

The Effects of Internal and Experience-Based Factors on the Perception of Lexical Pitch
Accent by Native and Nonnative Japanese Listeners

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy
in the Graduate School of The Ohio State University

By

Seth Joshua Goss

Graduate Program in East Asian Languages and Literatures

The Ohio State University

2015

Dissertation Committee:

Mineharu Nakayama, Advisor

Mari Noda

Ludmila Isurin

Copyright by
Seth Joshua Goss
2015

Abstract

This dissertation explores the influence of both learner-internal cognitive resources and experience-based factors on the perception of Japanese lexical accent. In Japanese, individual words carry a pitch pattern as part of their phonological form, and these accent patterns form a basic prosodic unit of the language. Cross-linguistic speech perception research has focused by and large on phonetic comparisons of a learner's native language (L1) with that of the second language (L2). Yet, a large degree of individual variation in L2 Japanese accent perception ability remains unaccounted for, suggesting that some learners—despite speaking the same L1 and possessing similar learning experiences—are better at attuning to new sound categories than others. Couple this with previous findings indicating that accent perception does not develop in parallel with Japanese proficiency level, and that immersion in the target-language environment seems to provide little advantage, and the need for a listener-focused investigation becomes apparent. The present research thus attempted to clarify this reported variation by looking at a range of learner-variables potentially involved in prosodic perception.

Three experiments investigated native and nonnative Japanese listeners' accent perception ability on correctness-judgment and categorization tasks. Experiment 1 looked at L1 Japanese listeners. Experiment 2 focused on advanced learners of Japanese from two language backgrounds, L1 Chinese and L1 Korean speakers. Experiment 3 measured

the development of accent perception over a semester of study in L1 English beginning learners of Japanese in two learning contexts, at-home and study-abroad. In all experiments, participants' phonological short-term memory capacity and acoustic pitch sensitivity were measured as learner-internal predictors of accent perception ability. Experience-based factors included L1 background, Japanese lexicon size and learning context. Multiple regression analyses were then used to determine the relative contribution of each of these factors on listeners' perception of lexical pitch accent.

The results revealed that for L1 Japanese speakers, who were very accurate at the correctness-judgment task, acoustic sensitivity predicted perception accuracy. Despite possessing robust knowledge of the phonological properties of their L1, this finding suggests that a basic cognitive resource is still active in L1 accent perception. With the advanced learners, two factors were significantly related to their ability to perceive accent—L1 background and L2 lexicon size. On the other hand, for beginning learners, who were still actively acquiring Japanese, the two basic cognitive resources predicted which individuals made the largest gains in perceptual ability over a semester of study. Thus, for L2 Japanese learners, I posit an acoustic-to-lexical continuum for accent acquisition, whereby beginners are reliant on memory and basic acoustic sensitivity to support their listening performance. However, as I observed with the advanced group, learners increasingly rely on their long-term knowledge of L2 word form in perceiving lexical accent. The pedagogical application of these results is discussed in terms of the increased use of form-focused correction of accent and visual supplements in the language classroom to aid lower-capacity learners' perceptual development.

This dissertation is dedicated to my family and friends.

Acknowledgments

Seeing a project the size of a dissertation through to the end cannot be done alone, and there are many to whom I am grateful. My sincerest thanks first go to my advisor Mineharu Nakayama, whose attentiveness to detail, and patience in allowing me to freely explore my research interests, not only enabled me to complete this project but also to mature as a researcher and a person. Through his mentorship alone, I've been able to turn something I love doing into a career. I am indebted to the members of my dissertation committee as well. Ludmila Isurin, who first exposed me to the fields of second language acquisition and bilingualism through her excellent seminars, showed me the importance of thinking about language acquisition holistically. I am also grateful to Mari Noda for enlightening discussions on Japanese-related topics ranging from idioms to onomatopoeia, website design to study-abroad programs.

Sincere thanks are due to Katsuo Tamaoka, Professor in the Graduate School of Languages and Cultures at Nagoya University, who hosted me in Japan for a year and was instrumental in the development of this dissertation. He kindly allowed me to use the resources of his psycholinguistics lab, and also worked closely with me on my project. With his help this project was greatly improved, and on a personal level, I developed substantially as a scholar under his tutelage. I also thank his lab members for their

assistance with various facets of my research, in particular Kyoko Hayakawa and Rinus Verdonschot.

There are many others in my own Department of East Asian Languages and Literatures to whom I owe gratitude, particularly Etsuyo Yuasa, who has been invaluable from day one in my development as a Japanese teacher and on the job search as well, and Debbie Knicely, who flawlessly handles the big task of keeping tabs on dozens of graduate students, myself included. Many fellow DEALL graduate students, past and current, have made my journey both stimulating and memorable, including Karen Curtin, Carlos Pimentel, Chris Stellato, Artem Vorobiev and Seth Wiener.

I am also deeply appreciative of the financial support I received both from DEALL and from the Japan Foundation's Doctoral Candidate Fellowship Program, without which this research would not have been possible. In particular I'd like to thank Nobutaka Saiki and Chie Yamamoto of the Japan Foundation's Kyoto office for their thorough support and also for organizing the wonderful cultural activities in and around the city, which really made for an ideal work-life balance while in Japan. Thanks are also due to Shuji Fujimura in the Tokyo office for handling many of the logistical issues inherent in an international move.

Lastly, I thank my parents for their encouragement and loving support during the long years of graduate school, and for seeing me off and welcoming me home after numerous trips abroad. And most of all, I thank my wife Sayo and son Louis for the love, patience and humor they show me daily.

Vita

2001.....	B.A. in Asian Studies, University of North Carolina, Chapel Hill
2009.....	M.A. in Japanese linguistics, The Ohio State University
2007 to present	Graduate Teaching Associate, Department of East Asian Languages & Literatures, The Ohio State University

Publications

- Goss, S. & Nakayama, M. (2011). Prosody and comprehension in oral reading by L2 learners of Japanese. *Second Language, 10*, 33-50.
- Goss, S. & Tamaoka, K. (2013). Working memory or pitch sensitivity? Pitch accent perception by native speakers of Japanese. *Proceedings of the Linguistic Society of Japan 146th Meeting*. 412-417.
- Goss, S. (2013). Review of “An Introduction to Japanese Linguistics, 3rd Edition” by Natsuko Tsujimura. *Journal of Japanese Linguistics, 29*, 51-53.
- Goss, S. & Tamaoka, K. (2015). Predicting lexical accent perception in native Japanese speakers: An investigation of acoustic pitch sensitivity and working memory. *Japanese Psychological Research, 57*(2), 143-154.

Fields of Study

Major Field: East Asian Languages and Literatures

Table of Contents

Abstract	ii
Dedication	iv
Acknowledgments.....	v
Vita.....	vii
List of Tables	xiv
List of Figures	xvi
Chapter 1: Introduction	1
1.1 Phonetic/phonological properties of Japanese lexical accent	4
1.2 Assumptions for the experiments.....	10
1.2.1 Experiment 1 – Native Japanese speakers.....	10
1.2.2 Experiment 2 – Advanced learners of Japanese	11
1.2.3 Experiment 3 – Beginning learners of Japanese.....	12
1.3 Organization of the dissertation	14
Chapter 2: Previous Studies	16
2.1 A psycholinguistic/processing account of lexical accent.....	16
2.2 Perception of lexical accent by L1 Japanese	24

2.2.1 Evidence from L1 listeners	24
2.2.2 Task design issues in previous studies	29
2.3 Accent perception/acquisition by L2 learners	34
2.3.1 Evidence from L2 learners	34
2.3.2 Individual variation in L2 accent perception	39
2.4 Learner-internal and experience-based factors in accent acquisition.....	42
2.4.1 Phonological short-term memory	43
2.4.2 Acoustic pitch sensitivity.....	46
2.4.3 L2 lexical knowledge.....	50
2.4.4 Influence from the L1 prosodic system	54
2.4.5 Learning context and instruction style.....	58
2.5 Looking ahead	59
Chapter 3: Experiment 1 – Accent Perception by Native Japanese Speakers	62
3.1 Introduction	62
3.2 Predictions	62
3.3 Methods	63
3.3.1 Participants	63
3.3.2 Materials	64
3.3.2.1 Phonological memory	64

3.3.2.2 Acoustic pitch sensitivity.....	66
3.3.2.3 Lexical accent perception tasks	67
3.3.3 Procedure	69
3.4 Results	71
3.4.1 Predictor variables	71
3.4.2 Accent perception tasks	73
3.4.3 Correlations between predictors and perception tasks	73
3.4.4 Regression analysis.....	74
3.4.5 Pattern analysis	76
3.4.6 Summary of results.....	78
3.5 Discussion	78
Chapter 4: Experiment 2 – Accent Perception by Advanced Learners of Japanese	83
4.1 Introduction	83
4.2 Predictions.....	84
4.3 Methods.....	86
4.3.1 Participants	86
4.3.2 Materials	87
4.3.2.1 Acoustic pitch sensitivity.....	87
4.3.2.2 PM	88

4.3.2.3 Lexical/grammatical proficiency tests	88
4.3.2.4 Pitch accent identification and categorization tasks	89
4.4 Procedure.....	91
4.5 Results	92
4.5.1 Predictor variables	92
4.5.2 Lexical accent perception tasks	94
4.5.3 Relationship between learner-variables and pitch perception	95
4.5.4 Multiple regression analyses	97
4.5.5 Analysis of individual accent patterns.....	99
4.5.5.1 Pattern type and PitchID performance.....	99
4.5.5.2 Pattern type and PitchCAT	102
4.7 Discussion.....	104
4.8 Summary of the experiment	111
Chapter 5: Experiment 3 – Accent Perception by Beginning Learners of Japanese	113
5.1 Introduction	113
5.1.1 Internal resources and beginning learners	114
5.1.2 Learning context and L2 acquisition	116
5.1.3 Accent acquisition over a semester of language study	117
5.2 Predictions.....	119

5.3 Methods	120
5.3.1 Participants	120
5.3.2 Materials	121
5.3.2.1 Acoustic pitch sensitivity.....	121
5.3.2.2 Nonword discrimination	122
5.3.2.3 Phonological memory.....	123
5.3.2.4 Pitch accent identification and categorization tasks	123
5.4 Procedure.....	126
5.5 Results	126
5.5.1 Predictor variables	126
5.5.2 Lexical accent perception tasks	128
5.5.3 Correlations between predictor variables and accent perception tasks	132
5.5.4. Multiple regression analyses	134
5.6 Summary of results.....	137
5.7 Discussion	137
Chapter 6: General Discussion and Conclusions	141
6.1 Summary of results from Experiments 1-3	141
6.2 Performance by perception task type	143
6.2.1 Correctness judgment task (PitchID).....	145

6.2.3 Categorization task (PitchCAT)	146
6.3 Cognitive resources and accent perception	149
6.3.1 Acoustic pitch sensitivity.....	149
6.3.2 Phonological memory and accent acquisition	153
6.4 Experience-based factors in accent perception	157
6.4.1 L1 background.....	157
6.4.2.1 Pattern analysis of accent categorization.....	159
6.4.2 L2 lexical knowledge.....	163
6.4.3 Learning context.....	164
6.5 Pedagogical application.....	166
6.6 Conclusion.....	171
References.....	173
Appendix A – Experiment 1 test stimuli.....	183
Appendix B – Experiment 2 test stimuli.....	184
Appendix C – Experiment 3 test stimuli.....	186
Appendix D – Background Questionnaire	188

List of Tables

Table 1. Descriptive statistics for all variables.	71
Table 2. Correlation coefficients for all variables.	74
Table 3. Multiple regression results for the PitchID and PitchCAT tasks.....	75
Table 4. Overall responses by accent pattern on both accent perception tasks.	76
Table 5. Overview of participants' background and Japanese learning experience.	87
Table 6. Descriptive statistics for all variables in L1 Chinese and Korean groups	93
Table 7. Correlations between predictor variables and accent perception tasks	96
Table 8. Multiple regression analysis for the entire cohort	98
Table 9. Correct/Incorrect response totals on the PitchID task for both L1 groups.	99
Table 10. Confusion matrix for Japanese pitch categorizations by L1 Chinese.....	103
Table 11. Confusion matrix for Japanese pitch categorizations by L1 Korean	104
Table 12. Participant data for both groups.....	120
Table 13. Results of learner-internal variables for combined cohort.....	127
Table 14. Comparison of learner-internal variables separated by group	127
Table 15. Comparison by group of accent perception tasks	131
Table 16. Accuracy gain scores over the 10-week testing period.....	132
Table 17. Correlation matrix for predictor variables and pitch accent gain scores	134
Table 18. Hierarchical multiple regression analysis for PitchID gain	135

Table 19. Hierarchical multiple regression analysis for PitchCAT gain	136
---	-----

List of Figures

Figure 1. Pitch waveforms for each of the four accent patterns in (1a-d).....	7
Figure 2. Phonetic to lexical continuum in speech prosody perception tasks.	33
Figure 3. Sample screen of the adaptive pitch test.....	66
Figure 4. Sample screen of the PitchCAT task	69
Figure 5. Mean accuracy scores on PitchID and PitchCAT tasks	94
Figure 6. Decision tree analysis of PitchID performance with both L1 groups	101
Figure 7. Sample screen of the PitchCAT task.	124
Figure 8. Group comparison of PitchID accuracy at Times 1 & 2	129
Figure 9. Group comparison of PitchCAT accuracy at Times 1 & 2	130

Chapter 1: Introduction

Learning to accurately perceive the sound structure of words in a second or foreign language (L2) is a complex process that lies at the heart of language acquisition. Speech perception involves not only the capacity to hear the physical, or acoustic, properties of spoken input, but also the ability to associate these sounds with representations stored in our long-term knowledge of the second language. In other words, learners are engaged in a process of analyzing the speech stream in order to build up a store of lexical knowledge as they gain proficiency in the language. Word-level prosodic features, such as pitch and tone, however, are frequently reported to be ongoing sources of perception difficulty for some L2 learners, despite their having made gains in overall proficiency (e.g., Lee, Murashima & Shirai, 2006; Nishinuma, Arai & Ayusawa, 1996; Shport, 2008; Ueyama, 2000; Wang, 2013). Japanese is one example of a language in which individual words carry a pitch accent pattern, leading it to be classified as a lexically-accented language (Beckman & Pierrehumbert, 1986; Tsujimura, 2014). Described generally, each syllable, or mora, within a word is associated with either a high or low pitch as a part of its phonological form. Considering the lexically-linked nature of pitch accent in Japanese, when learning vocabulary, L2 Japanese learners must acquire a word's accent pattern as they would its segmental features, such as consonants and vowels. Lexical pitch accent then being an inherent property of Tokyo-Standard

Japanese¹ (Tanaka & Kubozono, 2012), L2 learners are exposed to the accentual system either explicitly, through instruction in Japanese language classes and their textbooks, or implicitly by naturalistic exposure in the L2 environment. Yet, language acquisition research has shown that Japanese learners have ongoing difficulty both perceiving and producing lexical accents (e.g., Lee, Murashima & Shirai, 2006; Shibata & Hurtig, 2008). Thus, the acquisition of lexical accent has often been characterized as a random process, marked by a large degree of variation between individual learners with the same amount of exposure to Japanese (Hirano-Cook, 2011; Shport, 2011; Taylor, 2011, 2012). Despite evidence for gradual gains in other areas of Japanese phonology, such as the acquisition of short and long vowel contrasts, accent perception does not appear to develop in parallel with proficiency level (Hardison & Motohashi Saigo, 2010; Shibata & Hurtig, 2008; Tsukada, 2012). This may partly stem from the inherent phonetic properties of lexical accent—which features subtle pitch variations in its accent contours (Kitahara, 2001; Warner, Otake & Arai, 2010) and an absence of intensity and durational cues found in stress-accented languages like English (Beckman & Pierrehumbert, 1986; Ueyama, 2000). However, cross-linguistic comparisons and phonetic feature analyses have not accounted for all of the variation (e.g., Bohn, 1995; Strange & Shafer, 2008), leaving the question unanswered as to what factors within *individual* learners may result in the observed disparity in perceptual ability.

¹ In the present study, I limit my focus to the accent system of Tokyo-standard Japanese. This variety of Japanese is spoken over wide areas of Japan, used primarily on national television broadcasts, and taught in second and foreign language classrooms. Despite its name, several regions outside of the Tokyo area are also classified as standard-dialect in terms of lexical accent (Shibatani, 1990).

In the present study I propose that a significant amount of this variation in accent perception ability by L2 Japanese learners can be accounted for by differences in learner-internal resources on the one hand, and “top-down” or experience-based factors on the other, neither of which to my knowledge have been experimentally tested on the acquisition of a lexically-accented language. Considering that pitch is an obligatory lexical feature in Japanese, I argue that learner-internal resources which support word learning also facilitate the acquisition of lexical pitch accent. Among the resources which have received attention in second language acquisition (SLA) literature are phonological short-term memory, shortened to phonological memory (PM) hereafter (Kormos & Safar, 2008; O’Brien, Segalowitz, Collentine & Feed, 2006; O’Brien, Segalowitz, Freed & Collentine, 2007; Sunderman & Kroll, 2009); acoustic, or nonlinguistic, pitch sensitivity (Hao, 2012; Wayland, Herrera & Kaan, 2010; Wong & Perrachione, 2007); and L2 lexical knowledge (Martin & Ellis, 2012; Speciale, Ellis & Bywater, 2004)—all factors which inherently vary from learner to learner. It is generally assumed that individual differences in these resources manifest as differences in acquisition outcomes (Ellis, 2004). Furthermore, learners at different levels of L2 proficiency may use different resources when perceiving lexical accent.

In this dissertation I present an analysis of multiple learner-variables and their relationship with L2 learners’ ability to acquire Japanese lexical accent. I divide these variables into the basic cognitive resources phonological memory and acoustic pitch sensitivity, and the higher-level, or experience-based, factors of L2 vocabulary knowledge and native-language experience with lexical tone. Learning context and

instruction style will also be considered in the latter experienced-based category, as these may also exert a considerable influence on learning outcomes. Multi-variable approaches to L2 acquisition research have been increasingly called for in the field of SLA, as they enable a more ecologically-valid view of the complexity of language acquisition in quantitative L2 research (e.g., Cunnings, 2012; Plonsky, 2013). Such approaches provide an alternative to the more traditional experimental (i.e., control/comparison group) designs, which is particularly useful when considering difficult-to-control factors such as language-learning experience, multiple proficiency levels and language background. In the next section I will provide an overview of the linguistic properties of Japanese lexical pitch accent.

1.1 Phonetic/phonological properties of Japanese lexical accent

The use of prosodic, or rhythmic, features to convey meaning, prominence or emphasis is virtually a linguistic universal (Fox, 2002). Yet, different languages accomplish these functions in acoustically different ways and over different units of speech. For instance, at the word level, we have the typical pattern of alternating strong-weak stress on syllables in English, which is realized as pitch increases and duration and intensity changes of vowel sounds. Contrast this with Mandarin Chinese, on the other hand, which features syllable-level tone variations that are mainly indicated by pitch movements over the duration of a syllable. In Japanese, variations in pitch over adjacent mora, the syllable-like timing unit of Japanese, form its fundamental level of prosody (Beckman & Pierrehumbert, 1986; Fox, 2002). In strictly acoustic terms, changes in fundamental frequency (F0), which to the listener are perceived as pitch variations, are

used to mark certain mora within a word as being more prominent than others. It is this word-level, prominence-marking function of accent, an inherent property of Japanese words, on which I will focus in this dissertation.

In Tokyo-standard Japanese, a sharp fall in pitch from H(igh) to L(ow) is the sole prosodic variation which marks lexical accent location. Sentence-final intonation, on the other hand, can feature pitch contours with varying degrees of rises and falls, as in those used to express pragmatic intentions such as, among many, surprise and confirmation (Eda, 2004). A labeling system for these pitch variations, called J_TOBI, has been adapted to transcribe Japanese prosody (Venditti, 2005).² In this system, lexical accent is notated by the shape H*+L, in which H* label marks the mora to which the high-pitch accent is linked, followed by an L drop to low pitch on the following mora. As we can see from this label, lexical accent requires two adjacent mora as its minimal unit in order to be phonetically realized. In other words, without the fall to the low tone, the listener has no way of discerning the location of an accent. The possible shapes of these accent patterns are therefore simpler than the syllable-level pitch contours found in tone languages, like Mandarin or Cantonese, which both feature large inventories of minimal tone pairs (Shibata & Shibata, 1990). In English, for that matter, Pierrehumbert (1980) identified six possible lexical pitch shapes that are linked to the stressed syllable in a word. Generally speaking, English speakers can consciously vary these accents to create changes in their emphasis or pragmatic intention, while in Japanese lexical accents remain unchanged regardless of sentential context or speaker intention.

² Short for “Japanese Tones and Breaks Indices”

Despite the comparatively simplistic inventory of pitch shapes, however, in Tokyo-standard Japanese pitch accent is not an optional feature as are, for instance, stress and intonation rises used to mark focus or pragmatic meaning in English.³ An illustration of this is the fact that even in phrase-level prosodic phenomena, lexical accents are maintained. For example, downstep, or the lowering of pitch height on each subsequent accented word in a phrase in a stair step-like manner, depends on the presence of lexical accent (Venditti, 2005). Considering examples of specific words, a 3-mora Japanese noun can have $n+1$ pitch patterns, which are realized over the duration of the word, with the high tone extending into the following postposition in the “unaccented” pattern (1d below). Note that capital letters indicate high-pitch mora. When presented in isolation, without a postposition, patterns (1c) and (1d) are identical because the location of the pitch fall is not vocalized until just prior to the postposition in (1c).

- (1) a. *megane ga* (MEgane-ga) ‘glasses + SUBJECT marker’
- b. *tamago ga* (taMAgo-ga) ‘egg + SUBJ’
- c. *otoko ga* (oTOKO-ga) ‘man + SUBJ’
- d. *sakana ga* (saKANA-GA) ‘fish + SUBJ’ (Tanaka & Kubozono, 2012)

With nouns, the location of an H to L fall in pitch, or the absence of such a drop in unaccented words (1d), is unpredictable in that one cannot guess the pitch pattern with certainty simply by knowing a word’s length or segmental structure (Tsujimura, 2014).

³ Although in English, words do have a fixed stress nucleus and stress-accent opposed pairs such as REbel/reBEL exist. Notice however, in this pair the non-stressed syllable is reduced to schwa [ə] as in [rəbel].

However, once the location of the pitch fall is known, the pitch of the remaining mora in a word can be deduced given that only one H*+L accent exists in any single word (Tanaka & Kubozono, 2012). Given that F0 modulations are the primary cue to Japanese lexical accent, the clearest way to visualize the phonetic basis of accent is by looking at their pitch contours. The figures below present the pitch contours for each of the words in (1).

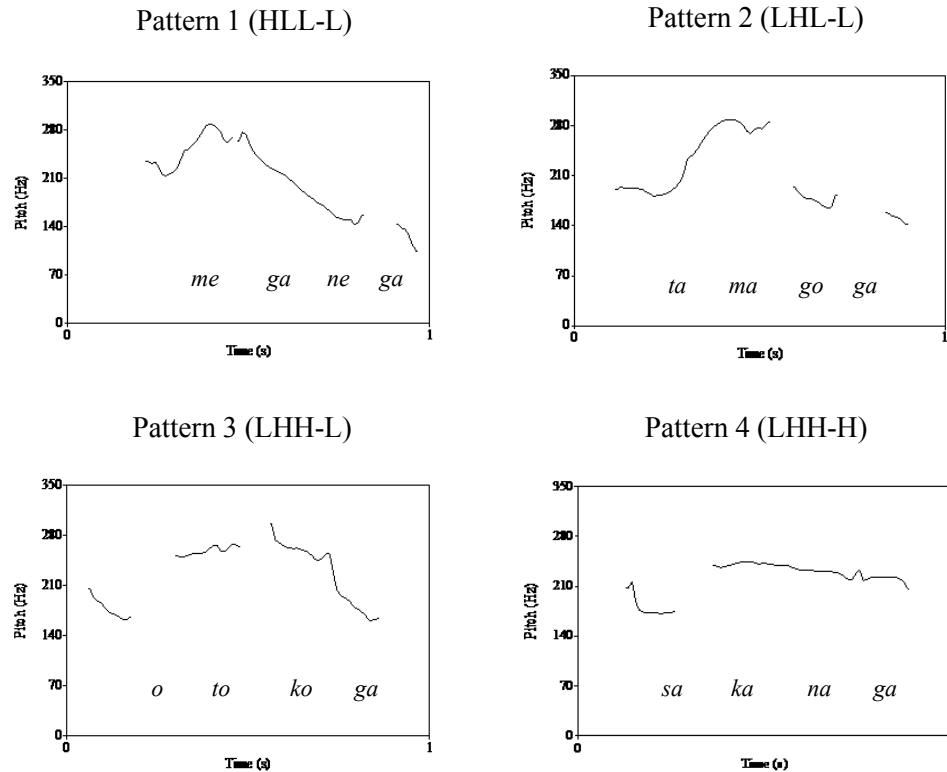


Figure 1. Pitch waveforms for each of the four accent patterns in (1a-d) above. Pattern 4 (1d) is generally referred to as the unaccented pattern because of the absence of a pitch fall within the word. Samples produced by a female speaker from a Tokyo dialect region.

In terms of the function of lexical accent, two main features are relevant when considering speech perception—the lexically (homophonic) distinctive function and the prominence-marking function of pitch (Beckman, 1986; Kitahara, 2001). First is the well-documented function of homophone distinction (e.g., *KAmi* ‘god’ vs. *kaMI* ‘paper’) in which pitch is used to differentiate a minimal pair (Vance, 2008). This function is roughly analogous to the primary role of tone in Mandarin Chinese. In Mandarin, tone plays a particularly important role in homophone distinction, as attested by the sheer number of tonal oppositions found in that language (for example, the often-cited four meanings of the syllable *ma*). Shibata & Shibata (1990), reported that around 14% of Japanese segmental homophones rely on pitch accent for perceptual disambiguation; contrast this with figures for Mandarin which indicate that upwards of 71% of words feature at least one tonally-contrastive pair. However, Kitahara (2001) noted that it is often the case that one member of a minimal pair in Japanese typically is much lower in frequency than the other, which may serve to reduce the distinctive function of accent in such instances. That is, the combined effect of frequency and sentence context likely serve to mitigate the reliance on pitch alone in perceiving Japanese, which may have implications on how both native and nonnative listeners perceive accents.

While only a subset of Japanese words require pitch for homophone distinction, as we have seen, all words feature some type of pitch variation, resulting in the prominence-marking function of accent applying to the entire lexicon, and not just to a subset of items. We can thus consider this function as mainly a lexical form-based property of Japanese, in that it is not extensively related to accessing word meaning, but

rather connected to a speakers' knowledge of how a word should be pronounced. In other words, pronunciation deviations from the standard accent, although not often resulting in one word being mistaken for another, will stand out as being spoken incorrectly. Not only this, but word recognition may be hindered to an extent, despite access still being possible. It is this form-based property of lexical accent, and how learners perceive and acquire it, that I am currently interested in. Pitch accent forms an important part of the metrical system of Japanese, which we can liken to stress accent in English. Although the two languages realize prominence-marking differently, the crucial point here is that this prominence is stored as a part of lexical form in a native speaker's long-term knowledge of words, that is, in the mental lexicon. Therefore, barring dialectal differences, native Japanese listeners are presumably very accurate at judging whether a word's pitch accent is misplaced, particularly for those of relatively high-frequency.⁴

In the present study I focus on how both L1 and L2 speakers perceive this prominence-marking function of Japanese lexical accent. Assuming that accentual information is a part of L1 Japanese speakers' lexical knowledge, research on how pitch is processed will be an informative starting point for my proposed experiments. In the next section, I will address several assumptions which were used to frame the experiments in this dissertation, organized around the three participant groups.

⁴ Similarly, a native English speaker can tell that the stress in classROOM, for example, is misplaced on second syllable, instead of the correct placement on the first syllable.

1.2 Assumptions for the experiments

1.2.1 Experiment 1 – Native Japanese speakers

In Tokyo-Standard Japanese, pitch accent is not an optional feature as is stress used to express pragmatic intention in English, for example (Vance, 2008). Therefore, L1 listeners should be able to make accurate and rapid decisions as to whether aurally-presented words are spoken with the correct or incorrect accent pattern (Otake & Cutler, 1999; Sekiguchi & Nakajima, 1999). However, some previous studies have shown that even L1 Japanese speakers from Tokyo-standard dialect regions have difficulty on accent perception tasks, but I argue that this is a result of using low-frequency lexical items presented in tasks that are biased toward an acoustic mode of perception, and that are also demanding on listeners' phonological memory capacity (Hirano-Cook, 2011; Shport, 2008). For example, a task in which participants are asked to identify the location of an accented syllable, or perform a same/different judgment on two aurally-presented words could likely be completed by someone who is highly sensitive to pitch rises and falls (i.e., trained musicians) (Wayland et al., 2010; Wong & Perrachione, 2007). Further, performance variation may arise when listeners must hold speech input in memory while making a complex categorization of a word's accent pattern. I consider this type of task to also involve a degree of metalinguistic knowledge of pitch accent, which native speakers do not necessarily possess. In fact, there is no standard way to transcribe Japanese pitch accent patterns, and training with the notation system may be necessary prior to the perception task. On the other hand, tasks requiring participants to decide whether a word was spoken with the correct or incorrect accent necessarily invoke prior

knowledge of the target words. That is, in making a correctness decision, listeners must compare the sound of a spoken word to the representation stored in their mental lexicon. Furthermore, a word's frequency, or even its subjective familiarity, may be directly related to its ease of recognition by Japanese listeners, whereby very commonly heard words are identified with greater accuracy than uncommon ones (Amano & Kondo, 1999; Ueno et al., 2014). In the experiments presented in this dissertation, I first attempt to clarify the conflicting findings regarding L1 lexical accent perception by employing frequency-controlled words in both an online correctness-decision task (in which reaction time is measured) and an offline categorization task, where participants categorize accents based on visual representations of accent patterns.

1.2.2 Experiment 2 – Advanced learners of Japanese

Previous research has suggested that L2 acquisition of lexical prosody follows a phonetic-to-lexical continuum, meaning that basic cognitive abilities such as acoustic, or non-linguistic, pitch sensitivity support the early stages of acquisition, while L2 lexical knowledge increasingly becomes the foundation for later lexical learning (e.g., Bent, Bradlow & Wright, 2006; Martin & Ellis, 2012; Wong & Perrachione, 2007). As L2 learners accrue knowledge of the target language, lexical decisions become not only more accurate, but more automatic as well (Segalowitz, Segalowitz & Wood, 1998; Speciale et al., 2004). Considering that earlier studies have shown contributions of the long-term lexical store in both L2 production and perception tasks, I assume that within a group of advanced Japanese learners, those with a larger L2 lexicon likely have more exemplars of word-level pitch accents. Furthermore, these representations are potentially more stable,

thus enabling higher accuracy and more rapid judgments on perception tasks. Hence, my first aim is to show that at the advanced level, those with greater Japanese lexical knowledge are more accurate on pitch accent perception tasks.

The second variable of interest with advanced learners is the effect of L1 background on lexical accent perception. I compare two groups of advanced Japanese learners from L1 backgrounds in which pitch differs in its phonemic status, namely, Chinese (+lexical tone) and Korean (-lexical tone) speakers. The presence of lexical-level phonemic tone in Chinese may enhance attentional focus to lexical accent in Japanese, making pitch variations more salient for L1 Chinese learners of Japanese (So & Best, 2010). In other words, since Chinese speakers use lexical prosody contrastively in their L1, I assume that they are more likely to focus their attention on phonologically similar contrasts in the L2. This line of reasoning leads me to argue that Chinese speakers should outperform Korean speakers, despite being closely matched for Japanese proficiency and overall learning experience, on lexical accent perception tasks if L1 experience with tone facilitates L2 perception, and if this influence continues to be active into the advanced proficiency level.

1.2.3 Experiment 3 – Beginning learners of Japanese

In the third experiment, I explore the acquisition *process* of lexical pitch accent over a semester by native English speakers in their second year of Japanese study. Drawing on research indicating that acoustic pitch sensitivity (Wayland et al., 2010; Wong & Perrachione, 2007) and phonological short-term memory (Hummel, 2009; Martin & Ellis, 2012; O'Brien et al., 2006) support lexical acquisition in various

populations of early-stage L2 learners, I argue that these two basic cognitive abilities are also tied to the acquisition of Japanese lexical accent. Considering the earlier discussion of the lexically-linked nature of Japanese, I test the prediction that basic capacities support the acquisition of prosodic lexical features as well. Beginning learners' PM capacity and acoustic pitch sensitivity were measured early in the semester, along with two accent perception tasks similar to those conducted on native Japanese speakers and advanced learners. The first test phase was followed by a second phase consisting of only the two accent perception tests, approximately ten weeks later in the semester.

Participants were selected from two learning contexts—one group studying Japanese at a university in the US and another group composed of study-abroad students at two universities in Japan—enabling me to make predictions about the role of learner-internal cognitive resources and learning context on the acquisition of lexical prosody.

