slightly. The timing of the gaze shifts were made consistent, while still being judged (by the author) as natural. All the eye shifts take 65 ms to complete, if they occur in a trial. All the trials have the same total duration of 4290 ms. See Figure 2 for an illustration of the different timings of the beginning of the trials. The three joint attention conditions were modeled as closely as possible to Shimada et al.’s study.
Figure 2. Diagram of the different gaze conditions created by adjusting the timing of stimuli in a trial

The following is a breakdown of the specific timings within each different gaze condition. In the gaze forward condition, the eyes of the ‘learning partner’ remained straight ahead the entire trial, meeting the gaze of the participant. 1155 ms into the start of a gaze forward trial, the images and the audio appeared simultaneously. The image remained present for 1980 ms, then disappeared and the trial ended 1155 ms later. This condition was a baseline condition for all the trials. In the initiating joint attention condition, the images appeared 495 ms into the trial. The video reaction was delayed for an additional 495 ms, and the gaze shift took 65 ms to move. The image persevered for
1980 ms. The eye gaze shifted forward after the image disappeared, and remained forward, taking 1155 ms before the trial ended. In the responsive joint attention condition, the eyes shifted 595 ms after the start of the trial, and remained for 495 ms, when the images appeared. 495 ms later, the sound played. During the 1980 ms that the image remained, the video partner’s gaze returned to the front, so the eyes were forward for the last 495 ms after the image disappeared.

In the simultaneous joint attention condition, the image appeared and the audio played 1155 ms into the trial. However, the video partner did not shift their eyes for another 495 ms. The eyes shifted back after 1980 ms, and the images disappeared.

6.4 Procedure

After giving their informed consent, participants were asked to sit inside a sound attenuated booth. Participants put on headphones and sat facing a 15” laptop. Before starting the task, the participants read through instructions and performed a practice session, which was exactly the same as the regular session, but the image and sounds were substituted with images from the Wingdings font and English non-words, instead of the non-images and Mandarin words in the main task. The content was also abridged, using only nine sound and image items. After the participants were trained in the procedure, they began the main task. The main task was composed of two phases, a learning and a teaching phase. The learning and teaching phases alternated, each occurring five times. (see Figure 3 for an illustration).
All 88 sound-image pairings were presented once in each phase. In the learning phase, the participants were asked to map non-images to audio. The participants viewed the computer monitor, which displayed the video ‘learning partner’ in the centre of the screen. The images appeared on the left and right side of the face of the learning partner. Aside from a fixation cross at the beginning and end of each trial, there were no instructions dictating where the participant must look. During the learning phase, the different gaze conditions were presented randomly. The location of the target image was randomly distributed on the right and left, with the target image being present on the left half the time (see Appendix C for examples of what the participant sees during the learning session). After all 88 sound-image pairings were presented, the learning phase ended. No response was required from the participant during the learning phases.

After every learning phase there was a testing phase. In the testing phase, participants listened to sounds and then indicated out of four images which image the sound had been paired with. While looking at a fixation cross, the participants were presented with the audio stimuli over headphones. After the sound ended, and a delay of 500ms, four non-images from the learning phase were displayed on the screen. The images included the correct answer, the image associated with the minimal pair of the correct answer, and two other random images. The participant indicated which image they thought was the correct image for the sound that they heard by clicking on the image
they chose with the mouse (See Appendix D for an example of a testing phase screen). After a 1000ms delay, the next trial was presented automatically. This continued until all the sounds from the learning phase were tested. If the participants were unsure about the answer they were asked to guess. Response time and accuracy data were gathered from the testing phase. After the testing phase ended, the participants repeated both phases again four more times, for a total of five repetitions for each phase, hereafter referred to as blocks. This procedure was modeled from Shimada’s study.

After the main task was completed, participants filled out three questionnaires, outlining their language history, musical background, and social tendencies (to identify autistic behaviours) (See Appendix E for the Music Questionnaire, Appendix F for the Language Questionnaire, and Appendix G for samples of the Autism Quotient Test). Finally, participants were tested on their reading span, using stimuli from Engle’s automated reading span task (Engle, 2010). In the reading span task, participants performed three practice spans of two trials each. A trial consisted of reading a sentence aloud along with an audio recording, and remembering a letter. The participant also had to make true/false judgments on the truthfulness of the sentences. At the end of each span, the participant was prompted to write down the letters. The actual task consisted of spans that varied between three to five trials. The task is scored on two variables: the number of perfectly recalled letter spans, and the total number of correctly recalled letters.
7. Results

7.1 Analysis

Mean percent correct values were calculated for each participant as a function of block, joint attention condition, and tone pairs. Chance performance was 25% given that participants had four alternative responses to choose from. Results were analyzed using a three-way repeated measures analysis of variance (ANOVA), with four levels for the joint attention variable, four levels for the tone pair variable, and five levels for the block variable. Significant main effects for each variable were observed, along with a significant two-way interaction between joint attention conditions and tone pairs.

