Treball de fi de màster Acadèmic

The Effect of Perceiving Pitch Gestures in the Acquisition of Mandarin Chinese Tones and Words

Nerea Suárez González

Màster: Lingüística Teòrica i Aplicada

Edició: 2015-2016

Directors: Dra. Pilar Prieto Vives

Mr. Santiago González Fuente

Any de defensa: 2016

Col·lecció: Treballs de fi de màster

Programa oficial de postgrau "Comunicació lingüística i mediació multilingüe"

Departament de Traducció i Ciències del Llenguatge

Index

Acknowledgments

This thesis is dedicated to my tutors, Dr. Pilar Prieto and Mr. Santiago González Fuente, for always advising me on how to overcome drawbacks with enthusiasm and devotion. As my tutor, Pilar Prieto deserves my warmest gratitude. Particularly, for her expertise in the field of SLA experimental research and for always being patient with me. As for Santiago, he has provided guidance at key moments in my work and has shared with all of us his highly skilled musical background. Moreover, I feel grateful for sharing time and ideas with Flor Baills in this project. My gratitude goes also to Joan Borràs for leading me in each step while conducting the statistical analysis.

Also, I would like to mention those having attended the oral presentation of the present study at Aarhus University and to show my special gratitude to Isabel Darcy for her generous comments and wise feedback. I am also thankful to the Department of Translation and Language Sciences of the Universitat Pompeu Fabra for allowing me the necessary facilities and electronic devices for conducting the experiment, and to the 50 Catalan students that have contributed their time to this project. Last but not least, I would like to express my most sincere joyful feeling for having had such a good friends along the way, both old and new ones.

I particularly dedicate this MA thesis to my parents. My words for them may be longer than the present thesis so I leave my gratitude to them for another time. I would like also to mention the support provided by Thalía Jimenez over these past months.

Sursum Corda

Abstract

Very few SLA studies have investigated the benefits of pitch gestures in Chinese language learning (see Chen 2013, Morett and Wang 2014).The present MA thesis is concerned with (a) assessing the contributions of observing pitch gestures in tonal contrast awareness, (b) exploring the potential reduction of the cognitive effort resulting in word recall facilitation triggered by observing the pitch gesture, and (c) with investigating the relative benefits of observing pitch gestures in those who already have well-developed auditory skills (i.e. musical training). Two between-subject experiments, with two experimental conditions (Gesture-Observe, No-Gesture), were conducted among 50 participants. Results indicate that observing pitch gestures significantly strengthens both participants' identification of Mandarin lexical tones and vocabulary learning. Nonetheless, for those with already good auditory skills, pitch gestures do not seem to enhance their tone identification accuracy. These findings have methodological implications in CSL (Chinese as a Second Language) pronunciation practices.

Keywords

SLA, multimodality, audio-visual perception, pitch gestures, Mandarin lexical tones, suprasegmentals, computer-assisted study.

1. Introduction

It is becoming increasingly clear that co-speech gestures (e.g., the hand, face and body movements that we produce while we speak) are an integral aspect of our language faculty and form an integrated system with speech both at the phonological (i.e. temporal) and semantic-pragmatic levels (e.g. McNeill 1985, 1992, 2000, 2005; Bernardis et al. 2006)[1]. Recent neurocognitive evidence further suggests that the brain integrates speech and gesture information in similar ways to speech alone (e.g. Bates & Dick, 2002, Özyürek, 2014)[2]. Moreover, there is ample evidence of the cognitive benefits of using co-speech gestures in different instructional settings like in math-solving contexts (e.g. Goldin-Meadow et al. 2009, Cook et. al 2008)[3]. To date there are no conclusive evidence regarding the real benefits of using gestures for linguistic purposes; whereas recent studies have suggested that gestures help reduce the cognitive load or processing cost (e.g. Goldin-Meadow et al. 2001, Wagner et al. 2004) and thus function as a compensatory and facilitatory device in the acquisition of a second language (e.g. Gullberg 1998, McCafferty 2002), other studies have suggested that when learning higher aspects of an L2 such as semantics, syntax or phonetics, observing (e.g. Kelly & Lee 2012) and producing gestures (e.g. Post et al. 2013, Hirata et al. 2014) only helps when the cognitive demands are low, otherwise becoming counterproductive and distracting.

A growing body of experimental research in second language acquisition has shown how co-speech gestures can be used as an effective tool that allows students to improve their language skills (see Gullberg 2006a, 2006b, 2010, 2014, Gullberg and McCafferty 2008, Gullberg, de Bot and Volterra 2008 for a review article on gestures in

[1] See also Clark, 1996, Goldin-Meadow 2003, Kendon 1980, 2004, 2014, Perniss et al. 2010, Perniss et al. 2014, Harris 2003, 2005, Lemke 2003, Levinson & Holler, 2014.
[2] See also papers in Özyürek & Kelly, 2007, Wu & Coulson, 2005, Calvert 2001.
[3] See also Radford 2002, 2005, 2015 for a theory of mathematical thinking objectification through gestures.

L1 and L2 acquisition). These studies also reveal that meaning-related iconic gestures are especially useful for learning novel words in comparison to meaning-related static images (e.g. Tellier, 2008), or in comparison with the same content only presented in speech (e.g. Kelly et al., 2009) or even with meaning-incongruent iconic gestures (e.g. Kelly et al.,2009). Goldin-Meadow, Nusbaum, Kelly, & Wagner (2001) suggested that "gesturing may prime a speaker's access to a temporarily inaccessible lexical item and thus facilitate the processing of speech" (p. 521) (an idea consistent with the Lexical Retrieval Hypothesis proposed by Krauss et al. 2000, see also Krauss et al. 1996 for a review). Similarly, recent studies have also shown the benefits of using hand beat gestures for the memorization of words in second language learning (e.g. Kushch et al 2015, So et al., 2012, Igualada et al., 2015) as well as for discourse comprehension (e.g. Krahmer and Swerts, 2007, Biau and Soto-Faraco, 2013).

1.1. Speech Perception in Tonal Languages

Tonal languages like Mandarin Chinese, as opposed to intonational languages like English or Catalan that convey prosodic, grammatical or emotional content along a sentence, use pitch (fundamental frequency: F0 henceforth) variations at word level (e.g., the so-called lexical tone contrasts) so as to distinguish meanings between otherwise segmentally identical words. In this regard, recent evidence in SLA research has demonstrated that speakers of tonal and non-tonal languages can be taught with equal success in the perception and production of L2 tonal systems (e.g. Hao 2012, Francis et al. 2008, So 2006). As for Mandarin Chinese, there is evidence that learners of non-tonal languages can be successfully trained to differentiate between Mandarin tones after a short perceptual tone training in the laboratory both in perception (e.g.

Wang et al. 1999, Wang et al. 2003a; Wong & Perrachione, 2007, So and Best 2010a, Burnham et al. 2001) and production (e.g. Wang et al., 2000, 2003b).

With regards to speech perception, it has been shown that visual information appears to enhance non-native speech perception in general (e.g. Hardison, 1999; Navarra & Soto-Faraco, 2004). In this sense, evidence for the benefits of visual cues for lexical tone languages was first presented by Burnham, Ciocca, and Stokes (2001). In that study, native Cantonese listeners were asked to identify minimal pairs differing in tone in AV, AO, and VO mode. Results showed that dynamic tones were identified significantly better than chance in the VO condition and that there was AV augmentation of Mandarin tone identification in noise but not when F0 information was filtered out[4]. Similarly, Smith and Burnham (2012) asked native Mandarin and native Australian English speakers to discriminate Mandarin minimal pairs differing in tones in five conditions: AO, AV, degraded (cochlear-implant-simulation) AO, degraded AV, and VO (silent video). Results again showed that the availability of visual speech information improved discrimination in the degraded audio conditions. More interestingly, in the VO condition, the tone-naïve English listeners outperformed native listeners (both above chance). Both studies suggested that visual speech information for tone is available to all perceivers regardless of their language background, but is possibly underused by normal-hearing tone language perceivers.

Following the same line of research, Burnham and colleagues (Reid et al. 2015) recently carried out two experiments focusing on Thai tone perception by native speakers of three tonal languages (Thai, Cantonese, and Mandarin), a pitch–accented language (Swedish), and a non-tonal language (English). Two experiments were carried out; the first one tested cross-language auditory-visual perception and involved a tone

[4] Similar results were found for Thai speakers (Mixdorff, Charnvivit, & Burnham, 2005).

discrimination test in auditory-only (AO), AV, and visual only (VO) conditions. The second experiment addressed the mechanisms by which language experience affects tone perception in three fundamental frequency equivalent contexts: speech, filtered speech, and violin sounds and also involved a discrimination task. Together the results of both experiments provide evidence for general augmentation of tone perception to be determined by both auditory and visual information (i.e. facial movements), by acoustic and linguistic contexts, and by universal and experiential factors across all languages.

Interestingly, there is large evidence documenting a direct relationship between the acoustic-physical production of tones and the visual information conveyed through the rigid motion of our heads (e.g. Burnham et al. 2006, Reid et al. 2015), eyebrow movements (e.g. Munhall et al. 2004, Yehia et al., 2002, Vatiktiotis-Bateson, 2000, Srinivasan and Massaro 2003) and the visible movements of the neck (e.g. Chen and Massaro 2008). For instance, Burnham et al. (2015) investigated the auditory-visual perception and production of Cantonese tones and phones by native and non-native speakers of lexical languages, comparing three types of training conditions: auditory-only (AO), visual-only (VO), and auditory-visual (AV) modes. Results showed that Cantonese perceivers performed above chance in the VO condition and further Functional Data Analysis (FDA) and Principal Component (PC) extraction analyses suggested that the salient PC to distinguish tones perceptually involve rigid motion of the head. This fact seems to be coherent with the finding that facial movement can be perceptually informative for suprasegmental prosodic identification (e.g. Cvejic, Kim & Davis, 2010, Krahmer & Swerts, 2007). Further research has demonstrated that this information can be perceptually informative in auditory learning of tonal contrast and effectively used after a short training session (e.g. Yehia et al 2002, Vatikiotis-Bateson et al. 2000). In this regard, Chen and Massaro (2008) tested native Mandarin speakers in visual lexical-tone identification before and after the learning phase. Results showed that when participants were taught to pay attention and to use this visible information (neck, head, and chain activity), visual tone identification improved significantly.