In sum, by testing three groups—L1 Japanese, advanced L2 learners and beginning L2 learners—my aim is to depict the process of, and clarify previous findings on, lexical accent acquisition across multiple stages of L2 learning. I will accomplish this by testing both learner-internal and experience-based factors which may contribute to how, and how well, Japanese lexical accent is learned. I hope to enhance the larger body of SLA literature by showing that L2 learners' "tuning-in" to the feature of pitch accent to some extent follows a phonetic-to-lexical continuum, whereby beginning learners rely on basic cognitive resources while advanced learners make use of top-down processing strategies. Furthermore, this focus on learner-internal differences in L2 speech perception will not only help to clarify the individual variation reported in previous studies on L2

Japanese learners, but will provide a cluster of abilities to target at greater depth in future studies. I must stress here, however, that I do not consider learner-internal differences as limiting factors on the potential to acquire a second language. In fact, SLA researchers have pointed out that by identifying how learners process their L2 at different stages of acquisition, we can suggest methods to assist learners based on their individual strengths and weaknesses (e.g., Brewer & Unsworth, 2012; Ellis, 2004; Hummel, 2009). Hence, I expect this research to contribute in such a way to Japanese as a foreign or second language education as well.

1.3 Organization of the dissertation

In Chapter 2, I review relevant studies on accent perception and acquisition by both native Japanese speakers and L2 learners, as well as on the potential role of learner-variables in lexical accent perception. Chapter 3 presents Experiment 1 which looks at how L1 Japanese speakers perceive lexical accent while considering the role of perception task type and cognitive abilities in speech perception. This experiment aims to establish a performance baseline for the accent perception tasks. Advanced learners are the focus of Chapter 4, in which I explore the relationship between L2 vocabulary knowledge, native-language background and how lexical accents are perceived. In Experiment 2 in this chapter, I specifically look at learners of Japanese from two L1 backgrounds, Chinese and Korean, who are immersed in the L2 environment and have attained a high level of Japanese proficiency. These groups enable me to examine the role of L2 lexical knowledge and native-language tone experience on accent perception ability. In Chapter 5 I report Experiment 3, a longitudinal study which tracks the

development of accent perception ability in beginner JSL/JFL learners. The central focus here is the relationship of learner-internal resources and learning context with how accent perception develops over a semester of Japanese study. I tested two learner groups over a semester in order to track the development of perception ability and to observe how basic cognitive abilities relate to performance gains on these perception tasks. Taken together, the experiments in Chapters 4 and 5 enable me to examine whether different resources are involved in the processing of lexical accent at different stages of acquisition. In Chapter 6, I synthesize and discuss the results of all three experiments, then describe their implications for the acquisition of L2 lexical prosody. Finally, I consider the applicability of these findings to Japanese language instruction and the wider field of SLA.

Chapter 2: Previous Studies

In this chapter I will set the stage for the three experiments that follow by reviewing the relevant literature on the perception and acquisition of lexical accent, as well as introducing each of the variables of interest to the current study. The organization is as follows. Background from psycholinguistic research on how lexical accent is processed in real time is in order first. I next compare the findings from both first and second-language accent perception research, along with an analysis of different perception task designs and how these may influence a listeners' perceptual accuracy. Lastly, and of central importance to the current experiments, I discuss how differences in several learner-variables, both internal cognitive capacities and language experience-based factors, may account for variation in the perception and acquisition of Japanese lexical pitch accent. Accordingly, I have given each of these factors a separate section in the review that follows.

2.1 A psycholinguistic/processing account of lexical accent

First I will review studies which provide insight on how lexical prosody is involved in the processing of spoken language, which in turn will shed light on the representation of prosody in the mental lexicon. How listeners use lexical prosody to process speech has been a question of interest in psycholinguistics over the past few decades (e.g., Cutler, 1986; Fodor, 2002; Levelt et al., 1999). The central goals of this

research have been to examine the role of prosody in word recognition, homophone distinction and syntactic parsing, as well as to determine the status of pitch and tone as phonemic units (Chen et al., 2002; Cutler & Chen, 1997; Cutler & Clifton, 1984). As experimental evidence amassed which showed that listeners crucially integrate suprasegmental information into their processing of spoken language, it became clear that prosody needed to be integrated into speech perception models. In fact, Fodor (2002) stated more than a decade ago that psycholinguistics can no longer ignore the role of prosody in speech processing, underscoring the need to move beyond accounts limited to only segmental information. Since then, experimental studies on the processing of prosody have expanded to include a variety of typologically different languages.

In Japanese, the status of pitch accent as a lexically-linked property has led both linguists and psycholinguists to assume that it is stored in a native speaker's mental lexicon as a part of a word's phonological form (e.g. Beckman & Pierrehumbert, 1986; Cutler & Otake, 1999; So & Best, 2010). Beckman & Pierrehumbert (1986) refer to the Japanese accent system as having "phonologized the tonal features of prominence to mark prominent syllables in the lexicon (p. 271)." As we saw in the examples in Chapter 1, prominent mora are marked with an H tone followed by an immediate drop to L pitch on the next mora, which is expressed exclusively by a decrease in fundamental frequency (F0). English words, by contrast, have lexicalized stress which is phonetically realized through vowel lengthening and intensity changes on certain syllables, usually perceived as a strong-weak prosodic pattern. This alternation of strong and weak syllables forms the basic prosodic rhythm of English. In Chinese, four syllabically-linked tones create a large

inventory of tonal minimal pairs, resulting, as we have seen, in tone playing a much greater role in lexical disambiguation. Each of these languages has been examined in terms of how their characteristic prosodic features affects listeners' real-time speech processing.

Early studies examining how L1 English listeners process lexical prosody, mainly using priming paradigms, failed to show facilitation of semantically related target-word recognition by stress-pattern information alone (e.g., Cutler & Clifton, 1984). In general terms, priming research examines the effect of a previously presented spoken word (the prime) on a target word in terms of whether a speed difference in task performance occurs relative to a baseline stimulus. For example, this would mean that the stress-accent minimal pair *PROject* (a noun) and *proJECT* (a verb) would facilitate the recognition of words related to *job* in the first instance, and *display* in the second. If stress were a critical factor in word recognition, then we would expect priming for only the semantically related target. However, it has been shown that targets are primed equally well by both stress patterns. Cutler (1986) explained this finding as a case of the low functional load of stress in English lexical access, and emphasized the primary role of segmental structure in accessing meaning. In the previous example this would mean that the segmental information [pradzekt], including the vowel quality change due to stress placement, is sufficient to activate words related to both the noun and verb form.

On the other end of the spectrum, we would expect tone in Chinese to carry a much larger functional load than stress in English in terms of word recognition, due to the sheer volume of words distinguished by tone alone. Chen, Chen & Dell (2002) used an

implicit priming task to test both segments and tones as units of word recognition with native speakers of Mandarin Chinese. They found that the syllable-plus-tone is a crucial unit in speech production planning for Mandarin speakers, separate from orthographic knowledge of Chinese characters. That is, syllable-plus-tone units primed targets to a greater extent than just the syllable alone. However, tone-only primes did not facilitate word production, suggesting that tone does not exist as a stand-alone sublexical (i.e., phonemic) unit in the L1 mental lexicon, and is likely processed similarly to stress in English. These results underscore the lexically-linked nature of tone in Chinese speakers' word knowledge.

What about the case of Japanese, which shares some similarities of both English stress and Chinese tone? Considering the obligatory nature and unpredictability of noun accent in standard Japanese, when native speakers of Japanese process spoken input, pitch accents are likely activated as part of a word's phonological form in their mental lexicon (Cutler & Otake, 1999; Masuda-Katsuse, 2006; Sekiguchi & Nakajima, 1999). But just how great of a role does pitch accent play in lexical access for L1 Japanese listeners? Sekiguchi & Nakajima (1999) used an auditory priming task to address the question of whether native Japanese listeners use accent to access word meaning in homophonic pairs. They found that when an auditory prime was congruently accented with a *kanji* visual target (e.g., *jidoo* [HLL] – 児童 'children'), response times were shorter than for incongruently accented pairs (e.g., *jidoo* [LHH] – 自動 'automatic'). This facilitative effect held even when the auditory primes were only fragments of words. Cutler & Otake (1999) conducted three experiments to further clarify the role of pitch

accent in spoken word recognition. They reported three results which argue for a critical role of pitch accent in word selection. First, using minimal pairs, the pitch pattern of isolated syllables provided enough information for listeners to accurately guess the word from which it was extracted. Second, in a gating task in which listeners heard one mora at a time, native listeners could accurately identify a word of the same pitch structure based on only the initial mora and its accompanying accent pattern (whether initial H or L). Third, pitch-plus-segment primes speeded recognition of the same word over segment-match only primes. To date, these studies provide the clearest experimental results on the role of lexical accent in word access, and suggest evidence of a link between pitch and segmental form in the mental lexicon. Both of the above studies showed that Japanese listeners make immediate use of pitch accent information when attempting to recognize spoken words.

By contrast, recent electroencephalographic (EEG) research, which measures real-time processing by looking at electrical activity generated in the brain, has cast some doubt on the role of accent in disambiguating minimal pairs in spoken language comprehension. Tamaoka, Saito, Kiyama, Timmer & Verdonschot (2014) recorded event-related brain potentials (ERPs) while L1 Japanese speakers listened to homophone pairs embedded in sentences (e.g., *ame* [LH] for ‘candy’ and *ame* [HL] for ‘rain’). They obtained ERP evidence, in the form of a meaning-mismatch indexing N400 waveform, only for semantically incorrect words in a given context, but not for incorrectly accented homophones (e.g., こんなに甘くておいしいブドウは初めてです。 ‘It’s (my) first time having budoo this sweet and delicious.’ *buDOO* (LHH) ‘grape’ vs. *BUdoo* (HLL)

‘martial arts’); a result which they interpreted as evidence for pitch accent’s minor role in homophone distinction.⁵

The previous three studies focused on the role of accent in semantically distinguishing words that share otherwise identical structure. How do native speakers perform on a task that looks at pitch accent as a general property of lexical form, or in other words, its prominence-marking function? This question is pertinent given the relatively limited number of minimal accent pairs in Japanese. Ueno et al. (2014) recently examined, as one part of their larger study, if Tokyo-standard Japanese speakers could perceive, then imitate aloud, spoken words containing both correct and incorrect accents in a combined perception/production task.⁶ They intentionally omitted homophones, so that none of the incorrectly-accented stimuli could be misinterpreted semantically as another word, in order to test if a word’s accent pattern and imageability related to processing accuracy. Considering that some accent patterns are more frequent than others, they predicted that participants would commit more errors on words with lower frequency patterns (LHL > HLL > LHH) and of low imageability, which they described as “an index of a word’s intrinsic strength of meaning” (p. 442). An example of a word ranked low on both pattern frequency and imageability is the LHL-accented *nakami* ‘contents.’ Their participants performed the repetition task with near-ceiling (99%) accuracy, regardless of accent pattern; however, participants were significantly slower to

⁵ Note, however, that an earlier study by Hayashi et al. (2001) reported contrary ERP findings which did indicate an N400 response by Japanese listeners for incorrectly-accented homophones. Thus, the brain-based evidence on pitch accent’s function remains inconclusive.

⁶ Ueno et al. (2014) also conducted a parallel-distributed processing (neural-network) simulation of how accent is processed, in conjunction with their behavioral study. These results supported their behavioral findings, reinforcing the effects of accent pattern typicality and imageability on lexical access in Japanese.

repeat the low-imageability words of the LHL pattern. In another task, their listeners were required to perform spoken corrections on misaccented words, which they did at well above 90%, but the least-frequent LHL pattern words tended to undergo “regularization” in that they were mistakenly corrected to a more frequent pattern (i.e., LHH). In light of my proposed study, Ueno et al (2014)’s findings indicate that L1 Japanese speakers are sensitive to pitch as part of a word’s phonological form in that they can accurately perceive, and repeat aloud, words of all accents types, but that not all patterns may be perceived or produced equally well.

A related question then arises as to whether incorrectly accented, but segmentally accurate, words are perceived as nonwords by native Japanese listeners. For example, in words that are accented on the wrong mora, but in which the incorrect accent does not result in its perception as an opposing member of a minimal pair (e.g., incorrect *SUshi* vs. correct *suSHI*), do listeners perceive these as nonwords, or just as a mispronunciation? The above studies suggest the latter to be true, but that perception would likely be hindered, particularly in less-than-ideal listening conditions, such as are often encountered in daily life. In other words, as Cutler (1986) has suggested for English, lexical access is still possible, but post-access retrieval of the word likely becomes more difficult due to a mismatch in accent pattern. As both Sekiguchi & Nakajima (1999) and Cutler & Otake (1999) demonstrated, the availability of correct pitch accent information facilitated lexical decisions, although accurate identification was still possible. This line of reasoning suggests the low functional load of pitch accent in lexical access (i.e.,

homophone distinction), and shifts the focus to the prominence-marking function of accent.

Summing up the findings from psycholinguistic research on the processing of prosody, it was noted that the realization of prosody varies by language, can be exploited in word recognition, and is likely represented as a part of word form in the mental lexicon. Yet, Japanese lexical accent does not appear to share equal priority as a clue to word recognition as do vowel and consonant segments. Where does that place lexical accent in terms of its function in word recognition? Although one recent study has questioned the role of accent in homophone distinction (i.e., Tamaoka et al., 2013), in the present research I consider accent a critical property of the Japanese lexicon in its central role as a phonologized marker of lexical prominence.

Applying these findings to the framework of the current research, I assume that in Japanese, pitch accent is a part of L1 listeners' word-form knowledge in that they are aware of the location of prominence, and therefore deviations in accent location, in spoken input. In fact, phonetic perception by adult L1 listeners is considered to be robust and automatic, and likely places little burden on one's cognitive resources (Strange & Shafer, 2008). It follows that native Japanese speakers should be able to reliably, and intuitively, identify if a word was spoken correctly or incorrectly in the same way that they are able to make decisions on segmental forms, such as vowel and consonant length contrasts. In other words, the H*+L accent is crucially linked to a word's segmental form in the mental lexicon. Although it may seem obvious that this is the case, considering that all words can be classified into one of the accent pattern types, perception research on L1

Japanese speakers has provided mixed results on listeners' ability to accurately judge the accents in spoken stimuli. In the next section, I will report several of these findings while considering possible causes of variation in accent perception.

2.2 Perception of lexical accent by L1 Japanese

2.2.1 Evidence from L1 listeners

In this section I will review studies which examine the perception of lexical pitch accent by native Japanese listeners from a Tokyo-standard dialect background.

Comparatively few perception studies focusing on native speakers exist, and have often included these listeners only as control groups. Furthermore, it has often been assumed that native listeners are homogenous in terms of performance on speech perception tasks. Yet, even among speakers of the standard dialect, which is spoken over wide regions of Japan and used on most television broadcasts, variation in perception accuracy has been reported on a variety of tasks. The experimental paradigms used with L1 Japanese listeners can generally be divided into three types of task, those formulated as accent location identification (Hirano-Cook, 2011), accent pattern discrimination and categorization (Sakamoto, 2010; Shport, 2008; 2011), and correctness judgments (Shibata & Hurtig, 2008). In reviewing these studies, one important question I will address is how the perception task type relates to performance, and how valid each of these tasks are in depicting L1 listeners' knowledge of lexical accent. This discussion will have clear implications for the study of accent perception by L2 learners; namely, if native speakers cannot perceive accents any higher than chance level, then it may be that we need to reconsider our expectations for L2 learners who are acquiring Japanese accent. I will

attempt to show in the experiments that such a reevaluation of L2 accent instruction is unnecessary. Where relevant, I will note which studies used naturally-produced words and which employed phonetically manipulated stimuli.⁷

Shport (2008) reported a perception task that required Tokyo standard dialect-speaking Japanese participants to listen to a set of three words—one of which differed in pitch accent—and decide which one was accentually the ‘odd-one-out’ of the triplet (AXY-discrimination). In this task, 16 participants had to cross out the differently-accented word after hearing a series of three words, two of which, either the X or Y stimulus, had the same pitch accent. Participants only attained 59% accuracy on this task, which while better than the JFL learners in the same study, was little above chance level. A criticism that can be leveled against this study is the use of low-frequency lexical items. For example, one AXY set consisted of the word-triplet of 月下 (*gekka*-HLL) ‘under the moon’, 琥珀 (*kohaku*-LHH) ‘amber’ and 短歌 (*tanka*-HLL) ‘Japanese poem’, some of which were likely unfamiliar even to L1 listeners, and may have contributed to the low accuracy. More important was that this test essentially measured listeners’ ability to phonetically discriminate between the pitch contours of the three stimuli—a task which in essence could be accomplished by any listener who is highly sensitive to pitch variations, without any knowledge of Japanese. This factor combined with the low-frequency stimuli may have promoted an acoustic mode of perception, whereby native

⁷ Phonetic manipulation is often used to regularize the phonetic qualities of test stimuli or to examine phonetic categorization boundaries over controlled degrees of pitch variation. Naturally-produced stimuli would involve no such manipulation. The use of manipulated versus natural stimuli is, of course, related to the particular goals of the researcher.

listeners were unable to bring to bear their lexical knowledge on the task. A possible account for the above finding is that L1 Japanese speakers default to a linguistic mode of perception when listening to spoken stimuli that are recognizable as real words. This holds true even for real, but unknown (i.e., very low-frequency) words; as long as listeners believe that the stimuli they are hearing are real, as opposed to nonwords, they remain in this linguistic listening mode. However, linguistic perception may fail on very low-frequency words, resulting in decreased accuracy on accent judgments because of the difficulty to switch to an acoustic mode of perception. Thus we can explain why language-naïve listeners, who perceive stimuli only as acoustic input, can ignore the lexical information and focus solely on the pitch variations in lexical perception tasks, potentially yielding similar performance to L1 listeners.

Hirano-Cook (2011) used an accent-prominence identification task in which participants heard 48 naturally produced four-mora nouns, 16 from each of four accent patterns (i.e., HLLL, LHLL, LHLH, LHHH), and then marked accent location (or no accent for words with no pitch fall) on a *hiragana*-transcribed answer sheet. Excluding listeners from non-standard dialect regions (4 in total), the remaining 12 participants scored an average of 70% on the identification test. This lower-than-expected result was potentially due to two factors: 1) Lexical frequency was not controlled across accent type, and 2) The stimuli were by-and-large of low-frequency. Examples of particularly low-frequency items include *rashingi* ‘compass/navigation device,’ *azumaya* ‘shelter,’ and *nanohana* ‘rapeseed blossom.’ The results of Shport (2008) and Hirano-Cook (2011) suggest that low-frequency words presented in pattern-comparison and prominence-

identification tasks yield low accuracy rates with adult native listeners of the same dialect. Such tasks discourage the use of lexical knowledge and emphasize an acoustic mode of perception, since again low-frequency words may be unfamiliar even to L1 speakers. Conversely, a perception task in which L1 Japanese listeners are explicitly required to discriminate or categorize nonword stimuli will likely evoke acoustic perception, to the benefit of performance.

Such accented nonword stimuli have been used in a perception task with L1 listeners. Sakamoto (2010) reported results of a categorization task in which Japanese listeners grouped 2-mora nonwords followed by a postposition (e.g., *mene mo*) into one of three accent categories. Participants performed at a mean accuracy of 91%, indicating that they were sensitive to phonetically-manipulated accent patterns even in the absence of meaningful lexical information. L1 Japanese listeners' performance, however, was no different from two comparison groups of L2 Japanese learners (high and low proficiency), which suggests that all participants performed this task through a reliance on basic auditory perception resources. Thus, in contrast with previous research using real-word stimuli, Sakamoto's study indicates that L1 speakers resort to an acoustic mode of listening when the spoken targets are clearly nonwords that lack representations in their mental lexicon.

On the other hand, Shibata & Hurtig (2008) conducted a task in which L1 Japanese listeners judged if aurally-presented words were spoken with the correct or incorrect accent pattern (2-alternative forced choice), which necessitated lexical knowledge to complete. They reported a mean accuracy of 97% on high-frequency 2- and

4-mora nouns followed by a postposition. Their aurally-presented stimuli, however, were accompanied by an onscreen presentation of the corresponding written word—a possible confound that may have contributed to the high accuracy. The researchers presumably did this to accommodate test items which featured HL/LH accentual minimal pairs such as [ame] (‘rain’ or ‘candy’) and [ishi] (‘will’ or ‘stone’). In other words, the presence of visual stimuli may have resulted in listeners using a different processing strategy, in which the spoken input is essentially redundant, than if the stimuli were only presented aurally. Regardless, this is one of the few studies which used an accent correctness judgment task and yielded near-ceiling performance among L1 listeners.

Looking at another high-accuracy result, Ueno et al. (2014), as discussed in the previous section, found in one of the behavioral portions of their study that L1 Japanese speakers could imitate spoken words, including their accent pattern, at near-perfect accuracy. Crucially, half of the words were spoken with an incorrect accent, which necessitated repetition based on the perceived pattern alone, while simultaneously suppressing word knowledge. Then, in the second phase of their task, listeners corrected these mistaken patterns by repeating them back with the accent pattern that they thought was correct, a task that did require word knowledge, which they did at a 90% or better accuracy for all targets. In sum, their study showed L1 Japanese speakers accurately perceived and produced lexical accents, irrespective of correctness, and were subsequently able to correct the errant patterns based on their long-term knowledge of the word. Difficulty in correcting accent patterns was only found on words with the least

frequent accent pattern (LHL) and low imageability (i.e., words not referring to easily-visualized objects).

From the above studies, it is clear that we must take into account the effects of task design and stimulus frequency on the reported results. For native listeners, tasks that tap basic-level acoustic perception will likely yield very different results from those which require lexical knowledge. In the next section I will further consider the effect of task design on perception accuracy, to set the stage for the discussion of L2 perception.

2.2.2 Task design issues in previous studies

A brief overview of perception task design and its effect on performance is due, considering the variation we have seen in native listeners' judgments of accents.

Language perception tasks can be generally grouped into two types, identification and discrimination (Strange & Shafer, 2008). Identification tasks can also be referred to as categorization tasks, as I will do in the present study, in that listeners must identify or label a stimulus as belonging to a discrete phonetic category. Applying this to Japanese accent, a frequently used task would require participants to identify the location of the accent within a word. Such tasks are often accomplished by having listeners categorize a single stimulus by its accent pattern, i.e., whether it is accented on the first or second mora, etc. We can also refer to such tasks in the Japanese context as prominence identification tasks in which participants identify accent location on a mark sheet with Japanese (or alphabetic) transcriptions of the spoken stimuli. The benefit of such a task is that the experimenter can see, through analysis of the error patterns, which stimulus types are easily confused and which are perceived with high accuracy. This gives one an idea

of where the perception difficulty lies—for example, we can address the question of whether phonetically similar accent patterns are easily mistaken for one another by comparing error patterns in a confusion matrix (So & Best, 2010). Although not as pronounced as is a same-different discrimination task, one criticism that can be levelled on this type of task in terms of how speech sounds are processed by a listener is that it tends to encourage an acoustic mode of listening (Strange & Shafer, 2008). For example, while L1 speakers are likely capable of identifying accent location based on their mental representation of a word's pronunciation even in the absence of spoken input, if a task is formulated in a way that participants are told to focus on where they hear the pitch fall, highly sensitive L2 listeners may be able to do so despite their lack of word knowledge. Hence, this task may encourage different strategies between L1 and L2 participants. Labelling conventions must also be taken into account when presenting identification tasks. With a feature like pitch accent, which lacks a standardized labeling system (Vance, 2008), task difficulty may also stem from listeners' unfamiliarity with the transcription system.⁸ In this sense, a degree of metalinguistic knowledge is called upon when learners are asked to apply a novel labelling system that they are likely unaccustomed to using. For instance, should we represent both the pitch rise and fall within a word, or just the location of the drop? Pre-task familiarization is critical when conducting identification tasks.

⁸ By comparison, consider pinyin tone markers ˉ ˊ ˋ ˋˊ and numbers (1, 2, 3, 4) in Mandarin Chinese as a relatively standardized tone marking system.

Discrimination tasks typically require listeners to compare a stimulus to two (or more) additional items and determine if the target matches either of the items (AX or ABX-discrimination). In an AX discrimination task, for instance, the “A” stimulus remains unchanged through a block of stimuli, with the second stimulus either matching (AA) or differing from (AB) the first item. Same-different discrimination tasks are thought to encourage acoustic processing, in that decisions can be made based on an immediate phonetic comparison of the two stimuli (Logan & Pruitt, 1995). That is, when listeners are mainly required to focus on acoustic differences in the stimuli, more basic cognitive abilities such as pitch sensitivity are called upon.

However, task demands can be amplified by increasing the number of stimuli (as in an ABX task, where X is the same as either A or B) or lengthening the inter-stimulus interval (ISI), thereby upping memory demands. In their most basic form though, discrimination tasks “reflect optimum auditory sensory abilities, and can even be performed by very low-proficiency or target-language naïve listeners” (Strange & Shafer, 2008; p. 161). Thus, discrimination tasks may primarily be a reflection of sensitivity to acoustic contrasts, and have little bearing on how listeners process words in context. We see this in the studies from a range of languages in which both native and non-native listeners perform similarly when making same/different discrimination decisions (e.g., Cutler & Chen, 1997; Shport, 2011).

A third type of task, similar to lexical decision tasks commonly used in psycholinguistic research is the correctness judgment task. Assuming a similar decision process as lexical decision, this task taps more directly into a listener’s lexical knowledge

in that accent judgments are made based on a comparison with, and accessing of, a word's representation in the mental lexicon. Presumably this is why standard-dialect Japanese listeners performed with high accuracy on correctness judgment tasks (e.g., Shibata & Hurtig, 2008; Ueno et al., 2014), but varied to a greater extent in their ability to identification or categorize words based on accent location (Hirano-Cook, 2011; Shport, 2008). Same-different phonetic discrimination tasks, on the other hand, usually yield high accuracy for both native and non-native listeners. In fact, discrimination tasks involving prosodic features in other languages, such as Mandarin and Cantonese, have also yielded similar results for both native and non-native listeners (e.g., Bent, Bradlow & Wright, 2006; Cutler & Chen, 1997; Hao, 2012). Cutler & Chen (1997) (Experiment 3) found that Cantonese-naïve listeners, all native speakers of Dutch, made same-different judgments on tone minimal pairs with essentially the same pattern as native speakers. Both the native and non-native groups displayed greater perceptual difficulty in proportion to the acoustic dissimilarity of the tone pairs, indicating that basic perceptual ability, and not Cantonese word recognition, determined task performance.

To summarize, the three tasks discussed above can be ordered into an acoustic-to-lexical processing continuum, following Wong & Perrachione (2007), in which discrimination tasks are at the acoustic end of the spectrum, followed by the more “ecologically valid” identification tasks (Strange & Shafer, 2008), and finally the correctness judgments task. It is important to add, however, that the categories below are not entirely mutually exclusive. That is, it is not the case that a listeners' acoustic perception ability is completely uninvolved in, say, a correctness judgment task. Based on

the tasks used in previous perception studies on both L1 and L2 listeners, we can posit the following continuum for listening tasks as a diagram in Figure 2.

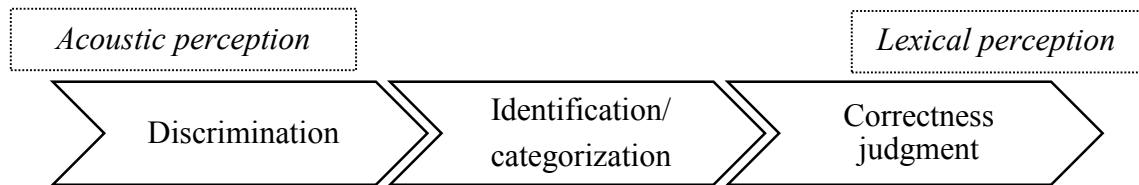


Figure 2. Phonetic to lexical continuum in speech prosody perception tasks.

A final issue which needs to be addressed when using real-word stimuli is lexical frequency. Especially relevant in correctness judgment tasks, very low frequency words may be judged more slowly and with greater variation by native speakers. In short, the accent pattern of words that are infrequently encountered might not be as strongly represented in the L1 lexicon. Word frequency can be checked in written and spoken language corpora, such as the commonly-used NTT database, which is a written corpus (Amano & Kondo, 1999; 2000), or the spoken-language Corpus of Spontaneous Japanese (Maekawa, 2003). In Japanese, lexical accent patterns themselves do not all share equal distributional frequency in the lexicon, and as such the less frequently encountered or “atypical” patterns may present greater perceptual difficulty, or at least yield lower accuracy on identification tasks (Ueno et al., 2014). Tanaka & Kubozono (2012) reported that for 3-mora words, which can have at most four accent patterns, the relative distribution within the lexicon for each pattern is: HLL-L (Pattern 1: < 40%), LHL-L

(Pattern 2: < 10%), LHH-L (Pattern 3: 5%), and LHH-H (Unaccented: 50%). Although beyond the scope of the current research, it remains to be seen if high-frequency words with uncommon pitch patterns are perceived differently than the reverse case.

The purpose of the above discussion was to consider the effects of task type and word and accent pattern frequency on speech perception. It is apparent that when designing an accent perception study one needs to control for these factors. First, task type may influence listeners' mode of perception, whether acoustic or linguistic-biased. Second, low frequency words may be particularly difficult for even native listeners to identify accent location or perform correctness judgments. By controlling these variables, I can conjecture that L1 Japanese should be able to consistently perceive accents with high accuracy. Furthermore, task design becomes even more important with L2 learners, as we will see in the following section. Questions to consider as the discussion moves to L2 accent perception in the next section include: What does it mean to "know" a word's accent pattern? Does knowledge constitute being able to identify the location of the accent, or the ability to determine if a word is spoken with the correct accent? In this sense, issues in perception task design are also intertwined with questions relating to the organization of the mental lexicon in second language acquisition research.

2.3 Accent perception/acquisition by L2 learners

2.3.1 *Evidence from L2 learners*

In this section I will review studies on how L2 Japanese learners both immediately perceive in experimental tasks, and acquire over a duration of study, lexical accent. Production studies will also be included in the discussion, mainly where they

address issues not reported on in the perception literature. We can broadly divide previous studies into those that have addressed L2 Japanese learners' perception and production ability from an SLA or pedagogical perspective, by addressing how accent is taught and acquired, or through the lens of phonetics/phonology, by identifying specific features of lexical accent which cause perception difficulty. SLA and language pedagogy researchers have noted three central issues regarding accent acquisition in L2 Japanese: 1) Accent perception does not develop in parallel with proficiency level, as does the perception/production of other phonological features (Lee et al., 2006; Shibata & Hurtig, 2008), 2) Pitch accent is particularly difficult to acquire from input alone, in the absence of form-focused instruction (Shport, 2008), and 3) Accent is often neglected in the JFL/JSL classroom (Eda, 2004; Short et al., 2013; Toda, 2001). Findings from linguistics, mainly in the areas of phonetics and phonology, can also be distilled into three points: 1) Perception difficulty arises from discrepancies between the L1/L2 prosodic systems (Shport, 2011; Ueyama, 2000), 2) Pitch accent lacks multiple phonetic cues, such as length or intensity changes, as indicators of accented syllables (Hirata, 1999), and 3) Accent carries a low functional load in semantic disambiguation (Kitahara, 2001; Tamaoka et al., 2014). The following discussion will incorporate these points as they relate to previous L2 findings, and set the stage for the examination of how learner-internal and experience-based factors may account for some of the reported variation in L2 accent perception and acquisition.

Research from the domain of second language acquisition has primarily looked at how L2 learners acquire lexical accent across proficiency levels or over a given period of

study or training. First, let's briefly compare the acquisition of lexical accent with another difficult, and equally well-studied, phonological property of Japanese, phonemic vowel and consonant length. Accuracy in the perception of vowel/consonant length contrasts appears to increase along with general proficiency, with several studies having reported a clear pattern of development (e.g., Hardison & Saigo-Motohashi, 2010; Shibata & Hurtig, 2008; Tsukada, 2012). Yet, current evidence does not indicate a similar developmental trend with lexical accent perception by L2 learners (Hirata, 1999). For example, in a cross-sectional study of three proficiency groups (Novice, Intermediate, and Advanced) of L1 English-speaking Japanese learners at an American university (below, JFL learners), Shibata & Hurtig (2008) reported that learners' identification of aurally-presented words containing phonemic vowel and consonant length contrasts (e.g., *oya* 'parent' vs. *ooya* 'landlord'; *kata* 'shoulder' vs. *katta* 'bought') clearly improved by level (Novice = 45%, Intermediate = 62%, Advanced = 75%), even though these length contrasts are not present in listeners' L1. By way of contrast, the same study found that none of the three levels differed significantly in either perceiving or producing the pitch accents in frequent 2- and 4-mora words, with each group scoring at around chance level (Perception task: Novice = 47%, Intermediate = 49%, Advanced = 56%; Production task: Novice = 37%, Intermediate = 42%, Advanced = 53%). However, the three learner groups did display clear differences in their standardized proficiency, as measured by the listening portion of the Level 3 Japanese Language Proficiency Test (Novice = 31%, Intermediate = 85%, Advanced = 93%).