Figure 4 shows the mean percent correct across the five training blocks, with the error bars indicating 95% confidence intervals.

![Figure 4. Mean proportion correct across training blocks](image)

Performance in Block 1 was 27.6% correct responses, close to chance, whereas by Block 5, performance was 44.4% correct. This pattern of results indicated participants improved by approximately 17% over the course of the experiment, an effect confirmed.
in the ANOVA, $F(4, 1) = 40.79$, $p = 5.678e-25$, $\eta^2 = .0707$. The lack of overlap in the error bars for the blocks indicates that all blocks were significantly different from each other, with the exception of blocks 4 and 5.

Figure 5 shows the percent correct responses for each of the joint attention conditions, and the baseline condition.

![Figure 5. Percent of correct responses for each joint attention condition](image)

Performance was consistent across joint attention conditions, with percent correct values centered around 38%. The exception was for the Responsive Joint Attention condition, in which performance was lower than the other joint attention conditions, with 34.04% correct, $F(3, 1) = 9.39$, $p = 9.989e-6$, $\eta^2 = .0071$.

Figure 6 shows the percent of correct responses for each of the tone pair conditions.
Performance was best for tone pairs 3 and 4, at 39.6% and 37.96% correct, respectively, with tone pairs 1 and 2 at 36.51% and 34.54% correct, respectively. This difference was significant, $F(3, 1) = 8.98$, $p = 1.645e-5$, $\eta^2 = .0071$. The confidence intervals indicate that performance for tone pair 3 was significantly higher than for pairs 1 and 2, but not significantly different from pair 4.

Figure 7 shows the percent correct responses for the interaction between joint attention conditions and tones.
This interaction was statistically significant, $F(9, 1) = 3.58, p = 2.499 \times 10^{-4}, \eta^2 = .0081$. In the initiating Joint Attention and Simultaneous Joint Attention conditions, tones are responded to with approximately equal accuracy. In the Gaze forward condition, tone 3 was identified significantly more accurately than any of the other tones. In the Responsive Joint Attention condition, tone 2 is identified significantly less accurately.
than the other tones. The pattern is complex and it is not clear what is responsible for the interaction.

Figure 8 shows a comparison of results by block, contrasting the top and bottom fifteen participants. This result is statistically significant. $F(1,4) = 41.11, p = 3.64\times10^{-21}$, eta-squared = .2330. No other variables interact with the group variable (i.e., top vs. bottom participants). What this suggests is that the main effects of the task are driven by the top performers. The top performers have significantly improved their scores in

![Figure 8. Top vs. Bottom Performance Across Blocks](image-url)
between blocks 1 and 2, blocks 2 and 3, and blocks 4 and 5, with barely significant results in blocks 3 and 4. When we contrast this result with figure 4, we see that the chief difference between the top performers in the word learning task and the rest of the performers is that they were still significantly improving their scores even in the later blocks. In figure 4, which represents all the participants, we see that the last three blocks do not continue to show the significant gains that the high scorers achieved. The major difference then, is that top performers can maintain their attention and continue to learn even into the later blocks.

7.2 THE RELATION BETWEEN JOINT ATTENTION AND AUTISM

A correlation analysis was conducted to determine if accuracy in the joint attention learning sessions was related to scores on an Autistic Quotient test, essentially to determine if autistic qualities reduced sensitivity to the social cues provided by the video learner. All participants scored in the low range on the measure of autistic characteristics, making such a comparison challenging, and no relationship was found. There are several contributing factors on why this study had atypical autism scores. Firstly, the eligible pool of participants were students taking a first year psychology or neuroscience course. While most university students may take such a course, psychology, and studies on word learning, may not appeal to people with autistic traits, who are generally present in greater numbers in fields such as mathematics. Furthermore, two-thirds of the participants were female, who are generally less likely to have autistic traits (Bouvet, Mottron, Valdois, & Donnadieu, 2013). Theoretically, this would suggest that the participants recruited are more likely to be responsive to the eye-gaze stimuli being presented. An additional reason for the lack of relationship could be that participants with
elements of Autism spectrum disorder compensate for their poor abilities to use pragmatic information with greater sensitivity in some areas of perception. It is unclear whether this enhanced perception would apply to tone (Yu, 2010).