1.2 *Studies in CSL (Chinese as a Second Language)*

Within the realm of SLA, less is known about the potential benefits of using co-speech gestures in pronunciation learning. It has been reported that when it comes to learning novel speech sounds, language learners benefit from training that includes speech and mouth movements compared to just speech alone (e.g. see Hardison 2003, 2005, Wang et al. 2008, Hirata and Kelly 2010, Calvert et al. 1999). The benefits of observing and producing beat gestures to mark the rhythm of speech in pronunciation improvement in non-native speech patterns have been explored by Gluhareva & Prieto (2015). A comparison between participants' observation in both pre-training and post-training speech samples demonstrated that beat gesture training significantly improved their accentedness ratings on the more discourse-demanding items.

Along this line of research, Hirata and Kelly (2010) carried out an experiment in which English learners were exposed to videos of Japanese speakers who were producing beat gestures showing Japanese long and short vowel phoneme contrast. Results showed that seeing lip movements (audio-mouth condition) during training significantly helped learners to perceive difficult phonemic contrasts while there was no categorical improvement in those conditions in which hand gesture was involved. Based on these results, authors commented that maybe the mere observation of hand gesture would not work at the segmental level. Even more recently, the same authors (Hirata et al. 2014) wanted to explore specifically whether hand gestures may have a role in the auditory learning of Japanese as a L2 at the segmental level. For this purpose, they carried out an experiment in which English natives were exposed to one of four training types: Syllable-Observe (audio-mouth-hands), Syllable-Produce, Mora-Observe and, Mora-Produce. Based on previous research (Hirata and Kelly 2010), they hypothesized that producing, instead of observing, beat gestures both for syllable gesture and mora gesture would empower auditory learning. However, results showed that there was no unique advantage of producing compared to perceiving in the overall amount of improvement for all groups and that the improvements depended on rate and word-initial position. Moreover, neither mora gestures nor syllable gesture seemed to be advantageous. Authors concluded that, at least at the beginning stage of L2 learning, manual gestures do not seem to be integrated with speech at the segmental phonology level.

Only a handful of studies have focused on the potential benefits of observing (also producing) pitch gestures[5] in the learning of Mandarin lexical tones and words by non-tonal speakers. In an instructional longitudinal study, Chen (2013) showed that learners perceiving and producing 'tonal gestures' (as he labelled them) seemed to have significantly superior communicative skills and also performed significantly better in tonal production with a higher frequency of accurate responses, regardless of the tonal and non-tonal background. Moreover, learners demonstrated a wider pitch range when producing Mandarin words together with gesture. Nonetheless, Chen's study was a classroom training study with no experimental control of the following items: (a) the materials used in the training, (b) a mix of the perception and production in training, and (c) no control of the participants' language background. In a more recent experiment, Morett and Wang (2014) aimed to investigate the contributions of three types of audio-visual training conditions to English-speaking learners' representations of Mandarin

[5] Hand gestures deploying movements and points of reference in space according to prosodic patterns such as height, duration and contour.

words. The first group saw and repeated metaphorical gestures depicting lexical tone pitch contours (pitch gesture) while hearing the Mandarin tones; the second group saw and repeated gestures conveying word meanings (semantic gestures) and the third one learnt without gestures (no gesture). In this study, 57 English speakers were tested on a Mandarin lexical tone identification task and on a word meaning association task. The results of this study showed that, in comparison with semantic gestures and no gestures, pitch gestures facilitated the learning Mandarin words differing in lexical tone with their meanings, but failed to enhance their lexical tone identification. These findings suggested that the visuospatial features of pitch gestures strengthen the relationship between English speakers' representations of Mandarin lexical tones and word meanings, supporting the predictions of spoken word recognition (e.g. Luce and Pisoni 1998, Elman, 2004)[6] and embodied cognition (e.g. Barsalou 1999, 2008)[7]. However, Morett and Wang's study cannot disentangle: (a) the potential effects of observing vs. producing pitch gestures (participants where both observing and producing the pitch gesture and the tone in the training), and, more importantly for the purpose of the present thesis, (b) the potential effects of training with pitch gestures on tone identification vs. word-meaning associations.

2. Research Questions (RQ) and Hypotheses (H)

Experiment 1

RQ1: Does observing pitch gestures help with learning Mandarin Chinese lexical tone contrasts?

[6] See Luce & McLennan 2005 for a recent review on contemporary models of spoken word recognition.
[7] See Adams 2010; Aizawa 2007; Chemero 2011; Shapiro 2011 for a recent debate over embodied cognition.

- **H1:** Although there is contradictory evidence as to whether manual gestures help with learning an L2 at the segmental level, there is independent evidence of a strong perceptual relationship between acoustic pitch and spatial movements for 'high' and 'low' (e.g. Connell et al. 2013[8], Cassidy 1993). Thus, *Hypothesis 1* is that observing pitch gestures conveying the tone height and contour may have a facilitatory effect in the learners' auditory perception and identification of the Mandarin tones.

Experiment 2

RQ2: Does observing pitch gestures help with learning Mandarin Chinese words?

- **H2:** Provided the large evidence suggesting that gestures can facilitate novel vocabulary learning (mentioned in previous sections) because they may reduce the cognitive processing cost (e.g. Wagner et al. 2004, Goldin-Meadow et al. 2001) and thus function as a compensatory and facilitatory device (e.g. Gullberg 1998, 2014, McCafferty 2002), *Hypothesis 2* would be that those participants observing pitch gestures will recall more vocabulary than those not receiving gestural input.

General Research Question

RQ3: Do individuals with previous musical instruction and/or better musical perception skills benefit equally from the gesture vs. no gesture condition?

- **H3:** There is substantial evidence revealing that musical experience enhances L2 sound perception and production in general (e.g. Sadakata and Sekiyama, 2011,

[8] The results of this study suggest that pitch perception is fundamentally audiospatial showing that English participants intuitively associated pitch information through a spatial movement in gesture representing height (consistent with older outcomes (e.g. Cassidy, 1993). Besides, these results also go in consonance with the proposed spatial conceptual metaphor of pitch through gesture (e.g. Casasanto, Phillips, & Boroditsky, 2003) in which the upward direction represents high-frequency pitch and the downward direction represents low-frequency pitch. This spatial conceptual metaphor of pitch is present in prelinguistic infants (e.g. Dolscheid, Hunnius, Casasanto, & Majid, 2012), indicating that it is language-independent.

Slevc and Miyake 2006). More importantly, musical experience has also been found to enhance the auditory perception and production of lexical tones, both for speakers of tonal languages (e.g. Wu et al. 2015) and non-tonal speakers (e.g. Tong Yee and Ching Tang 2016, Alexander et al. 2005, Gottfried 2007), and tone word learning (e.g. Cooper and Wang 2012, Wong and Perrachione 2007). Thus, we hypothesize that those participants with previous musical instruction (better auditory skills) will get the best results, no matter the condition they are in, because they will rely more on acoustic rather than visual cues.

3. Overall Experimental Design

The present study consists of two experiments which share the experimental design, that is, a between-subject study in which participants were randomly assigned to one of the conditions (No-Gesture Condition or Gesture-Observe Condition (henceforth NG and GO).The first one investigates the benefits of observing pitch gestures in tone identification learning (e.g. tone learning phase in Figure 1) and the other investigates the benefits of observing pitch gestures in vocabulary learning (e.g. vocabulary learning phase in Figure 1). Before and after the experiment we also took an external measure of each participant's memory span and of each participant's musical skills (pitch and melody).

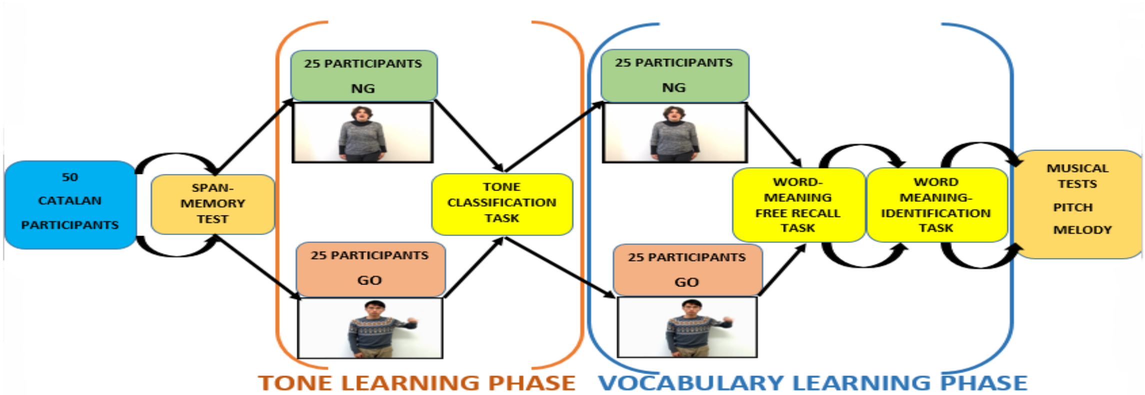

Fig. 1: Overall experimental design. Tone Learning Phase (Exp. 1) followed by a Vocabulary Learning Phase (Exp. 2).

This design aimed to analyse the relative improvement of the experimental group, GO, in comparison to the control or NG group. The participants were tested individually and it took one hour per participant to complete both experiments (Exp.1 and Exp. 2). The new section will briefly address the external measure we took from each participants' memory span and musical skills (pitch and melody) and the following sections will focus on the two experiments individually in methodological terms, each presenting its respective materials.

3.1.Span Memory Task

The typical way of measuring Short Term Memory capacity (STM), also called memory span, is to present a sequence of numbers, letters, or words aloud to a person/child at the rate of about 1 second per item (see Appendix A). For this, audios of a Catalan speaker were recorded. The researcher asked participants to listen to and to repeat the sequence of words verbally, either in the order each item was presented or in a backward order (e.g. Wilde & Strauss, 2002). The length of the sequence was continually increased until the person was correct only 50% of the time. A typical adult can remember a list of about 5 to 9 unrelated words. Yet, when these words are organized as a normal sentence, memory span increases to 15 or 16 words (e.g. Baddeley, Vallar, & Wilson, 1987 cited in Aben et al., 2012). The results of this task showed that all participants remembered an appropriate number of words.