Shport (2008) noted a similar pattern of perception ability with intermediate learners, who only attained a mean accuracy of 46% on an AXY discrimination task (described above in Section 2). This same group of learners' accent production accuracy was only slightly higher at 56%. Hirano-Cook (2011) compared JFL learners from five proficiency levels, plus a group of L1 Japanese speakers, on their ability to identify accent location in 4-mora Japanese words. While all L2 groups, except the highest proficiency level (Level 5), who correctly identified accent location in 58% of the targets, were significantly lower than L1 listeners (70%), only one of the JFL groups (Level 3) performed significantly better than the lowest proficiency group (51% vs. 38%; $p < .05$). This suggests that the numerical variation did not, in fact, show a clear trend of development across levels. Note that the small number of Level 5 participants ($n = 5$) likely prevented this group from attaining statistical significance with the lower levels. Thus, the apparently gradual acquisition across proficiency reported by Hirano-Cook was not counter-evidence to other studies' results (e.g., Shibata & Hurtig, 2008).

Sakamoto (2010) proposed that Japanese learners follow an experience-dependent continuum in accent acquisition, beginning with the ability to phonetically perceive pitch variations, followed by the assignment of these variations to linguistic categories (i.e., accent patterns), which can finally be applied to individual lexical items. She first found that the ability to phonetically distinguish pitch patterns was high for both inexperienced and experienced English-speaking JFL learners, neither of which differed from L1 Japanese. Yet, on a categorization task, which required learners to assign phonetically manipulated accented nonwords to existent pitch categories, Sakamoto reported a

percentage difference between the inexperienced (74%) and experienced groups (80%). Japanese proficiency was not the sole predictor of categorization performance, however, since she also found that several beginning learners displayed categorical perception similar to native speakers. Unfortunately, this study stopped short of addressing the final, crucial transition from categorical perception of accent in nonwords to lexical perception, as no real words were included in the experiment. Although Sakamoto repeatedly suggests that the ability to perceive lexical accent is a direct product of L2 experience, her results, and study design, do not bear out this assumption.

The above studies all focused on native speakers of English. However, the difficulty in Japanese accent acquisition does not appear to be limited to speakers of a non-tonal language. Lee, Murashima & Shirai (2006) tested the prosodic production ability of three Cantonese speakers over a two-year period, finding that even though these learners spent a year in Japan and attained a high level of Japanese proficiency on a standardized test, they failed to show improvement in their accent production accuracy, as judged by two native listeners, on simple words over the initial testing session (T1 = 70.8% vs. T2 = 64.4% vs. T3 = 71.8%). Perception was not tested in their study, so it remains to be seen whether L1 Chinese speakers' difficulty was merely at the level of articulation and not in their listening ability.

Studies considering how the L1 phonemic status of lexical prosody influences L2 perception have shown that phonetic discrimination is relatively unaffected by the native language, but that linguistic categorization is influenced in various ways (e.g., Shport 2011; Ueyama, 2000). Shport (2011) found that Japanese-naïve listeners could

discriminate same/different nonword pairs by accent well above what would be expected from random guessing (L1 English: 72% vs. L1 Japanese: 85%), although the Japanese listeners were more sensitive to pitch fall location. Yet, the same nonnative participants could not use F0 fall, the critical cue to accent location, to categorize words by accent pattern. Training did improve learners' ability to attend to F0 fall, but no gain was found in their identification of the unaccented pattern (LHH). In this sense, Japanese learners must attune to two cues when making judgments on accents—whether a pitch fall exists or not; and if a fall is present, on which mora this is found. Ueyama (2000), looking at prosodic transfer between English and Japanese (and vice versa), found that L1 English JFL learners showed positive transfer when producing F0 increases on accented words, which she claimed is due to the use of pitch in English stress accents. That is, participants were able to increase their pitch on accented mora when producing Japanese. However, both beginner and advanced learners had difficulty with suppressing the duration of accented mora in a word, which should be isochronic (i.e., equal in duration) with the unaccented mora, leading her to state that there is “no systematic correlation between [the correct application of] vowel duration, a feature they need to suppress, and L2 proficiency” (p. 40). In other words, JFL learners were capable of modulating pitch in a native-like way, but tended to mistakenly apply English duration cues on accented mora, leading to non-standard productions.

2.3.2 Individual variation in L2 accent perception

Despite the apparent difficulty in perceiving and producing L2 prosodic contrasts, it seems that some learners are more capable than others at acquiring this lexical feature.

Shport (2011) noted a large degree of individual variation among L1 English speakers, who were all absolute beginners, in her study on Japanese pitch accent perception. Individual scores ranged from 43% to a high of 88% for identifying Pattern 1 (HLL) accents. In an earlier study by Nishinuma, Arai & Ayusawa (1996), native English speakers with two years of Japanese study ($N = 54$) displayed wide variation on pitch accent perception, with the lower third of learners scoring an average of 42% correct identification, while the top third averaged 73%. In a review of accent perception studies on L2 Japanese learners from a variety of L1s, Ayusawa (2003) reported that the ability to perceive accent location ranged from random-chance level to a high of 87%. The abovementioned study by Hirano-Cook (2011) also found that on an accent prominence identification task the top group ($n = 13$) scored an average of 79%, while the bottom group ($n = 25$) attained only a markedly low 36% accuracy. However, her analysis combined learners of different Japanese levels, so it is unclear whether this variation was a function of proficiency, or if other individual factors were involved. Despite the rather arbitrary performance division criteria used in Hirano-Cook's study, her results still showed large variation that may not be attributable solely to Japanese proficiency. Studies on lexical prosody in languages like Chinese have also noted individual variation in tone perception, despite learners being on equal footing in terms of their target language experience. In a Mandarin tone identification task conducted by Bent et al. (2006), L1 English participants' accuracy scores ranged from 26% to a high of 88%.

If we take a narrow account of prosodic transfer, then for L1 English speakers, phonemic vowel/consonant length, as well as lexical pitch accent in Japanese, both of

which are absent from English, should present a comparable degree of acquisition difficulty (Ueyama, 2000). In cross-linguistic speech perception studies on several languages, the presence or absence of word-level prosodic features in a listener's native language has been shown to influence—both positively and negatively—one's success in acquiring lexical accent (e.g., Best & Tyler, 2007; So & Best, 2010). The reported discrepancy in Japanese learners' perception accuracy for length contrasts versus lexical accent is likely related to some extent to the larger functional load carried by length contrasts, as opposed to that of accent. Yet, as demonstrated by the individual perception and production variation discussed above, we may be overlooking other important factors contributing to accent acquisition if we limit our analysis to L1 prosodic transfer and the role of accent in semantic disambiguation as the lone sources of acquisition difficulty. Although researchers (e.g., Bohn, 1995; Shport, 2011; Strange & Shafer, 2008) have suggested that factors besides L1 influence, such as attention, memory and learning context, are likely sources of individual variation in prosodic acquisition, these have not been explored in detail.

Anyone who has taught or observed a foreign language classroom is aware of individual differences in performance, but little research has focused on the sources of variation in Japanese learners. Taylor (2011) recently examined the production ability of Japanese pitch accent in two learner levels, described as less versus more learning experience, and again found a high degree of individual variation within both levels, leading her to the conclusion that the acquisition of pitch accent by Japanese learners is a random process. Random here most likely entailing the lack of development across

proficiency levels, high degree of individual variation, and seeming resistance to instruction or learning context. The current study aims to identify and characterize the learner variables, both cognitive mechanisms and experience-based factors, which may support the perception of lexical pitch accent.

2.4 Learner-internal and experience-based factors in accent acquisition

An area left unaddressed in previous L2 prosodic acquisition studies, but which may account for a significant proportion of individual variation, is the effect that both learner-internal capacities and experience-based factors have on perception ability, and ultimately acquisition outcomes. Individual differences which listeners bring to the task of language perception, often referred to as “subject variables,” may support speech perception (Bohn, 1995; Strange, 2011). Strange & Shafer (2008) defined speech perception as “an internal mental process by which the perceiver recognizes incoming stimulus events as instances of mental categories” (p. 159). Hence, we can assume that cognitive abilities such as working memory and acoustic sensitivity are involved in the perception of spoken input, while accessing these mental categories is likely related to a listener’s stored knowledge of the target language. I will use this internal/external dichotomy to classify the learner variables introduced in the following subsections. Because pitch accent is a property of Japanese words, cognitive abilities which support word form acquisition may also operate during the acquisition of lexical accent (Belin, Fecteau & Bedard, 2004; Wong & Perrachione, 2007). In the present study, I will investigate the relationship between four learner variables, which I divided into the basic cognitive abilities phonological short-term memory (or phonological memory, PM) and

acoustic pitch sensitivity, and higher-level, or experience-based factors—L2 lexical knowledge and L1 tone experience. Further, as others have pointed out, different capacities may operate at different proficiency levels (e.g., Hummel, 2009). Learning context will also be considered as an influence on the longitudinal development of perception ability in beginning learners. Each of the variables are discussed at length in separate sections below.

2.4.1 Phonological short-term memory

The first factor of interest is the component of the working memory (WM) system responsible for maintaining and manipulating sound input, the phonological loop. A considerable body of evidence has emerged over the past three decades on the role of WM in the acquisition of language in both L1 and L2. The phonological loop was proposed by Baddeley (1986; elaborated in Baddeley, Gathercole & Papagno, 1996) as the component of working memory that is responsible for the short term storage and maintenance of verbal input in language processing. The more demanding the input, the more that the phonological loop is called upon to support retention of verbal information (Ellis, 1996; Just & Carpenter, 1992). Originally applied to children's L1 acquisition, the model has been extended to account for the learning of an L2 in adulthood (Gathercole, 2009; Hummel, 2009; Kaushanskaya, 2012). To first clarify terminology, WM is generally considered to be the cognitive resource involved in the brief storage and manipulation of information and is typically measured by reading or listening span tasks. On the other hand, phonological short-term memory, below shortened to phonological memory (PM), refers to short-term sound storage capacity and is indexed by nonword

repetition or recognition tasks (Szmalec et al., 2013). Martin & Ellis (2012) state that “most researchers attribute the connection between PM and language learning to the importance of the phonological loop for forming stable, long-term mental representations of novel phonological material” (p. 383). In the present study I will adhere to this characterization of phonological memory in relation to the processing and acquisition of lexical accent.

Previous research (e.g. Cheung, 1996; Hummel, 2009; O’Brien, Segalowitz, Collentine & Freed, 2006; Speciale, Ellis & Bywater, 2004) on verbal WM/PM provides us clues to a relationship between vocabulary acquisition and memory resources in adult L2 learners. Speciale et al. (2004) tested the notion that individual differences in phonological store capacity may influence the acquisition of word form in the L2. They found a positive correlation ($r = .35$) between PM capacity and L2 Spanish receptive vocabulary ability over a 10-week period in beginning learners, but as learning progressed further, long-term knowledge of L2 sound regularities became the main factor facilitating lexical acquisition. Martin & Ellis (2012) reported that PM, in the form of nonword repetition ability, predicted 14% of the variance among learners’ ability to acquire the vocabulary of an artificial language, which was based on natural language properties. WM was found to separately contribute to vocabulary acquisition, explaining an additional 10% of performance variation. They concluded that PM and WM independently contribute to the acquisition of L2 word form. Additionally, O’Brien et al. (2006) investigated the relationship between PM and L2 proficiency in learners of Spanish, finding that PM showed a significant positive correlation with gains in

productive vocabulary in lower level learners, who are still compiling long-term knowledge of L2 phonology.

It has been widely reported in the memory literature that PM is most influential in the early stages of language acquisition, both in L1 and L2 (Baddeley et al., 1998; Kormos & Safar, 2008; Martin & Ellis, 2012). This is likely due to the intertwining of the capacity to retain language-like nonwords with the ability to process the phonological regularities of the target language. In other words, there is a reciprocal relationship between one's knowledge of a language's sound structure and the ability to retain language-like sound sequences (i.e., nonwords) (Martin & Ellis, 2012). As vocabulary knowledge increases with language experience, the role of PM as a predictor for language learning gradually diminishes, suggesting that the function of PM is critically linked to the compilation of knowledge of a language's phonological regularities. Yet, a study by Hummel (2009) showed that even in advanced learners of English, PM continued to be a significant predictor of L2 proficiency. The relationship of long-term lexical knowledge with perception ability will be discussed further in a later section.

Of interest to the present study is the fact that none of the languages examined to date in the L2 memory literature featured pitch as a lexical property (e.g., stress-accented languages such as English and Spanish). It remains to be seen whether the prosodic properties of word form acquisition are mediated to any extent by a learner's phonological store capacity. "Memory" of a word's accent pattern may play a role in the immediate performance on accent perception tasks, given that the task places sufficient demands on PM capacity. This ability to store a word's prosodic form may be a predictor

of one's ability to acquire accent as a lexical property. I assume that PM supports not only the acquisition of a word's segmental form, but also its prosodic form, as exemplified by the lexical accent patterns found in Japanese. This prediction can be surmised in the following way. If the phonological store is involved in L2 learners' ability to maintain speech input in the phonological loop, which research (e.g., Baddeley et al., 1998; Engel de Abreu & Gathercole, 2012; Martin & Ellis, 2012) has shown supports the consolidation of long-term L2 knowledge, then PM capacity may facilitate the acquisition of Japanese lexical accent. The basic assumption here is that learners with a greater PM capacity are able to process lexical form more efficiently, including pitch accent, with the predicted outcome being better performance not only on controlled perception tasks in the case of advanced learners, but also in perception gains in lower-level learners still in the process of forming long-term lexical representations.

2.4.2 Acoustic pitch sensitivity

Speech perception can be generally divided into two modes—an acoustic (phonetic) and a phonological (linguistic) mode of perception. The acoustic mode has been characterized as being cognitively demanding, slower and representative of how language-naïve, or even beginning L2 learners, perceive foreign speech sounds (Strange, 2011). The phonological mode, by contrast, typifies native language perception in that it is automatic and relatively cognitively undemanding (Strange & Shafer, 2008). In general, theories of speech perception posit a dissociation of basic auditory perceptual ability and the perception of language. Two findings in support of this assumption are 1) “Tone-deaf” listeners acquire the ability to produce L1 sounds normally, both segmental

and suprasegmental (Liu et al., 2010; Nan, Sun & Peretz, 2010) and 2) Language-naïve listeners can typically discriminate foreign language phonetic contrasts with accuracy similar to native listeners (Bent et al., 2006). This is in stark contrast, as we have seen earlier, to performance on tasks requiring lexical knowledge or linguistic categorization, which elicit large differences between native and non-native listeners.

Yet, in contrast to the previously-noted compartmentalization of perceptual resources, some recent research has made a case for the involvement of basic, nonlinguistic auditory capacities in the perception and acquisition of non-native prosodic features (Perrachione, Lee, Ha & Wong, 2011; Wayland, Herrera & Kaan, 2010; Wong & Perrachione, 2007). In other words, while the human auditory perceptual system may be a general-purpose processing mechanism, used for perceiving sounds like musical pitch, it may also support the acquisition of linguistically relevant pitch categories. If this is the case, we would expect to find some overlap in the ability to perceive nonlinguistic sound contrasts as well as language-relevant categories.

Before discussing L2 perception, let's look at evidence for a link between acoustic and linguistic perception in L1 listeners, among whom we would not expect a connection to exist. Within the domain of cognitive science, Liu, Patel, Fourcin & Stewart (2010) provided evidence for shared structures for music and language perception in the brain. Native English speakers in their study identified as “tone-deaf” (*congenital amusica*) lagged behind matched normal participants in their ability to correctly identify intonation patterns in their native language, despite the fact that all participants reported normal communicative ability in English. Nan et al. (2010) added

support to these findings by demonstrating that even native speakers of a tone language (Mandarin Chinese) who reported tone-deafness also displayed impaired perception of L1 tone patterns, again despite having normal production ability. Within a normal-hearing population, Chan, Ho & Cheung (1998) found that pitch sensitivity gained through musical training improved adult Chinese speakers' verbal memory for native vocabulary. In their study, participants with music training scored 16% higher than non-musician controls on a lexical recall (i.e., verbal memory) task featuring orally-presented words, suggesting shared memory resources for nonlinguistic pitch and tone-carrying words in a tonal language.

Considering the varied outcomes in nonnative perception ability discussed earlier, we would suspect the connection between acoustic perception and L2 speech perception to be much stronger than for L1 speakers. Wong & Perrachione (2007) tested the assumption that L2 learners acquire the prosody of a foreign language through a “bottom-up” process, which begins with acoustic analyses of the target-language relevant phonetic features. They found that English-speaking listeners who were initially more sensitive to non-linguistic pitch variations were more successful at acquiring Mandarin Chinese tones. In fact, initial pitch sensitivity predicted approximately 50% ($R^2 = .50$) of the variance in level of attainment after a series of training sessions. Shport (2011) reported that the number of instruments played ($r = .47$, $p < .05$) and years of music experience ($r = .45$, $p < .05$) correlated with post-training test scores on a Japanese pitch accent identification task. Namely, listeners with music training tended to benefit more from a series of perceptual training sessions. Acquisition of absolute pitch (AP), in the musical

sense, by tone language speakers likely involves the same processes as acquisition of second language tone. Deutsch et al. (2009) demonstrated a connection between tone language perception and AP in music conservatory students, albeit in the opposite direction as the above study. In their study, listeners considered “very fluent” in a tone language displayed a higher preponderance of AP ability as compared to matched participants who were exposed to a tone language at birth (i.e., both parents were L1 tone language speakers), but who no longer spoke the language fluently because of their long residence in the US. They concluded that since L1 speakers of tone languages acquire tone category labels based on arbitrary pitch variations linked to words (i.e., the four Mandarin tones), they are also at an advantage to acquire AP, which involves applying similarly arbitrary labels to musical notes, such as ‘A’ to notes with a pitch of 440 Hz. Thus, the authors argue that AP is acquired in a similar manner as a child would acquire an L2 tone language. Extending this line of reasoning for music and tone/pitch acquisition to L2 learners, I can predict that adults with greater perceptual ability for non-speech tones might also possess an increased sensitivity in acquiring L2 pitch accent.

Nevertheless, Wayland et al. (2010) reported that, although the pitch-sensitive learners (i.e., trained musicians) in their study were initially better at identifying Mandarin tones, non-musicians were generally able to catch up to them after a series of training tasks. Their results suggest that the early advantage for pitch-sensitive learners may not be permanent, and that adult learners’ perceptual systems are still sufficiently pliable to acquire difficult L2 prosodic contrasts.

The aforementioned findings suggest a connection between general auditory sensitivity and perception of the prosodic features of language. As we saw early in the above discussion, auditory sensitivity and facility with speech sounds are typically considered separate abilities in the human perceptual system. It is, however, quite possible that L2 learners rely on these basic abilities to support perception of prosodic features such as pitch and tone, as the reviewed evidence has shown. Bohn (1995) described phonetic features as differing in “psychoacoustic salience” in that temporally cued contrasts, such as segmental length or voicing contrasts, are more acoustically prominent than spectrally-cued contrasts, like tone and pitch. Applied to Japanese, this would indicate that phonemic vowel length is more easily perceived than are the less-salient lexical accents, which are realized solely through F0 variations. Considering this difference in perceptual salience, the acquisition of lexical accent may be particularly reliant on basic perceptual capacities. Thus, the present study investigated the effect of general acoustic perception acuity on L1/L2 pitch accent perception.

2.4.3 L2 lexical knowledge

In contrast with PM and acoustic perception discussed above, knowledge of the phonological regularities of the L2, as represented in the long-term lexical store, may also contribute to lexical accent perception ability. Vocabulary knowledge, or the accumulated representations of word form in the mental lexicon, has been shown to underpin L2 grammar learning (e.g., Service & Kohonen, 1995; Williams & Lovatt, 2003), facilitate automaticity of lexical decisions (Segalowitz, Segalowitz & Wood, 1998) and provide a basis on which further L2 lexical learning occurs (Martin & Ellis, 2012). In other words,

it is difficult to overstate the contribution of word form knowledge to L2 processing and acquisition. Yet, the learning of phonological form, particularly in the early stages of L2 acquisition, is a frequent source of difficulty for learners, and more so for features that are not present in the L1. At this stage of acquisition, word meaning often takes precedence over form, even in the presence of form-focused instruction (Doughty & Williams, 1998). That is, semantic/conceptual information is readily available to L2 learners as part of their world knowledge via their L1, whereas lexical form is often highly language-specific. As we have seen, the Japanese lexical accent system is a case of a formal property that learners must attend to in spoken input, but is also one which has been shown to be an ongoing source of perception difficulty. If L2 Japanese learners set as their goal the attainment of language ability which enables them to interact with native speakers in a manner that makes their counterparts comfortable, in terms of both perception and production, then formal properties of language must be a central focus of vocabulary acquisition.

If we consider the two abovementioned cognitive capacities as supporting early-stage lexical form acquisition, then we can assume that once a sufficient store has been established, this accrued long-term knowledge of phonological regularities becomes crucially involved in further word learning, as well as perception task performance. Such a relationship would entail that learners with a larger L2 lexical store are not only more accurate at making decisions on lexical form, such as judging lexical accent accuracy, but also more automatic in their judgments. This is because a measure of L2 lexical knowledge would likely serve as an estimate of how strongly knowledge of lexical form

has been established in a learner's L2 mental lexicon. Previous research has demonstrated that as L2 learners accumulate knowledge of the target language, lexical processing becomes not only more accurate, but also more automatic (e.g., Segalowitz et al., 1998). Speciale et al. (2004) conducted two experiments which considered broadly the contributions of PM and long-term lexical knowledge to second language acquisition. In their second experiment on novice learners of Spanish, they demonstrated the importance of long-term phonological sequence knowledge on learners' ability to make judgments on newly encountered real words as well as Spanish-like nonwords. Long-term lexical knowledge also became interlinked with the short-term store capacity in the L2, in such a way that those who performed better on measures of L2 receptive vocabulary displayed a larger L2 WM capacity. Gathercole, Service, Hitch, Adams & Martin (1999) noted that 40% of performance variation among L1 learners in a word/nonword repetition task was a product of long-term lexical knowledge. If we take their word/nonword repetition tasks to be roughly equivalent to the processing and learning of new L2 vocabulary, then this finding suggests that those with a larger long-term store are more efficient at vocabulary acquisition. In fact, Martin & Ellis (2012) found that when learning an artificial language, composed of what are essentially nonwords, participants' initial vocabulary scores from the first training session strongly correlated with the ability to comprehend new vocabulary items in a generalization test held several days after the training.

In terms of Japanese phonological structure, L2 lexical knowledge entails an understanding of the phonological regularities of the language, including mora structure, phonemic length contrasts, and most pertinent to this study, accent pattern. For example,

knowledge of the word *kuruma* involves a representation of its moraic structure (*ku-ru-ma*; three mora), meaning ('car') and accent pattern (LHH). Let's suppose a task in which learners heard the word *kuruma* and were asked to decide if the accent pattern was correct or incorrect. This decision would require a comparison between the spoken input and the representation of the word in the learner's long-term lexical store. Besides the obvious prerequisite knowledge of the word itself, success on this form-based judgment would entail that the word's pitch accent has been correctly learned and accessed from long-term memory. If we take a correctness judgment task to be an indicator of whether a word's accent pattern has been acquired, then a measure of overall L2 lexical knowledge should predict the ability to judge individual items. In other words, the larger a learner's L2 lexicon, the stronger the representations of lexical accent categories in long-term memory, which may facilitate accent judgments. Relatedly, word familiarity, considered to be a subjective measurement of word knowledge, may also be tied to a listener's ability to perform lexical decisions, in such a way that highly familiar words are recognized more accurately and rapidly (Amano & Kondo, 1999). These familiarity ratings, which have been assessed through judgments by L1 Japanese speakers, could potentially relate to the accuracy of accent identification as well, necessitating their control in designing a perception task.

The above discussion has highlighted how L2 vocabulary knowledge may be connected to the perception of lexical pitch accent. Considering that identifying whether an accent pattern is correct or incorrect requires the listener to compare spoken input with a representation in long-term memory, the more instances of words in memory, the more

accurate a listener's judgment is likely to be. In the present study I am concerned specifically with the knowledge of L2 word form and its relation to accent perception, namely, the ability to perceive and make decisions on the correctness and location of lexical accents. To my knowledge no data exists which shows a relationship between a learners' L2 long-term lexical store and their ability to perform immediate form-based lexical decisions in a perception task. Having noted above the direct contributions of the long-term lexical store to L2 vocabulary acquisition, I assume that learners with greater lexical knowledge will draw on their L2 lexicon when processing speech, resulting in higher accuracy on accent perception tasks. In Japanese specifically, this implies that a greater lexical store means that there are more exemplars of pitch accent patterns in long-term memory, and thus this knowledge can be applied to the judging of spoken stimuli.

2.4.4 Influence from the L1 prosodic system

Experience with the prosodic features of one's native language may influence, often in profound ways, how one perceives and ultimately acquires L2 prosodic categories (So & Best, 2010; Strange, 2011; Ueyama, 2000). Long-term exposure from an early age to the sounds, both segmental and suprasegmental, of one's native language narrows the perceptual space for speech sounds (Werker & Curtin, 2005). This narrowing is beneficial for acquiring one's native language, but often detrimental to the later acquisition of non-native contrasts. Several models of speech perception have yielded hypotheses on how established L1 perceptual routines influence L2 learning. The automatic selective perception model (ASP), proposed by Strange (2011) is one current framework which tries to account for adult L2 learning by considering three factors: 1)

the acoustic salience of a given sound contrast, 2) listener's attention level, and 3) memory resources. While this model has not been applied to Japanese lexical prosody, some researchers have considered how L1 word-level prosody affects L2 Japanese perception. Ueyama (2000), for example, found that on a speech production task, L1 English learners of Japanese applied English stress characteristics, such as intensity and duration increases, when reading aloud in Japanese. In contrast, Japanese EFL learners in the same study were able to apply English-like stress patterns to the sample productions. She accounts for this as a case of L1 English speakers' inability to suppress native-language prosodic features (i.e., lengthening and intensification of stressed syllables). Japanese EFL learners, on the other hand, did not have to suppress their L1 cue to lexical accent (pitch boost), but instead had the easier task of applying novel (i.e., not in the L1) English stress features. Further, So & Best (2010) noted that L1 Japanese listeners outperformed English listeners on a Mandarin tone identification (labelling) task, despite none of the listeners having any knowledge of Mandarin. The authors argue that the phonemic use of pitch in Japanese to distinguish between minimal pairs enhanced listeners' sensitivity to these contrasts in comparison with L1 English speakers. How about in the case of two groups of L2 listeners with knowledge of Japanese, one from a tonal language, the other from a non-tonal L1? I will next consider Chinese and Korean learners of Japanese, whose languages feature very different prosodic and grammatical systems.

Chinese and Korean learners of Japanese have been noted to attain a similarly high level of Japanese proficiency in a short period of study (Tamaoka, 2014). Thus, they

quickly develop skills to handle a variety of linguistic tasks in Japanese, making L1 speakers of these languages ideal for cross-linguistic comparison in terms of the contrasting differences in the prosodic systems of each language. Mandarin Chinese is a tone-accented language, while the standard dialect of Korean (Seoul Korean) features a non-tonal, stress-like system of accentuation (Jun, 1998). Syntactically, Korean is close to Japanese in both its standard word order (SOV) and use of case-markers and particles, while Chinese is a structurally-distant SVO language, but shares a great deal of orthographic overlap. English, by comparison, is both prosodically and syntactically remote from Japanese, and is stress-accented and SVO in its order.

In both Chinese and Japanese, tone and pitch, respectively, have phonemic status, as illustrated by the presence of minimally contrastive accent pairs. It must be noted, however, that this semantic distinguishing function of tone minimal pairs is far more common in Chinese than in Japanese (Kitahara, 2001; So & Best, 2010; Tamaoka et al., 2014). To contrast specific pitch shapes between the two languages, the L to H pitch rise in Japanese is phonetically similar to Mandarin Tone 2 (rising tone), and the H to L fall—the locus of the accent in Japanese and the crucial phonetic feature in identifying accent—is similar to Mandarin Tone 4 (falling tone). The frequent use of tone at the syllable and word level in their L1 provides Mandarin speakers with rich experience using lexical prosody, which they bring to the task of acquiring Japanese lexical accent.

Standard (Seoul) Korean, on the other hand, is considered a non-tonal language, in that pitch (F0) is not used contrastively, nor is F0 variation alone used to mark

prominent syllables within a word as it is in Japanese (Jun, 1998).⁹ Therefore, for comparison in the current study, I assume that pitch is not represented in L1 Korean speakers' mental lexicon as a part of phonological form in the same way as pitch accent is in L1 Japanese speakers' mental lexicon. Namely, there is a difference in phonemic status of pitch between standard Japanese and Seoul Korean, and this difference in status may result in increased perceptual difficulty for Korean learners of Japanese, and hence more difficulty in perceiving accent correctness and location.

In the current study, rather than phonetically analyzing the similarities/differences of specific L1-L2 pitch accent shapes, I consider L1 experience with tone as a general contributor to the acquisition of Japanese lexical accent, albeit one that may aid perception in a pattern-specific manner. The presence of phonemic tone in Chinese may enhance learners' attention to lexical accent in Japanese, making pitch variations more salient for Chinese speakers (So & Best, 2010). In other words, speakers of a language in which a given feature, in this case lexical prosody, is used contrastively, are more likely to focus their attention on phonologically similar contrasts in the L2. Therefore, it is predicted that Chinese speakers will outperform Korean speakers if L1 experience with lexical tone facilitates L2 perception, and if this influence continues into the advanced proficiency level.

⁹ There is a pitch-accented dialect of Korean, Northern Kyungsang Korean, which uses lexical pitch similarly to standard Japanese (for discussion see Jun et al., 2006). However, none of the Korean participants tested in the current research spoke this dialect.

2.4.5 Learning context and instruction style

In the final section, I will consider learning context, also including a discussion of the effect of explicit, form-focused instruction on L2 lexical perception. First of all, learning context, broadly divided includes those studying in their home countries (JFL, at-home) and those learning Japanese in Japan (study-abroad, JSL). We can generally assume that learners in the JSL context have access to increased L2 speech input in comparison to JFL learners. Whether learners take advantage of this extra input, however, is another issue beyond the scope of the current research. In fact, intensive or immersion-like curricula in at-home contexts may provide similar opportunities for increased L2 input. Findings on the advantages of study-abroad versus at-home learning environments on the acquisition of vocabulary have been mixed (e.g., Collentine, 2004; Collentine & Freed, 2004).

Sunderman & Kroll (2009) attempted to account for this variation in learning outcomes by looking at learner-internal resources in their “resource threshold hypothesis.” In terms of L2 acquisition of word form in the target-language environment, Sunderman & Kroll (2009) reported little advantage for certain learners who spent a semester abroad studying Spanish. Namely, those without a certain level of WM capacity were unable to benefit from the increased aural input in terms of their L2 comprehension and production development. Although I have proposed above that WM/PM are likely involved in the acquisition of lexical accent from aural input, the view that those without a certain level of memory capacity cannot benefit from extra input regardless of learning environment seems rather pessimistic in outlook. I feel, instead, that instruction style—

what is done or not done in the L2 classroom—may trump learning context, and compensate for the over-reliance on WM. The low WM-capacity learners in Sunderman & Kroll's study may not have been able to benefit from increased L2 input, but the type of instruction these learners received was not clearly discussed. In other words, increased input in the *absence* of explicit, or form-focused, instruction or correction of accent patterns, in the case of the current research, does not likely benefit learners universally in terms of their ability to acquire lexical accent. In particular, considering the previously discussed difficulty L2 learners display in both producing and perceiving Japanese lexical accent, explicit feedback and correction may be essential to the acquisition of this feature. This being the case, I would expect learners in a JFL context who receive explicit feedback on accent to outperform higher-input JSL learners who lack such correction, and that the influence of instruction should be separate from the contribution made by PM capacity. To test this hypothesis, in the present study I examined two groups of Japanese learners, at-home and study-abroad, twice in a semester of language study.

2.5 Looking ahead

In the preceding literature review I examined research on Japanese lexical accent from a wide range of fields including psycholinguistics, second language acquisition and language pedagogy. Critical points from the previous research to bear in mind as we move to the experiments can be summarized as follows. 1) Native Japanese speakers can accurately perceive lexical accent, but this performance may greatly vary by task type. 2) Lexical accent is difficult to both perceive and produce for L2 Japanese learners, and it does not develop in parallel with overall language proficiency. 3) A considerable degree

of individual variation has been reported in accent perception studies. 4) Sources of this variation have been suggested, but not explored in depth. 5) Both learner-internal and experience-based factors identified in previous research, but not tested in Japanese, may account for a significant amount of this variation.