7.3 The Relation between Working Memory and Word Learning.

Figure 9 shows the relationship between working memory performance and overall word-image learning. A marginally significant correlation between WM span score and mean overall word-image learning was obtained, $r = .22$, $p = .1250$, indicating a weak positive relationship between these variables. Although consistent with previous research, the correlation between working memory and performance was not a strong one. Looking at a top and bottom split of participants based on working memory score also found no significant effects, nor any interactions. A comparison of means across blocks shows that the top 15 participants based on working memory had a significantly higher mean for the last block than the bottom 15. This is illustrated in figure 10.
Figure 9. Interaction of WM and Word Learning
8. Discussion

8.1 Joint Attention as a Learning Tool

The results of the present study were consistent with previous research that suggested joint attention influences memory and attention, and that the type of gaze interaction affects what parts of the brain become more active in response. Responsive
joint attention was shown to be the context associated with the worst sound-image matching abilities. Every other gaze condition, including gaze forward (no eye movement) had superior performance to Responsive Joint Attention. This is consistent with Kim and Mundy's (2012) results, where they found the initiating joint attention context was associated with better learning abilities. While Kim and Mundy only contrasted initiating joint attention and responsive joint attention, this study also simulated simultaneous joint attention, as well as a basic gaze forward condition. The results indicating that these additional contexts produced similar scores to the Initiating Joint Attention context was unexpected. That the gaze forward condition, which had no eye shifts toward the target images, produced the same results as the conditions where gaze shifts occur, which would suggest that the eye shifts were irrelevant to the learning process. However, the fact that the responsive joint attention condition did have significantly lower scores suggests that in some cases, participants were affected by the video learner’s behavior.

The results suggest that the responsive joint attention condition, rather than simply being less effective in adults, is actively detrimental to facilitating learning. The responsive joint attention condition may have been a distracting element instead of a supportive one, as the video partner’s behavior may not have been what the participant was expecting, therefore distracting participants from their task. This would suggest that social context is a double-edged sword, potentially improving individual’s performance by congruent facilitation of action, or impairing performance by pulling attentional processes away from the task at hand.

The other findings were somewhat puzzling. The effects of tone and interactions among tone and attention condition were unexpected. It is not clear why some tone pairings would benefit from specific joint attention conditions. One could speculate that
the tone pairings vary in their difficulty and that some attention conditions were more effective in facilitating learning the tone-picture pairings than other attention conditions. Unfortunately, this account is more a redescription of the results than an actual explanation of an underlying mechanism that explains the results.

In Shimada et al.’s (2012) study, simultaneous joint attention was the variable that had the most significant influence, and then only in the last block. Their study also asked participants to match sound and image pairs, but for the Japanese speakers, the target language was English. However, their experiment used a live interaction paradigm, and English is not as unfamiliar to Japanese speakers as Mandarin is to English ones. In addition, there has been no assessment of whether there are cultural differences in how adults may utilize joint attention. Attitudes about others, as well as cultural differences in behavior may influence how, or indeed whether, joint attention is still used. It seems unlikely that infants would already be responding strongly to such social pressures.

Overall, joint attention appears to still influence cognitive processes, which may benefit word learning, though the difficulty of the L2 linguistic feature relative to the learner’s L1 could temper its effectiveness. In the responsive joint attention condition, where performance was significantly worse than any other condition, tone pair 2, which was shown to be the most challenging, was matched with the least success of any tone in any condition. If joint attention can be a distraction, it makes particularly challenging elements even more difficult than their counterparts.

### 8.2 Memory and Attention

It can be argued that the difference in the top performers is not about memory capacity, but about the WM abilities, and in the same vein, their attentional abilities. As
Engle might argue, holding more items in one’s head simply means a greater ability to control attention, not some fixed memory store (Engle, 2002).

8.3 BILINGUALISM

Based on the responses to the language questionnaire, eleven of the 51 participants could be classified as bilingual. These candidates would be predicted to have higher than average working memory and attention skills, if not from natural talent, then possibly through the experience of acquiring a second language. Because of the unbalanced nature of the two groups, a Welch two sample t-test was used. The results neared significance with the bilinguals (M=32.5) performing marginally worse than the monolingual speakers (M=38.5); t(17.33)=-1.986, p = 0.06. This would suggest that multilingualism is not predicated on above average cognitive abilities, but possibly the right kind of exposure. The additional language experience of the multilinguals was not sufficient to improve their performance on the tones beyond that of the monolinguals.