3.2.Music Perception skills (pitch and melody)

There is substantial evidence revealing that musical experience enhances the auditory perception and production of lexical tones, both for speakers of tonal languages (e.g. Wu et al. 2015) and non-tonal speakers (e.g. Tong Yee and Ching Tang 2016, Alexander et al. 2005, Gottfried 2007), and tone word learning (e.g. Cooper and Wang 2012, Wong and Perrachione 2007). We thought that an external measure of musical perception skills that could be correlated with individual performance or the different conditions would be of importance especially for the experimental control (based on Law and Zentner's 2012 proposal). For this reason, we customized a modular version of the brief musical proms developed by the department of psychology of the University of Innsbruck so as to test participants' skills in pitch and melody discrimination (available at https://www.uibk.ac.at/psychologie/fachbereiche/pdd/personality_assessment/proms/take-the-test/modular-proms.html). Thus, we selected pitch and melody proms. Each of them consisted of 18 questions and each took 5-7 minutes to complete. This software retrieved a personalized performance feedback that could be downloaded (see the results of this task in footnote number 12).

4. *Experiment 1: Mandarin Chinese Tone Learning*

In order to assess the contribution of observing vs. not observing pitch gestures in the perception of Mandarin tones, we designed a training phase in which participants watched a training video and after that, they were presented with a tone classification task.

4.1 Methodology

Participants

A total of 50 undergraduate and graduate students (age: M= 19,78, SD= 3,26, 16 males), mostly recruited at the Communication Campus at the Universitat Pompeu Fabra, participated in this experiment. All participants were selected on the basis of being dominant in Central Catalan (reported more than 75% of daily use). None of them reported having had neither previous knowledge nor exposure to Mandarin Chinese and they reported little or no prior training in phonology or phonetics. All participants had normal or corrected-to-normal vision and normal hearing and they all were screened to be right-handed [9]. Participants with previous musical training (as screened by a questionnaire) were not excluded because, although it has been reported that musical training affects auditory learning of foreign languages (e.g. Sadakata and Sekiyama, 2011) we also wanted to explore the effects of gestures in those with already good auditory skills (52,4% counterbalanced among conditions). We also took an external measure on the participants' musical skills (pitch and melody); see previous section. Participants were unaware of the real purpose of the study and they received 10 euros for their participation in the study.

Materials

This section will deal with (a) the stimuli chosen as target words and the later audiovisual stimuli that we created with the recordings of these words and, (b) the audio recordings that participants listened to in the tone classification task.

Audio-visual materials for the pitch habituation and the pitch training phase

A total of thirty-six monosyllabic Mandarin words (18 minimal pairs differing only in tone) were chosen for creating the audio-visual materials for the vocabulary learning phase; see Table 1. Words were selected so that all minimal pair words shared the same grammatical category. There were a total of 5 pairs of verbs, 10 pairs of nouns and 3 pairs of adjectives. All words conformed to the Catalan phonotactic restrictions.

[9] The GO condition involved observing the right hand in gesture on the computer screen (videos were flipped according to previous literature).

Mean log frequency of the target words per million words was matched for the Catalan target words (MlogCatalan = 1.27) using NIM (Guasch, Boada, Ferré, & Sánchez-Casas, 2013).

Chinese Word	Pseudo-transcription	Catalan Translation	English Translation	Relative _F	Log	Absolute _F	Letter s	P O S
tǔ	[thu]	natiu	native	0,956	0,291	49	5	A
chí	[txi]	retardat	retarded	2,185	0,503	112	8	A
tū	[thu]	calb	bald	2,517	0,546	129	4	A
fǔ		acariciar	to caress	2,653	0,563	136	9	V
lì		beneficiar	to profit	3,258	0,629	167	10	V
lǔ		brusc	rough	4,058	0,704	208	5	A
mà		insult	insult	5,307	0,8	272	6	N
tì	[thi]	pentinat	hairstyle	5,326	0,801	273	8	N
chī	[txi]	tonto	foolish	5,97	0,843	306	5	A
dī	[ti]	taxi	taxi	6,087	0,85	312	4	N
má		lli	light	8,507	0,978	436	3	N
lí		acomiadar	to greet	8,585	0,982	440	9	V
fá		castigar	to punish	8,624	0,983	442	8	V
fā		compartir	to share	9,326	1,014	478	9	V
gù	[ku]	sòlid	solid	10,516	1,061	539	5	A
fú		sorgir	to arise	10,809	1,072	554	6	V
bǐ	[pi]	llapis	pencil	10,848	1,074	556	6	N
chǔ	[txu]	estalviar	to save	14,633	1,194	750	9	V
bō	[puo]	onada	wave	17,209	1,26	882	5	N
lù		ocupat	occupied	21,091	1,344	1081	6	A
tá	[tha]	pila	battery	21,267	1,348	1090	4	N
mǐ		arròs	rice	24,818	1,412	1272	5	N
mì		mel	honey	30,203	1,494	1548	3	N
pò	[phuo]	trencar	to break	30,964	1,505	1587	7	V
tǎ	[tha]	torre	tower	43,977	1,653	2254	5	N
pō	[phuo]	tirar	to throw	47,938	1,69	2457	5	V
gè	[ke]	peça	piece	48,348	1,693	2478	4	N
bó	[puo]	oncle	uncle	58,845	1,777	3016	5	N
ná		agafar	to take	60,171	1,787	3084	6	V
bí	[pi]	nas	nose	61,498	1,796	3152	3	N
gē	[ke]	cançó	song	63,391	1,809	3249	5	N
nà		rebre	to receive	79,897	1,908	4095	5	V
gǔ	[ku]	antic	antique	80,677	1,912	4135	5	A
dī	[thi]	escala	stairs	101,924	2,013	5224	6	N
chū	[txu]	sortir	to leave	168,554	2,229	8639	6	V
dǐ	[ti]	fons	background	252,119	2,403	12922	4	N

Table 1. Mandarin words and their Catalan and English translations used in Exp. 1 and Exp. 2. The pseudo-transcription of the lexical tones is also included. The darker rows correspond to the six minimal-pairs used in the second experimental study.

As for the video recordings, a male (native) Mandarin speaker and a female teacher thereof were recorded using a Panasonic AG-HMC41 professional video camera against a white background in a quiet room at the Universitat Pompeu Fabra. The video clips for all recordings showed the speaker's face and the upper half of the body so that participants could see the hand movements and the face at the same time. They were asked to produce the 18 minimal target pairs either with or without the pitch gesture for the two conditions (e.g., the Gesture-Observe and the No-Gesture Conditions). The two pictures in Figure 3 show a screenshot from the video corresponding to the Gesture-Observe condition (left panel) and to the No-Gesture condition (right panel).

Fig. 3: Screenshot from the video corresponding to the Gesture-Observe condition (left panel) and to the No-Gesture condition (right panel). The red arrow conveys the hand motion performed concurrently with pronunciation of the Mandarin word (here, T3).

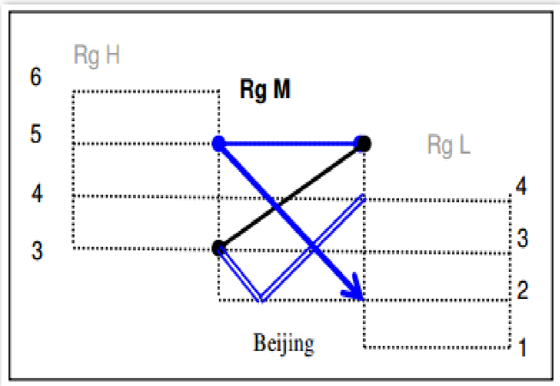

Fig. 2: The Single-register Tonal Model (Xiaonong, 2012, pg. 10)

We controlled for the speaking rate and the gesture of both speakers by recording each of them producing each pair of minimal pair one after the other (both with and without gesture). We also controlled the target gestures by instructing both teachers to imagine four vertical points in their bodies (eyes, shoulders, belly button and hips) which would correspond to the 4-scale Mandarin tonal model recently proposed by Xiaonong (2012)[10]; see Figure 2. The female speaker was also video recorded for the habituation video, e.g., to introduce some previous knowledge on Chinese lexical tones. Furthermore, the videos were digitally flipped using Adobe Premier Pro CC 2015 software so that participants could mirror the gestures with their own right hand (consistent with Kelly et al. 2014) and the audio quality was enhanced using the same software. The audios were extracted and later modified using Audacity software so that there was no background noise in the final stimuli and were sampled at 44,100 Hz.

Eventually, in order to control over primacy and recency in the presentation of the audio-visual stimuli, six video sequences were created for each training condition (NG vs. GO); each presented with a different randomization of the 18 minimal pairs so that at the end, each participant went through the 36 trials only once (see Table 2). None of these sequences shared the same combination of stimuli in the same position. In order to emphasise differences in tone, words from each minimal pair were presented one after the other (see also Kelly et al 2014 and Morett and Wang 2014). Adobe Premiere Pro CC 2015 software was used so that the inter-stimulus intervals were at intervals of 2 seconds for the control group and 3 seconds for the experimental group.

Group 1	T1/T2	T4/T3	T4/T1	T2/T3	T1/T3	T4/T2
Group 2	T1/T3	T4/T2	T3/T2	T1/T4	T1/T2	T4/T3
Group 3	T4/T1	T2/T3	T3/T4	T1/T2	T2/T4	T3/T1
Group 4	T3/T2	T1/T4	T3/T1	T2/T4	T3/T4	T2/T1
Group 5	T2/T4	T3/T1	T2/T1	T4/T3	T4/T1	T2/T3
Group 6	T3/T4	T2/T1	T4/T2	T1/T3	T3/T2	T1/T4

Table 2. Stimuli randomization per groups in both training conditions.

[10] Compare to the 5-scale model proposed by Chao, 1968.

Procedure

Participants were tested individually in a quiet room at the Universitat Pompeu Fabra in Barcelona and went through two phases; the learning phase and the task phase. Upon arrival, participants were randomly assigned to one of the two training conditions: No-Gesture Condition and Gesture Observe Condition (n = 25 in each condition; 8 males). In front of each participant was a laptop computer equipped with earphones, and next to the computer there was a sheet of paper so as to answer the tone classification task. Participants were left alone in the experimental room and were asked to call the experimenter back when they had completed the task. Subjects were instructed to be silent the whole time and to carefully listen to the pitch training video, to pay attention to the gestures conveying the melodic movements in the case of the experimental group.

In the pitch training video, within each trial, the orthographic form of the Mandarin word together with the tone mark that indicates its tonal properties was presented in written from (see Figure 4) previous to being presented by the teacher in audiovisual format[11].

Fig 4: Screenshot showing two minimal pairs with their corresponding tone mark (T1 and T3 on the right and T2 and T4 on the left).

[11] It has been reported that the visual illustration of lexical tones' pitch contours (i.e. tone marks) in combination with the orthographic representation of words containing them facilitate their acquisition (e.g. Liu et al. 2011).