The current study attempts to resolve some of the previous questions on L2 accent perception and acquisition by considering the effects of perception task type and multiple learner variables, both internal and experience-based. I will briefly describe three perception experiments that focus on different listener populations. Chapter 3 examines the perception of accent by Tokyo-standard L1 Japanese speakers with the aim of clarifying previously-reported discrepancy by looking at the effects of task type and listener variables on perception accuracy. Further, this L1 listener group will establish a useful reference point for the subsequent studies with nonnative listeners. In Chapter 4, I focus on two groups of advanced-proficiency L2 learners, native Chinese and Korean speakers. These groups were closely matched in terms of their Japanese proficiency to make possible a clearer examination of the effects of learner-variables on accent perception performance. Lastly, Chapter 5 is a longitudinal study on beginning-level English-speaking Japanese learners in two contexts, an “at-home” (JFL) group in the US and a study-abroad (JSL) group in Japan. This experiment again considers the effects of learner variables, along with learning context and instruction style, but targets the acquisition *process* in a longitudinal design over a semester of Japanese study.

In sum, taking a wider view of the experiments that follow, these three studies enable me to see how individual differences affect accent perception cross-sectionally

over a native-to-beginner continuum, as well as longitudinally at the beginner level. This is important because the learner-variables discussed above, such as phonological memory, may function differently depending on target language experience (Hummel, 2009). Specifically, learners may rely on basic cognitive resources like PM and acoustic pitch sensitivity during the early stages of L2 acquisition (Wong & Perrachione, 2007), while long-term knowledge of Japanese, as evidenced by the size of the L2 lexicon, is likely important in higher-proficiency groups (Martin & Ellis, 2012). Further, Chapter 5 provides a finer-grained view of the same group of learners over a semester of study. Thus, on the one hand this dissertation examines how multiple language-related factors operate across levels of Japanese proficiency, including L1 listeners, and also within the same individuals over a duration of study.

Chapter 3: Experiment 1 – Accent Perception by Native Japanese Speakers

3.1 Introduction

This chapter presents the design, procedures and results of an experiment on native Japanese speakers' perception of lexical pitch accent. Two perception tasks were tested with L1 listeners, an accent pattern-correctness judgment and a categorization task. Prior to these tasks, phonological memory capacity and acoustic pitch sensitivity were measured as predictors of perception accuracy. I specifically consider how task type and basic cognitive abilities relate to the perception of lexical accent in real words, with the aim of clarifying some of the discrepancy seen in the perception findings reviewed in the previous chapter.

3.2 Predictions

Four predictions guided the design of Experiment 1. First, I predicted that adult L1 Japanese listeners, who speak Tokyo-standard Japanese, can reliably judge the correctness of spoken word accents. Second, categorizing words by accent pattern, which involves greater processing demands and metalinguistic knowledge, will yield more individual variation as we have seen in previous studies on L1 listeners. Third, because Japanese pitch accent fundamentally consists of lexically-linked contrasting high and low tones, listeners who possess greater acoustic pitch sensitivity will perform better on both perception tasks. Although this assumption runs somewhat counter to the prevailing

views in L1 speech perception, it is worth investigating since data on individual variation in adult L1 prosodic perception is sparse. Fourth, phonological memory span will be implicated in the memory-resource demanding accent categorization task, which requires listeners to briefly store spoken input in order to make a decision, but not in the speeded correctness judgment task.

3.3 Methods

3.3.1 *Participants*

Thirty native Japanese speakers (19 female and 11 male) with a mean age of 19.2 years ($SD = .78$) at a research university in Japan volunteered for this experiment.¹⁰ The participants indicated on a questionnaire that they were from the Tōkai region in central Japan, which is classified as a Tokyo-standard accent region (Shibatani, 1990). Participants reported normal hearing, and none had an extended length of musical training (≤ 3 years) or fluency in a tone language. Their main L2 exposure was to English, which all had studied mandatorily in junior high and high school (approx. 6 years), plus an additional one to two years at university. Three participants did report studying Chinese as an elective language at their university, but their length of study was no longer than 1.5 years.¹¹ Thus, musical training and tone language fluency can be ruled out as potential contributors to pitch perception ability. All participants were presented with a small payment at the conclusion of the test session.

¹⁰ All participants were undergraduates of various majors recruited from a mandatory English class.

¹¹ A typical elective foreign-language class at a Japanese university meets twice a week for approximately 90 minutes. Thus, we can assume that these learners' were not above the beginning proficiency level.

3.3.2 Materials

3.3.2.1 Phonological memory

The serial nonword recognition (SNWR) task was first used by O'Brien, Segalowitz, Freed & Collentine (2006) as a measure of phonological short-term memory in their investigation of the development of L2 Spanish production abilities, and it provides an alternative to the conventional nonword repetition tasks which require vocalization of nonword stimuli. Some have claimed that eliminating the speech production aspect may enable a more accurate measure of PM capacity in that it removes the demands of articulation and is less susceptible to long-term memory effects (e.g., O'Brien et al., 2007). In contrast with O'Brien et al. (2006)'s English-based nonwords, in the current study stimuli for the SNWR task were nonwords based on the Japanese mora system, of the structure /(C)VCV/. Martin & Ellis (2012) argued that PM tasks that use nonwords which are more word-like (i.e., those that adhere to the phonotactic structure) in the target language are better predictors of L2 vocabulary acquisition, which is the focus of Experiments 2 and 3. The same assumption was held when developing the current Japanese-segment based SNWR task in that stimuli used only legal mora and an existent accent pattern. These nonword stimuli were recorded by a male speaker of Tokyo-standard Japanese, a trained linguist, and were produced with an unaccented Low-High pitch pattern (see example (1d) in the previous chapter). I chose to record the nonword stimuli with an existent accent pattern, since my aim was to measure the relationship between PM capacity and the processing of lexically-accented words, on the assumption that pitch-carrying nonwords will better predict performance on real-word

perception tasks (Yuzawa & Saito, 2006). Stimuli were recorded in a sound-attenuated recording booth with a high-quality microphone. Prior to constructing the task, pitch patterns of the nonwords were checked in Praat to confirm their adherence to the LH pitch pattern. Audio files for the PM task were then compiled in Audacity sound editor as follows. For each trial, two lists consisting of the same number of nonwords were created with a 1,500 ms pause separating the lists (refer to (1) below). Nonword lists increased from 4 items in the practice phase to 5, 6, and 7 nonwords for the test phase. The inter-stimulus interval (ISI) between the nonwords in each list was regularized to 750 ms. There were 8 trials at each of the 3 list lengths (i.e., 5, 6 and 7 nonwords), yielding a total of 24 trials. Two types of trials were then created from these lists, “same” trials, meaning that the nonwords in both lists were presented in an identical order, and “different” trials in which the order of two of the nonwords was switched. However, the first and last items of the list were never switched. Half of the trials ($n = 4$) at each list length were same trials and the other half different trials. Participants’ task was to decide if both lists were in the same or different order, requiring them to keep track of the serial order of the nonwords in order to make this decision. Note that participants were not required to use pitch accent to make judgments on the nonword stimuli.

- (1) Sample nonword trials at the 5-nonword list length. All nonwords were produced with an L-H accent pattern. Switched items are indicated in italics.

(Same) gohe zuka imyo heji baro <1.5 s pause>

gohe zuka imyo heji baro

(Different) chida *ruge hami* jare kebu <1.5 s pause>

chida *hami ruge* jare kebu

3.3.2.2 Acoustic pitch sensitivity

An adaptive pitch test, which increased in difficulty based on participants' performance, was used as the measure of acoustic pitch sensitivity. This task is web-based and is similar in format to commonly used AX discrimination tasks (e.g., Wayland et al., 2010), in that the first pitch stimulus (A) remained constant, while the second stimulus (X) varied by a predetermined parameter (Mandel, 2009). Figure 3 is a sample screenshot of the adaptive pitch test.



Figure 3. Sample screen of the adaptive pitch test, which is a type of AX tone discrimination task (Mandel, 2009).

In this task specifically, the first tone of the two-item pair was a 500 Hz pure tone, with the second tone differing by set intervals of hertz (96, 48, 24, 12, 6, 3, 1.5 Hz, etc.). The between-stimulus pitch difference either increased or decreased based on participants' responses. For example, at the 12 Hz interval, the first tone was 500 Hz and the second tone 512 Hz. Each tone was 250 ms in length and tones were separated by a 500 ms pause. Loudness was adjusted to an adequate level for each participant prior to the task. 40 stimulus pairs were presented, and participants heard each pair only once. The resulting score in Hz represented the degree to which a listener could reliably distinguish the paired tones, with greater sensitivity indicated by a lower score.

3.3.2.3 *Lexical accent perception tasks*

Two lexical perception tasks using Japanese sentences were the dependent measures of perception and categorization ability. First, participants judged if a noun followed by a postposition was spoken with the correct or incorrect accent pattern (below, PitchID). Second, for the correctly-accented items only, they categorized the noun-plus-postposition into one of four pitch contours in a four-alternative forced choice (AFC) task (below, PitchCAT). A total of 32 test words were each embedded in sentence-initial position, and all sentences were of the structure [N + postposition + V] (e.g., 一人で行く, *hitori de iku*, 'I go alone'). Half of these stimuli ($n = 16$) were spoken with the correct accent and the other half with an incorrect accent. Stimuli were composed of a mixture of 3 and 4-mora nouns featuring one of four accent patterns (for 3-mora words: HLL-L, LHL-L, LHH-L and LHH-H; 4-mora words: HLLL-L, LHLL-L, LHHH-L,

LHHH-*H*).¹² Words were controlled by frequency across each of the four accent types using the NTT database of approximately 300 million words (Amano & Kondo, 1999; 2000). The mean normative frequency for all stimuli was 36.6 tokens per million ($SD = 52.6$) and by pattern as follows: Pattern 1 ($M = 35.7$, $SD = 54$), Pattern 2 ($M = 15.1$; $SD = 15.1$), Pattern 3 ($M = 47.7$, $SD = 67.5$), Pattern 4 ($M = 46.3$; $SD = 61.4$).¹³ Although the normative frequencies appeared to vary between patterns, this difference was not significant for any of the four patterns ($F(3, 28) = .633$, $p = .600$). Word familiarity, which has been shown to influence both recognition accuracy and speed, was also checked for each stimulus word in the above corpus. Amano & Kondo (1999) measured lexical familiarity by having native Japanese speakers rate words, both in their spoken and written forms, on a 7-point Likert scale from 1 (not familiar at all) to 7 (very familiar). The mean familiarity rating for all stimuli was 6.19 ($SD = .32$), indicating the words were subjectively very familiar to L1 speakers. Further, no difference was found when comparing pattern type by familiarity rating ($F(3,28) = .035$, ns). Lastly, all carrier sentences were controlled to a mean length of 6.7 mora ($SD = .52$). Refer to Appendix A for a complete list of test words and their carrier sentences.

Stimuli were sampled at 44,100 Hz by a male native speaker of Tokyo-standard Japanese with extensive training in accent production.¹⁴ No acoustic modification was

¹² Although there are actually five possible patterns for 4-mora words, I excluded the LHHL-L pattern, which is infrequent (< 10% of 4-mora words), in order to limit the categorization answer choices to four (Tanaka & Kubozono, 2012).

¹³ The high SD s in Patterns 1, 2 & 4 are due to the presence of one or two very high-frequency words in each of these categories. The majority of the other words clustered nearer to the mean frequency.

¹⁴ The same speaker produced the stimuli for the accent perception task in all three experiments. Note that male speakers generally have a smaller pitch range due to the physiological structure of vocal tract (van Heuven & Haan, 1999). Thus, it has been claimed anecdotally that accent location may be more

made to the test stimuli, however, pitch contours were checked in Praat to confirm that accent location matched the intended pattern (Boersma & Weenink, 2008). I found that the pitch contours of all words adhered to the standard accent patterns found in the NHK accent dictionary (NHK, 1998). Participants were instructed to focus on the sentence-initial [N + postposition] in making their decisions. After a correctly-accented word was presented and the correctness judgment made, four pitch graphs corresponding to the four possible pitch contours were immediately displayed on-screen. Figure 4 shows an example of a categorization graph for a 3-mora stimulus. Participants' task was to select the pitch graph which they felt most closely fit the accent pattern of the spoken target.

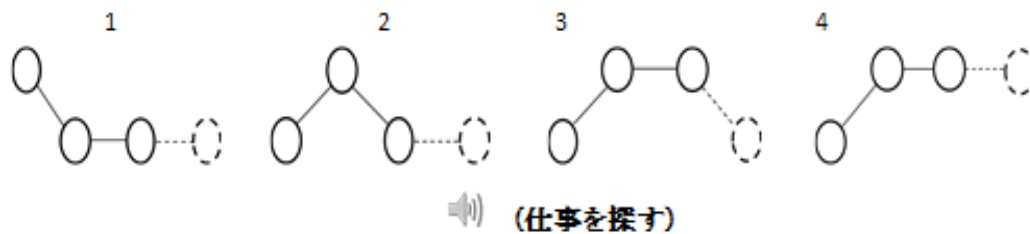


Figure 4. In the PitchCAT task, listeners categorized the noun-plus-postposition (*/sigoto o sagasu/* ‘(I) search for a job’) into one of four pitch contours by selecting a visual representation of the pitch pattern.

3.3.3 Procedure

Participants completed four tasks in the following order: (1) acoustic pitch sensitivity test, (2) SNWR task (phonological memory measure), (3) the combined lexical

acoustically perceptible in female speech, but this conjecture lacks experimental documentation for Japanese lexical accent.

accent correctness judgment (PitchID), and (4) categorization (PitchCAT) tasks. A short break was provided after the PM task to alleviate participant fatigue due to the multi-part experiment. Besides the web-based adaptive pitch test, the stimuli were presented in E-Prime 2.0 experimental software on a desktop computer. Responses were made via a 4-button response box, with the PM and PitchID tasks requiring a 2-button Yes/No answer, and the PitchCAT task using a 4-button response.¹⁵ Reaction time (RT) was measured on the PitchID task, and participants were instructed to respond as rapidly as possible after the spoken stimulus ended for this task only. Stimuli for tasks (2) – (4) were presented in random order across participants. Tasks were presented aurally with noise-cancelling headphones in a quiet room. All tasks were preceded by a short practice phase to familiarize participants with the procedures, with further explanation provided by the researcher if participants were unclear about the task procedures. Due to the complexity of the 4-choice PitchCAT task in particular, participants were familiarized with a printout of the pitch contour graph (Figure 4) similar to those in the main categorization task during the practice phase. All participants stated that they were clear with the categorization process prior to beginning the main experiment. Including the short break, participants took approximately 30 minutes to complete all experimental tasks.

¹⁵ The switch between a 2-button vs. a 4-button response may have made the categorization task more cognitively demanding. Note, however, that reaction time was not measured for the 4-button PitchCAT task, meaning that participants were not under a time constraint to switch hand positions between tasks.

3.4 Results

3.4.1 Predictor variables

Descriptive results for all tasks in their order of presentation are shown in Table 1. Data from all 30 participants were included in the analyses. Pitch sensitivity scores reflected the average threshold at which participants were able to distinguish two pure tones. Thus, the lower the numerical score, the more sensitive the listener was to pitch variations. Scores were \log_2 -converted for analysis in order to regularize the Hz intervals and obtain a more normal distribution. PM was scored by a weighted measurement, assigning points to correct responses based on the length of the sets, with the longer sets (i.e., those containing more nonwords) yielding a higher point total (see O'Brien et al., 2007 for scoring procedure). First, I will describe the results of the predictor variables.

Table 1. Descriptive statistics for all variables.

Variable	Max	Mean	<i>SD</i>	Range	
Pitch Sensitivity	--	2.49	1.44	5	-0.91
Phonological Mem.	144	84.07	17.34	50	117
PitchID	32	29.77	1.63	26	32
PitchCAT	16	9.20	3.39	3	15
PitchID RT (<i>ms</i>)	--	2,215	233	1,816	2,683

Notes: Max = maximum score possible; RTs include stimulus duration.

The average *log*-converted pitch sensitivity score was 2.49, which is equivalent to an approximately 5.5 Hz difference between the unchanging first tone (A) and the second tone (X). This result closely mirrors a previous study by Wayland et al. (2010), who found non-musician participants capable on average of reliably discriminating pure tones from frequencies starting at around 5 Hz. Pitch variations below 10 Hz are generally considered to require attentional focus to distinguish reliably, indicating that participants were focused on the task (Mandel, 2009). Further, I noted a wide degree of individual variation, with the most-sensitive listener displaying reliable discrimination ability at an interval as low as 0.5 Hz (i.e., 500 Hz vs. 500.5 Hz), and the least-sensitive learners requiring a 30 Hz-plus difference to distinguish the pairs.

The PM task also produced a wide distribution of scores that were generally skewed in a positive direction, attesting to the difficulty of this task. When converted to percentages, the average score was 58%, with the minimum score being 35% and the maximum reaching 81%. Despite the demanding nature of this PM measure, the distribution of scores is what we would expect from a task that measures individual variation in listeners' capacity to briefly retain spoken input in short-term phonological memory. It is interesting to note also that the mean score is slightly below previous results on L1 English-speakers' ability to retain English-like nonwords. O'Brien et al. (2007) reported a mean accuracy of 64% on their SNWR task, which was scored using an identical weighted point system. The lower accuracy in the present study may have resulted from the fact that the Japanese-based nonwords were all two mora in length, as opposed to the single-syllable nonwords (e.g., 'charg' or 'jeel') in O'Brien et al.'s task.

Thus, our participants had to retain lists of two-syllable nonwords, which likely placed greater demands on the phonological loop capacity.

3.4.2 Accent perception tasks

Moving to the real-word perception tasks (Table 1), I found that participants successfully judged 93% of the spoken noun-plus-postposition targets as having either the correct or incorrect accent pattern (PitchID). However, only 61% of the correctly-accented words were accurately categorized by visual pitch pattern on the 4-choice categorization task (PitchCAT). Because the PitchCAT task required a 1-of-4 response, it is worth noting this percentage is above that expected from chance guessing, which would be 25%. Reaction time (RT) in milliseconds from the onset of the audio until the button-press response in the PitchID task was also measured as a predictor for PitchCAT performance. RTs separated by more than 2.5 standard deviations from the mean were excluded from the analysis. The average RT appears quite long at 2.2 s, but note that this figure includes the duration of the audio stimulus. Despite the inclusion of the audio in the decision time, however, since the length of the audio stimuli were strictly controlled to 6.7 mora ($SD = .52$), I feel that RT is still a valid index of decision time on this task.

3.4.3 Correlations between predictors and perception tasks

I next examined the correlation coefficients for the two predictor variables and the PitchID and PitchCAT tasks, which are presented in Table 2. Significant negative correlations were found between pitch sensitivity in both the PitchID ($r = -0.42, p < .05$) and PitchCAT tasks ($r = -0.42, p < .05$), indicating that those with a lower acoustic sensitivity threshold were more accurate on the spoken word perception tasks. The

negative correlations are due to the fact that as Hz sensitivity scores dropped, indicating greater sensitivity, lexical accent perception ability tended to increase. However, no significant correlations were found for PM capacity on either task. Reaction time on PitchID was also negatively correlated ($r = -0.39, p < .05$) with participants' performance on the PitchCAT task, suggesting that listeners who made faster correctness judgments were in turn more accurate at categorizing accent patterns. Here we can take RT as a rough indicator of participants' certainty of their correctness judgments. Note also that PitchID accuracy was not correlated with PitchCAT score, a finding which will be discussed later.

Table 2. Correlation coefficients for all variables.

Variable	1	2	3	4	5
1. PM	--				
2. Pitch sensitivity	.28	--			
3. PitchID	.18	-.42*	--		
4. PitchCAT	-.11	-.42*	.01	--	
5. PitchID RT	.04	-.20	-.06	-.39*	--

*Note: * $p < .05$*

3.4.4 Regression analysis

A multiple regression analysis was next conducted with pitch sensitivity and PM entered together as predictors for PitchID performance. Regression coefficients enable us to see the degree to which each predictor (i.e., independent variables) accounts for the

variation observed in participants' responses to the PitchID task. As shown in Table 3, these two variables accounted for 27% ($R^2 = .27$) of the variance in participants' ability to judge accent correctness, with pitch sensitivity being identified as a highly significant predictor ($\beta = -0.51, p < .01$), and PM capacity approaching significance ($\beta = 0.32, p = .073$). Likewise, I created a separate regression model for PitchCAT with three variables of interest—pitch sensitivity, PM and PitchID RT. Table 3 shows that these three variables accounted for 25% ($R^2 = .25$) of the variance in accent categorization, with participants' pitch sensitivity scores again a significant predictor ($\beta = -0.35, p < .05$) and RT approaching significance ($\beta = -0.31, p = .067$). Phonological memory capacity, however, failed to attain significance as a predictor ($\beta = 0.05, p = .875$). Overall, the two regression models showed that participants' performance on the adaptive pitch test, which measured basic acoustic sensitivity, explained a significant amount of variation in the ability to perceive and categorize real-word accent patterns.

Table 3. Multiple regression results for the PitchID and PitchCAT tasks.

Task	Predictors	R	R^2	β	t
PitchID		.52	.27*		
	Pitch Sensitivity			-.51	-2.99**
	PM			.32	1.87
PitchCAT		.49	.25*		
	Pitch Sensitivity			-.35	-1.93*
	PitchID RT			-.31	-1.72
	PM			.05	.27

Note: ** $p < .01$, * $p < .05$.

3.4.5 Pattern analysis

I next wished to explore listeners' error patterns on the perception tasks, as specific accent types may have influenced overall performance (Hao, 2012; So & Best, 2010). Table 4 provides raw accuracy scores by accent pattern for both the PitchID and PitchCAT tasks. On the PitchID task, accuracy on Pattern 3 words (82.5%) was markedly lower than for the other three accent patterns. By comparison, only a single error was recorded among all 240 responses to Pattern 4 targets (LHH-H), suggesting that this pattern was the most easily-perceived of the four accent types. Categorization accuracy was lower overall on the 4-choice PitchCAT task, which is partially attributable to the task design. Yet, initial-mora accented Pattern 1 (HLL-L) words were categorized at a relatively-high 72.5% accuracy, but noticeably dropped on the remaining three patterns.

Table 4. Overall responses by accent pattern on both accent perception tasks.

Task	Target	Correct		Incorrect		Total
		Responses	Percent	Responses	Percent	
PitchID	Pattern 1	220	91.67%	20	8.33%	240
	Pattern 2	235	97.92%	5	2.08%	240
	Pattern 3	198	82.50%	42	17.50%	240
	Pattern 4	239	99.58%	1	.42%	240
PitchCAT	Pattern 1	87	72.50%	33	27.50%	120
	Pattern 2	64	53.33%	56	46.67%	120
	Pattern 3	59	49.17%	61	50.83%	120
	Pattern 4	67	55.83%	53	44.17%	120

Note: PitchID consisted of 120 correctly-accented and 120 incorrectly-accented items. PitchCAT featured only the correctly-accented items, yielding 120 total items.

Although the predicted correlation between PM capacity and the more demanding PitchCAT task was not found, I next wanted to determine whether PM related to accuracy on any specific accent pattern in this task. Previous research has suggested that words featuring an existing pitch pattern are more easily recalled by native Japanese listeners than those with no pitch variation (Yuzawa & Saito, 2006). Thus, I wanted to examine the possibility of an opposite connection, namely, if the ability to hold unaccented nonwords in the short-term phonological store is related to the ability to store, and categorize, real words of the same pattern. When analyzed by pattern, a moderate correlation ($r = .476, p < .01$) was found only for Pattern 4 (unaccented, or LHH-H) words, suggesting that memory span for unaccented nonwords related to categorization ability of real words of the same pitch pattern. To further explore performance on the unaccented stimuli in the PitchCAT and PitchID tasks, I then separated participants into three even groups ($n = 10$) according to the highest ($M = 102.9, SD = 9.43$), middle ($M = 83.8; SD = 5.26$) and lowest ($M = 65.5; SD = 8.68$) PM scores ($F(2, 27) = 54.61, p < .001$). A one-way ANOVA revealed that the PM groups differed significantly in their categorization of Pattern 4 words ($F(2,27) = 4.049, p < .05$), with the High-span group ($p < .05$) and Mid-span group ($p = .064$) outperforming the Low-span group. By contrast, no group difference was found for Pattern 4 words on the PitchID task ($F(2,27) = 1.00, ns$), likely due to a performance ceiling effect for this pattern (99.5% mean accuracy). Thus, the group analysis suggested that serial recognition of unaccented nonwords on the PM task was predictive of categorization performance for Pattern 4 words only, which were spoken with the same Low-High unaccented pitch pattern.

3.4.6 Summary of results

To summarize the results overall, L1 Japanese speakers were highly accurate at making correctness judgments on the accent patterns of sentence-embedded spoken words, while greater individual variation was observed in the four-choice categorization task. Of particular interest was the finding that listeners with greater sensitivity to nonlinguistic pitch contrasts on the adaptive pitch test were superior at both judging and categorizing lexical accents. This suggests that even among native listeners, one's acoustic sensitivity threshold is involved in the perception of lexical accent. Although PM was not predictive of overall accuracy, when stimuli were analyzed by pattern, listeners with a higher PM span for unaccented nonwords were better at categorizing real words with the same accent pattern (i.e., Pattern 4). It remains to be seen why participants' ability to retain nonwords was only involved in a pattern-specific manner, rather than in their general categorization performance. Next, I will consider these results in relation to my predictions as well as previous findings from perception studies.

3.5 Discussion

The present study examined four predictions regarding Japanese speakers' ability to perceive lexical pitch accent. I first considered the effect of perception task type on accuracy scores for lexically-accented words. Then, I measured acoustic pitch sensitivity and phonological memory capacity as two possible predictors of perception task accuracy.

First, native Japanese speakers were accurate ($M = 93\%$) at determining whether high frequency noun-plus-postposition targets were spoken with the correct or incorrect

pitch accent (PitchID). This finding corroborates with previous perception research indicating that pitch and tone are accessed along with word form by native listeners (e.g., Otake & Cutler, 1999; Shibata & Hurtig, 2008). Contrast this with the categorization task (PitchCAT), on which listeners were only able to accurately categorize 61% of the targets. Despite participants reporting familiarity with task instructions, accuracy rates were similar to those reported in previous studies using accent-pattern comparison or categorization tasks (e.g., Hirano-Cook, 2011; Shport, 2008). Although the task design makes direct comparison of the results problematic (1-of-2 correctness judgment vs. 1-of-4 categorization), we can interpret the performance discrepancy and resulting lack of correlation between the two tasks as suggestive of the differences in processing demands placed on L1 listeners by the two tasks, despite having well-developed native lexical stores. In other words, accent correctness judgments were intuitively possible for Japanese speakers, but categorization, in which participants' held a spoken stimulus in memory while selecting a visual representation of the accent pattern, likely placed a high processing load even on L1 listeners. Furthermore, categorization required a metalinguistic judgment in the form of assigning a label to a pitch contour, which may be difficult for L1 speakers since it requires them to visualize the word's accent pattern separate from its phonological form in long-term memory. By comparison, Ueno et al. (2014) recently found that L1 Japanese speakers could repeat aloud single, lexically-accented words at near-perfect accuracy on an immediate imitation task, echoing participants' high scores on the speeded PitchID task. Thus, the present findings suggest that the low categorization accuracy and wide variation (ranging from 23% to 100%)

were potentially by-products of the inability to assign an abstract visual label to a word's accent pattern, rather than a reflection of participants' reduced perceptual ability *per se*. Whether or not the use of a different labelling system would have affected performance is unclear, but it is worth noting that despite the use of various systems in previous studies, categorization accuracy has generally been highly variable. Alternative schemes will be discussed again in the final chapter. Accordingly, some previous studies' conclusions that pitch is an unimportant feature of spoken Japanese, based mainly on L1 listeners' poor performance on categorization tasks of similar design, are likely unwarranted (e.g., Hirano-Cook, 2011; Shport, 2008). Thus, we've seen that task type can greatly influence results and needs to be carefully considered when designing accent perception studies.

Second, the multiple regression analyses showed that acoustic pitch sensitivity significantly predicted accuracy for the PitchID and PitchCAT tasks. In other words, listeners who were highly sensitive to subtle, nonlinguistic pitch variations performed better on both lexical accent perception measures, regardless of the differences in task demands pointed out above. The fact that a non-linguistic auditory capacity predicted lexical perception ability is in line with speech perception models suggesting that speech is processed in a bottom-up fashion, whereby spoken words are processed by listeners starting with an acoustic analysis of the input, prior to the availability of higher-level lexical or contextual information (e.g., Norris, McQueen & Cutler, 2000; Warrier & Zatorre, 2004). This finding also provides evidence for shared perceptual resources for pitch and tone in both speech and nonspeech sounds. Although all participants reported normal hearing and no extensive music training, both of which may relate to pitch

perception (Deutsch et al., 2009), a significant degree of variation was found that could be explained by differences in general acoustic sensitivity. This finding runs contrary to models which assume that adult perception of native language phonological categories is the product of early-age L1 experience and unrelated to basic auditory capacities (e.g., Iverson & Kuhl, 1996; Strange & Shafer, 2008). Although I do not dispute that language experience in infancy is the primary source of adult perceptual acuity in the L1, in the case of lexical accent, the phonetic similarity (i.e., LH/HL pitch variations) between the tones in the adaptive pitch test and accent perception tasks in the present study may have contributed to the predictive power of this nonlinguistic capacity. Moreover, since pitch accent is considered to have a relatively low psychoacoustic salience (Bohn, 1995), in that its phonetic cue of F0 modulation is not as prominent as, for example, the length contrasts found in Japanese vowels and consonants, a nonlinguistic perceptual capacity may enhance perception of this relatively subtle feature.

Lastly, the role of PM in accent perception was negligible in terms of overall task performance. The fact that the serial recognition task required participants to focus on the linear order of the nonword stimuli, rather than explicitly on their sound properties, may account for this finding. In fact, measures of PM, like the serial nonword recognition task used in the current experiment, have been linked more to *development* of language ability in both L1 and L2 learner populations, rather than immediate task performance (O'Brien et al., 2007). Yet, I did find that PM capacity significantly predicted performance on the categorization of Pattern 4 (LHH-H pitch) words, which were identical in accent pattern with the stimuli used in the nonword recognition task. Note that the common acoustic

feature of the Pattern 4 words and the nonwords was the absence of a pitch fall.

Interestingly, this echoes the connection between pitch accent contributing to recall ability noted in previous research (Yuzawa & Saito, 2006), albeit in the direction that memory for nonwords is predictive of word perception only for targets of the same accent pattern, suggesting a possible pattern-specific effect for pitch memory. Future studies should employ PM measures featuring a variety of accent patterns to confirm if indeed pitch memory is selectively predictive of lexical accent pattern. In addition, further exploration of the role of PM in tone languages like Mandarin, in the form of tasks which specifically target pitch memory, is needed to shed light on this issue.

In sum, Experiment 1 showed that L1 Japanese listeners from the Tokyo-standard dialect region can reliably judge the correctness of the accent patterns of spoken stimuli, but that categorizing words by their accent pattern in a task requiring phonetic comparison with visual pitch contours proved to be more difficult. Importantly, acoustic pitch sensitivity, a general cognitive ability, significantly predicted performance variation on both experimental tasks. In addition, PM capacity predicted perception ability only in an accent pattern-specific manner on the memory-intensive categorization task. This pattern of findings suggests that even among adult L1 Japanese listeners, specific task demands contribute to performance variation, and domain-general cognitive resources continue to support the perception of speech sounds involving lexical-level accents.

Chapter 4: Experiment 2 – Accent Perception by Advanced Learners of Japanese

4.1 Introduction

This chapter presents the second experiment, in which I explore four factors that potentially contribute to lexical accent perception by second language learners of Japanese. Two of the factors are domain-general cognitive abilities, phonological short-term memory (PM) and acoustic pitch sensitivity; and two are variables that I classify as experience-based or “top-down” factors, Japanese vocabulary knowledge and L1 experience with lexical tone. Although L1 tone knowledge can be considered as deriving from both experience-related and internal sources, I chose to group it into the former category for the following reasons. First, we must consider the fact that languages differ in their prosodic systems, and that the prosodic features in the L1 are known to influence the perception of L2 sounds. This immediate influence on L2 perception indeed stems from the established, or internal, knowledge of one’s L1 prosodic system. However, the narrowing of the perceptual space during L1 acquisition of Mandarin Chinese lexical tones is the result of repeated exposure to the language, and is typically gained from the L1 speech environment. This gives rise to the experience-dependent view of L1 tone knowledge adopted in the current study.