8.5 LIMITATIONS

As this study reproduced joint attention via a video learning partner, it is possible that the effects are not as robust as would be obtained with a live interaction paradigm. The influence of the different variables of the video learning partner, from the size of the image on the screen, the timing of the movement, and the prominence of the actions can only be conjectured. Utilizing an eye-tracking device would have, at minimum, been a tool for determining whether participants were behaving as expected. In a more ideal set up, such a device could have been used in the same way as Kim and Mundy’s (2012) study, making the avatar responsive to the participants’ actual behavior, as opposed to
relying merely on timed videos, as in this study. While this study is using tests and behavior to surmise what may be occurring in the brain, an EEG study would be better at determining if there truly are differences in cognitive behavior under the different joint attention conditions.

Because of the results achieved in the gaze forward condition, it would appear that it may not have been the most suitable to represent a baseline task. When setting up the study, it was thought that the absence of a face for a quarter of the items would have been a distraction in itself, which would have interfered with the desired effects.

Interpretation of the interaction between tone pairings and the other variables may have been more straightforward if the perception of the tones could be separated from the learning task. In the present study perception of the tones is confounded to some extent with the learning task in which the tones are paired with pictures. Although the pairing of tones and pictures was necessary to examine word learning, it likely produces additional cognitive load not present in a purely perceptual task. Considering the relative difficulty associated with the perception of the tones may have aided interpretation of the some of the tone-related word learning effects found in the present experiment.

8.6 Time Variability in the Separate Conditions

Although stimuli in all the joint attention conditions were presented for the same amount of time after the image, past studies measured the differences in the fixation time of the eyes on the target stimuli. In both Shimada et al. (2012) and Kim and Mundy (2012), eye-tracking devices measured the differences in fixation periods on the stimuli, and found no relationship between fixation period and matching ability. No eye-tracker was used in this study to determine the fixation of the subject’s gaze. If this study were to
be repeated, using an eye-tracker may be beneficial simply to ensure that the participants are attending to the task, as well as to monitor their responsiveness to the learning partner. Furthermore, setting up the experiment so that the learning partner truly reacts to the subjects’ movements through the feedback of an eye-tracker would increase the realism of the joint attention conditions. The timed responses used in this study have a risk of being too automatic, which would undermine the assumption of agency in the learning partner.

9. Conclusion

In general, my results are consistent with the view that language learning occurs using general-purpose cognitive mechanisms that include attention and working memory. Joint attention is an integral part of word learning, showing effects even when the attention partner is a video. In addition, although the effect was not strong, those with greater working memory capacity were able to learn the tones more effectively. Research on joint attention holds promise for expanding our understanding of language learning and making it a more efficient and perhaps more effective process.
References


Yu, A. C. L. (2010). Perceptual compensation is correlated with individuals’ “autistic” traits: implications for models of sound change. *PloS One, 5*(8), e11950. doi:10.1371/journal.pone.0011950
### Appendix

#### Appendix A: Samples of Images from Kroll and Potter, 1984

![Image 1]

![Image 2]

![Image 3]

#### Appendix B: Mandarin words in Pinyin

(The number identifies the tone pattern. 1 represents a high, level tone. 2 represents a mid high, rising tone. 3 represents a low tone that drops slightly, then rises above middle range. 4 represents a high tone that quickly drops to the low range.)

<table>
<thead>
<tr>
<th>bie2</th>
<th>duan2</th>
<th>kang2</th>
<th>shuai2</th>
</tr>
</thead>
<tbody>
<tr>
<td>bie3</td>
<td>duan3</td>
<td>kang3</td>
<td>shuai3</td>
</tr>
<tr>
<td>chun2</td>
<td>beng2</td>
<td>ming2</td>
<td>pou2</td>
</tr>
<tr>
<td>chun3</td>
<td>beng3</td>
<td>ming3</td>
<td>pou3</td>
</tr>
<tr>
<td>su2</td>
<td>hui2</td>
<td>fei2</td>
<td>cui1</td>
</tr>
<tr>
<td>su3</td>
<td>hui3</td>
<td>fei3</td>
<td>cui4</td>
</tr>
<tr>
<td>bao1</td>
<td>bing1</td>
<td>nian1</td>
<td>diao1</td>
</tr>
<tr>
<td>bao4</td>
<td>bing4</td>
<td>nian4</td>
<td>diao4</td>
</tr>
<tr>
<td>kou1</td>
<td>zhuang1</td>
<td>shai1</td>
<td>fang1</td>
</tr>
<tr>
<td>kou4</td>
<td>zhuang4</td>
<td>shai4</td>
<td>fang4</td>
</tr>
<tr>
<td>tuo1</td>
<td>bei3</td>
<td>chua1</td>
<td>hua13</td>
</tr>
<tr>
<td>tuo4</td>
<td>bei4</td>
<td>chua4</td>
<td>hua4</td>
</tr>
<tr>
<td>ding3</td>
<td>chui3</td>
<td>pang3</td>
<td>min3</td>
</tr>
<tr>
<td>ding4</td>
<td>chui4</td>
<td>pang4</td>
<td>min4</td>
</tr>
<tr>
<td>geng3</td>
<td>cou3</td>
<td>rua3</td>
<td>nong3</td>
</tr>
<tr>
<td>geng4</td>
<td>cou4</td>
<td>rua4</td>
<td>nong4</td>
</tr>
<tr>
<td>zha03</td>
<td>gui1</td>
<td>zu1</td>
<td>bian1</td>
</tr>
<tr>
<td>zha04</td>
<td>gui2</td>
<td>zu2</td>
<td>bian2</td>
</tr>
<tr>
<td>nai1</td>
<td>kun1</td>
<td>liao1</td>
<td>ceng1</td>
</tr>
<tr>
<td>nai2</td>
<td>kun2</td>
<td>liao2</td>
<td>ceng2</td>
</tr>
<tr>
<td>shang1</td>
<td>duo1</td>
<td>mao1</td>
<td>chou1</td>
</tr>
<tr>
<td>shang2</td>
<td>duo2</td>
<td>mao2</td>
<td>chou2</td>
</tr>
</tbody>
</table>
Appendix C: Sample Screen Shots of Learning Trials