Tone Classification Task and Stimuli

The aim of this task was to detect whether participants were able to identify the four Mandarin tones after a short training session. The stimuli for this task were eight real syllables or pseudo-syllables respecting Chinese phonotactic rules, four of them untrained (see Table 3). Participants heard these words one by one (and only once) and were instructed to write down the target Mandarin word with its corresponding tone mark on a separate paper ("*You are going to hear a syllable only once. Write down which syllable together with its associated tone mark.*"). Participants were always tested individually after each training phase through the web page Survey Gizmo questionnaire software on the contents of the training video. The stimuli were automatically randomized by the program across participants in both conditions. All the audios for this task were recorded with different speakers from the ones appearing in the videos. The sex of the speakers was also counterbalanced among the items to ensure that the performance reflected learners' ability to identify Mandarin lexical tones across word tokens, rather than their recall of the specific token produced during the learning phase. Participants received no feedback on their performance at any time. The task was self-paced and took about 10–12 min to complete.

Speaker Sex	**trained**	**untrained**
Male	mi T4 fu T3	te T2 nu T4
Female	txi T2 pi T1	la T1 txe T3

Table 3. Items in the tone classification task spoken by female and male

4.2. Results

A total of 400 responses were obtained (25 participants x 2 conditions x 8 tone identification questions) and submitted to statistical analysis by means of a Generalized Linear Mixed Model (henceforth GLMM) with Responses as the dependent variable (Correct vs. False) and with CONDITION (two levels: NG vs. GO) and MUSICAL_SKILLS (two levels: Musicians vs. Non-Musicians) [12], as well as their interaction (CONDITION*MUSICAL SKILLS) as fixed factors. Also, Sex, Age and the Item were set as random factors. Results revealed a significant main effect of Condition [F(1, 396) = 6.202, p < .05], revealing that the GO experimental group outperformed the NG control group (see Figure 3). Moreover, Bonferroni's post-hoc pair-wise comparisons revealed a significant difference between Condition among Non-Musicians [F(1,396) = 8,183 (p< .01)], revealing that (1) Musicians obtained very similar scores in both conditions (NG and GO) and (2) that Non-Musicians had significantly higher scores in GO vs. NG condition (see Figure 5).

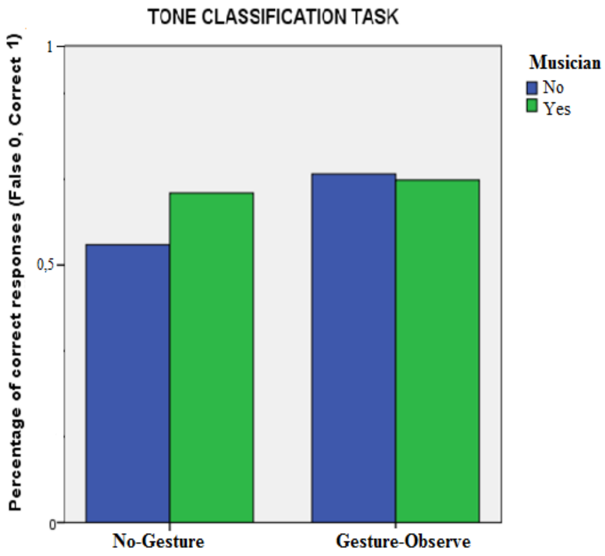

Figure 5: Performance on the word-meaning free recall task for participants assigned to the GO Condition and to the NG Condition.

4.3. Discussion

All in all, results show that observing pitch gestures significantly improves Mandarin tonal contrast learning in those learners with no previous musical instruction. These results seem to provide further evidence for those theories that claim that pitch

[12] These labels, musician or non-musician, were selected on the basis on participants' answer in the questionnaire. They were asked whether they had previous musical education or not (Yes vs. No) and, if so, we asked the years thereof (counterbalanced among conditions; i.e. 47,6% of musicians in NG condition). Later, we took an external measure of pitch and melody abilities to confirm the above mentioned variable. GLMM statistical analysis was run with MUSICIANS (two levels: Yes vs. No) as the dependent variable and with the results of the MELODY and PITCH tests as fixed factors. Also, Sex, Age and the Item were set as random factors. Results revealed a significant main effect of MELODY [F(17, 340) = 2,258, p < .05] and PITCH [F(15, 340) = 2,036, p < .05] abilities for those who responded "Yes" in the questionnaire thus allowing us to take this variable as reliable.

perception is fundamentally audiospatial (e.g. Connel et al. 2013, Cassidy 1993) as well as those supporting the spatial conceptual metaphor of pitch (Casasanto, Phillips, & Boroditsky, 2003). These results confirm previous longitudinal studies in the field of CSL (Chen 2013) and also confirm that a brief training with only perceiving tones can be effective (see e.g., Morett and Wang 2014).

Furthermore, these results also indicate that for those students with already good auditory skills, pitch gestures do not seem to enhance the auditory learning of Mandarin tones. In this regard, our results coincide with those studies that have demonstrated that musical experience enhances L2 sound perception and production in general (e.g. Sadakata and Sekiyama, 2011, Slevc and Miyake 2006) and more specific with studies showing that musical experience enhances the auditory perception and production of lexical tones by non-tonal speakers (e.g. Tong Yee and Ching Tang 2016, Alexander et al. 2005, Gottfried 2007). These results of the present thesis seem to coincide with previous studies in that subjects with previous musical instruction – with well-developed auditory skills – are used to relying strongly on acoustic cues and, therefore, observing the pitch gesture does not seem to help them as an informative cue in learning the Mandarin tonal contrast. Nonetheless, observing pitch gestures does not seem to be counterproductive or distracting neither to musicians nor to non-musicians. Thus, our results do not coincide with those studies that have suggested that when learning higher aspects of an L2, observing (e.g. Kelly & Lee 2012) and producing gestures (e.g. Post et al. 2013, Hirata et al. 2014) do not have a facilitatory role.

5. Experiment 2: Vocabulary learning

In order to assess the contribution of observing pitch gestures on vocabulary learning in Mandarin Chinese, the same subjects undertook vocabulary training plus two vocabulary tasks. These tasks aimed to analyse the amount of improvement of the experimental group, which received gestural input, in comparison to the control group, which did not receive gestural input.

5.1 Methodology

Participants

The same subjects that participated in Experiment 1.

Materials

Stimuli for the vocabulary training videos

To-be-learned words consisted of six minimal pairs of words differing only in tone - which were selected from the first study (see the darker rows in Table. 1).

Group 1	1	2	3
Group 2	1	3	2
Group 3	2	1	3
Group 4	2	3	1
Group 5	3	1	2
Group 6	3	2	1

Table 4. Order of appearance of the three different blocks per group in both conditions.

Three blocks (see Table 5), of the twelve individual words each, were organized in a randomized order (see Table 4) so as to create a unique block with three different repetitions of the trials along the three blocks. As in Experiment 1, in order to emphasise differences in tone, words from each minimal pair were presented in consecutive trials. This process was repeated six times so that, at the end, each group went through the same Catalan target word (and heard the Mandarin word) three times in a different order. This process was repeated for both conditions.

1st Block	puo T1	puo T2	mi T3	mi T4	tha T3	tha T2	ke T1	ke T4	ma T2	ma T4	ti T1	ti T3
2nd Block	ke T4	ke T1	tha T2	tha T3	ti T1	ti T3	ma T4	ma T2	puo T2	puo T1	mi T4	mi T3
3rd Block	ma T2	ma T4	ti T3	ti T1	mi T3	mi T4	puo T1	puo T2	tha T2	tha T3	ke T1	ke T4

Table 5. Syllables and tones used in the randomization of vocabulary stimuli per blocks.

Stimuli for the vocabulary tasks

The stimuli for the two vocabulary tasks were audio recordings of the twelve trained words during the vocabulary learning phase. Again, the voices of other male and female speakers were recorded and counterbalanced among the items.

Procedure

Experiment 2 shared a similar procedure to the one in Experiment 1. First, a vocabulary training phase was given. After that, the task phase was divided into two counterbalanced blocks: a tone identification task and word-meaning association task.

First, a slide with the target Catalan word was illustrated in the screen (see Figure 6).

Figure 6: Procedure in the vocabulary training phase showing a minimal pair (2 trials)

Vocabulary tasks

The vocabulary tasks took 20 min approximately for each participant to complete. For each task, the stimuli described above were automatically randomized by Survey Gizmo program across participants in both conditions.

Word-meaning free recall task

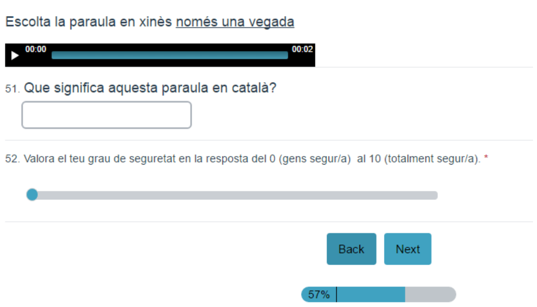

Fig. 7. Screenshot from Survey Gizmo.

The first task consisted of a word-meaning free recall task in which participants were asked to listen only once to the 12 items individually and to write down their target

Catalan translation (“*In this task you are going to hear twelve Chinese words. Write down their Catalan translation.*”); see Figure 7. The test was self-paced, but participants were told not to go back to previous answers once they moved on to the later trials.

Word-meaning association task

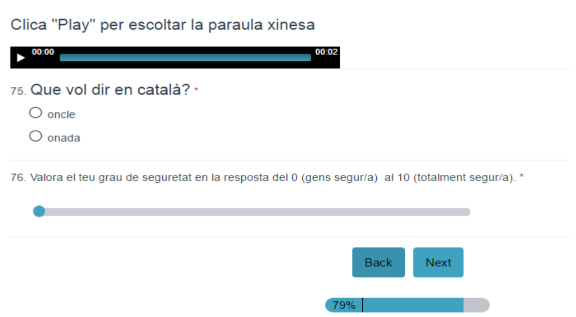

Fig. 8. Screenshot from Survey Gizmo

In this task, upon hearing each word, participants guessed its meaning by choosing one of the two options in the survey, one of which was the meaning of the presented word and one of which was the meaning of the other word in that minimal pair; both displayed on the screen (“*You are going to hear twelve Chinese words*. *What is the Catalan meaning for this word?”);* see Figure 8.