As with Experiment 1, two perception tasks that varied in the locus of their decisions were presented to L2 learners; one task required a speeded correctness

judgment of a word's accent, and the other a categorization of this accent pattern. Thus, my first goal is to compare the perception ability of advanced-proficiency learners who have spent extended time in Japan with that of the native Japanese speakers tested in the previous chapter. However, my central aim is to account for previously-reported variation in L2 accent perception at the advanced level by contrasting the relative contribution of domain-general cognitive resources with language-specific knowledge on accent perception tasks. To accomplish this, I closely matched two groups of advanced L2 Japanese learners, speakers of a tone language (Mandarin Chinese) and speakers of a non-tone language (Korean speakers), on their Japanese proficiency level. Language proficiency was assessed by both lexical and grammatical knowledge tests, as well as an evaluation of learners' overall exposure to Japanese. I then measured learners' performance on the domain-general predictor variables, which were compared with accent perception ability on the two lexical perception tasks. The following questions guided this study: Do differences in basic cognitive resources manifest as performance differences at the advanced level? Alternatively, does L2 lexical knowledge uniquely contribute to learners' ability to perceive pitch accent? Does L1 tone experience enhance perception of lexical accent relative to a proficiency-matched non-tone L1 group?

4.2 Predictions

Experiment 2 addresses four predictions regarding advanced learners' perception of lexical pitch accent. In general terms, I predicted that at the advanced proficiency level the two experience-based factors, L2 vocabulary knowledge and L1 tone experience, will be more closely predictive of accent perception ability than basic cognitive resources.

Specific predictions are as follows. First, learners with greater L2 lexical knowledge, but not grammatical knowledge, will be more accurate at perceiving accents due to the lexically-focused nature of the perception tasks. Bear in mind that correctness judgments, which formed part of the perception task as in Experiment 1, necessarily require a comparison of the perceived input with its representation in the listener's long-term lexical store. Second, tone experience in the L1 will afford an advantage to accent perception, such that Chinese speakers will outperform proficiency-matched Korean speakers.

Third, considering that in Experiment 1 native Japanese speakers' acoustic pitch sensitivity predicted their accent perception ability, I cannot rule out a similar relationship in advanced learners either. Yet, at the advanced proficiency level, L2 lexical knowledge is likely the stronger predictor of accent perception ability, since I assume that it is more reliable for a listener to utilize their stored lexical knowledge in making correctness judgments, thus relegating basic perceptual resources to a secondary role. Fourth, similar to the result found with L1 Japanese speakers, at the advanced proficiency level PM capacity is not predicted to have any influence on the perception and brief retention of the short spoken stimuli. This likely being a result of L2 learners using their long-term lexical knowledge, instead of relying on the relatively limited capacity of the short-term store, in making accent judgments.

4.3 Methods

4.3.1 Participants

A total of 61 advanced learners of Japanese, 31 L1 Chinese speakers and 30 L1 Korean speakers, participated in Experiment 2. All participants were recruited from the undergraduate and graduate community at Nagoya University, a large research university in central Japan, and ranged in age from 21 to 36 years old. None of the Korean speakers reported fluency in Chinese or a length of Chinese study as an elective longer than 1.5 years. Participants were compensated for their participation with a small payment.

Particular care was taken to control participants' Japanese learning background, as we can see in Table 5. Questionnaire data indicated that the L1 Chinese group had an average length of study (LOS) of 68 months, while the L1 Korean group's LOS was 63 months, a difference that was not significant ($t(59) = 0.859$, *ns*). A majority of the participants in both groups reported studying Japanese in their home country for at least two years (> 24 months) prior to their arrival in Japan. Neither did the groups differ by length of residence (LOR), with the Chinese group having lived in Japan for an average of 27 months, and the Korean group an average of 28 months ($t(59) = 0.611$, *ns*). Japanese usage was also assessed by a questionnaire on which participants indicated the amount of time spent speaking and listening to Japanese. Self-reported Japanese usage per week revealed highly similar figures, with Chinese speakers reporting an average of 10.6 hours compared to Koreans' 10.4 hours, again yielding no significant difference ($t(59) = 0.140$, *ns*). These three measures (LOS/LOR/Usage) clearly indicate the similarity of the two learner populations in terms of their experience with Japanese both

at home and in Japan. Results of the proficiency tests presented in the next section will further reveal the homogeneity, and by extension comparability, of the two learner groups.

Table 5. Overview of participants' background and Japanese learning experience

	L1 Chinese (<i>n</i> = 31)	L1 Korean (<i>n</i> = 30)	<i>t</i> -test
Age (<i>years</i>)	21-36	21-36	--
Length of Study (<i>mos.</i>)	68	63	<i>t</i> = .611, <i>ns</i>
Length of Residence (<i>mos.</i>)	27	28	<i>t</i> = .178, <i>ns</i>
Usage/week (<i>hrs.</i>)	10.6	10.4	<i>t</i> = .140, <i>ns</i>

4.3.2 Materials

The following tasks were presented to all participants in this study. The first three tasks—the adaptive pitch test, serial non-word recognition (the measure of phonological memory capacity) and L2 lexical knowledge—formed the predictor variables. Native language was used as a categorical variable to separate the participants into tone and non-tone L1 groups. Participants' scores on the accent pattern correctness judgment and categorization tasks constituted the dependent variables. Only the portions of the tasks which differed design-wise from those in Experiment 1 will be described below.

4.3.2.1 Acoustic pitch sensitivity

The materials used in the adaptive pitch test were identical to those in Experiment 1.

4.3.2.2 PM

The serial nonword recognition test (SNWR), identical to that in Experiment 1, was used to measure PM capacity in this experiment. Although I am aware that there is a debate in the literature as to the role of L2 WM/PM span, and how it differs from L1 memory span, in language processing (for an overview see Juffs & Harrington, 2011 and Szmalec, Brysbaert & Duyck, 2012), I selected an L2-based memory measure on the one hand because the participants lacked a common first language, but on the other because I assumed that the advanced-level participants were capable of handling the demands of an aurally-presented L2 memory task.¹⁶

4.3.2.3 Lexical/grammatical proficiency tests

Participants' L2 lexical and grammatical knowledge were measured with two proficiency tests, a 48-item lexical knowledge test (for details see Miyaoka, Tamaoka & Sakai, 2011) and a 36-item grammatical knowledge test (Miyaoka, Tamaoka & Sakai, in press). Both tests were created with vocabulary and grammar items selected from the widely-used Japanese Language Proficiency Test (Japan Foundation, 2004), and are considered to be reliable measures of their respective target abilities. These tests were validated in the two studies above, which used them to control participants' proficiency level prior to comparing their performance on a target feature. In the present study, these tests were first used to determine how similar the two L1 groups were in terms of overall Japanese proficiency, which is a necessary prerequisite to the comparison of the groups'

¹⁶ Refer to Kaushanskaya & Yoo (2013) for a comparison of bilinguals' L1 and L2 phonological and working memory performance.

accent perception ability. It should be noted that the L2 lexical knowledge test, but not the grammar test, was used as a predictor of pitch accent perception ability on the following tasks, in order to test the purported link between long-term lexical knowledge and vocabulary recognition (Martin & Ellis, 2012). Lastly, I must point out the notion of word familiarity, a subjective measure of word knowledge (Amano & Kondo, 1999; 2000), as it relates to the test of Japanese lexical knowledge. Although word familiarity based on L1 Japanese speakers' ratings was controlled in each experiment, it can be argued that due to the heterogeneity of the L2 learners' experiences (i.e., differences such as acquisition order, instruction style and learning context) these ratings are not applicable to L2. In fact, the test of lexical knowledge can possibly be construed as a measure of familiarity rather than lexicon size. I assume, however, that the current lexical knowledge test is a relatively *objective* measure of L2 lexicon size. Furthermore, L1 word familiarity ratings can not only be generally applied to advanced L2 learners, but that the high-frequency words used on the accent perception task largely nullified the influence from potential deviations between L1 and L2 familiarity, particularly since high-frequency words are typically rated highly on familiarity (Amano & Kondo, 1999).

4.3.2.4 *Pitch accent identification and categorization tasks*

Forty-eight frequency controlled nouns were selected as stimuli, and the length of the carrier sentences was set to 6-7 mora for 3-mora target nouns as in 八時に起きる /*hatizi ni okiru*/ '(I) wake up at 8 o'clock' and 7-8 mora for 4-mora targets as in 地下鉄に乗る /*tikatetu ni noru*/ '(I) ride on the subway' ($M = 6.7$ mora; $SD = .52$). Half of the 48 stimuli contained 3-mora target nouns, and the other half 4-mora nouns. 32 of these 48

stimuli were identical to those used in Experiment 1. The stimuli were further divided into four pitch accent patterns, yielding the following accent types for 3-mora nouns—Pattern 1: HLL+L, Pattern 2: LHL+L, Pattern 3: LHH+L, Pattern 4: LHH+H; and for 4-mora nouns: HLLL+L, LHLL+L, LHHH+L, LHHH+H. Additionally, 24 of the noun stimuli were spoken with the correct pitch accent and 24 with an incorrect accent pattern. All stimuli were sampled at 44.1 kHz and were presented to participants with no phonetic manipulation or normalization. The same male speaker as in Experiment 1 produced all stimuli in the current experiment.

Stimuli were controlled by frequency across each of the four accent types based on the NTT corpus, as in Experiment 1 (Amano & Kondo, 1999; 2000). To account for the added stimuli in Experiment 2, mean normative frequencies were recalculated for all stimuli. The total frequency was 30.3 tokens per million ($SD = 46.3$), and by pattern as follows: Pattern 1 ($M = 32.1$, $SD = 45.8$), Pattern 2 ($M = 14.4$; $SD = 13.1$), Pattern 3 ($M = 37.3$, $SD = 59$), Pattern 4 ($M = 37.4$; $SD = 54.7$). Although Pattern 2 was again lower in frequency than the other three patterns, this difference was not significant ($F(3, 44) = .650$, $p = .587$). Word familiarity was also verified in the above corpus (Amano & Kondo 1999). The mean familiarity rating for all stimuli was 6.19 ($SD = .29$), indicating the words were subjectively very familiar to L1 Japanese speakers, and thus assumed to be familiar to advanced L2 learners as well. Further, no difference was found when comparing pattern type by familiarity rating ($F(3,44) = 2.204$, $p = .101$), however, the difference between Pattern 1 and Pattern 3 did approach significance ($p = .085$), with Pattern 3 ($M = 6.3$) words being of slightly higher familiarity than Pattern 1 ($M = 6.02$).

Lastly, all carrier sentences were controlled to a mean length of 6.9 mora ($SD = .80$).

Refer to Appendix B for a complete list of target words.

The PitchID and PitchCAT tasks were combined into one task using shared stimuli as follows. For the correctly-accented items only ($n = 24$), immediately following the PitchID judgment, on which reaction time was measured, participants were required to categorize the noun-plus-postposition into one of four pitch contours by selecting a graph which corresponded to the pitch contour (see Figure 4 in the previous chapter). Stimuli sentences were only heard once by the listener, and participants were never required to categorize a word that was spoken with the incorrect pitch accent. Lexical pitch patterns are, of course, abstract in the sense that there is no standard way to orthographically transcribe pitch contours, making this categorization task a type of metalinguistic judgment. I consider this knowledge, coupled with the online correctness judgment in the PitchID task, an overall measure of the strength of the representations of lexical pitch accents in a learner's mental lexicon.

4.4 Procedure

With the exception of the vocabulary and grammar proficiency tests, all tasks were presented to participants on a computer with headphones in a soundproofed room. Stimuli for the serial nonword recognition (PM) task and the PitchID/PitchCAT tasks were presented in E-Prime, with correct or incorrect judgments made via a 2-button response for the PM and PitchID tasks. Reaction time was recorded for the PitchID task and participants were asked to respond as rapidly as possible. The PitchCAT task employed a 4-button response in which each button corresponded to one of the four pitch

patterns, and did not require a speeded response. Experimental tasks were presented in the following order: 1) Adaptive pitch test, 2) Serial nonword recognition (PM), 3) PitchID/PitchCAT tasks. A short break was provided after the PM task. Each task began with a brief practice phase in order to familiarize participants with the procedure. Due to the complexity of the PitchCAT task in particular, participants were familiarized with sample pitch contour graphs (Chapter 3, Figure 4) similar to those in the main categorization task during the practice phase. All participants stated they understood the categorization process prior to beginning the main experiment. Upon completing the experimental tasks, the written lexical and grammar knowledge tests were administered to all participants. Including the proficiency tests, participants required about 1 hour 15 minutes to complete all experimental tasks. At the conclusion of the study, participants were asked to fill out a background questionnaire, and were provided a small payment for their participation.

4.5 Results

4.5.1 *Predictor variables*

The descriptive results for all variables are presented in Table 6 below. On the adaptive pitch test, which measured acoustic sensitivity, the L1 Korean group outperformed the L1 Chinese group in their ability to discriminate nonlinguistic tone pairs. A t -test confirmed that this difference was significant ($t(59) = 3.25, p < .01$). Conversely, L1 Chinese displayed a greater capacity to process spoken nonwords in the PM task than L1 Koreans ($t(59) = 2.74, p < .01$). On the lexical knowledge test, which estimated the size of participants' L2 lexicon, there was no significant difference ($t(59) =$

0.552, *ns*) between L1 Chinese ($M = 37.94$, $SD = 4.93$) and L1 Korean speakers ($M = 38.70$, $SD = 5.87$). Likewise, the measure of grammar knowledge revealed no significant difference ($t(59) = 0.903$, *ns*) between Chinese ($M = 31.94$, $SD = 3.05$) and Korean speakers ($M = 32.60$, $SD = 2.67$).

Table 6. Descriptive statistics for all variables in L1 Chinese and Korean groups

Variable	L1 Chinese ($N = 31$)					L1 Korean ($N = 30$)			
	Max	M	SD	Range		M	SD	Range	
Pitch sensitivity	--	24.66	19.65	1.88	50	9.83	11.65	0.45	50
Phonological memory	144	88.19	17.50	66	137	74.57	21.28	31	121
Lexical test	48	37.94	4.93	27	45	38.70	5.87	28	47
Grammar test	36	31.94	3.05	23	36	32.60	2.67	25	36
PitchID	48	37.16	3.03	31	42	28.40	5.49	21	43
PitchCAT	24	14.45	4.93	6	24	11.83	3.48	5	21

Notes: Pitch sensitivity measured in Hz; $*p < .05$; $**p < .01$; $***p < .001$

The results of these latter two language proficiency measures, combined with the L2 learning background questionnaire data discussed in the Methods section, indicate highly similar groups in terms of Japanese ability and learning experience. By controlling between-group Japanese proficiency, this enabled me to examine the role of within-group variation on the predictor variables and how these related to the accent perception tasks. The three variables of interest—pitch sensitivity, phonological short-term memory and

L2 lexical knowledge—were subsequently examined as predictors of performance on the accent perception tasks discussed below.

4.5.2 Lexical accent perception tasks

Figure 5 displays the results by L1 group for the PitchID and PitchCAT perception tasks. Here we see a clear difference in accuracy scores between the L1 Chinese and Korean groups on both tasks.

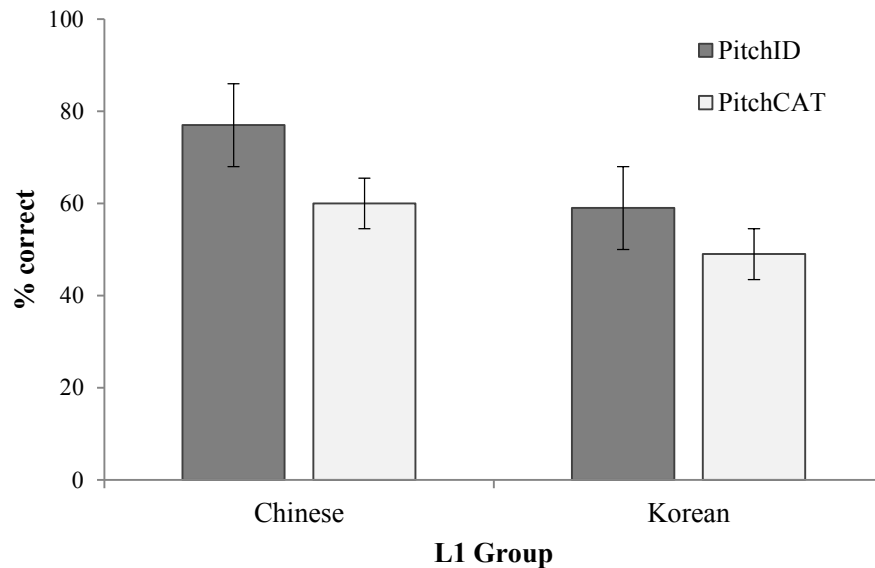


Figure 5. Mean accuracy scores (with standard error bars) on PitchID and PitchCAT tasks

On the PitchID task, group accuracy differed significantly, with the L1 Chinese group attaining a mean accuracy of 77.5%, while the L1 Korean group displayed a mean of 59.2% ($t(59) = 7.74, p < .001$). The same trend was found on the PitchCAT task, in

which learners assigned visual representations of pitch contours to a word's accent pattern, with the L1 Chinese posting a mean accuracy of 60% and Korean speakers a mean of 49% ($t(59) = 2.39, p < .05$). Bear in mind that the categorization task required a 1-out-of-4 answer choice, so I consider learner performance, although low percentage-wise, to be well above the level of random chance (i.e., ~25%).

4.5.3 Relationship between learner-variables and pitch perception

In order to explore the relationship between the predictor variables and accent perception, I next calculated correlation coefficients, which are presented in Table 7. This table combines the results of participants from both L1 groups ($N = 61$). When using L1 as a categorical variable, I found a strong correlation between native language and performance on the PitchID task ($r = .71; p < .01$), as well as a low-moderate correlation on PitchCAT ($r = .29; p < .05$). Furthermore, Japanese lexical knowledge correlated with both PitchID ($r = .32; p < .05$) and PitchCAT performance ($r = .34; p < .01$). As predicted, L2 grammatical had little relation to performance on either of the lexical perception tasks. However, PM capacity and acoustic pitch sensitivity, the two domain-general cognitive abilities, did not yield significant correlations on either of the accent perception tasks. The combined data of the L1 groups indicated a relationship between both L1 background and L2 lexical knowledge and performance on the lexical accent perception tasks. The analysis also revealed that listeners who could accurately judge accent correctness on the PitchID task also showed a strong tendency to perform well at accent categorization ($r = .52; p < .01$). Reaction time (RT) on the PitchID task yielded a significant negative correlation with accuracy on PitchCAT ($r = -.26; p < .05$), suggesting

that listeners who made rapid decisions on accent correctness were also more successful at categorizing target words by pattern.

Table 7. Correlations between predictor variables and lexical accent perception tasks for L1 Chinese and Korean groups combined ($N = 61$)

Variable	1	2	3	4	5	6	7
1. L1	–						
2. Lexical knowledge	-.07	–					
3. Grammar knowledge	.12	.58**	–				
4. Pitch sensitivity	.39**	.02	-.01	–			
5. Phonological memory	.34**	.01	.04	.22	–		
6. PitchID	.71**	.32*	.12	.32	.09	–	
7. PitchID RT	-.12	-.21	.01	-.05	.15	-.09	–
8. PitchCAT	.29*	.34**	.17	.11	.09	.52**	-.26*

Notes: RT = reaction time; * $p < .05$; ** $p < .01$

I next separated the participants by L1 group to examine the relationship between lexical knowledge and accent perception ability. Within the L1 Chinese group, I found a strong positive correlation between L2 lexical knowledge and performance on the PitchID task ($r = .64, p < .01$). Likewise for L1 Korean speakers, those with greater Japanese lexical knowledge displayed a strong tendency to perform accurately on the PitchID task ($r = .49, p < .01$). Similar results emerged on a comparison of lexical knowledge and PitchCAT performance, with significant correlations again being noted for both groups (Chinese: $r = .41, p < .05$; Korean: $r = .38, p < .05$). From these results we can tentatively say that despite the apparent advantage for native speakers of Chinese,

who have rich experience using tone phonemically in their L1, the size of a learner's L2 lexicon, as measured by the lexical knowledge test, is integrally tied to their ability to perceive L2 lexical accent. In order to determine the relative contribution of the predictors on performance variation, I next conducted multiple regression analyses for both perception tasks.

4.5.4 Multiple regression analyses

A multiple regression analysis, with four variables entered into the model, was performed to identify predictors for the PitchID task. As reported in Table 8, with this regression model, which accounted for 64.2% ($R^2 = 0.64$) of the total variance in performance, I identified both L1 ($\beta = 0.74, p < .001$) and L2 lexical knowledge ($\beta = 0.37, p < .001$) as significant contributors to the ability to make correctness judgments on spoken target word-plus-postposition units. However, PM ($\beta = -0.18, ns$) and acoustic pitch sensitivity ($\beta = 0.03, ns$) did not emerge as significant factors in this model.

Table 8. Multiple regression analysis for the entire cohort (N = 61)

Predictor variable	R^2	β	t
PitchID	0.64		
L1		0.74	9.35***
Lexical knowledge		0.37	4.72***
Phonological memory		-0.18	-2.18
Pitch sensitivity		0.03	0.35
PitchCAT	0.22		
Lexical knowledge		0.36	3.12**
L1		0.32	2.78**
Pitch sensitivity		-0.03	-0.22
Phonological memory		-0.02	-0.18

Note: *** $p < .001$; ** $p < .01$

Next, a second regression analysis was conducted to identify predictors on the PitchCAT task. The lower half of Table 8 shows that this model accounted for 22.0% ($R^2 = 0.22$) of the total variance, with L2 lexical knowledge ($\beta = 0.36, p < .01$) and L1 ($\beta = 0.32, p < .01$) both contributing significantly to the model. As with the model for PitchID, neither acoustic pitch sensitivity ($\beta = -0.03, ns$) nor PM ($\beta = -0.02, ns$) were related to perception ability. To summarize, both native language background, in this case tone-language experience, and the size of an individual's L2 lexical knowledge independently explained a significant proportion of the variance in the participants' ability to accurately perceive and categorize lexically-accented Japanese vocabulary.

4.5.5 Analysis of individual accent patterns

4.5.5.1 Pattern type and PitchID performance

Considering that previous research has shown that not all accent types may be perceived with equal accuracy (e.g., Nishinuma et al., 1996; Toda, 2001), I next evaluated the test stimuli by pitch pattern to identify perception difficulty stemming from particular lexical accent patterns. Table 9 shows the accuracy scores separated by pattern on the PitchID task for both L1 groups.

Table 9. Correct/Incorrect response totals on the PitchID task for both L1 groups.

Target	L1 Chinese (<i>N</i> = 31)				L1 Korean (<i>N</i> = 30)			
	Correct		Incorrect		Correct		Incorrect	
Pattern 1	316	84.95%	56	15.05%	199	55.28%	161	44.72%
Pattern 2	311	83.60%	61	16.40%	234	65.00%	126	35.00%
Pattern 3	181	48.66%	191	51.34%	157	43.61%	203	56.39%
Pattern 4	344	92.47%	28	7.53%	262	72.78%	98	27.22%

When errors in correctness judgments were broken down by accent pattern, we can see that Chinese speakers made the majority of their errors on Pattern 3 (LHH+L pitch) words. In fact, more than half of the totals errors ($n = 191$, or 57% of all errors) occurred on this pattern. Considering that half of the stimuli were presented with the correct accent and the other half the incorrect, I next examined the effect of correctness on perception accuracy. Chinese listeners made 80.65% of their errors (271 out of 336 total errors) on the incorrectly accented stimuli, which they erroneously accepted as

correct. The remaining 19.45% of errors ($n = 65$) were on correctly accented items that were mistakenly rejected. Korean speakers, on the other hand, showed a more dispersed response pattern, although again Pattern 3 words yielded the lowest accuracy score among the four patterns. Similar to L1 Chinese, rejecting incorrectly accented items was again problematic—74% of errors (435 out of 588 total errors) occurred on incorrect words—which listeners accepted as correct. The remaining 26% of errors ($n = 153$) were on correctly pronounced stimuli, which listeners incorrectly rejected.

I next conducted a decision-tree analysis to visually explore the PitchID task responses by both accent pattern and L1 group (Tamaoka & Ikeda, 2010). A decision tree enabled me to map participants' response data onto a multi-tiered chart, which is separated into branches, or nodes, according to the two factors of interest, accent pattern and L1. A Chi-squared goodness-of-fit test was used to divide the responses into these nodes, with statistically-different factors (i.e., accent patterns or L1 group) being assigned to separate nodes. It is important to note that factors which do not differ significantly are combined into a single node in the tree. This method allows for the easy visualization of the variables on which participants' performance diverged. The decision tree in Figure 6 shows in the uppermost node the total number of correct and incorrect responses made by the combined participant groups ($N = 61$). The next level down the tree displays responses by accent pattern in their order of accuracy, from left to right. Here we observe that the accent patterns are separated into three nodes based on statistical differences in accuracy rates, which yielded the accuracy order Pattern 4 (82.79%) > Pattern 1 = Pattern 2 (72.40%) > Pattern 3 (46.17%). Note here that these percentages are the combined

accuracy rates of *both* participant groups. Finally, in the lowest level of branches, responses are broken down by L1 group. We can interpret the split below Pattern 4 (nodes 4 & 5) as reflecting a significant difference between Chinese and Korean speakers on their judgments of these words, as well as on their accuracy for Patterns 1 and 2 (nodes 6 & 7). However, this breakdown reveals that the groups did not differ on Pattern 3 accuracy, which was equally difficult for both groups. In sum, this analysis enabled me to view the correctness judgment results on a wider scale by separating perception scores by both accent pattern and L1 group.

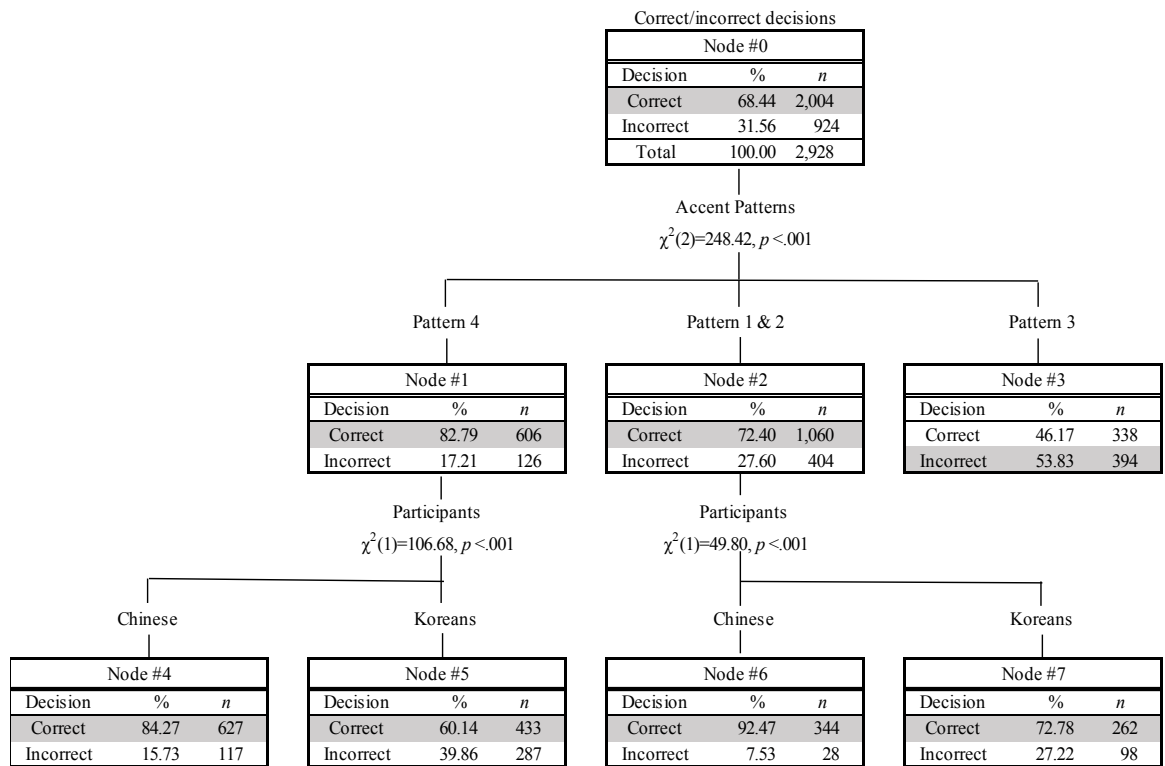


Figure 6. Decision tree analysis for performance on the PitchID task by L1 Chinese and Korean speakers

4.5.5.2 *Pattern type and PitchCAT*

I next considered the effect of accent pattern on accuracy in the visual accent categorization task. For the results of the 4-choice PitchCAT task, I constructed a confusion matrix in Table 10 to pinpoint the type of categorization errors listeners made when selecting accent pattern graphs (So & Best, 2010). The “Target” column on the left indicates the actual lexical pitch pattern heard by the listeners, while the numbered “Response Frequency” row at the top displays responses the listeners selected from among the four accent graphs. For example, on Pattern 1 targets, L1 Chinese listeners correctly categorized 69.35% (or 129 items) as Pattern 1, while only 10.22% (19 items) were mistaken for Pattern 2. Contrast this with Pattern 3 targets, in which L1 Chinese only selected the correct graph on 42.47% of the targets.

Here we see a clear tendency for the L1 Chinese group to misidentify Pattern 3 as Pattern 4 (42.47% versus 40.32%), with a nearly equal percentage of responses made between these two accent types when categorizing Pattern 3 stimuli. However, note that for target Pattern 4 items, Chinese listeners showed a relatively high level of accuracy, with the majority of responses (69.89%) being correct. A Chi-squared goodness-of-fit test of revealed a significant difference (far-right column in Table 10) between the categorization accuracy of the target stimuli, in gray, and the incorrect answer choices presented in the same row.

Table 10. Confusion matrix for accent pattern categorizations by L1 Chinese speakers

Target	Response Frequencies and Percentage								Chi-squared test
	Pattern 1		Pattern 2		Pattern 3		Pattern 4		
Pattern 1	129	69.35%	19	10.22%	14	7.53%	24	12.90%	$X^2=196.24, p<.001$
Pattern 2	10	5.38%	110	59.14%	36	19.35%	30	16.13%	$X^2=123.59, p<.001$
Pattern 3	10	5.38%	22	11.83%	79	42.47%	75	40.32%	$X^2=81.74, p<.001$
Pattern 4	13	6.99%	13	6.99%	30	16.13%	130	69.89%	$X^2=112.00, p<.01$

Notes: $N = 186$ total responses per pattern. Gray-shaded cells indicate correct responses.

Next, a decision tree analysis was conducted to evaluate the response patterns made by the Chinese participants. Although the raw percentages of the accuracy rates differed between the correctly-categorized targets, as noted in the gray-colored cells in Table 10 above, the analysis revealed the accuracy order to be Pattern 4 = Pattern 1 = Pattern 2 > Pattern 3 ($X^2(2) = 38.28, p < .001$). That is, Patterns 4, 1 and 2 were categorized with significantly higher accuracy than Pattern 3, but that the accuracy of these three patterns did not differ from one another.

In Table 11, we see that L1 Korean speakers were accurate in their categorization of Pattern 1 stimuli. Yet there was a gradual decrease in accuracy across the remainder of the accent patterns, with the lowest scores appearing on Pattern 4 words.

Table 11. Confusion matrix for accent pattern categorizations by L1 Korean speakers

Target	Response Frequencies and Percentage								Chi-squared test
	Pattern 1		Pattern 2		Pattern 3		Pattern 4		
Pattern 1	131	72.78%	21	11.67%	17	9.44%	11	6.11%	$X^2=220.3, p<.001$
Pattern 2	15	8.33%	94	52.22%	48	26.67%	23	12.78%	$X^2=84.31, p<.001$
Pattern 3	17	9.44%	43	23.89%	84	46.67%	36	20.00%	$X^2=53.11, p<.001$
Pattern 4	31	17.22%	39	21.67%	62	34.44%	48	26.67%	$X^2=11.78, p<.01$

Notes: $N = 180$ total responses per pattern. Gray-shaded cells indicate correct responses.

I again examined the categorization patterns displayed by Korean speakers using a series of Chi-squared tests. For this group, I found a markedly different accuracy order in comparison with L1 Chinese, namely, Pattern 1 > Pattern 2 = Pattern 3 > Pattern 4 ($X^2(3)=76.56, p<.001$). Thus, for the Korean group, the initial-mora accented Pattern 1 was categorized significantly better than Patterns 2 and 3, while Pattern 4, unlike what we saw with the L1 Chinese, was by far the most problematic.

4.7 Discussion

In the present chapter I explored both domain-general and language-specific variables as predictors of lexical accent perception in two native-language groups of advanced L2 Japanese learners. Acquiring the ability to accurately perceive lexical accent patterns presents a challenge to L2 Japanese learners from a range of proficiency levels and native language backgrounds (Shibata & Hurtig, 2008; Lee et al., 2006). I predicted that with advanced learners, domain-general cognitive resources such as acoustic sensitivity and PM capacity would not be implicated in the perception of lexical accent

patterns, or will be involved to a much lesser extent, than long-term knowledge of L2 vocabulary and L1 tone experience. These findings paint a complex picture of the abilities involved in learners' perception of lexically-accented speech, with three relevant results emerging. First, experience using tone phonemically by L1 Chinese speakers greatly facilitated their accent perception ability in comparison to the proficiency-matched L1 Korean speakers. Second, L2 lexical knowledge, and by extension knowledge of Japanese phonological regularities, predicted performance regardless of a listener's native language. Finally, perception accuracy varied in relation to specific accent patterns, L1 group and task type.