Appendix D: Sample Screen shot of testing trial

(Subject has indicated their selection)
Appendix E: Musical Training Questionnaire

Do you have any musical training? Yes   No

State your musical training, including age and duration, and type of schooling

<table>
<thead>
<tr>
<th>Age</th>
<th>Duration of training</th>
<th>Instrument</th>
<th>Schooling</th>
</tr>
</thead>
<tbody>
<tr>
<td>___</td>
<td>___</td>
<td>___</td>
<td>___</td>
</tr>
<tr>
<td>___</td>
<td>___</td>
<td>___</td>
<td>___</td>
</tr>
<tr>
<td>___</td>
<td>___</td>
<td>___</td>
<td>___</td>
</tr>
<tr>
<td>___</td>
<td>___</td>
<td>___</td>
<td>___</td>
</tr>
</tbody>
</table>

Rate your musical skill level

beginner   intermediate   good   excellent   professional

Appendix F: Language History Questionnaire

Self Report Language History questionnaire

Please indicate your date of birth: ___/___/____
Your place of birth:____________________
Please list, to the best of your knowledge, where you have lived during your life. Please list Country, Province, and City and at what age you were when you lived there:

<table>
<thead>
<tr>
<th>Age</th>
<th>Country</th>
<th>Province</th>
<th>City</th>
</tr>
</thead>
<tbody>
<tr>
<td>___</td>
<td>___</td>
<td>___</td>
<td>___</td>
</tr>
<tr>
<td>___</td>
<td>___</td>
<td>___</td>
<td>___</td>
</tr>
<tr>
<td>___</td>
<td>___</td>
<td>___</td>
<td>___</td>
</tr>
<tr>
<td>___</td>
<td>___</td>
<td>___</td>
<td>___</td>
</tr>
<tr>
<td>___</td>
<td>___</td>
<td>___</td>
<td>___</td>
</tr>
<tr>
<td>___</td>
<td>___</td>
<td>___</td>
<td>___</td>
</tr>
</tbody>
</table>

Is English your first language?
While in elementary or primary school, did you receive any instruction in a language other than English?  Yes    No

If Yes, which language? ______________________
If yes, approximately how many hours a day were you exposed to this language?
____________

Did any parent/guardian/daycare provider have a first language other than Canadian English?  
Yes    No

If Yes, which language? ______________________

Did any parent/guardian/daycare provider speak a language other than Canadian English to you on a regular basis?  Yes    No

If Yes, which language? ______________________

**Appendix G: Sample Questions from the AQ Test**

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>E1. I am willing to take risks.</td>
<td>definitely agree</td>
<td>slightly agree</td>
<td>slightly definitely disagree</td>
</tr>
<tr>
<td>E2. I like playing board games.</td>
<td>definitely agree</td>
<td>slightly agree</td>
<td>slightly definitely disagree</td>
</tr>
<tr>
<td>E3. I find learning to play musical instruments easy.</td>
<td>definitely agree</td>
<td>slightly agree</td>
<td>slightly definitely disagree</td>
</tr>
<tr>
<td>E4. I am fascinated by other cultures.</td>
<td>definitely agree</td>
<td>slightly agree</td>
<td>slightly definitely disagree</td>
</tr>
</tbody>
</table>

52