5.2 Results

Word-meaning free recall task

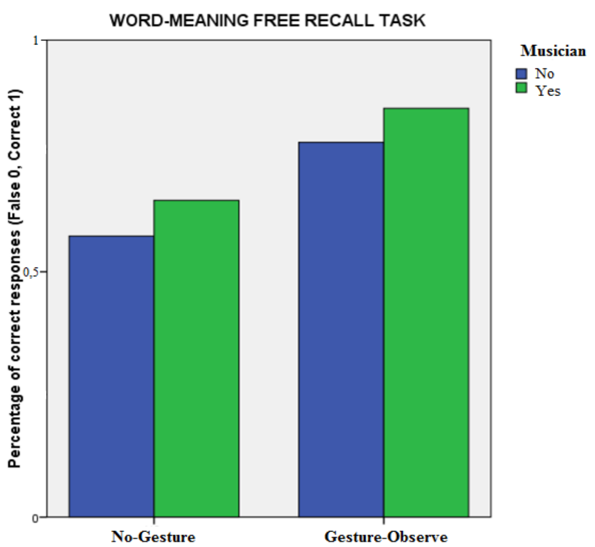

Figure 9: Performance on the word-meaning free recall task for participants assigned to the GO Condition and to the NG Condition.

A total of 600 responses were obtained (25 participants x 2 conditions x 12 questions) and submitted to statistical analysis by means of a GLMM with Responses as the dependent variable (Correct 1, False 0) and with CONDITION (two levels: NG vs. GO) and MUSICAL_SKILLS (two levels: Musicians vs. Non-Musicians), as well as their interaction (CONDITION*MUSICAL SKILLS) set as fixed factors. Also, Sex, Age and the Item_number were set as random factors. Results revealed a significant main effect of Condition [F (1,596) = 15,847, p < .01] confirming that the GO experimental group outperformed the NG control group (see Figure 9). Moreover, Bonferroni's post-hoc pair-wise comparisons showed that musicians in both conditions had significantly higher scores [F (1, 596) = 8,571, p < .01]; see Figure 9.

Word-meaning association task

A total of 600 responses were obtained (25 participants x 2 conditions x 12 questions) and submitted to statistical analysis by means of a GLMM with Responses as the dependent variable (Correct 1, False 0) and with CONDITION (two levels: NG vs. GO) and MUSICAL_SKILLS (two levels: Musicians vs. Non-Musicians), as well as their interaction (CONDITION*MUSICAL SKILLS) as fixed factors. Sex, Age and the Items were set as random factors. Results show a significant main effect of Condition [F (1,596) = 29,630, p < .01], revealing that the GO experimental group outperformed the NG control group (see Figure 6). Interestingly, results also showed a significant interaction between Condition*Music_Skills (see Figure 10). Bonferroni's post-hoc pair-wise comparisons showed a statistically significant difference between musicians and non-musicians in

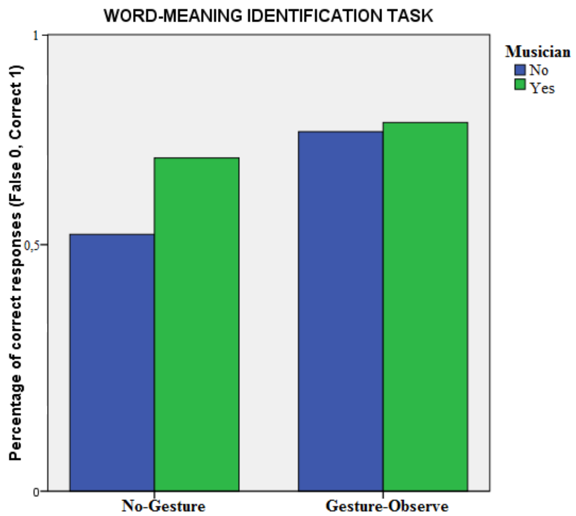

Figure 10: Performance on the word-meaning free recall task for participants assigned to the GO Condition and to the NG Condition.

the NG condition [F (1, 569) = 29,295, p < .05], but not in the GO condition; see Figure 10.

5.3. Discussion

Our results demonstrate that a short vocabulary training observing pitch gestures enhance L2 students' vocabulary learning in a tonal language like Chinese; at least at an initial stage of learning (see also Chen 2013, Morett and Wang 2014). In general, these results agree with recent results regarding Mandarin and Thai VO tone perception (e.g. Burnham et al. 2015, Burnham, Ciocca, and Stokes 2001, Smith and Burnham, 2012). In those studies, non-tonal (English) language speakers outperformed the tonal speakers in the VO condition. Thus, non-tonal participants attended more to visual information when faced with a foreign rather than a native speaker or language (e.g. Chen & Hazan, 2009; Sekiyama & Burnham, 2008). Together these results show that since non-tonal speakers are not accustomed to use F0 information at the segmental, they are intuitively driven to rely more on visual cues.

Furthermore, the results about the effects of musical ability indicate a contrast between the performance of subjects with and without musical expertise. As shown in Experiment 1, for those students with already good auditory skills, pitch gestures do not seem to enhance the auditory learning of Mandarin tones. These results seem to be coherent with those studies that have shown that musical experience enhances the non-tonal speakers' tone word learning (e.g. Cooper and Wang 2012, Wong and Perrachione 2007). As in Experiment 1, this fact suggests that subjects with previous musical experience are accustomed to relying strongly on acoustic cues, and that pitch gestures do not seem to help them much in learning the Mandarin tonal contrast.

6. General Conclusions and Discussion

The aim of the present thesis was to explore (a) whether observing pitch gestures helps with learning Mandarin tonal contrasts, (b) whether observing pitch gestures will help with novel vocabulary acquisition and (c) whether participants with already good skills (i.e. musical training) will beneficiate from observing pitch gestures. In order to assess these questions, two between-subject experiments were carried out among 50 Catalan participants. All in all, the results of both experiments revealed that observing pitch gestures enhances the auditory perception of the Mandarin tones and promotes novel vocabulary learning of those participants without previous musical edication, thus confirming previous observations in the field of CSL (Chen 2013, Morett and Wang 2014). These results seem to provide further evidence for those theories that claim that pitch perception is fundamentally audiospatial (e.g. Connel et al. 2013, Cassidy 1993) as well as those supporting the spatial conceptual metaphor of pitch (Casasanto, Phillips, & Boroditsky, 2003). This may be due to the fact that gestures are the most appropriate expressive modality for conveying spatially extended content (e.g. Goldin-Meadow 1999; Lemke 1999b).

Moreover, our results seem to support those studies suggesting that gestures help reduce the cognitive load or processing cost (e.g. Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001; Wagner, Nusbaum, & Goldin-Meadow, 2004), thus functioning as a compensatory and facilitatory device in the acquisition of a second language (e.g. Gullberg 1998, McCafferty 2002). In contrast, our results do not seem to agree with those studies that have suggested that when learning higher aspects of an L2, such as semantics, syntax or phonetics, observing (e.g. Kelly & Lee 2012) and producing gestures (e.g. Post et al. 2013, Hirata et al. 2014) only help when the cognitive demands are low. Even to those participants with previous musical education pitch gesture observation did not seem to become counterproductive and distracting. Our results suggest that observing pitch gestures help at the lower phonological level and with novel vocabulary learning, at least for learning tonal languages at an initial stage. Interestingly, it has been reported that word retention in L1 is impacted by the richness of stimuli accompanying the word (e.g. Craik and Tulving, 1975; Engelkamp and Zimmer, 1994) and that a word network consisting of many components, i.e., visual, aural, kinetic, olfactory, etc. stores and retrieves information more efficiently than a small network (McClelland, 1985; Macedonia and Klimesch, 2014).

Among the more recent findings in L2 acquisition, several studies have suggested that gesturing might help L2 learners internalise new knowledge (e.g. Van Compernolle and Williams 2011, Lee 2008, McCafferty, 2004, Negueruela et al., 2004). In this regard, there is a large tradition of research, now labelled 'embodied cognition', which shows that producing actions leads to better learning and memory than just observing them (e.g. Saltz and Donnenwerth-Nolan, 1981, Barsalou, 2008, Fischer & Zwaan, 2008)[13]. This theory argues that gesturing, but not speaking, thus solidified learning because it is in itself a high-level cognitive process which externalizes thinking (e.g. Goldin-Meadow et al. 2010, Roth 2001, McCafferty , 2002,2004, 2006, 2008), a claim that is supported by recent experiments (e.g, Goldin-Meadow 2014, Ramirez Verduro 2006, Macedonia and Klimesch 2014, Morett, 2014, Macedonia et al. 2011). Furthermore, recent neuroimaging work has demonstrated that imitating gestures activates a more distributed network of neural regions than simply observing gestures highly impacting retention (e.g. Montgomery, Isenberg, & Haxby, 2007; Engelkamp & Dehne, 2000); see Macedonia, 2014 for a review of behavioural research on the positive

[13] See also Asher's "Total Physical Response" theory, 1969, 1977.

effects of gestures on memory); see the MA thesis of the UPF student Florence Baills to see the results of producting pitch gestures.

Furthermore, our results indicate that for those students with already good auditory skills, pitch gestures seem not to enhance the auditory learning of Mandarin tones. Yet, those with previous musical education performed better in the three tasks because they relied on acoustic cues. These results seem to be coherent with those studies that have demonstrated that musical experience enhances L2 sound perception and production in general (e.g. Sadakata and Sekiyama, 2011, Slevc and Miyake 2006) and, in particular, with those studies that have shown that musical experience enhances the auditory perception and production of lexical tones, both tonal (e.g. Wu et al. 2015) and non-tonal speakers (e.g. Tong Yee and Ching Tang 2016, Alexander et al. 2005, Gottfried 2007). In this regard, a growing body of evidence suggests that brain mechanisms governing music and language processing interact.