The first finding, presented earlier in Figure 5, was that Chinese speakers as a group were more accurate than Koreans at judging the correctness of lexical accents in spoken Japanese noun-plus-postposition units¹⁷ (PitchID: 77.5% versus 59.2%). A parallel result was also obtained on the 4-choice pitch categorization task (PitchCAT: 60% versus 49%). Although the two L1 groups were closely matched on standardized L2 vocabulary and grammar proficiency measures, the Chinese speakers were significantly more accurate at performing correctness judgments than the Korean speakers. I conclude that this result is an instance of an L1 phonemic feature facilitating L2 perception ability. In particular, Chinese contains word-level tones that indicate phonemic variations—for example, Tone 2 (falling) and Tone 4 (rising) have been noted to approximate the falling (H-L) and rising (L-H) pitch contours found in Japanese (So & Best, 2010). Although

¹⁷ This unit is referred to as a *bunsetsu* (文節) in Japanese, and is generally defined as a content word followed by a functional item such as a case marker or a postposition.

these pitch shapes are not phonetically identical in the two languages, and are used over different timing units (i.e., single syllable in Chinese versus two mora in Japanese), their phonemic use in Chinese likely increases the salience of word-level Japanese accents for Chinese speakers than for Koreans. Additionally, the fact that Chinese speakers focus on multiple phonetic cues—pitch height and movement—when perceiving tones in their L1 potentially aids their perception of pitch accent cues in Japanese. Furthermore, the majority of L1 Chinese speakers' errors were confined to their incorrect rejection and miscategorization of a single accent type (Pattern 3). Pattern 3 words feature a High-Low pitch fall, as do Patterns 1 and 2, but this drop is located on the word-final mora and is dependent on the presence of a following particle. Although a specific Chinese tone (Tone 4) may have facilitated perception of the accent fall in Pattern 1 words, Patterns 2 and 3 feature both a pitch rise on the initial mora *and* a pitch fall later in the word. Considering this, familiarity with the H-L pitch fall alone may not have aided in Patterns 2 and 3 perception. Thus, because the presence of pattern-specific facilitation was unclear in the Chinese group, the most plausible account for their overall high performance on non-Pattern 3 is that experience with L1 lexical tone draws attention to the general properties of the Japanese accent system during acquisition.

Although the data indicate a marked advantage for the tonal-L1 Chinese group, I found that L2 lexical knowledge was closely tied with performance in both L1 groups on the perception tasks. The regression analysis in Table 8 showed that L2 lexical knowledge significantly predicted learners' ability to judge ($\beta = .37, p < .001$) and categorize ($\beta = .36, p < .01$) lexically-accented words. Importantly, even for the Korean

speakers, those with greater overall lexical knowledge were more accurate on both accent perception tasks. In other words, within-group variation in L2 lexical knowledge accounted for accent perception performance regardless of L1. If we consider that neither L2 grammar knowledge, length of study or residence in Japan were correlated with perception ability, thereby ruling out the possibility that task performance was a function of general L2 proficiency, we can conclude that accent perception is specifically tied to learners' word knowledge. Possessing robust lexical knowledge in one's L2 provides a wealth of exemplars of word form, and in Japanese in particular, the results suggest that lexical knowledge is a good indicator of accent perception ability. Furthermore, this finding ties in with theories suggesting a phonetic-to-lexical continuum of lexical acquisition, which propose that L2 learners increasingly utilize lexical strategies, as opposed to relying on basic perceptual resources, to aid speech perception as their L2 knowledge increases (Wong & Perrachione, 2007). The present results suggest that those from either L1 group who possess a high level of lexical knowledge have largely internalized lexical accents, and are thus situated at the lexical end of this proposed spectrum. The next chapter will continue to explore this line of phonetic to lexical development.

This finding also echoes what SLA researchers have said about the role of long-term memory in lexical learning tasks by non-beginner L2 learners. Namely, the stable, long-term phonological representations take a much more prominent role in language processing as learners progress in their L2 proficiency (e.g., Martin & Ellis, 2012; Speciale et al., 2004). Rather than “falling back” on domain-general cognitive abilities

such as acoustic perception and PM, the advanced learners in the present study likely made use of specific knowledge of the Japanese lexical accent system to judge the correctness of and categorize spoken words. Although it must be noted that for Pattern 3 words in particular, which require knowledge of the word-plus-particle unit, this lexical perception account may not hold up. In sum, the results from the current experiment support my prediction that learners who possess greater lexical knowledge of Japanese, but who are not generally more proficient in the language, also have more robust representations of L2 pitch accent, enabling greater accuracy on both perception tasks.

Yet in contrast with the two previous experience-based factors, the basic cognitive abilities of acoustic pitch sensitivity and PM were not significant predictors of accent perception. With advanced learners, acoustic sensitivity in non-lexical contexts no longer mediated performance, or at least was overshadowed by the experience-based factors, on the two perception tasks. This finding is in stark relief with the finding in Experiment 1 that L1 Japanese speakers' accent perception ability was accounted for to a significant degree by acoustic sensitivity. I will further discuss this discrepancy in the final chapter, and will only address the findings within the context of the current experiment here.

The majority of studies examining basic acoustic sensitivity and lexical perception have focused on target-language naïve (i.e., absolute beginners) or low proficiency learners, finding that sensitivity to pitch gained from music training, for example, was related to perception accuracy on non-native pitch and tone contrasts (e.g., Wayland et al., 2010; Wong & Perrachione, 2007). In these studies, listeners with little to no knowledge of L2 lexical form likely had no alternative but to rely on basic perceptual

resources when making decisions on target-language sound categories like Mandarin tones or Japanese pitch patterns. To their ears, even real-word stimuli were perceived as meaningless acoustic input, much in the same way that participants in the current study perceived the stimuli in the adaptive pitch test. Moreover, the pitch sensitivity task measured listeners' acuity in distinguishing subtle variations in pure tones, which although roughly approximated the bitonal (L-H/H-L) pitch variations found in the Japanese prosodic system, undoubtedly tapped an acoustic, rather than linguistic, mode of perception (Strange & Shafer, 2008). This view explains why the non-tone L1 Korean speakers were able to outperform L1 Chinese on the acoustic sensitivity measure. In other words, no advantage was derived from the L1 when discriminating pure tones. In the next chapter I will further address the role of basic acoustic sensitivity in beginning-level learners' perception of accent.

The main observation regarding memory and accent perception is that advanced learners were no longer relying on their phonological short-term store capacity when making decisions on relatively short spoken stimuli. I assumed that PM may be implicated in the two-part accent perception task, in which learners had to first judge, then following a short pause, categorize spoken stimuli according to visual representations of the pitch contours. Since listeners had to categorize targets based on a memory trace of the spoken stimuli following a pause, the PitchCAT task in particular was predicted to tax PM capacity. The results, however, suggested that learners were likely accessing long-term knowledge of pitch patterns in the L2 mental lexicon in making categorizations, rather than resorting to the phonological loop for short-term

maintenance of the stimuli. In fact, PM capacity is generally considered to be most closely involved in lexical acquisition in beginning-level learners¹⁸, in which it has been shown to be predictive of the acquisition of vocabulary, grammar and speech-fluency over a period of learning or training (e.g., Gathercole, 2009; Kaushanskaya, 2012; Martin & Ellis, 2012; O’Brien et al., 2007). Hence, the next experiment will introduce a longitudinal design to examine the role of PM on the acquisition process.

Finally, the analysis of individual accent patterns revealed that of the four accent types tested, certain patterns were perceived more accurately than others. On the PitchID task, I noted an increased difficulty on Pattern 3, which was especially pronounced with the L1 Chinese group, who were highly accurate on the other three patterns. L1 Korean speakers also had difficulty perceiving Pattern 3, although their errors were more diffusely distributed throughout the other accent types. Why the difficulty with Pattern 3? Pattern 3 words depend on the following postposition (for examples see Chapter 1, Section 1) for their pitch fall to become audible—in isolation three-mora Pattern 3 words carry the same (L-H-H) pitch contour as Pattern 4 words. Previous studies have reported high accuracy on Pattern 4, or “unaccented” words (e.g., Nishinuma et al., 1996; Sakamoto, 2011), but to my knowledge no research has noted such a stark difference between Patterns 3 and 4, deriving from the presence of a postposition. This may have been primarily due to the use of tasks in which stimuli were either presented in isolation or embedded in sentences like *watashi wa X to itta* (‘I said X.’), where target words directly preceded the verb. Perceiving the difference between these patterns requires

¹⁸ cf. Hummel (2009) for evidence of the role of PM in language proficiency at the advanced level.

listeners to focus their attention beyond the boundary of the word, and onto the pitch of the following postposition, with a failure to do so resulting in the inability to distinguish these patterns.

In terms of pitch categorization, lower accuracy on Pattern 3 words was again observed in both listener groups. The L1 Chinese group displayed fairly high accuracy on the other accent types, as shown in the confusion matrix in Table 10. Korean speakers, on the other hand, made few errors with Pattern 1 categorizations, but gradually declined in accuracy on Patterns 2, 3, and 4 (Table 11). It appears that for Korean speakers, rather than a specific pattern causing categorization difficulty, any non-initially accented (i.e., Pattern 1) word resulted in increased error rates. I conjecture that this is the result of an over-simplification of accent patterns by Korean speakers into an initial-accent versus non-initial accent dichotomy. In other words, if the pitch fall was not detected on the initial mora, then the word was randomly categorized as one of the other three patterns, resulting in low accuracy outside of Pattern 1 words.

4.8 Summary of the experiment

Summing up the results as a whole, I found that the L1 Chinese group was able to perform both the PitchID and PitchCAT perception tasks with higher accuracy than the L1 Korean group. Next, the regression model showed that L2 lexical knowledge significantly predicted performance in both groups, and its contribution was separate of the advantage afforded by L1 tone experience. Finally, not all accent patterns were perceived equally well by L2 learners, resulting in both L1-group and accent pattern-specific errors. In the next chapter I will shift the focus to beginning learners of Japanese.

In contrast with the advanced learners, basic cognitive abilities are assumed to be predictive of the development of perception ability over a semester of Japanese study. Further, I introduce learning context, whether at home or abroad, as a factor which is likely involved in accent acquisition.

Chapter 5: Experiment 3 – Accent Perception by Beginning Learners of Japanese

5.1 Introduction

In this chapter I present the results of the third and final experiment of this dissertation. In this experiment I examined the development of accent perception ability in beginning learners of Japanese studying in two contexts, at home and in Japan. The same learner-internal variables as in Experiments 1 and 2 were again used as predictors of pitch accent perception, however, two additional factors—learning context and instruction style were also introduced. Specifically, I contrast a group of L1 English speakers studying Japanese in their home country (JFL group), with a group of L1 English study-abroad learners in Japan (JSL group). Moreover, Experiment 3 is a longitudinal study in which learners' perception ability was tested at two times during a semester of Japanese study. Thus, the central focus of this experiment is to examine the influence of the learner-internal capacities measured in the prior chapters on the gains made in accent perception over a semester of study, while considering the role of learner-external factors.

In the next three subsections, I will briefly review how the design of Experiment 3 relates specifically to the acquisition of lexical accent in beginning Japanese learners in light of the findings from the previous two experiments, in addition to the L2 lexical acquisition literature in general. I will then discuss the predictions for this study, followed

by the presentation of the experimental design in sections 5.3 and 5.4. Experimental results will be discussed in section 5.5, concluding with a discussion of the findings in the final section.

5.1.1 Internal resources and beginning learners

In Chapter 2, I made the case based on previous research that basic cognitive capacities may account for some of the previously-observed variation in the acquisition of lexical pitch accent. Generally speaking, these basic cognitive resources are considered to play a greater role in early-stage acquisition than in advanced proficiency L2 learners (Gathercole, 2009; O'Brien et al., 2006; 2007; Wong & Perrachione, 2007). In fact, previous studies on lexical acquisition in beginning L2 learners have suggested that phonological memory span and basic acoustic perception are active during early-stage learning, but diminish as one acquires a stable long-term store of the L2's phonological regularities. Cheung (1996), for example, found that PM capacity predicted vocabulary acquisition only in a low-proficiency group of L2 English learners, but not with more experienced learners. Similar results have also been reported by O'Brien et al. (2006) and Kormos & Safar (2008), with PM effects either significantly weakening or disappearing altogether as language proficiency increases. Juffs & Harrington (2011) summarized this pattern of findings as indicating that "PM capacity and phonological sequence learning are initially separable constraints on L2 vocabulary acquisition (p. 151)." In other words, PM acts as a support mechanism for learners in the early stages of acquiring the phonological patterns of their L2, but not in those who have established stable lexical representations in long-term storage (i.e., advanced proficiency learners). In Experiment

2 in the current study, I found that lexical knowledge was a significant predictor of accent perception in advanced learners of Japanese, but neither acoustic pitch sensitivity nor PM capacity had any effect in this group. This finding seems to align with previous research suggesting a diminished role of basic cognitive resources in lexical perception.

Acoustic pitch sensitivity may also support the learning of L2 prosodic features, as we saw in several of the studies reviewed in Chapter 2 on the L2 perception of Chinese tones. In particular, research suggests a connection between basic auditory sensitivity in nonlinguistic contexts and the perception of prosodic features in a way that highly pitch-sensitive learners are initially able to perceive tones more accurately than non-musicians (Deutsch et al., 2009; Wayland et al., 2010; Wong & Perrachione, 2007). However, it is crucial to note that following a period of perceptual training in the tone language, the supportive role of auditory sensitivity may be lessened, as evidenced by the lack of a performance difference between musicians and non-musicians (e.g., Wayland et al., 2010). In a similar vein, it was demonstrated in Experiment 2 in Chapter 4 that basic acoustic sensitivity had no predictive role of real-word accent perception in a sample of high-proficiency learners of Japanese. Rather, in the case of these learners their degree of L2 lexical knowledge was the central factor involved in accent perception. Considering previous findings suggesting an advantage for pitch-sensitive listeners in the perception of lexical prosody, the experiment in this chapter will test the prediction that acoustic pitch sensitivity is related to the development of accent perception ability in Japanese learners at low proficiency levels.

5.1.2 Learning context and L2 acquisition

The effect of learning context on lexical acquisition has been explored in detail (e.g., Freed, 1995; Sunderman & Kroll, 2009). However, data on the role of context in acquiring a specific lexical feature, in this case pitch accent, is ostensibly lacking. In terms of vocabulary acquisition, it is generally assumed that immersion in the L2 environment provides increased access to input, and hence greater exposure to exemplars of spoken words and their phonological patterning (Collentine, 2004). With Japanese lexical accent, greater repetitions of accent patterns in the input may serve to reinforce long-term representations of accent more efficiently than those without access to the increased L2 input available in the study-abroad setting. Whether or not learners focus their attention on and effectively utilize this increased input is beyond the scope of the current research, and needs to be assessed on an individual basis. Sunderman & Kroll (2009) considered the role of study-abroad on L2 acquisition in light of the observation that not infrequently, learners return from the L2 environment displaying little or no improvement in their productive vocabulary skills. They proposed a “resource threshold hypothesis” in which insufficient working memory resources in some learners may hinder their ability to benefit from the increased spoken input available in the L2 setting. Their findings suggested that in fact learners with poorer WM capacities were unable to utilize the greater L2 input to the same degree as high WM capacity learners, lagging behind the high-WM group in their ability to perform picture translation and naming tasks. Their study, however, did not consider the fact that the type of instruction delivered to learners in each context may have differed.

Although the present study does not attempt to establish a minimum “threshold” of learner-internal capacities necessary to acquire pitch accent from input in a given context, I analyze learning context in terms of a higher input (study abroad, or JSL) and a lower input (at home, or JFL) context. Bear in mind that the “higher” and “lower” distinctions are relative to the current study, based on participants’ reported time spent speaking and listening to Japanese as assessed by a questionnaire. Furthermore, tied to learning environment in this study is the fact that the JFL group received instruction and oral correction on lexical accent, while the JSL group did not. This division by learning context will be used as a predictor of perceptual development over the semester. Sunderman & Kroll’s results are another piece of evidence that access to increased input may not be universally beneficial to L2 lexical acquisition in all learners, and hence appeal to the need to consider learner-internal capacities together with learner-external factors (i.e., learning context and instruction style) as variables in L2 acquisition. Unique to the design of Experiment 3 is that learners in the study-abroad group reported no overt correction or feedback on lexical accent, while the at-home group did receive correction/instruction despite reporting lower overall exposure to and time spent using Japanese.

5.1.3 Accent acquisition over a semester of language study

The question of how lexical perception ability develops in individual learners during a language course has been insufficiently addressed, and has typically been approached through cross-sectional designs based on language proficiency level. However, there are relevant perception studies which can lend us clues. In a study on L2

Spanish learners, O'Brien et al. (2006; 2007) found that receptive vocabulary knowledge did increase over a 10-week period abroad relative to at-home learners who were matched for proficiency prior to the semester. The only study that employed a longitudinal methodology to investigate Japanese lexical accent was, to my knowledge, Lee et al. (2006). They looked at changes in production ability of lexical accent over a two-year period by native Cantonese speakers studying partly in Japan and partly in their home country. They found no clear evidence of gains in production ability despite learners' having made significant improvement in overall Japanese ability, as indexed by a standardized proficiency test. However, their small case study only included three learners, making it difficult to generalize to a larger population of L2 Japanese learners.

In this dissertation, Experiments 1 and 2 measured accent perception at a single time point, whereby limiting my consideration of the involvement of cognitive capacities in speech perception to their immediate influence on listening task performance. Yet, the central role of PM/WM, at least in the domain of L2 acquisition research, is widely thought to be that of a support mechanism for language *acquisition* (Gathercole, 2009; Martin & Ellis, 2012). Considering that the task of the L2 learner is to extract target-language relevant sound patterns from the input, and to use these phonological patterns for further learning, then the role of one's storage capacity for sounds logically comes into play. What previous research has suggested then, is that learners who are more efficient at processing speech input show greater development in the L2 (Martin & Ellis, 2012). Thus, one fundamental difference between Experiment 3 and the previous two

experiments is the fact that I focus on how cognitive capacities relate to *gains* in perception accuracy.

5.2 Predictions

Experiment 3 addresses four predictions regarding L1 English-speaking beginning learners' perception of Japanese lexical pitch accent. The first two predictions concern learner-internal factors and how they relate to accent perception gains. First, phonological memory capacity will predict gains made in perception accuracy over a semester of Japanese study. Specifically, learners with a greater capacity to retain spoken input in their phonological loop are predicted to be more efficient at extracting pitch accent information as a property of a word's phonological form, and will thus show greater perceptual development. Second, acoustic pitch sensitivity will predict gains in accent perception over a semester. Beginning learners who are initially sensitive to pitch variations in non-linguistic contexts will also be more sensitive to lexical accent contrasts due to the fact that the perceptual resources used on both tasks are likely shared to an extent. Third, considering previous findings showing little advantage of study-abroad context on accent acquisition, I predict that context alone will not have an effect of accent perception gains. Closely tied to context is my fourth prediction. Namely, with a difficult phonological feature like lexical accent, the presence of overt oral correction/feedback on accent (i.e., instruction style) will outweigh context alone as a predictor of perceptual development. Thus, JFL learners will display greater overall ability to perceive lexical accents.

5.3 Methods

5.3.1 Participants

Forty L1 English speakers were recruited for this study: 19 American undergraduates at Ohio State University (JFL) and 21 study-abroad students at two Japanese universities (JSL) in central Japan. Three JFL learners were removed from the analysis because they did not attend the second test session, and one JSL learner was excluded because of insufficient vocabulary knowledge. Thus, 36 participants (20 male, 16 female) were included in the final analysis. Since the JSL group consisted of learners from a variety of Japanese language curricula, in contrast with the largely homogenous learning background of the JFL group, these learners were tested on their knowledge of the target vocabulary at the conclusion of the experiment. Participant data is presented in Table 12 below.

Table 12. Participant data for both groups

	JSL (n=20)	JFL (n=16)	<i>t</i> -test
Age	20-34	19-26	--
LOS (<i>hrs</i>)	237	220	$t = 1.77, p = .085$
Usage/wk (<i>hrs</i>)	19	14	$t = 2.15, p < .05$
LOR (<i>weeks</i>)	22	--	--

Notes: LOS = Length of study at time of first test session; LOR = Length of residence (in Japan)

Learners in both contexts were enrolled in a second-year Japanese class at the time of the experiment and reported a comparable mean length of instructional hours

(JFL: 220 hrs vs. JSL: 237 hrs; $p = .085$). Importantly, learners in the two groups differed in their usage of Japanese per week, which was measured by their reported time spent speaking and listening to Japanese, and was assessed by a questionnaire (see Appendix D). Also critical to the present study was the fact that the JFL group received explicit correction and feedback on lexical accent, while none of the participants in the JSL group reported having experienced instruction or correction on lexical accent. However, JSL group participants all indicated that they were aware of the existence of the pitch accent system in Japanese. Based on the author's classroom observation of the JFL group, explicit correction means that when learners made an error on producing an accent, the instructor would point out the error, model the correct accent, then have the learner reproduce the mistaken utterance. Such corrections are made with regularity and in front of the entire class, making this input accessible to all learners. Thus, accurate perception of this corrective feedback is important in forming representations of lexical accent patterns. JSL learners reported no such exchanges regarding accent in their Japanese language classes. Lastly, all participants reported normal hearing. All were volunteers and were compensated for their time with a small payment.

5.3.2 Materials

5.3.2.1 Acoustic pitch sensitivity

The materials used in the adaptive pitch test were identical to those in Experiments 1 and 2.

5.3.2.2 Nonword discrimination

A second measure of acoustic sensitivity, in the form of a nonword discrimination task, required a same-different judgment on 72 pairs of Japanese mora-based nonwords that varied by accent pattern alone (e.g., *NEmate* (HLL) vs. *neMATE* (LHH)). Although the use of nonwords in this task meant that listeners had to primarily rely on acoustic perception to differentiate the pairs, the task is more “word-like” in that the nonwords carried real Japanese pitch patterns of the same three types as featured on the real-word accent perception tasks. Thus, this task compliments the adaptive pitch test, but may carry more predictive power than pure-tone perception ability. Similar AX discrimination tasks have been used in previous studies on accent/tone perception (e.g., Bent et al., 2006; Wayland et al., 2010), and it has been generally pointed out that this type of same/different discrimination is relatively undemanding—as evidenced by high performance by non-tone language speakers—on cognitive resources in comparison with three stimuli ABX or categorization tasks. In AX discrimination tasks, difficulty is usually a function of the degree of phonetic similarity of a given stimulus pair since it involves no comparison with established lexical knowledge. For example, in Chinese tone-based tasks, high flat tones are easily distinguished from falling tones due to their high degree of dissimilarity, but tones with rising contours (i.e., T2 & T3) are more frequently confused among both experienced and inexperienced listeners (Hao, 2012; So & Best, 2010). In sum, this test provides another measure of acoustic pitch sensitivity, but with more target-language like stimuli than the pure tones used in the adaptive pitch test.

5.3.2.3 Phonological memory

The serial nonword recognition task, which measured phonological memory capacity, was identical to that used in Experiments 1 and 2.

5.3.2.4 Pitch accent identification and categorization tasks

Thirty-six nouns were selected as test stimuli from an introductory Japanese textbook (*Japanese: The Spoken Language, Vols. 1 & 2*, Jorden & Noda, 1987; 1988). These vocabulary had already been covered in the JFL group's language courses, which ensured that the learners were familiar with the words and had heard them in their spoken form. By contrast, the JSL group came from a variety of Japanese learning backgrounds, none having used this particular textbook. Thus, a post-experiment questionnaire, in which participants were asked to translate the target nouns into English, was used to assess knowledge of the stimuli. As reported above, one JSL participant was excluded due to unfamiliarity with some of the test stimuli. Target words were embedded in sentence-initial position of carrier sentences that were set to a length of 6-7 mora for 3-mora target nouns as in 手紙を書く /*tegami o kaku*/ '(I) write a letter' and 7-8 mora for 4-mora targets as in 飛行機に乗る /*hikooki ni noru*/ '(I) take an airplane.'¹⁹ Half of the stimuli ($n = 18$) were 3-mora target nouns, and the other half 4-mora nouns.

The pitch identification and categorization tasks were slightly simplified for the beginning proficiency groups by including words from only three of the four accent patterns used in the previous two experiments. Specifically, Pattern 3 words, whose pitch

¹⁹ These mora counts include the stimulus word.

accent is only audible when followed by a particle, were not included on the perception tasks. This made it unnecessary for learners to take into account the effect of the postposition on accent fall location. Thus, the three accent patterns used in this experiment were, for 3-mora nouns, Pattern 1: HLL, Pattern 2: LHL, Pattern 4: LHH; and for 4-mora nouns: HLLL, LHLL, LHHH. The LHHL pattern used in Experiments 1 and 2 was excluded in order to limit the categorization responses to three choices. This pattern is the least frequent pattern among 4-mora words (Tanaka & Kubozono, 2012), and thus was felt to be the most suitable candidate for elimination. Note that in order to maintain labelling consistency with Experiments 1 and 2, I will continue to refer to the unaccented pattern as Pattern 4, despite their only being three answer choices for the PitchCAT task. Additionally, half of the noun stimuli ($n = 18$) were spoken with the correct pitch accent and 18 with an incorrect accent pattern. All stimuli were produced by the same male native speaker of Tokyo-standard Japanese who created the stimuli in the previous two experiments. The stimuli were sampled at 44.1 kHz and presented to participants with no phonetic manipulation or normalization.

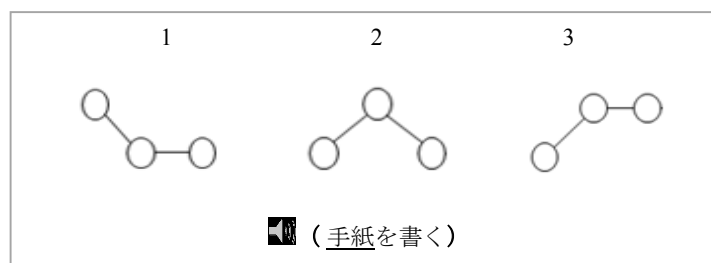


Figure 7. In the PitchCAT task, listeners categorized the underlined sentence-initial noun (/tegami o kaku/ ‘(I) write a letter’) into one of three pitch contours by selecting a visual representation of the pitch pattern. Participants did not see the Japanese script in this task.

Although target words had to be selected from the limited number of vocabulary known to this beginning group, frequency was controlled as tightly as possible across accent patterns as in the previous experiments (Amano & Kondo, 1999; 2000). The mean frequency for all test stimuli was 58.6 tokens per million ($SD = 73$), and by pattern it was as follows: Pattern 1 ($M = 55.5$, $SD = 71.6$), Pattern 2 ($M = 39.2$; $SD = 67.5$), Pattern 4 ($M = 81.1$, $SD = 73.6$). Although Pattern 2 was again lower in frequency than the other two patterns, a one-way ANOVA confirmed that this difference was not significant ($F(2, 33) = .977$, $p = .387$), which was verified on post-hoc tests. Thus, test stimuli were considered comparable in their lexical frequency across the three accent patterns. Word familiarity was again verified in the NTT corpus (Amano & Kondo 1999). The mean familiarity rating for all stimuli was 6.38 ($SD = .22$), indicating the words were subjectively very familiar to L1 Japanese speakers. Further, no difference was found when comparing words from each accent pattern type by their familiarity rating ($F(2,33) = 1.982$, $p = .154$). Lastly, all stimulus sentences including target nouns were controlled to a mean length of 6.8 mora ($SD = .71$). Refer to Appendix C for a complete list of test stimuli.

The PitchID and PitchCAT tasks were combined into one task using shared stimuli as follows. For the correctly-accented items only ($n = 18$), immediately following the PitchID judgment, participants were required to categorize only the target noun into one of three pitch contours by selecting a graph which corresponded to the pitch contour (see Figure 7 above). Stimuli sentences were only heard once by the listener, and

participants were never required to categorize a word that was spoken with the incorrect pitch accent.

5.4 Procedure

Participants in both the JSL and JFL groups took part in two test sessions during a semester of Japanese study. Early in the semester (T1), each participant took the adaptive pitch test, nonword discrimination test, and PM test, in addition to the PitchID and PitchCAT tasks. Approximately ten weeks later (T2), learners again performed the PitchID and PitchCAT tasks with test stimuli rearranged in a pseudo-random presentation order to mitigate the possibility of stimulus-order memorization. All perception tasks were administered in a quiet room via headphones on a computer running SuperLab 4.5 stimulus presentation software, with responses made through a standard keyboard. Including a short break, the test session at T1 took approximately 45 minutes, while the T2 session was typically completed in 15 minutes.

5.5 Results

5.5.1 Predictor variables

Results of the three measures of the learner-internal cognitive abilities are presented in Tables 13 and 14. Table 13 shows the data for the entire Cohort (i.e., both groups combined), while in Table 14 the results are separated by learning-context group.

Table 13. Results of learner-internal variables for combined Cohort ($N = 36$)

Variable	Max	Mean	<i>SD</i>	Range	
Pitch sensitivity	--	2.75	1.70	5	-0.91
Nonword disc.	72	67.97	3.32	58	72
PM	144	83.33	17.04	50	117

Note: Pitch sensitivity means were *log*-converted from raw Hz scores.

We can see that both groups displayed a very similar pitch sensitivity threshold on both the acoustic pitch test ($t(34) = .081, p = .936$) and the nonword discrimination task ($t(34) = 1.48, p = .148$). Average phonological memory scores were also highly similar between groups ($t(34) = .48, p = .634$).

Table 14. Comparison of learner-internal variables separated by group

Variable	JSL ($n = 20$)		JFL ($n = 16$)		<i>t</i> -test
	Mean	<i>SD</i>	Mean	<i>SD</i>	
Pitch sensitivity	2.73	1.88	2.77	1.53	<i>ns</i>
Nonword disc.	67.25	3.35	68.88	3.16	<i>ns</i>
PM	82.1	17.22	84.88	17.23	<i>ns</i>

Note: Pitch sensitivity means were *log*-converted from raw Hz scores.

Since my intent was to track the development of perception ability in this experiment, it is important that the two groups be comparable on these internal capacities at the outset of the experiment, considering previous research showing a fairly clear

initial advantage for musicians and other highly pitch-sensitive populations (e.g., Wayland et al., 2010; Wong & Perrachione, 2007). In sum, the results confirmed that in fact the groups did not differ on these predictors. Keep in mind that it is the within-group, or individual, variation in these capacities that I am most interested in.

5.5.2 Lexical accent perception tasks

Figures 8 and 9 present the results of the real-word accent perception tasks in graph form. Here I am concerned with two results. First is the comparison between the two groups' perception accuracy at both test times, and second whether each group made accuracy gains over the testing period. Figure 8 compares the performance of the JSL and JFL groups on the PitchID task at both test times. We can see that the JFL group, who received instruction on lexical accent in their language classes, was more accurate at both T1 (51.3% vs. 46.5%) and T2 (57.8% vs. 52.3%) than the JSL group, and that this difference was statistically significant at both times (T1: $t(34) = 3.45, p < .01$; T2: $t(34) = 2.85, p < .01$).

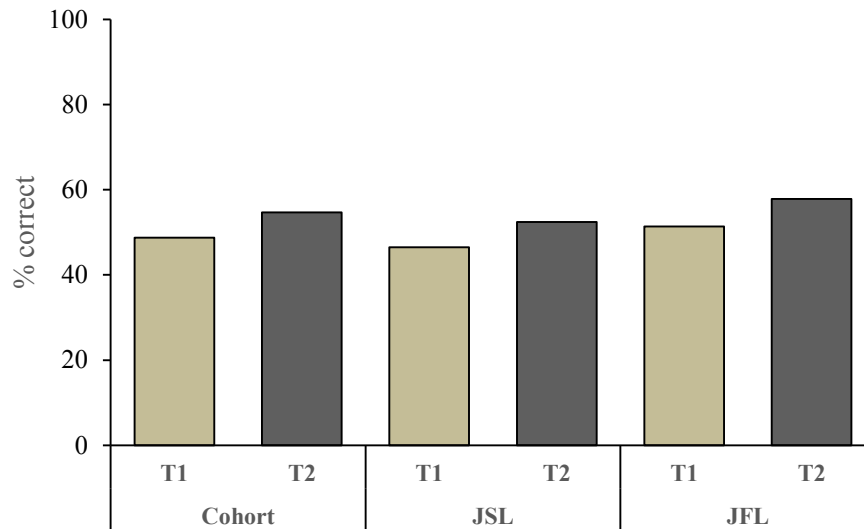


Figure 8. Group comparison of PitchID accuracy at Times 1 & 2

The PitchCAT results are shown in Figure 9. While the JFL group was again more accurate at this task at both times (T1: 52.4% vs. 45%; T2: 54.5% vs. 49.7%), this difference was negligible (T1: $t(34) = 1.46, p = .154$; T2: ($t(34) = 1.05, p = .298$), likely due to the smaller number of stimuli in this task.