On one hand, it has been reported that musicians demonstrate perceptual enhancement in language-specific abilities such as phonological processing (Anvari, Trainor, Woodside, & Levy, 2002), verbal memory (Chan, Ho, & Cheung, 1998; Franklin et al., 2008), pitch discrimination (e.g. Bidelman & Krishnan, 2010, Strait et al 2010) sensitivity to prosodic cues (e.g. Thompson, Schellenberg, & Husain, 2004), degraded speech perception (e.g. Bidelman & Krishnan, 2010; Parbery-Clark, Skoe, Lam, & Kraus, 2009), lexical tone identification (e.g. Delogu, Lampis, & Olivetti Belardinelli, 2006; Delogu, Lampis, & Olivetti Belardinelli, 2010; Lee & Hung, 2008, Burnham, Booker and Reid 2014) and tone word learning (e.g. Cooper and Wang 2012, Wong and Perrachione 2007). Inversely, neurophysiological enhancement of musical pitch processing in native speakers of a tone language has been also reported. Bidelman et al. (2011a) found that brainstem responses evoked by musical intervals were both more accurate and more robust in native speakers of Mandarin Chinese compared to English-speaking non-musicians, suggesting that long-term experience with linguistic pitch may transfer to subcortical encoding of musical pitch. Following the same line of research, Bidelman et al. (2011b) investigated whether or not linguistic expertise (e.g., speaking a tone language) positively influences music-related neural processing and its perception by examining brainstem responses of English-speaking musicians/non-musicians and native speakers of Mandarin Chinese in the perceptual discrimination of musical pitch. Results showed that, relative to non-musicians, both tone language (Chinese) and musical experience had stronger brainstem encoding of the defining pitches of musical sequences, regardless of whether the signal is specific to music or language per se (Bidelman et al. 2011a, Bidelman et all 2011b). In contrast, two behavioural pitch discrimination tasks revealed that neither Chinese nor non-musicians were able to discriminate subtle changes in musical pitch with the same accuracy as musicians. Together, these studies demonstrate a potential for linguistic pitch experience to carry over into nonlinguistic (i.e., musical) domains (e.g. Bidelman, Gandour, et al., 2011a, Bidelman, et al. 2011b, Pfordresher & Brown, 2009, Krishnan, Gandour, Bidelman, & Swaminathan, 2009, Marie, Kujala, & Besson, 2010). Similarly, Krishnan et al.'s (2010) study using Thai tones suggested that tone language speakers (Thai and Mandarin) have developed more sensitive brain stem mechanisms for representing pitch (reflected by tracking accuracy and pitch strength) than non-tonal language speakers (English).

All in all, the results of the present thesis indicate that a short training observing pitch gestures help both with tone learning and novel vocabulary acquisition in a tonal language; at least at an initial stage for those students without musical education. Moreover, our results suggest that subjects with previous musical education are used to rely on acoustic cues and that pitch gestures do not seem to help them in learning Mandarin tonal contrasts. Yet, they show better scores than non-musicians both in tone identification and vocabulary learning. These findings have methodological implications in CSL (Chinese as a Second Language) pronunciation practices.

References

Aben B., Stapert S., & Blokland A. (2012). About the distinction between working memory and short-term memory. *Frontiers in Psychology*, *3* (301), 1-9.

Adams, F. (2010). Embodied cognition. *Phenomenology and the Cognitive Sciences, 9* (4), 619-628.

Aizawa, K. (2007). Understanding the embodiment of perception. *Journal of Philosophy*, *104* (1), 5-25.

Alexander, J. A., Wong, P. C. M., & Bradlow, A. R. (2005). Lexical tone perception by musicians and non-musicians. In *Proceedings of Interspeech, Eurospeech, 9th European Conference on Speech Communication and Technology*, Lisbon, Portugal (2005).

Anvari, S. H., Trainor, L. J., Woodside, J., & Levy, B. A. (2002). Relations among musical skills, phonological processing and early reading ability in preschool children. *Journal of Experimental Child Psychology, 83*(2), 111–130.

Asher, J. J. (1977). Learning Another Language through Actions. *The Complete Teacher's Guide Book* (6th ed.). Los Gatos: Sky Oaks Productions, Inc.

Asher, J. J. (1969). The Total Physical Response: Approach to Second Language Learning. *The Modern Language Journal, 53* (1), 3-17.

Baddeley, A. D., Vallar. G., & Wilson. B. (1987). Sentence comprehension and phonological memory: Some neuropsychological evidence. In M. Coltheart (Ed.), *Alienlion and performance XII: The psychology of reading.* London: Lawrence Erlbaum Associates Ltd.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22* (4), 577-660.

Bates, E., & Dick, F. (2002). Language, gesture and the developing brain. In B. J. Casey & Y. Munakata (eds.), *Special issue: Converging Method Approach to the Study of Developmental Science*. *Developmental Psychobiology*, *40*, 293- 310.

Bernardis P., & Gentilucci M. (2006). Speech and gesture share the same communication system. *Neuropsychologia*, *44*, 178-190.

Biau, D. J., & Soto-Faraco S., (2013). Beat gestures modulate auditory integration in speech perception. *Brain Language*, *124* (2), 143-52.

Bidelman, G. M., Krishnan, A., & Gandour, J. T. (2011a). Enhanced brainstem encoding predicts musicians' perceptual advantages with pitch. *European Journal of Neuroscience*, *33*(3), 530–538.

Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2011b). Musicians and tone-language speakers share enhanced brainstem encoding but not perceptual benefits for musical pitch. *Brain and Cognition*, *77*, 1–10.

Bidelman, G. M., & Krishnan, A. (2010). Effects of reverberation on brainstem representation of speech in musicians and non-musicians. *Brain Research, 1355,* 112–125.

Bluhme, H. and Burr, R. (1971). An audio-visual display of pitch for teaching Chinese tone. *Studies in Linguistics*, *22*, 51-57.

Burnham, D., Kasisopa, B., Reid, A., Luksaneeyanawin, S., Lacerda, F., Attina, V., Rattanasone, N., Schwarz, I., Webster, D. (2014). Universality and language-specific experience in the perception of lexical tone and pitch. *Journal of Applied Psycholinguistics*, *Cambridge Journals Online, 36* (6), 1459-1491.

Burnham, D., Booker, R. and Reid, A. (2014). The effects of absolute pitch ability and musical training on lexical tone perception. *Psychology of Music*, *43* (6), 881-897.

Burnham, D., Reynolds, J., Vatikiotis-Bateson, E., Yehia, H., Ciocca, V., Morris, R.H., Hill, H., Vognali, G., Bollwerk, S., Tam, H., & Jones, C. (2006). The perception and production of phones and tones: The role of rigid and non-rigid face and head motion. In Yehia, H. (Eds), *Proceedings of the 7th International Seminar on Speech Production*, Brazil, 1-8.

Burnham, D., Ciocca, V., & Stokes, S. (2001). Auditory–visual perception of lexical tone. In P. Dalsgaard, B. Lindberg, H. Benner, & Z.-H Tan (Eds.), *Proceedings of the 7th Conference on Speech Communication and Technology*, *EUROSPEECH*, Scandinavia (pp. 395–398).

Burnham, D., Lau, S., Tam, H., & Shoknecht, C. (2001). Visual discrimination of Cantonese tone by tonal but non-Cantonese speakers, and by non-tonal language speakers. In D. Massaro, J. Light, and K. Geraci (Eds.) *Proceedings of the International Conference on Auditory-Visual Speech Processing*, Sept. 2011, (pp. 155-160).

Calvert G., A. (2001). Crossmodal processing Brain: Insights from Functional Magnetic Neuroimaging Studies. *Oxford University Press, 11*(12), 1110-1123.

Calvert G. A., Brammer M., Bullmore E, Campbell, R., Iversen, S. D., & David, A. (1999). Response amplification in sensory-specific cortices during crossmodal binding. *NeuroReport*, 10, 2619-2623.

Casasanto, D., Phillips, W, & Boroditsky, L. (2003). Do we think about music in terms of space? Metaphoric representation of musical pitch. In R. Alterman & D. Kirsh (Eds.), *Proceedings of 25th Annual Conference of the Cognitive Science Society*. Boston, MA. (pp. 216-221).

Cassidy, J. W. (1993) Effects of various sightsinging strategies on nonmusic majors' pitch accuracy. *Journal of Research in Music Education*, 71, 293-302.

Chan, A. S., Ho, Y., & Cheung, M. (1998). Music training improves verbal memory. *Nature*, *396*, 128.

Chao, Y.R. (1968). A Grammar of Spoken Chinese. University of California Press, Berkeley.

Chemero, A. (2011). *Radical Embodied Cognitive Science*. A Bradford Book.

Chen, C.M. (2013). Gestures as tone markers in multilingual communication. In Istvan Kecskes (ed.), *Trends in Applied Linguistics Series: Research in Chinese as a Second Language*, 9, 143-168.

Chen Y., & Hazan V. (2009) Developmental factors and the non-native speaker effect in auditory-visual speech perception. *Journal of the Acoustical Society of America*, *26* (2), 858-65.

Chen, H. T., & Massaro, D. W. (2008) Seeing pitch: Visual information for lexical tones of Mandarin-Chinese. *Journal of the Acoustical Society of America*, *123* (4), 2356-2366.

Clark, H. H. (1996). Using language. Cambridge: Cambridge University Press.
Compernolle, R. A. & Williams, L. (2011). Thinking with your hands: Speech-gesture activity during an L2 awareness-raising task. *Language Awareness 20*(3), 203-219.

Connel, L., Cai, Z. G., & Holler, J. (2013). Do you see what I'm singing? Visuospatial movement biases pitch perception. *Brain and Cognition, 81*, 124-130.

Cook, S.W, Mitchell Z, Goldin-Meadow, S. (2008). Gesturing makes learning last. *Cognition*, *106*, 1047–58.

Cooper, A, & Wang, Y. (2012). The influence of linguistic and musical experience on Cantonese word learning. *Journal of the Acoustical Society of America, 131* (6), 4756-69.

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *104*, 268-294.

Cvejic, E., Kim, J., & Davis, C. (2010). It's all the same to me: prosodic discrimination across speakers and face areas. In M. Hasegawa-Johnson (Ed.) *Proceedings of the 5th International Conference on Speech Prosody*, [Hotel] Doubletree Magnificent Mile, Chicago, May 11-14, 2010.

Delogu, F., Lampis, G., & Olivetti Belardinelli, M. (2010). From melody to lexical tone: Musical ability enhances specific aspects of foreign language perception. *European Journal of Cognitive Psychology, 22*(1), 46–61.

Delogu, F., Lampis, G., & Olivetti Belardinelli, M. (2006). Music-to-language transfer effect: May melodic ability improve learning of tonal languages by native nontonal speakers? *Cognitive Processing, 7*(3), 203–207.

Elman, J. L. (2004). An alternative view of the mental lexicon. *Trends in Cognitive Sciences, 8*, 301–306.

Engelkamp J., & Dehne D M. (2000). Item and order information in subject-performed tasks and experimenter-performed tasks (2000). *Journal of Experimental Psychology: Learning, Memory, and Cognition. 26* (3), 671-82

Engelkamp, J. & Zimmer, H. D. (1999). Levels-of-processing effects in subject-performed tasks**.** *Memory & Cognition, 27* (5), 907-914.

Fischer M. H., & Zwaan, R. A. (2008). Embodied language: a review of the role of the motor system in language comprehension. *The Quarterly Journal of Experimental Psychology, 61*(6), 825-50.

Francis, A., Ciocca, V., Ma, L., & Fen, K. (2008). Perceptual learning of Cantonese lexical tones tone by tone and non-tone language speakers. *Journal of Phonetics, 36*, 268-294.

Franklin M. S., Moore K. S., Yip C. Y., Jonides J., Rattray K., Moher J. (2008). The effects of musical training on verbal memory. *Psychology of Music, 36*, 353–365.

Gluhareva, D. Prieto, P. (2016, in press). Brief training with beat gestures favours L2 pronunciation, *Language Teaching Research*.

Goldin-Meadow, S. (2014). How gesture works to change our minds. *Trends in Neuroscience and Education, 3*, 4-6.

Goldin-Meadow, S. & Beilock, S. L (2010). Action’s influence on thought: The case of gesture. *Perspectives on Phychological Science, 5,* 664-674.

Goldin-Meadow, S., Cook S. W., & Mitchell Z.A. (2009). Gesturing gives children new ideas about math. *Perspectives on Psychological Science, 20,* 267-72.

Goldin-Meadow, S. (2003). *Hearing Gesture: How Our Hands Help Us Think.* Cambridge, MA: Harvard University Press.

Goldin-Meadow, S., Nusbaum H, Kelly S. D., & Wagner S. M. (2001). Explaining math: gesturing lightens the load. *Perspectives on Psychological Science, 12,* 516–22.

Goldin-Meadow, S. (1999). The role of gesture in communication and thinking. *Trends in Cognitive Sciences, 3* (11), 419-429.

Gottfried, T. L. (2007). Music and language learning: Effect of musical training on learning L2 speech contrasts. In O.-S. Bohn and M. J. Munro (Eds.) *Language Experience in Second Language Speech Learning*: In honour of James Emil Flege (pp. 221–237). Amsterdam: John Benjamins.

Guasch, M., Boada, R., Ferré, P., & Sánchez-Casas, R. (2013). NIM: A Web-based Swiss Army knife to select stimuli for psycholinguistic studies. *Behavior Research Methods, 45*, 765-771.

Gullberg, M. (2014). Gestures and second language acquisition. In Müller, C., Cienki, A., Fricke, E., Ladewig, S. H., McNeill, D. & Tessendorf, S. (Eds.) *Body, language, communication: an international handbook on multimodality in human interaction*, *2*. (pp. 1868-1875). Mouton de Gruyter.

Gullberg, M. (2010). Methodological reflections on gesture analysis in second language acquisition and bilingualism research. *Second Language Research, 26* (1), 75-102.

Gullberg, M., de Bot, K. and Volterra, V. (2008): Gestures and some key issues in the study of language development. *Gesture,* 8, 149-79.

Gullberg, M., McCafferty, S. G. (2008) Introduction to Gesture and SLA: Toward an Integrated Approach. *Studies in Second Language Acquisition, 30*, 133-146. Cambridge University Press.

Gullberg, M. (2006a). Handling discourse: Gestures, reference tracking, and communication strategies in early L2. *Language Learning, 56* (1), 155-196.

Gullberg, M. (2006b). Some reasons for studying gesture and second language acquisition. *International Review of Applied Linguistics in Language Teaching (IRAL), 44*, 103-124.

Gullberg, M. (1998). *Gesture as a communication strategy in second language discourse: a study of learners of French and Swedish*. Lund, Sweden: Lund University Press.

Hao, Y.C. (2012). Second language acquisition of Mandarin Chinese tones by tonal and non-tonal language speakers. *Journal of Phonetics 40*(2):269–79.

Hirata, Y., and Kelly, S., Mansala, M. (2014). Effects of Hand Gestures in Auditory Learning of Second-Language Vowel Length Contrasts. *Journal of Speech, Language, and Hearing Research, 57*, 2090-2101.

Hirata, Y., and Kelly, S. (2010). Effects of lips and hands on auditory learning of second-language speech sounds. *Journal of Speech Language and Hearing Research, 53*, 298-310.

Hunnius, S., Casasanto, D., & Majid, A. (2012). The sound of thickness: Prelinguistic infants' associations of space and pitch. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th annual meeting of the Cognitive Science Society*, 306–311.

Igualada, A., Esteve-Gibert, N., Prieto, P. (submitted). Beat gestures improve word recall in 3- to 5- year- old children. *Journal of Experimental Child Psychology.*

Kelly, S. D., Hirata, Y, Manansala, M., & Huang, J. (2014). "Exploring the role of hand gestures in learning novel phoneme contrast and vocabulary in second language". *Frontiers in psychology, 5,* 673.

Kelly, S. D., & Lee, A. (2012). When actions speak too much louder than words: Gesture disrupts word learning when phonetic demands are high. *Language and Cognitive Processes, 27*, 793-807.

Kelly, S. D., McDevitt, T. & Esch, M. (2009). Brief training with co-speech gesture lens a hand to word learning in a foreign language. *Language and Cognitive Processes, 24* (2), 313-334.

Kendon, A. (2014). Semiotic diversity in utterance production and the concept of 'language'. *Philosophical Transaction of the Royal Society*. B 369, 20130293. http://dx.doi.org/10.1098/rstb.2013.0293

Kendon, A. (2004). *Gesture: Visible action as utterance*. New York: Cambridge University Press.

Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. In M. R. Key (Ed.), *Relationship of verbal and nonverbal communication*, (pp.207-227). The Hague: Mouton.

Klimesch W. (1994*). The Structure of Long-Term Memory: A Connectivity Model of Semantic Processing*. Hillsdale, NJ: Erlbaum.

Krahmer, E., & Swerts, M. (2007). The effects of visual beats on prosodic prominence: Acoustic analyses, auditory perception and visual perception. *Journal of Memory and Language*, *57* (3), 396-414.

Krauss, R. K., Chen, Y., & Gottesman, R. F (2000). Lexical gestures and lexical access. A process model. In D. McNeill (Ed.), *Language and gesture* (pp. 261-238). New York: Cambridge University Press.

Krauss, R. M., Chen, Y., & Chawla, P. (1996). Nonverbal behaviour and nonverbal communication: What do conversational hand gestures tell us? In M. Zanna (Ed.), *Advances in experimental social psychology* (pp. 389-450). San Diego, CA: Academic Press.

Krishnan A, Gandour J.T., Bidelman G. M. (2010). The effects of tone language experience on pitch processing in the brainstem. *Journal of Neurolinguistics, 23,* 81–95.

Krishnan, A., Gandour, J. T., Bidelman, G. M., & Swaminathan, J. (2009). Experience-dependent neural representation of dynamic pitch in the brainstem. *Neuroreport, 20*(4), 408–413.

Kushch, O. Igualada, A. & Prieto, P. (submitted, April 2016). "Gestural and prosodic prominence favour second language novel word acquisition", *Language and Cognitive Processes.*

Law, L.N.C., & Zentner, M. (2012). Assessing Musical Abilities Objectively: Construction and Validation of the Profile of Music Perception Skills. *PLoS One, 7*(12): e52508.

Lee, C. Y., & Hung, T. H. (2008). Identification of Mandarin tones by English-speaking musicians and nonmusicians. *Journal of the Acoustical Society of America, 124* (5), 3235–3248.

Lee, J. (2008) Gesture and private speech in second language Acquisition. *Studies in Second Language Acquisition 30*(2), 169-190.

Lemke, J. L. (1999). Meaning making in the conversation: Head spinning, heart winning, and everything in between. *Human Development, 42,* 87–91

Liu, Y., Wang, M., Perfetti, C. A., Brubaker, B., Wu, S., & MacWhinney, B. (2011). Learning a tonal language by attending to the tone: An in vivo experiment. *Language Learning, 61*, 1119-1141.

Luce P. A., McLennan CT. (2005). Spoken word recognition: The challenge of variation. In: Pisoni D. B., Remez R. E. (Eds.) *The handbook of speech perception*. Blackwell; Malden, 591–609.

Luce P.A,. & Pisoni D.B. (1998) Recognizing spoken words: The neighbourhood activation model. *Ear and Hearing, 19*(1), 1-36.

Macedonia M. (2014). Bringing back the body into the mind: gestures enhance word learning in foreign language. *Frontiers in Psychology. 5,* 1467.

Macedonia, M., & Klimesch, W. (2014). Long-term effects of gestures on memory for foreign language words trained in the classroom. *Mind, Brain, and Education. 8,* 74–88.

Macedonia M., Müller K., & Friederici A. D. (2011). The impact of iconic gestures on foreign language word learning and its neural substrate. *Human Brain Mapping. 32*, 982–998.

Marie, C., Magne, C., & Besson, M. (2010b). Musicians and the metric structure of words. *Journal of Cognitive Neuroscience, 23*(2), 294–305.

McCafferty, S. G. (2008). Material foundations for second language acquisition: Gesture, metaphor, and internalization. In S. G. McCafferty & G. Stam (Eds), *Gesture: Second language acquisition and classroom research* (pp: 3–24) London: Routledge.

McCafferty, S. G. (2006). Gesture and the materialization of second language prosody. *International Review of Applied Linguistics, 44,* 195–207.

McCafferty, S. G. (2004). Space for cognition: Gesture and second language learning. *International Journal of Applied Linguistics, 14,* 148–165.

McCafferty, S. G. (2002). Gesture and creating zones of proximal development for second language learning. *Modern Language Journal, 86*, 192–203.

McClelland, J. L. (1985). Distributed models of cognitive processes: Applications to learning and memory. *Academic Sciences. 444*, 1–9.

McNeill, D. (2005). *Gesture and thought*. Chicago, IL: University of Chicago Press.

McNeill, D. (2000). *Language and Gesture*. New York: Cambridge University Press.
Montgomery K., J., Isenberg N., Haxby J.V. Communicative hand gestures and object-directed hand movements activated the mirror neuron system. (2007). *Social Cognitive and Affective Neuroscience, 2* (2), 114-22.

McNeill, D. (1992). *Hand and mind: what gestures reveal about thoughts*. Chicago, IL: University of Chicago Press.

McNeill, D. (1985). So you think gestures are nonverbal? *Psychological Review, 92*, 350-371.

Morett, L. (2014). When Hands Speak Louder Than Words: The Role of Gesture in the Communication, Encoding, and Recall of Words in a Novel Second Language. *The Modern Language Journal, 98*(3), 834–853.