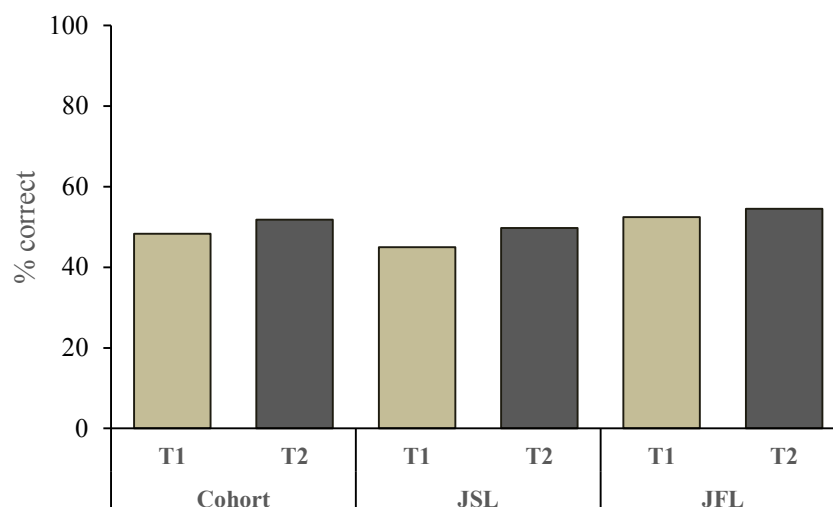


Figure 9. Group comparison of PitchCAT accuracy at Times 1 & 2

Table 15 displays the results of the accent perception tasks at Times 1 and 2. Here I considered whether the learner groups in both contexts made gains in accuracy on the accent perception tasks over the semester. To test this, I performed paired-samples t-tests on the combined Cohort ($t(35) = 4.97, p < .01$), as well as the JSL ($t(19) = 3.462, p < .01$) and JFL ($t(15) = 3.489, p < .01$) groups separately, finding that there was a significant improvement on the PitchID task. On the PitchCAT task, however, the slight development in each group did not attain significance.

Table 15. Comparison of the accent perception tasks by L1 group

Variable	Max	Entire group ($N = 36$)		JSL ($n = 20$)		JFL ($n = 16$)	
		Time 1	Time 2	Time 1	Time 2	Time 1	Time 2
PitchID	36	17.53 (1.73)	19.72 (2.25)**	16.75 (1.74)	18.85 (1.93)**	18.5 (1.15)	20.81 (2.19)**
PitchCAT	18	8.69 (2.78)	9.33 (2.44)	8.1 (2.43)	8.95 (2.31)	9.44 (3.08)	9.81 (2.59)

Notes: ** $p < .01$; SD in parenthesis.

Next, I compared perception accuracy *gain* scores between the JSL and JFL groups over the 10-week testing period. In Table 16 below we can see that both groups improved in their accuracy to a comparable degree on the PitchID and PitchCAT tasks.

Table 16. Accuracy gains over the 10-week testing period

	JSL ($n = 20$)	JFL ($n = 16$)	t -test
PitchID	5.8%	6.5%	$t = .341, ns$
PitchCAT	4.7%	2.0%	$t = .556, ns$

Two t -tests confirmed that the gain scores did not differ by group for PitchID ($t(34) = .341, p = .735$) or PitchCAT ($t(34) = .556, p = .582$), suggesting that a similar degree of improvement was achieved in both learning contexts. Up to this point we have seen that accent perception of sentence-embedded words was difficult for both groups, as evidenced by the low accuracy scores. However, it appeared that the JFL group had a slight advantage in perception accuracy at both times. Yet, both groups displayed similar gains over the testing period. In the next section, I will address the role of learner-internal capacities and how these abilities may account for individual variation in accent perception.

5.5.3 Correlations between predictor variables and accent perception tasks

The two learner groups were first combined, then correlation coefficients were calculated for the learner-internal capacities (e.g., PM) and the gain scores in accent

perception. Note that since my central interest was in the development of accent perception over the semester, I used the gain scores calculated for each participant in the correlation matrix in Table 17. First, I found that pitch sensitivity and nonword discrimination, the two measures of acoustic perception ability, were negatively correlated ($r = -.526$, $p < .01$), reflecting the fact that learners who could distinguish subtle pitch variations on the adaptive pitch test were able to discriminate more of the nonword accent pairs. This indicates that these two tests, which are considered to test basic acoustic perception, were indeed measuring the same ability (Strange & Shafer, 2008). Next, PM capacity was correlated with the gain scores in the accent correctness judgment task ($r = .424$, $p < .01$), but not the categorization task. On the other hand, PitchCAT performance, which is thought to be more biased to an acoustic mode of perception than the PitchID task, was related to the two measures of acoustic pitch sensitivity. Finally, PitchID and PitchCAT gains were correlated, indicating that learners who improved on one task tended to develop on the other, despite these two tasks apparently being supported by different cognitive resources.

Table 17. Correlation matrix for predictor variables and pitch accent gain scores ($N = 36$)

Variable	1	2	3	4	5	6
1. Learning context	–					
2. Pitch sensitivity	.014	–				
3. Nonword Disc.	.148	-.526**	–			
4. PM	.002	.056	-.121	–		
5. PitchID Gain	.058	.278	.076	.424**	–	
6. PitchCAT Gain	.095	-.444**	.397*	.222	.378*	–

Notes: ** $p < .01$; * $p < .05$

5.5.4. Multiple regression analyses

Two hierarchical multiple regression analyses were performed to determine the extent to which each of the four predictor variables—learning context, acoustic sensitivity, nonword discrimination ability and phonological memory—influenced learners’ gains in accent correctness judgment and categorization. Hierarchical regression enables us to identify the unique contribution of each independent variable, by controlling for the variance due to the other variables (entered in previous steps, hence “hierarchical”). The regression model for PitchID gain is shown in Table 18. Learning context was entered into the model in Step 1, but did not account for a significant amount of variation. In other words, as we saw in the previous section, both JSL and JFL groups improved to a similar extent. The two measures of acoustic sensitivity were added to the model next, and accounted for a total of 8.2% of the variance, but did not attain significance. In the final step, with the inclusion of PM ($\beta = .423$; $p < .01$) the model gained an additional 17.6% ($p < .01$) of explanatory power, as is shown in the ΔR^2

column ('change in R^2 '). This indicates that a significant amount of the variation in learners' perceptual development can be explained by differences in their PM capacity, with higher capacity learners tending to improve more over the semester. In total, the regression model for PitchID with all four predictors entered accounted for 26.1% ($R^2 = .261$; $p < .05$) of the total variance in learners' development of their ability to judge the correctness of spoken lexical accents.

Table 18. Hierarchical multiple regression analysis for PitchID gain

Predictors	ΔR^2	β	t
Step 1	--		
Learning context		.058	.341
Step 2	.082		
Learning context		.33	.185
Nonword disc.		.086	.411
Pitch sensitivity		.277	1.59
Step 3	.176**		
Learning context		.017	.106
Nonword disc.		.145	.757
Pitch sensitivity		.33	1.79
PM capacity		.423	2.72**

Notes: $R^2 = .261$; ** $p < .01$

Table 19 presents the regression analysis for gains made in categorizing words by their accent pattern. The same four predictors as in the previous analysis were entered in four steps into the model. First we see that the addition of PM in Step 2 did not account for a significant amount of variation in the model. However, by adding nonword

discrimination ($\beta = .365$; $p < .05$), which was correlated with PitchCAT gains, I was able to explain an additional 12.3% ($p < .05$) of the variation in learners' development. The final step of the model includes acoustic pitch sensitivity ($\beta = -.378$; $p < .05$), and yielded a further 7.9% ($p < .10$) of predictive power to the analysis. That the two acoustic sensitivity measures accounted for a significant amount of variation in PitchCAT gains suggests that basic auditory capacities are tied to learners' ability to assign category labels to the spoken stimuli. Altogether, the four predictors in the regression analysis for PitchCAT explained 27.9% ($R^2 = .279$; $p < .05$) of the variation in learners' development of the ability to categorize lexical accents by pattern.

Table 19. Hierarchical multiple regression analysis for PitchCAT gain

Predictors	ΔR^2	β	t
Step 1	--		
Learning context		.095	.556
Step 2	.069		
Learning context		.117	.696
PM capacity		.264	1.57
Step 3	.123*		
Learning context		.022	.136
PM capacity		.212	1.32
Nonword disc.		.365	2.21*
Step 4	.079 [†]		
Learning context		.073	.456
PM capacity		.211	1.36
Nonword disc.		.177	.937
Pitch sensitivity		-.378	-2.31*

Notes: $R^2 = .279$; * $p < .05$; [†] $p < .10$

5.6 Summary of results

Looking at the overall results of Experiment 3, I first noted that beginning-level Japanese learners had difficulty perceiving lexical pitch accents, as shown in their low accuracy scores on both perception tasks. Yet, the JFL group, who received explicit correction in the classroom, did perform slightly better than the JSL group. My main interest, however, was in the development of perception ability, and regarding this I found that both groups made slight gains in accent perception. Most critically though were the findings that both phonological short-term memory capacity and acoustic pitch sensitivity were significant predictors of perception accuracy gains over the 10-week testing period, albeit on different perception tasks. In the next section I will discuss these results in the context of previous findings on individual differences in L2 perceptual development.

5.7 Discussion

In Experiment 3, I examined how learner-internal cognitive resources related to the acquisition of lexical pitch accent in two groups of English-speaking beginning Japanese learners. I will address each of the three predictions below. I first predicted that instruction style rather than learning context will result in greater perceptual accuracy for lexical accents. Previous literature has shown that accent perception accuracy by Japanese learners in both at-home and study abroad settings is low (e.g., Nishinuma et al., 1996; Shibata & Hurtig, 2008), but the two contexts have never been compared in the same study.

The results showed that the JFL group, who received explicit correction on lexical accent as part of their Japanese language instruction, perceived lexical accents at slightly higher accuracy than the JSL group, who lacked such feedback. Although the JSL group had increased exposure to spoken input while studying in Japan, the JFL learners judged the correctness of and categorized the target stimuli with greater accuracy at both test times. This finding indicates that explicit feedback on lexical accent may be more beneficial than exposure alone in the reportedly higher-input study-abroad context. Yet, both groups' accuracy improved to a similar degree over the 10-week period, suggesting that adult learners' perceptual acuity for L2 phonological categories remains pliable. The overall low accuracy underscores previous findings which show that lexical accent is a particularly difficult feature for beginners to internalize as a lexical feature, due to its low acoustic salience, the need for targeted training, etc.

The next two predictions can be grouped together under the assumption that differences in learner-internal abilities manifest as differences in acquisition outcomes for pitch accent perception. Namely, those with a greater capacity to retain nonwords in phonological memory and those who are sensitive to subtle differences in nonlinguistic pitch are better equipped, at least initially, to acquire pitch accents from spoken input. Two regression analyses explained these predictions regarding learner-internal resources. First, PM capacity significantly predicted accuracy gains on the correctness judgment task, accounting for 17.6% of the unique variance in perceptual development. Learners with a higher capacity to temporarily store and process Japanese mora-based nonwords on a serial nonword recognition task showed greater development over the semester in

their ability to acquire pitch as a part of lexical form. Previous research on both verbal and phonological memory and SLA suggests that high PM span learners are better at the consolidation of L2 phonological form, thus the relation of PM to the correctness judgment task in the current study (e.g., Hummel, 2009; Martin & Ellis 2012; Sunderman & Kroll, 2008). In fact, PM has been linked to a range of areas of L2 acquisition. Hummel (2009) found that differences in PM capacity accounted for 13% of the variation in L2 English learners' proficiency. Martin & Ellis (2012) also reported that PM capacity explained 14% of the variance in learners' ability to acquire the vocabulary of an artificial language, which replicated the task of L2 vocabulary learning. In order to accurately perform the correctness judgment in the present experiment, a learner had to compare the spoken input to their own representation of the word in their long-term store. Thus, it appears that high-PM learners were better at internalizing spoken input, whether from spoken input alone in the study abroad context, or from the combination of classroom input and correction in the case of the JFL learners, thus enabling more accurate lexical judgments. The present result extends findings showing that working memory and phonological short-term store capacity relate to L2 lexical acquisition to include lexical prosodic features in a pitch accent language.

Second, basic acoustic perception ability, as measured by the adaptive pitch test and nonword discrimination task, predicted gains made on the PitchCAT task over the semester. In fact, the scores on the two tests combined to explain roughly 20% of learners' categorization development. That is, learners who were highly sensitive to subtle, nonlinguistic pitch contrasts made greater gains at assigning visual accent

category labels to spoken Japanese words. Despite a widely-assumed dissociation between speech and nonspeech perception (Strange & Shafer, 2008), I found that a non-language capacity predicted gains made on a language task involving accent decisions. This can potentially be attributed to 1) the inherent difficulty in the pitch accent cues which made the categorization task perceptually demanding, and 2) the similarity between the bitonal (L-H/H-L) pitch variations in Japanese words and the tones featured in the adaptive pitch and nonword discrimination tasks. In sum, this pattern of findings suggests that basic cognitive resources may in fact play an important supportive role in low-proficiency adult second language perceptual learning. The next chapter will consider together the findings from each of the three experiments, then provide pedagogical strategies in light of the findings that basic cognitive resources support the development of accent perception ability.

Chapter 6: General Discussion and Conclusions

6.1 Summary of results from Experiments 1-3

The experiments reported in the previous three chapters were designed to investigate whether a combination of learner-internal capacities and experience-based factors could explain the previously-reported variation in the perception Japanese lexical pitch accent. The over-arching goal of this research was to identify and measure the degree to which individual variation in these sets of factors could account for variation in natural speech perception tasks. Three listener groups were tested. Experiment 1 focused on native Japanese speakers in order to not only establish a baseline of accent perception ability, but also to identify the role of internal cognitive resources on L1 speech perception. Experiment 2 employed a similar design and looked at accent perception in advanced learners of Japanese in two language groups in which tone status differs phonemically, namely, Chinese and Korean. In Experiment 3, the focus was shifted to the development of accent perception ability over a semester of study in beginning Japanese learners from an L1 English background.

Previous research, along with direct observation of learners, has revealed that the ability to accurately perceive Japanese lexical accent is 1) particularly resistant to development in line with overall proficiency level; 2) often unchanged even following study abroad; 3) difficult to perceive regardless of the status of tone in a learner's L1; 4)

marked by a wide degree of individual variation. I addressed these previous findings by first asking if differences in learner-internal resources, phonological short-term memory and acoustic sensitivity, and experience-based factors, L2 lexical knowledge and L1 tone experience, contributed to the perception and acquisition of lexical accent in Japanese. Furthermore, I considered perception task type and learning context as additional factors that influence the complex process of L2 speech perception.

Stated broadly over the three studies, the chief findings included 1) a decline in the ability to judge the correctness of spoken accents across the three experimental groups: Native speakers, advanced and beginning-proficiency Japanese learners 2) the involvement of experience-based factors, L1 tone experience and L2 vocabulary knowledge, at the advanced proficiency level, and 3) the supportive role of PM and acoustic pitch sensitivity in accent perception in beginning-level Japanese learners. Cross-linguistic speech perception models have generally focused on the effect that differences in segmental-level acoustic properties have on L2 acquisition, leaving speech prosody under-addressed (Best & Tyler, 2008; Hao, 2012). Moreover, phonemic comparisons have failed to capture the variation resulting from individual differences that learners bring to the task of speech perception, and have also neglected external variables like learning context and instruction style. In the current research I attempted to systematically investigate the role of learner-variables on the perception of Japanese lexical prosody. My aim was to characterize how multiple factors, including those which vary by individual and are not necessarily language-specific, like PM and acoustic sensitivity, and others which are obtained through exposure to language, related to speech

perception. I will next discuss each of the perception-related factors investigated in the experiments—moving from a consideration of perception task type, to an analysis of cognitive resources, then on to a look at experience-based factors. Finally, I will discuss suggestions for how these findings can be implemented in the Japanese language classroom.

6.2 Performance by perception task type

Two accent perception tasks were used in each of the three experiments. The first task (PitchID) required listeners to judge the correctness of the pitch accent of a sentence-embedded spoken stimulus, and the second involved categorizing these words (PitchCAT) by their accent on the basis of visual pitch patterns. I will first consider the effect of perception task type in each of the three listener groups—L1 Japanese speakers, advanced learners, and beginning learners of Japanese.

A substantial literature exists on the effects of perception task type on perception accuracy. Tasks that require phonetic discrimination of sound contrasts are generally reported to be performed accurately even by target-language naïve listeners who have no comparable feature in their L1 (e.g., Bent et al., 2006; Hao, 2012; Wang, 2013). On the other hand, categorization tasks that require listeners to group sounds based on their goodness-of-fit with a target language category frequently yield very different results between experienced and inexperienced listeners (e.g., So & Best, 2010). Speech perception models have been constructed which make explicit predictions on the relative ease or difficulty in the perception of a target L2 feature, depending on its status in the L1 (Best & Tyler, 2008; Kuhl & Iverson, 1995). For example, the /r/ and /l/ phonemes in

English share one phonetic category in Japanese [r] (alveolar flap), and inexperienced L2 English listeners from an L1 Japanese background often label these as belonging to a single category (Aoyama et al., 2004).

Although not used as a test of categorical perception, correctness judgment tasks, akin to lexical decision tasks used in psycholinguistic research, are another way to probe listeners' knowledge of lexical form. In terms of accent perception, correctness judgment tasks can potentially reveal whether a listener has established an accurate representation in long-term memory of the accent pattern as a part of a word's phonological form. Correctness judgment tasks can therefore serve as an insightful probe of the mental representation of a word's form, as lexical knowledge is crucial to perform this task. Categorization as used in the current study, on the other hand, may invoke psychoacoustic perception to an extent, in that matching accent patterns to visual representations could in theory be accomplished by language-naïve listeners who are highly sensitive to pitch variations, and who are also adept at using visual notation to assign the perceived word to a category label (i.e., those with musical training). Used in combination, these tasks enabled me to first test if a listener had lexicalized a word's accent pattern in their long-term word store, and then to see how they assigned category labels to the word's pitch contour. All three experiments used sentence-embedded real words as stimuli, simulating the task that listeners encounter when listening to spoken Japanese.

6.2.1 Correctness judgment task (*PitchID*)

In Experiment 1, native Japanese speakers were accurate ($M = 93\%$, $SD = 1.63$ (5.1%)) at determining whether naturally produced, high frequency noun-plus-postposition targets were spoken with the correct or incorrect pitch accent. This finding corroborates with previous perception research indicating that pitch and tone are accessed simultaneously with word form by native listeners, rather than processed sequential to segmental information (e.g., Chen, Chen & Dell, 2002; Otake & Cutler, 1999). Turning to the advanced learners in Experiment 2, I found that native Chinese speakers made correctness judgments with an average accuracy of 77% (*range*: 65% - 88%, $SD = 3.03$ (6.3%)), while proficiency-matched L1 Korean speakers did so with a mean of 59% (*range*: 44% - 90%, $SD = 5.49$ (11.4%)). This large disparity between L1 groups on accent perception will be discussed in detail later. In Experiment 3, L1 English learners of Japanese in their second year of study (both JSL and JFL) perceived accents at a mean accuracy of 55% in a second testing session at the end of a semester of language learning (*range*: 44% - 69%, $SD = 2.25$ (6.3%)), little above a chance level of 50%. A previous perception study reported a 47% average with novice-level L1 English speakers on a similar accent correctness-judgment task (Shibata & Hurtig, 2008), affirming the difficulty learners encounter in this type of task. However, the present experiments were primarily concerned with accounting for the variation in both perception performance and perceptual development (in the case of Experiment 3), which was observed in the learner-variable analyses. In general, I noted a proficiency-related decline across groups; yet when considering that the advanced-proficiency L1 Korean group only performed at 59%

accuracy, a strong proficiency-based account does not hold up well. Although L1 Korean speakers had attained a level of Japanese directly comparable to the L1 Chinese speakers, their ability to perceive lexical accent noticeably lagged behind this latter group. This suggests, along with the observed range in individual scores, grounds for exploring other possible sources of performance variation on this task.

6.2.3 Categorization task (PitchCAT)

The categorization task was an attempt to see how listeners grouped naturally-produced stimuli based on visual representations of the pitch fall location. Classic categorization tasks typically manipulate a phonetic parameter by set intervals, and have participants classify the stimuli based on perceived representativeness of a target category (Logan & Pruitt, 1995). The current task differed from this paradigm in that listeners assigned visual labels to acoustically unaltered stimuli. The task specifically required listeners to identify the location of a pitch fall within a word, or the lack thereof, by means of schematized accent contours. Although there is no standard notation system for accent in Japanese, several studies have incorporated similar representations of pitch accent in the testing of L2 perception (e.g., Hirata, 1999; Shport, 2011).

Categorization accuracy was generally low in all groups, reflecting overall task difficulty and the presence of four response choices, as opposed to two in the PitchID task. Note that Experiment 3 featured a simplified three-choice paradigm for beginning learners, but nonetheless yield similar scores. Although for this reason it is difficult to draw conclusions based on a comparison of correctness judgment with categorization, I can compare the groups on their categorization performance. In Experiment 1, native

speakers of Japanese categorized the noun-plus-postposition targets with 58% accuracy (*range*: 19% - 94%, *SD* = 3.39 (21%)). By comparison, the averages for the advanced learner groups were 60% for L1 Chinese (*range*: 25% - 100%, *SD* = 4.93 (21%)) and 49% for L1 Koreans (*range*: 21% - 88%, *SD* = 3.48 (14.5%)). I can explain the fact that L1 Japanese and advanced L2 learners performed similarly on this task in a couple of possible ways: 1) Unfamiliarity of task type to L1 Japanese listeners, but relatively more familiar to L2 learners by way of accent notations found in language textbooks; 2) Pitch patterns are not represented separately of the word in L1 speakers' knowledge of lexical form, making the abstraction of pitch patterns into discrete categories difficult. All participants reported understanding the visual notation used to represent pitch patterns (see Figure 4, Chapter 3) in the practice phase prior to the task, which required associating schematized mora (as circles) with high-low accent contours. A variety of labelling schemes have been used in previous research, and the lack of a standardized labelling system likely added to task difficulty (see Sakamoto, 2010 for a sample of these systems). Perhaps additional practice with the labelling scheme would have boosted performance, at least in the L1 listener group. As Strange (2011) has pointed out, factors such as task demands, instructions and stimulus structure may partly determine what kinds of perceptual information listeners attend to and base their judgments on (p. 457). With L1 Japanese listeners in Experiment 1, for instance, a possible alternative labelling scheme would have been to use Chinese characters to represent High (高) and Low (低) pitch mora. This approach might have been more intuitive for native listeners due to their familiarity with characters. Yet, the locus of the task—assigning H and L pitch to

individual mora—would remain the same, leading me to doubt the possibility of a marked performance increase resulting from a change in labelling schemes. The central point here is that because the accent nucleus is tied to the mora at the lexical level, it is important to use a system that represents individual mora, as was done with the circles employed in the current experiments (i.e., Chapter 3, Fig. 4).

Assuming that listeners *were* familiar with the labelling system, then an alternate explanation for the difficulty is required. We know from previous research that lexical accents in both Chinese and Japanese are crucially linked to segmental form in the lexicon (Chen et al, 2002; Otake & Cutler, 1999). In light of this fact, it is therefore possible that assigning an abstract visual label to a pitch category is universally difficult for L1 Japanese speakers. In fact, in previous research on which L1 speakers were required to label accent location, accuracy was markedly low even among native Japanese speakers ($M = 58\%$) (e.g., Shport, 2008). Furthermore, the acoustic phenomenon of peak delay, in which the pitch accent peak occurs slightly later than the mora to which it is attached, may also make visual categorization problematic for the accented patterns (Shport, 2011). Correctness judgments, on the other hand, necessarily required listeners to take into account the entire word when making their decision, and thus yielded high accuracy among native listeners. The chief contribution of the categorization task, as we shall see in a later section, was in the identification of which accent patterns were more difficult to classify and how the categories were confused with one another.

6.3 Cognitive resources and accent perception

Two domain-general cognitive resources, acoustic pitch sensitivity and phonological short-term memory, were measured as predictors of accent perception ability in all three experiments. It has been argued that basic processing resources, which all humans possess albeit in varying capacities, support acquisition at early stages of both L1 and L2 learning (Perrachione et al., 2011; Speciale et al., 2004; Strange, 2011). As language familiarity increases and learners establish long-term representations of the language's phonological patterns, these effects likely diminish or disappear entirely (Masoura & Gathercole, 2005; O'Brien et al., 2007). Although adults possess a highly-routinized sensitivity to language-relevant sound contrasts in their L1, the findings from the present experiments indicate that domain-general cognitive capacities still hold sway to an extent on Japanese lexical accent perception. Next, I will discuss the present findings from this perspective on language acquisition.

6.3.1 Acoustic pitch sensitivity

Acoustic sensitivity, also referred to as psychoacoustic perception, has been widely investigated for its relation to the perception and acquisition of suprasegmental features (e.g., Deutsch et al., 2009; Hao, 2012; Wayland, et al., 2010; Wong & Perrachione, 2007). At the core of these studies is the assumption that basic perceptual resources which support the perception of nonlinguistic melody and pitch may also be active in the perception or learning of tonal features of language. A great deal of evidence has shown that regardless of how prosody is realized in one's native language, listeners can, at a minimum, perceive acoustic differences in a variety of languages on par with

native listeners (Eda, 2004; So & Best, 2010). Discrimination tasks have shown that target-language-naïve listeners can acoustically distinguish whether a pair of Mandarin Chinese tone contrasts is the same or different with relatively high accuracy when tested in ideal listening conditions, such as a lab environment (Bent et al., 2006). Parallel results have been found for discriminating Japanese pitch accent contrasts (Shport, 2011). In fact, beginning Japanese learners in the present Experiment 3 were able to make same-different judgments on accented nonwords at an average accuracy of 95%. Yet as we saw in the previous section, when required to make judgments on linguistically-relevant features, a clear separation emerges between L1 and L2. This line of evidence would therefore suggest that basic auditory perceptual resources play little role in the perception of categorically-meaningful prosodic contrasts.

But yet in stark contrast to this reasoning, in Experiment 1 acoustic sensitivity was predictive of accent perception by native Japanese listeners. A regression analysis confirmed that basic acoustic sensitivity accounted for 27% of the performance variation in the correctness judgment task and 25% in the categorization task. Shouldn't basic psychoacoustic perception be unrelated to L1 perception, which is thought to be highly routinized and robust? An examination of the measure of acoustic sensitivity, the adaptive pitch test, may be insightful here. Looking at the basic perception thresholds, indexed by the smallest difference in Hz at which a listener could reliably distinguish two tones, all listener groups except the L1 Chinese ($M = 22$ Hz) performed in the sub-10Hz range. This indicates, for example, that listeners could accurately differentiate a 400 Hz reference tone from a 410 Hz tone. The L1 Japanese speakers discriminated pure tone

pairs at 5.5 Hz, while the L1 Korean listeners did so at 9.5 Hz. The L1 English listeners, with the least experience with spoken Japanese among the groups tested, discriminated the pairs at an average of 7.3 Hz. These numbers are similar to those in other studies of acoustic perception. For example, Wayland et al. (2010) found that at a 7 Hz frequency interval, both musician and non-musician participants discriminated pure tones with around 70% accuracy. Listeners' accuracy dropped sharply when the pitch difference fell below a 4 Hz interval. Bent et al. (2006) reported that their participants' acoustic thresholds were 3.5 Hz for L1 English speakers and 2.6 Hz for L1 Chinese.

In the current study, the adaptive pitch test measured the threshold at which listeners could reliably distinguish a pair of pure tones, with the sole acoustic cue being F0 variation. Compare this with the typical AX discrimination tasks used in some previous studies (e.g., Bent et al., 2006; Sakamoto, 2010; Shport, 2011), which were similar to the nonword discrimination task in Experiment 3. Both native and nonnative listeners performed with high accuracy on these tasks, with native listeners typically achieving ceiling-level accuracy. One possible account for the finding that pitch sensitivity is predictive of even L1 listener's accuracy on accent perception tasks is that the adaptive pitch test used in the current study provided a more precise measure of listeners' acoustic sensitivity threshold. On an AX discrimination task featuring manipulations of acoustic parameters of word-like stimuli, L1 listeners are likely to hit a performance ceiling, whereas there was no such limit on the adaptive pitch test. This in turn yielded greater distribution of perception scores which were correlated with language-based perception tasks.

Acoustic perception was found to have no relation to lexical perception in the advanced learners, while it was predictive with the beginner group. With the low-proficiency L1 English group, the regression model including both measures of acoustic sensitivity accounted for roughly 20% of the unique variance in categorization task development (as shown in the ΔR^2 column in Table 19), although these predictors were not significant in the correctness judgment task. If we consider the categorization task to be the less lexical of the two tasks—in that categorizations can be made by comparing the spoken pitch contour to the musical notation-like visual representations—then this finding is reasonable.

Disregarding the L1 results for the moment, when looking at just the L2 learner groups in Experiments 2 and 3, this pattern of results suggests a previously noted trend for the perception of prosody to develop on a phonetic-to-lexical, or bottom-up, continuum (Norris et al., 2000; Wong & Perrachione, 2007). The implication is that low-level learners make use of domain-general perceptual resources when perceiving L2 suprasegmentals, prior to the establishment of lexical or so-called categorical perception. The finding that vocabulary size predicted perception accuracy in the two advanced-proficiency groups supports such a developmental pathway.

The fact that acoustic perception remained active in L1 Japanese can be reconciled by the earlier point on the fine-grained measurability of the adaptive pitch test. However, explaining the lack of a relationship in advanced L2 listeners in the same terms is difficult to resolve, unless separate accounts are provided for L1 and L2 speech perception mechanisms. Further, why did the Chinese listener group, who use tone

extensively in their L1, perform significantly *worse* on the adaptive pitch test than all other groups? Future work is necessary to resolve these findings, although my central claim still holds—basic acoustic sensitivity is active to a significant degree in the perception of lexical pitch accent.

6.3.2 Phonological memory and accent acquisition

Memory resources have been closely linked to the acquisition of language. Considering that one important task of a language learner is to establish a connection between meaning and phonological form, which is derived from spoken input, the involvement of memory follows logically from this. In fact, the phonological loop component of the WM model has been referred to as a ‘language learning device’ (e.g., Baddeley et al., 1998; Ellis, 1996; Gathercole, 2009). This has been reinforced by research findings from multiple areas of L2 lexical acquisition (e.g., Cheung, 1996; Hummel, 2009; Masoura & Gathercole, 2005; O’Brien et al., 2006, 2007). Language learners who possess a greater capacity to process and store spoken input in the phonological loop appear to be at an advantage in their acquisition of a word’s phonological form. The current research argued that the function of phonological memory capacity on lexical acquisition extends to the learning of prosodic form as well, as in the case of Japanese lexical pitch accent. Furthermore, I assume that PM’s primary role among L2 learners can be seen in the vocabulary acquisition *process* (Gathercole, 2009; O’Brien et al, 2007), rather than in one-time task performance.

This prediction was tested explicitly in Experiment 3. The results suggested that beginning learners are relying to some extent on their short-term phonological store in the

acquisition of lexical accent. Specifically, beginning Japanese learners' ability to judge the correctness of pitch accents was predicted by their capacity to retain nonwords based on the target language's phonotactics. Importantly, a regression model showed that PM accounted for 19% of the *development* of perception ability over a semester of Japanese study, rather than their performance at a given time. It is PM's role in language development in which I am chiefly interested, as the results align with the previous findings on PM and language acquisition (Gathercole, 2009; O'Brien et al., 2007). Although working and phonological memory are no doubt active during the processing of sound input at any given time, it was found that its predictive role in language acquisition was centered on the gains made in accent perception. By comparison, in Experiments 1 and 2, this capacity was not related to accent perception on any of the perception tasks presented at a single test session.

Additionally, PM's function as a support mechanism for language learning has been generally confined to early-stage language learning, when long-term representations of lexical form are not well-established, either in L1 and L2 acquisition. My findings in Experiment 3 echo this pattern, in that low-proficiency Japanese learners' performance on the serial nonword recognition (SNWR) task predicted their *gain* scores over the semester. In a study on L2 English learners from an L1 Chinese background, Cheung (1996) found that nonword repetition ability predicted lexical gains in the low proficiency group only. O'Brien et al. (2007) also observed the influence of PM capacity on spoken fluency in early-stage learners of Spanish. As Speciale et al. (2004) argue, PM is likely an active constraint on L2 vocabulary learning at the early stages of acquisition, but

diminishes as learners accrue long-term knowledge of the phonological structures of the language.

Returning to the current findings, PM capacity was active in learners in both the at-home and study-abroad contexts. Learners who can extract pitch information from the input could also store this as a part of lexical form, enabling larger performance gains on the correctness judgment task. If a language learner's task is to form a long-term representation of a word in memory, then apparently those who are more efficient in general at processing L2-like sounds, in the form of nonwords in the current study, are better at doing so with real words as well. Hence the reason for their greater development in recognizing correct and incorrect instances of spoken words. As mentioned in an earlier section, the crux of the correctness judgment task is the comparison of the spoken input with the word's representation in long-term memory. The categorization task, on the other hand, seemed to be more akin to phonetic categorization, and was not related to PM capacity in any of the experiments.