Morett, L. & Chang, L.Y. (2015) Emphasising sound and meaning: pitch gestures enhance Mandarin lexical tone acquisition, *Language, Cognition and Neuroscience, 30*, 3, 347-35.

Munhall, K. G., Jones J. A., Callan, D. E., Kuratate, T., & Vatikiotis- Bateson, E. (2004). Visual prosody and speech intelligigibility: Head movement improves auditory speech perception, Psychological. *Science. 15*, 133-137

Negueruela, E., James P. L., Susan R. J., & Gelabert, J. (2004). The "private function" of gesture in second language speaking activity: A study of motion verbs and gesturing in English and Spanish. *International Journal of Applied Linguistics, 14*(1), 113-147.

Özyürek A. (2014).Hearing and seeing meaning in speech and gesture: insights from brain and behaviour. *Philosophical Transactions of the Royal Society*. B 369: 20130296. http://dx.doi.org/10.1098/rstb.2013.0296

Özyürek, A. & Kelly, S.D. (2007). Special isssue 'Gesture, brain, and language'. *Brain and Language, 101*(3), 181-184.

Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009b). Musician enhancement for speech-in-noise. *Ear and Hearing, 30*(6), 653–661.

Perniss P, Vigliocco G. (2014). The bridge of iconicity: from a world of experience to the experience of language. *Philosophical Transactions of the Royal Society*. B 369: 20130300. http://dx.doi.org/10.1098/rstb.2013.0300

Perniss P., Thompson T., & Vigliocco, G. (2010). Iconicity as a general property of language: evidence from spoken and signed languages. *Frontiers in Psychology*. 1, 1–15.

Pfordresher P. Q. & S. Brown (2009). Enhance production and perception musical pitch in tone languages speakers. *Attention, Perception, & Psychologists, 71*(6), 1385-1398.

Post, L.S., Van Gog, T., Pass, F., & Zwaan, R. A. (2013). Effects of simultaneously observing and making gestures while studying grammar animations on cognitive load and learning. *Computers in Human Behaviour, 29*, 1450-1445.

Radford, L. (2015). The Epistemological Foundations of the Theory of Objectification. *Isonomia*, 127-149.

Radford, L. (2005). Why do gestures matter? Gestures as semiotic means of objectification. In Helen L. Chick, Jill L. Vincent (Eds.), Proceedings of the 29th Conference of *the International Group for the Psychology of Mathematics Education,* 1, 143-145.

Radford, L. (2002). The seen, the spoken and the written. A semiotic approach to the problem of objectification of mathematical knowledge. *For the Learning of Mathematics*, *22*(2), 14-23.

Ramírez Verdugo, D. (2006). A Study on Intonation Awareness and Learning by Non-native Speakers of English. *Language Awareness Journal, 15* (3), 141-159.

Reid, A., Burnham, D., Kasisopa, B., Reilly, R., S., Attina, V., Rattanasone, N. Y., & Best, C. (2015). Perceptual assimilation of lexical tone: the roles of language experience and visual information. *Attention, perception & psychophysics*, 77 (2), 571-591. http://europepmc.org/articles/PMC4335097

Roth, W. M. (2001). Learning science through technological design. *Journal of Research in Science Teaching, 38*, 768-790.

Sadakata, M., & Sekiyama, K. (2011). Enhanced perception of various linguistic features by musicians: a cross-linguistic study. *Acta Phychologica, 138,* 1-10.

Saltz, E., & Donnenwerth-Nolan, S. (1981). Does motoric imagery facilitate memory for sentences? A selective interference test. *Journal of Verbal Learning & Verbal Behaviour*, *20*(3):322–332.

Sekiyama K., & Burnham D. (2008). Impact of language on development of auditory-visual speech perception. *Developmental Science, 11*(2), 306-320.

Slevc, L. R. & Miyake, A. (2006). Individual differences in second language proficiency: does musical ability matter? *Psychological Science, 17*, 675-681.

Smith, D., & Burnham, D. (2012). Facilitation of Mandarin tone perception by visual speech in clear and degraded audio: Implications for cochlear implants. *Journal of the Acoustical Society of America*, *131*, 1480-1489.

So, C. K. (2006). Perception of non-native tonal contrast: Effect of native phonological and phonetic influence. In Proceedings of the 11th *Australian International Conference on Speech Science & Technology*, 438-443

So, W. C., Sim, C., & Low, W. S. (2012). Mnemonic effect of iconic gesture and beat in adults and children. Is meaning important for memory recall? *Language and Cognitive Processes, 5*, 665-681

Soto-Faraco, S., Navarra, J., & Alsius, A. (2004) Assessing automaticity in audiovisual speech integration: evidence from the speeded classification task. *Cognition 92* (3), B13-B23.

Srinivasan, R. J., & Massaro, D. W. (2003).Perceiving prosody from the face and voice: Distinguishing statements from echoic questions in English. *Language Speech, 46*, 1-22.

Strait D. L., Kraus N., Parbery-Clark A., & Ashley R. (2010). Musical experience shapes top-down auditory mechanisms: evidence from masking and auditory attention performance. *Hearing Research*, 261.

Tellier, M. (2008). The effect of gestures on second language memorisation by young children. *Gesture*, 8 (2), 219-235.

Thompson, W. F., Schellenberg, E. G., & Husain, G. (2004). Decoding speech prosody: Do music lessons help? *Emotion, 4,* 46–64.

Tong, X. & Tang, Y. C. (2016) Modulation of Musical Experience and Prosodic Complexity on Lexical Pitch Learning. In *Enhanced Proceedings of Speech Prosody Convention*.(https://drive.google.com/file/d/0B5tX0NnTCf-IdXJwejUwNlNGTnc/view)

Vatikiotis-Bateson, E., Kuratate, T., Munhall, K. G., & Yehia, H. C. (2000). The production and perception of a realistic talking face", In O. Fujimura, B. D. Joseph and B. Palek (Eds.), *Proceedings of LP'98, Item order in language and speech* (pp. 439-460), Prague: Karolinum Press (Charles University).

Wagner, S., M., Nusbaum, H. & Goldin-Meadow, S. (2004). Probing the mental representation of gesture: Is hand waving spatial? *Journal of Memory and Language, 50*(4), 395-407.

Wang, Y., Behne, D. M., & Jiang, H. (2008). Linguistic experience and audio-visual perception of non-native fricatives. *The Journal of the Acoustical Society of America, 124* (3), 1716-1726.

Wang, Y., Jongman, A., & Sereno, J. A. (2003a). Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *Journal of the Acoustical Society of America, 113*, 1033-1044.

Wang, Y., Sereno, J.A., Jongman, A., and Hirsch, J. (2003b). fMRI evidence for cortical modification during learning of Mandarin lexical tone. *Journal of Cognitive Neuroscience*, 15(7), 1-9

Wang, Y., Spence, M. M., Jongman, A., & Sereno, J. A. (1999). Training American listeners to perceive Mandarin tones. *Journal of the Acoustical Society of America 106*, 3649-3658.

Wang, Y., Sereno, J.A., Jongman, A., and Hirsch, J. (2000). Cortical reorganization associated with the acquisition of Mandarin tones by American learners: An fMRI study. Proceedings of the 6th *International conference on Spoken Language Processing, 2*, 511-514.

Wilde N., & Strauss, E. (2002) Functional equivalence of WAIS – III/WMS – III digit and spatial span under forward and backward recall conditions. *The Clinical Neuropsychologist, 16,* 322–330.

Wong P. C. M., & Perrachione T. K. (2007). Learning pitch patterns in lexical identification by native English-speaking adults. *Appllied Psycholinguistics, 28,* 565–585.

Wu, H., Ma X., Zhang L., Liu Y., Zhang Y., & Shu H. (2015) Musical experience modulates categorical perception of lexical tones by native Chinese speakers, *Frontiers in Psychology 6*,436.

Wu, Y. C. & Coulson, S. (2005). Meaningful gestures: Electrophysiological indices of iconic gesture comprehension. *Psychophysiology, 42*(6), 654-667.
Xiaonong, Zhu (2012) .Multi-Registers and Four Levels: a New Tonal Model. *Journal of Chinese Linguistics, 40* (1), 1-18

Yehia, H. C., Kuratate, T., & Vatikiotis-Bateson, E. (2002). Linking facial animation, head motion and speech acoustics. *Journal of Phonetics 30*, 555-568.

Appendix A

5 words in a row list
1. Pastis, cadira, autobús, granota, anell
2. Pallasso, clau, corbata, dit, coll
3. Ull, cotxe, mitjó, jersei, rellotge
4. Taula, texans, poma, mà, ordinador
5. Manta, llum, síndria, tassa, quadre
6. Bolígraf, samarreta, armari, disc, cor

6 words in a row list
1. Diari, parella, plat, música, catifa, gos,
2. Quadern, pilota, opció, flor, cama, pistola
3. llençol, porta, bufanda, cullera, borsa, patata
4. Cremallera, ganivet, pel-lícula, gorra, gat, petó
5. Maduixa, tele, forquilla, escuma, llavis, roda
6. Treball, abric, mosca, estufa, full, taronja

7 words in a row list
1. Riba, esperança, granja, tennis, ferro, muntanya, sol
2. Sang, somni, semana, cavall, edat, soroll, pastanaga
3. Pebrot, habitació, exemplar, ovella, illa, sopa, amic
4. Xifra, raó, vidre, conversa, any, cercle, home
5. Neu, casa, fill, situació, sistema, orella, espectacle
6. Noia, minut, pregunta, joc, sal, idea, arbre

8 words in a row
1. nou, ocell, riu, diners, cançió, nom, viatge, cos
2. tanca, equip, energía, meló, pijama, gent, cinturó, balcó
3. ciutat, ala, nen, postres, euro, estrella, escola, interès
4. oceà, cultura, ratolí, milla, llapis, all, pestanya, sopar
5. demà, historia, sort, paraula, pare, foscor, lunar, marca
6. paraigua, aigua, avi, divorci, cabell, carta, misteri, teclat

9 words in a row
1. nombre, bigoti, porc, línea, ball, terra, escriptor, pendents, coma
2.botiga, salsa, assecador, maleta, coixí, signatura, oferta, celles, tapa
3. brúixola, polsera, vaixell, camí, profesor, endoll, sabata, assumpte, paella
4. ascensor, matrimoni, móbil, libre, pera, xampany, tovalló, bronca, exemple
5. plàtan, colònia, so, veí, escuma, poeta, empresari, veu, deute
6. article, lladre, cabra, núvol, espia, caixa, compra, esborrany, onada.