What role does the specific PM measure used in the current study have in the findings? If we consider the serial nonword recognition (SNWR) task to be an index of how well learners can “construct well-defined phonological representations from the acoustic input” (Engel de Abreu & Gathercole, 2012; p. 982), we have a potential account for why higher-PM learners in the early stages of Japanese acquisition improved more at judging the correctness of lexical accents. The phonological loop appears to be crucial to the learning of the phonological form of words. Baddeley et al. (1998) posited that “[t]he phonological loop appears to provide a critical input to the construction of the

more permanent phonological structures that are stored in the mental lexicon” (p. 163). The current results show that it is also active in the learning of accent patterns for words whose segmental form is already known. Namely, beginning Japanese learners who can efficiently and accurately process L2-like phonological strings are also better at extracting pitch information from spoken input. Memory for novel phonological strings in a sense mimics the process of learning vocabulary in a second language (Ellis, 1996; Martin & Ellis, 2012; Williams & Lovatt, 2003). I extend current implications on the role of PM to include the acquisition of lexical-level prosody, which is an inherent phonological property of the Japanese lexicon.

Lastly, when examining performance on the PM task by the listeners in all three experiments, I unexpectedly found that the group means were highly similar. Even though the task used nonwords constructed from actual Japanese mora and featured a L(ow)-H(igh) accent pattern, the results showed that L1 listeners were not at an advantage on this task. In fact, their average performance was nearly identical to that of the beginning-proficiency Japanese learners. This finding contradicts the assumption of many researchers (e.g., Gathercole, Willis, Emslie & Baddeley, 1991; Kaushanskaya & Yoo, 2013; Speciale et al., 2004) that as one’s long-term lexical store increases, so will the capacity to store nonwords in the phonological loop. As such, L1 Japanese listeners, with their robust lexical knowledge, should have clearly outperformed beginning learners. The performance variation on the SNWR task, it seems, was found among individuals *within* each group, in which some learners, regardless of their experience with Japanese, showed a much greater ability to process nonword phonology in PM. It appears

then that the SNWR task tapped a more domain-general, rather than language-specific mode of perception, whereby L1 Japanese speakers held no advantage over nonnative listeners. This pattern of results needs further empirical elaboration.

6.4 Experience-based factors in accent perception

Along with the cognitive factors just examined, experience with lexical-level prosody in one's native language and its influence on perception—as well as learners' interaction with the L2 during the acquisition process—might mold one's perceptual system in significant ways. Therefore, a second focus of the present research was to consider the potential for influence which language learning experience, including experience with the L1, has on Japanese lexical accent perception.

6.4.1 L1 background

The powerful effect that one's native language has on the acquisition of subsequent languages has been demonstrated repeatedly. This influence manifests in a wide range of domains including syntax, lexicon, and morpho-phonology (e.g., So & Best, 2010; Ueyama, 2000). In the second experiment, I compared two L2 learner groups from L1 backgrounds in which the phonemic status of tone differs. Namely, Mandarin Chinese extensively uses four tonal patterns to distinguish homophones, whereas no such role of lexical-level pitch variation is present in standard Korean. Despite the L1 Chinese group being the least sensitive to non-linguistic pitch contrasts on the adaptive pitch test, these listeners were the most accurate at making decisions on lexically-relevant spoken accents, performing significantly better on both perception tasks. In fact, native language was by far the most significant predictor of perception accuracy—despite both language

groups being matched for proficiency—accounting for approximately 50% of the variance in performance on the regression model in Experiment 2.

It has been noted, however, that just by virtue of speaking a tone language as one's L1 does not necessarily enhance the perception of non-native tones (So & Best, 2010; Wang, 2013). Cross-linguistic speech perception models such as PAM-L2 were devised to account for the relative difficulty of the perceptual assimilation of non-native speech sounds (Best & Tyler, 2008). For example, L1 Cantonese speakers, whose language uses tone phonemically, displayed difficulty in categorizing Mandarin tones which were similar, but not identical to their native tone contours (Hao, 2012). Arguing from the predictions of PAM-L2, the authors concluded that Cantonese speakers were assigning two Mandarin tones (Tone 1 and Tone 4) to a single category in their native language, hence the difficulty with these tones.²⁰ However, this study as well as many others used target-language naïve listeners in a replication of what is essentially learning a language from zero (Strange, 2011). Yet it is reasonable to expect that those with some degree of target language experience can bring this knowledge to bear on L2 prosodic acquisition. In the context of the current study, L1 Chinese speakers likely found Japanese pitch accents to be highly salient features of word form, as is tone in their own L1, thus their attention to these nonnative pitch categories was heightened. Furthermore, Japanese language education in China, by virtue of the ubiquity of tone in their own L1,

²⁰ Hao (2012)'s finding echoes the often-reported segmental assimilation of English /r/ and /l/ by L1 Japanese speakers.

likely places greater emphasis on the instruction of lexical accent patterns in the Japanese classroom than is done in non-tonal language environments such as the US.²¹

6.4.2.1 Pattern analysis of accent categorization

A discussion of the role of L1 prosody on L2 speech perception would be incomplete without considering listeners' specific patterns of perception errors. The categorization task required learners to use visual representations of pitch contours to classify accents into four categories (or three in Experiment 3). In Experiment 2, it was revealed that the two groups of advanced Japanese learners differed in their accuracy of lexical categorization. Some of this variation can be explained in terms of L1 perceptual biases, while some is best characterized as stimulus-related difficulty. That is, certain accent patterns may be globally more difficult than others for a number of reasons, as I will describe next.

Certain tones are more difficult than others due to their acoustic properties. The Mandarin Chinese tones 2 and 3 have been noted to be difficult for L2 listeners to perceive regardless of their experience with the language or L1 tone background, due to the phonetic similarity of these tones (Hao, 2012; So & Best, 2010; Wang, 2013). In Japanese, the acoustic cue to pitch fall is a sharp drop in F0 over two adjacent mora, so rather than the pitch *contour* being meaningful, it is the location of this drop that is vital. In the case of unaccented words, termed "Pattern 4" in the current study, the lack of a pitch fall was the crucial cue to identifying this pattern on the perception tasks. The general difficulty faced by L2 Japanese learners has been cited as stemming from 1) their

²¹ Personal communication (X.J. Chu).

inability to identify the location of the pitch fall, and 2) in the case of unaccented words (i.e., Pattern 4), recognizing the lack of such a drop (Shport, 2011). Previous research on L2 Japanese perception has produced mixed findings regarding this point, with some studies reporting a perceptual preference for the unaccented pattern (e.g., Arai, 1997; Nishinuma et al., 1996; Sakamoto, 2010), while others higher accuracy on the accented patterns (e.g., Shport, 2011; Toda, 2001).

Are some accent patterns more easily categorized than others? The categorization task results enabled me to see how listeners from different native languages assigned labels to the accent patterns of frequent nouns. Looking at the data from Experiments 1 and 2, I noted that Pattern 1 words, in which the High pitch is on the first mora and falls on the second mora, were the most accurately categorized. Native Japanese listeners categorized these words (HLL) with a 72.5% average accuracy. Correct responses decreased markedly on the remaining three patterns, in the order of Pattern 4, Pattern 2 and Pattern 3. As noted earlier, Pattern 3 words require the listener to take into account the following case particle in making their categorization. Specifically, whether the pitch falls just prior to the particle, or whether the high pitch extends onto the particle, in which case the correct category would be the unaccented Pattern 4. Shifting the focus to Experiment 2, L1 Chinese speakers showed an acuity similar to L1 listeners for categorizing Pattern 1 words at 69.4%. However, their categorization of Pattern 4 items was also accurate, at 70%. A decline in accuracy was apparent on Pattern 2 words (59%), with Pattern 3 being the least accurate (42.5%). L1 Korean listeners shared the preference for Pattern 1 categorizations (72.8%), but unlike the other two listener groups, their

accuracy declined along the numerical order of the patterns, with the unaccented pattern being the least accurate.

The heightened accuracy on Pattern 1 words by all three listener groups can be explained by the comparatively high perceptual salience of the initial high-pitched accented words (Strange, 2011). Minematsu & Hirose (1995) found that L1 Japanese speakers required the least amount of acoustic information to identify Pattern 1 words in a gating task, in which a word's phonetic information was revealed in short intervals. Phonetic analyses of pitch height in lexical accents have shown that the F0 peak is highest on initial-mora accented words, with the other accented patterns being produced with a slightly reduced pitch range than initial-mora accents (Beckman & Pierrehumbert, 1986). Thus, the slope of the pitch fall from the first to second mora is steepest on Pattern 1 words (refer to Chapter 1, Fig. 1), resulting in its pitch cue likely being more audible than the other patterns. Among the accented words, the early pitch fall on Pattern 1 words appears to make it the most acoustically prominent and thus most easily categorized.

As noted earlier, the presence of a case particle forms an important grouping in Japanese accentual phrases, with the unaccented particles assimilating into the pitch contour of preceding unaccented (Pattern 4) words (Vance, 2008). The fact that Pattern 3 words were often confused with those of Pattern 4 by both groups in Experiment 2 suggests that learners are either conflating these two patterns into one, or ignoring the particle altogether. In fact, native Japanese speakers also had the most difficulty in both judging and categorizing Pattern 3 items. However, it is a stretch to say that for L1

Japanese speakers, they are failing to account for the case particle's role in accentual grouping, necessitating an alternative account for this difficulty.

A final and underexplored factor for consideration in the analysis of individual accent patterns is the potential effect of accent pattern frequency on perception. Frequency effects on lexical perception have gained wide attention in psycholinguistic literature, with the general, and commonsense, observation being that frequent words are more accurately and rapidly retrieved than infrequent ones (e.g., Miwa et al., 2013; Ueno et al. 2014). However, the relative frequency of individual accent patterns themselves on speech perception accuracy has only begun to gain attention in Japanese-focused research. Ueno et al. (2014) found that L1 Japanese participants displayed lower accuracy and slower reaction times in identifying what the authors considered an 'atypical' accent pattern (Pattern 2, or the LHL pattern). In fact, these L1 listeners' errors in a word repetition task tended to be such that they repeated atypically (but correctly) accented words *incorrectly* by changing these to a more frequent pattern (i.e., repeating Pattern 2 words with a Pattern 1 accent). In the current study, lexical frequency was controlled across all accent types, so we can assume that the words in each pattern category were encountered in daily life by our participants with relatively equal rate. Yet, Tanaka & Kubozono (2012), among others, have shown that pitch accent patterns themselves can diverge widely in their commonness, depending on a word's length. For the words used in the current study the order of frequency for 3-mora words is Pattern 4, Pattern 1, Pattern 2, Pattern 3; and for 4-mora words is Pattern 4, Pattern 2, Pattern 1, Pattern 3. Thus, in the current study the most 'atypical' accent pattern would be the problematic

Pattern 3. How do listeners' results compare to this frequency order? In general, it appeared that the higher-frequency Pattern 1 and 4 words yielded noticeably higher accuracy rates than the mid-accented Pattern 2 and 3 words. The current experiments were not designed specifically to test the effects of accent pattern frequency, but my data suggest the possibility of pattern frequency or typicality, similar to what Ueno et al. (2014) found, as one constraint on accent perception. Pattern frequency is a rich area for follow-up study in L2 accent perception and production.

6.4.2 L2 lexical knowledge

Among the advanced learners in Experiment 2, Japanese lexical knowledge accounted for a significant amount of variation in accent perception ability. Considering that grammatical knowledge, measured together with vocabulary as an index of L2 proficiency, was not related to accent perception, I can rule out a general effect of proficiency on the present findings. Furthermore, the two groups were controlled for length of residence in Japan, length of study, and frequency of daily Japanese usage, none of which was significantly linked to accent perception. The stimuli used in the perception tasks were all frequent vocabulary items selected from beginner-level word lists; words that highly-proficient L2 speakers can rapidly and accurately recognize. Hence, the L2 vocabulary measure was not a measure of whether learners knew the individual stimuli, but an overall measure of Japanese lexicon size.

Why would such a general measure of L2 lexical knowledge be predictive of perception of a specific lexical feature like pitch accent? Martin & Ellis (2012) argue that grammar acquisition is underpinned by the creation of a “storehouse” of words and

phrases (i.e., the mental lexicon) from which grammatical rules are generalized (p. 382). Long-term knowledge of L2 phonological regularities is also implicated in developing automaticity in lexical processing and subsequent vocabulary learning (Masoura & Gathercole, 2005; Speciale et al., 2004; Williams & Lovatt, 2003). Furthermore, Ueno et al. (2014), discussed in the previous section, noted the contribution of a ‘typicality’ effect on accent perception and production, suggesting a direct link between long-term knowledge of the distributional properties of accent and accent perception. Generalizing the assumption of lexical knowledge as a support for language perception to accent task performance, I can posit that the long-term store of L2 vocabulary provides a source of exemplars against which listeners can compare spoken input, and thus base their lexical decisions upon. The logic here is that learners with a larger L2 lexicon also have a larger phonological store of lexical pitch categories, in the sense that they possess more tokens of each accent category. This observation highlights the importance of learning pitch accent as an inherent lexical property, linked to a word’s segmental structure. In sum, individual differences in cognitive capacities, such as PM, are active in low-proficiency learners, but this gives way to the long-term store as the central determinant of further lexical learning. Experiment 2 provides an argument that the same can be said for the acquisition of Japanese accent categories as well.

6.4.3 Learning context

Experiment 3 introduced learning context and instructional style as potential influencers on the development of perception ability in low-level Japanese learners. In this experiment, study-abroad learners, who reported greater access to Japanese input and

more frequent daily usage, were compared with proficiency-matched at-home learners in the US. The other key difference between these two groups was that the at-home group regularly received explicit feedback and correction on their production of accent in their language classroom, despite reporting lower Japanese usage than the study-abroad group. Of interest was whether the higher-input study-abroad group would improve more in their perception of accent over a semester, despite not receiving explicit correction. Considering the mixed findings on the effect of study-abroad on lexical acquisition (e.g., Collentine, 2004; O'Brien et al., 2007; Sunderman & Kroll, 2009), I assumed that regardless of learning context overt correction is vital for phonological form learning, particularly due to the subtle acoustic cue that marks pitch accent location. I therefore predicted that the at-home Japanese learner group would improve more on the perception tasks. I found, however, that both groups improved to a similar degree over the semester, but that the at-home group both started and finished with higher perception accuracy. This partially supported the prediction that learners who receive explicit feedback on lexical accent are better at accent perception.

Sunderman & Kroll (2009) presented evidence for what they termed the 'resource threshold hypothesis', in which low WM capacity learners may be less capable or even unable of benefiting from the increased input afforded by the L2 context while studying abroad. They found that WM limitations, as measured by a reading span task, particularly hampered the development of language production skills in low-span learners in a study-abroad context. My findings showed that PM was an important contributor to perceptual development in *both* the JFL and JSL contexts, whereby even among learners who

received explicit feedback on their accent production, higher-resource learners tended to display an advantage. It appears then that both PM resources and classroom correction are important factors in L2 perceptual development. Keep in mind that in-class feedback involved the instructor modelling the correct accent pattern prior to eliciting a learner's production. Thus, accurate perception of this feedback is a crucial phase leading up to articulation. When instruction was considered as a factor, I found that the at-home group who received this feedback on accent modestly outperformed the study abroad group at both testing times. While both groups improved similarly percentage-wise over the semester, and these gains were linked to PM resources on the PitchID task and acoustic sensitivity on the PitchCAT task, the group with access to form-focused instruction on accent had both higher starting and ending points. This suggests then that drawing learners' attention to accent through overt correction can indeed be compensatory for low-PM learners.

6.5 Pedagogical application

In the previous sections I discussed the key findings on how both cognitive resources and experience-related factors impact lexical accent perception in Japanese. When asked to judge the correctness of spoken accents, L1 Japanese listeners could do so reliably in a task that replicates what listeners are doing unconsciously everyday—matching spoken input to their mental representation of words. However, accuracy on the perception tasks declined, often drastically, for the L2 Japanese learners, a finding consistent with the Japanese perception literature. I acknowledge that the presence of relatively few homophonic pairs, coupled with the low acoustic salience of accent, which

is dependent on spectral cues only (i.e., F0 variations), makes it an inherently difficult feature of Japanese to acquire.

The current study's main contribution was to show that a significant degree of this variation could be accounted for by a set of factors that vary from both individual to individual, and group to group. We are then left with the question of how to best interpret these results in light of the teaching of Japanese to L2 learners. One relevant finding was that at the beginning level, basic cognitive resources were significant predictors of accent perception gains. The ability to store phonological input and one's acoustic sensitivity threshold seem to affect perceptual development, although differently depending on the type of perception task. Yet, these factors diminished and yielded way to lexical perception strategies in the advanced learners, as evidenced by the fact that L2 vocabulary size, and not overall proficiency or length of residence in Japan, was significantly correlated with perception accuracy in the two groups in Experiment 2. Such a pattern suggests a degree of initial constraint on perception ability that is linked to one's cognitive capacities.

In spite of this apparent "constraint," taking a language teacher's perspective, I do not consider this finding as evidence of limitations on an individual's ultimate capacity to acquire Japanese phonology. I will next argue that by having described the sources of this variation, better strategies for the teaching of accent and perceptual training can be developed. Hummel & French (2010) have addressed this issue in their article on working memory and L2 learning. I adopt suggestions from their argument in my points below regarding PM and accent acquisition.

Low-PM capacity learners may have difficulty developing their L2 lexicon from spoken input alone (Speciale et al., 2004; Sunderman & Kroll, 2009). The speech stream is rapid, and if accent is not focused on as a part of a word's phonological form, then this group in particular may be at a disadvantage. In fact, previous research has reported that in an ESL classroom using a strong version of the communicative method, learners' PM capacity accounted for as much as 72% of their listening skills (French, 2009). Hammond (1995) pointed out that a long-standing, and possibly misguided, assumption of communicative language teaching is that providing a correct spoken model to students is sufficient for them to acquire L2 phonology to a reasonable degree. The present findings on the involvement of individual resources in accent acquisition make this "input-is-sufficient" assumption problematic. Considering the above-discussed relation between PM and lexical learning, teaching methods that emphasize or compensate for varying PM capacities are important. Hummel & French (2010) suggest the use of visual input to supplement spoken language as a general strategy for vocabulary instruction. In terms of Japanese instruction, visual input for accent patterns could be delivered in the form of pitch graphs, like those used in the current experiments, hand gestures to indicate pitch fall, or the use of internet-based programs or software such as those developed by Hirata (1999) and Short, Hirose & Minematsu (2013) for lexical pitch, and Eda (2004) for sentential intonation. Morett & Chang (2014) employed hand gestures in the teaching of Mandarin Chinese tones to L1 English speakers, finding that conveying tone contours this way helped learners to discriminate the meanings of minimal tone pairs. Shport (2011) found that talker-variability as well as access to immediate feedback is valuable in

training accent perception.²² This same approach of immediate feedback can also be conducted in the classroom, without disrupting the communicative goals of the lesson. For example, modeling the correct accent can be done briefly at the end of a learner's speaking turn, but the correction should not stop there—having the learner then use the word in a meaningful context prevents the communicative activity, and overall flow of the lesson, from breaking down.

Although discounted in some language teaching approaches, drill-like repetition or explicit correction of accent can potentially benefit L2 Japanese learners' development of accent perception. Form-focused drills are important in strengthening long-term knowledge of the relatively subtle prosodic features like pitch accent and tone (Doughty & Williams, 1998). Due to Japanese lexical accent being relatively low in perceptual salience, especially to learners from non-tone L1 backgrounds, feedback conducted early and often in the language course would likely give learners the best chance to begin building more robust lexical representations. It is interesting to note that corrective feedback and practice in the L2 classroom mainly focuses on the production of accent, although the JFL learners in Experiment 3 derived *perceptual* development from classroom production training. A learner's production of accent is of course what the language instructor actually hears, but in order to benefit from correction, learners must be made perceptually aware of, and also accurately perceive, word accent in the spoken input. Setting aside arguments for or against the primacy of perception over production, I

²² Although note that Sadakata & McQueen (2014) recently found that a high-variability training paradigm for Mandarin tones was only beneficial to high-aptitude (i.e., tone category-sensitive) listeners, but not to a low-aptitude group.

feel that native-like perception is a vital component of language acquisition, and a key skill from moving from lower proficiency levels to more advanced L2 competence. If the goal is to produce language users who interact comfortably in the target culture, especially in Japan where native-nonnative speaker interactions are comparatively infrequent, then developing learners' perceptual ability should be a task of foremost importance in the Japanese language classroom.

Lastly, it is impossible to implement corrective feedback into the L2 classroom without well-trained language instructors. How might the current research findings on L2 accent perception be applied to the training of Japanese teachers? At first glance, instructors may feel that because individual differences in basic cognitive capacities were found to be predictive of perceptual development in low-proficiency learners, that there is little one can do in the classroom itself. Yet, if instructors are made aware of, and critically reflect upon, the means by which they provide input, they can adapt their own practices to include input that is easily accessible to learners of varying memory and auditory resource capacities. The awareness that visual feedback or repetition may be effective with these learner populations is one readily transmittable finding to teacher-training. Using immediate corrective feedback, which some may have discounted amidst the shifting teaching methodologies over the years, may also prove beneficial to low-PM span learners (Hummel & French, 2010). Moreover, familiarizing both new and experienced Japanese teachers with available technology for accent training, some of which is freely accessible online like the Online Japanese Accent Dictionary (OJAD; Hirano et al., 2013), is one method for expanding instructors' pedagogical toolboxes.

6.6 Conclusion

In this dissertation, I have tried to characterize some of the complexity underlying the perception of lexical accent both by native and nonnative Japanese listeners. The experimental results showed that the ability to accurately perceive, along with the process of acquiring, accent is influenced by an individual's cognitive resources and language learning experiences to a greater extent than was previously considered. The construct of aptitude, under which the cognitive abilities I tested can be subsumed, has been measured as a predictor of L2 acquisition for decades (Dornyei, 2005; Skehan, 1989). Yet, until recently, few studies in the field of SLA have considered the relationship of aptitude measures on the learning of a language's prosodic features. The findings from the present experiments have shed further light on how individuals vary in their success at L2 prosodic perception, in a variety of learning contexts and across multiple proficiency levels. From the earliest stages of learning, the general cognitive mechanisms of PM and acoustic sensitivity are active in the development of perception of L2 lexical form. It appears then that L2 learning, at least to a degree, is akin to the acquisition of other higher-level cognitive abilities, such as playing a musical instrument. Learning music, for instance, involves not only the memorization of sequences of musical pitch, in which WM/PM may play a role (Berz, 1995), but also the ability to physically distinguish these notes (Deutsch et al., 2009), similar to the acoustic sensitivity task. The discovery of an overlap of cognitive resources and L2 learning in the current study suggests that language may not have a particularly privileged place in higher-level cognition, at least in its earliest developmental stages as in the beginners in Experiment 3. When turning the

focus to highly-experienced L2 learners in Experiment 2, however, it was found that aptitude no longer related to lexical accent perception in any significant way. Experience with L2 lexical form was the key factor in this group's perception ability. Thus, the current research demonstrated that, although with listeners from various L1 backgrounds, as L2 acquisition progresses learners increasingly make use of language-specific knowledge to aid speech perception.

Although the present study is an initial description of how listeners' individual abilities contribute to lexical accent perception, I feel that I have captured a side of L2 phonological acquisition insufficiently addressed by many cross-linguistic perception studies. That is, too narrow a focus on the phonetic properties of a given pair of languages can potentially mask other sources that contribute to the relative ease or difficulty of L2 perceptual development. Lastly, I must also acknowledge the undeniable fact that aptitude alone is not deterministic in any absolute sense, in that L2 learners rely on a range of factors as they progress in their language studies. The current research tried to capture this aspect of acquisition as well, by showing that experienced-based factors like L2 lexical knowledge, learning context and instruction style are active contributors to the development process. It is my hope that future work will continue to explore the complexity inherent not only in the second/foreign language acquisition of Japanese lexical accent, but also in the learning of prosodic features of other languages.

References

- Amano, N., & Kondo, K. (1999). *Lexical properties of Japanese Vol. 1*. Tokyo: Sanseido.
- Amano, N., & Kondo, K. (2000). *Lexical properties of Japanese Vol. 7*. Tokyo: Sanseido.
- Aoyama, K., Flege, J. E., Guion, S. G., Yamada, T., & Akahane-Yamada, R. (2004). Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /r/ and /l/. *Journal of Phonetics*, 23, 233–250.
- Arai, M. (1997). The results of a longitudinal survey on the perception of the Tokyo accent: American learners living in Kyoto. *Spoken Japanese language education: Looking to the 21st century* (pp. 73-79). Tokyo: The National Language Research Institute.
- Ayusawa, T. (2003). Acquisition of Japanese accent and intonation by foreign learners. *Journal of the Phonetic Society of Japan*, 7, 47-58.
- Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, 105, 158-173.
- Beckman, M. (1986). *Stress and non-stress accent*. Dordrecht: Foris.
- Beckman, M., & Pierrehumbert, J. (1986). Intonational structure in Japanese and English. *Phonology Yearbook*, 3, 255-309.
- Belin, P., Fecteau, S., & Bédard, C. (2004). Thinking the voice: Neural correlates of voice perception. *Trends in Cognitive Sciences*, 3, 129–135.
- Bent, T., Bradlow, A., & Wright, B. (2006). The influence of linguistic experience on the cognitive processing of pitch in speech and nonspeech sounds. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 97-103.
- Berz, W. (1995). Working memory in music: A theoretical model. *Music Perception: An Interdisciplinary Journal*, 12, 353-264.

- Best, C., & Tyler, M. (2007). Nonnative and second-language speech perception. In O.S. Bohn & M. Munro. (Eds.) *Language experience in second language speech learning*. Amsterdam: John Benjamins, pp. 13-34.
- Boersma, P., & Weenink, D. (2008). Praat: Doing phonetics by computer.
- Bohn, O.S. (1995). Cross-language speech perception in adults: First language transfer doesn't tell it all. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 279-304). Baltimore: York Press.
- Brewer, G., & Unsworth, N. (2012). Individual differences in the effects of retrieval from long-term memory. *Journal of Memory and Language*, 66, 407-415.
- Collentine, J. (2004). The effects of learning contexts on morphosyntactic and lexical development. *Studies in Second Language Acquisition*, 26, 227-248.
- Collentine, J., & Freed, B. (2004). Learning context and its effects on second language acquisition. *Studies in Second Language Acquisition*, 26, 153-171.
- Chan, A., Ho, Y., & Cheung, M. (1998). Music training improves verbal memory. *Nature*, 396, 128.
- Chen, J., Chen, T., & Dell, G. (2002). Word-form encoding in Mandarin Chinese as assessed by the implicit priming task. *Journal of Memory and Language*, 46, 751-781.
- Cheung, H. (1996). Nonword span as a unique predictor of second-language vocabulary learning. *Developmental Psychology*, 32, 867-873.
- Cunnings, I. (2012). An overview of mixed-effects statistical models for second-language researchers. *Second Language Research*, 28, 369-382.
- Cutler, A. (1986). Forbear is a homophone: Lexical prosody does not constrain lexical access. *Language and Speech*, 29, 201-220.
- Cutler, A., & Chen, H. C. (1997). Lexical tone in Cantonese spoken-word processing. *Perception & Psychophysics*, 59, 165-179.
- Cutler, A. & Clifton, C. (1984). The use of prosodic information in word recognition. In H. Bouma & D.G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 183- 196). Hillsdale, N.J.: Erlbaum.

- Cutler, A., & Otake, T. (1999). Pitch accent in spoken-word recognition in Japanese. *Journal of the Acoustical Society of America*, 105(3), 1877-1888.
- Deutsch, D., Dooley, K., Henthorn, T., & Head, B. (2009). Absolute pitch among students in an American conservatory: Association with tone language fluency. *Journal of the Acoustical Society of America*, 125, 2398-2403.
- Dornyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Mahwah, NJ: Erlbaum.
- Doughty, C., & Williams, J. (1998). *Focus on form in classroom second language acquisition*. New York: Cambridge University Press.
- Eda, S. (2004) Processing of intonation patterns in Japanese: Implications for Japanese as a foreign language. Doctoral dissertation. Ohio State University, Columbus.
- Ellis, N. (1996). Sequencing in SLA: Phonological memory, chunking and points of order. *Studies in Second Language Acquisition*, 18, 91-126.
- Ellis, N. (2004). The processes of second language acquisition. In B. VanPatten, J. Williams, S. Rott, & M. Overstreet (Eds.), *Form-meaning connections in second language acquisition* (pp. 49-76). Mahwah, NJ: Erlbaum.
- Engel de Abreu, P., & Gathercole, S. (2012). Executive and phonological processes in second-language acquisition. *Journal of Educational Psychology*, 104, 974-986.
- Fodor, J. D. (2002) Psycholinguistics cannot escape prosody. In *Proceedings of the SPEECH PROSODY 2002 Conference*, Aix-en-Provence, France.
- Fox, A. (2002). *Prosodic features and prosodic structure: The phonology of suprasegmentals*. Oxford: Oxford University Press.
- Gathercole, S. (2009) Working Memory and Language. In M. Gaskell (Ed.) *Oxford Handbook of Psycholinguistics*. Oxford: Oxford University Press, pp. 757-769.
- Gathercole, S. E., Willis, C., Emslie, H., & Baddeley, A. (1992). Phonological memory and vocabulary development during the early school years: A longitudinal study. *Developmental Psychology*, 28, 887-898.
- Hammond, R. (1995). Foreign accent and phonetic interference: The application of phonetic research to the second language classroom. In F. Eckman, D. Highland, P. Lee, J. Mileham, & R. Weber. (Eds.) *Second language acquisition theory and pedagogy*. Lawrence Earlbaum, pp. 293-303.

- Hao, Y. (2012). Second language acquisition of Mandarin tones by tonal and non-tonal language speakers. *Journal of Phonetics*, 40, 269-279.
- Hardison, D. & Saigo, M. (2010). Development of perception of second language Japanese geminates: Role of duration, sonority, and segmentation strategy. *Applied Psycholinguistics*, 31, 81-99.
- Hayashi, R., Imaizumi, S., Mori, K., Niimi, S., Ueno, U., & Kiritani, S. (2001). Elicitation of N400m in sentence comprehension due to lexical prosody incongruity. *Neuroreport*, 12, 1753–1756.
- Hirano, H., Nakamura, I., Minematsu, N., Suzuki, M., Nakagawa, C., Nakamura, N., Tagawa, Y., Hirose, K., & Hashimoto, H. (2013). A free online accent and intonation dictionary for teachers and learners of Japanese. *Proceedings of Interspeech*, 1875-1876.
- Hirano-Cook, E. (2011). Japanese pitch accent acquisition by learners of Japanese: Effects of training on Japanese accent instruction, perception and production. Doctoral dissertation. University of Kansas, Lawrence.
- Hirata, Y. (1999). Acquisition of Japanese rhythm and pitch accent by English native speakers. Doctoral dissertation. University of Chicago, Chicago.
- Hummel, K. (2009). Aptitude, phonological memory, and second language proficiency in nonnovice adult learners. *Applied Psycholinguistics*, 30, 225–249.
- Hummel, K., & French, L. (2010). Phonological memory and implications for the second language classroom. *Canadian Modern Language Review*, 66, 371–391.
- Iverson, P., & Kuhl, P. (1996). Influences of phonetic identification and category goodness on American listeners' perception of /r/ and /l/. *Journal of the Acoustical Society of America*, 99, 1130-1140.
- Jincho, N., Namiki, H., & Mazuka, R. (2008). Effects of verbal working memory and cumulative linguistic knowledge on reading comprehension. *Japanese Psychological Research*, 50, 12-23.
- Jorden, E. & Noda, M. (1987). *Japanese: The Spoken Language, Part 1*. New Haven: Yale University Press.
- Jorden, E. & Noda, M. (1988). *Japanese: The Spoken Language, Part 2*. New Haven: Yale University Press.

- Juffs, A. & Harrington, M. (2011). Aspects of working memory in L2 learning. *Language Teaching*, 44, 137-166.
- Jun, S-A. (1998). The accentual phrase in the Korean prosodic hierarchy. *Phonology*, 15, 189-226.
- Jun, J., Jungsun, K., Lee, H., & Sun-Ah, J. (2006). The prosodic structure and pitch accent of Northern Kyungsang Korean. *Journal of East Asian Linguistics*, 15, 289-317.
- Just, M. & Carpenter, P. (1992). A capacity theory of comprehension: Individual differences in working memory. *The Psychological Review*, 99, 122-149.
- Kaushanskaya, M. (2012). Cognitive mechanisms of word learning in bilingual and monolingual adults: The role of phonological memory. *Bilingualism: Language and Cognition*, 15, 470-489.
- Kaushanskaya, M., & Yoo, J. (2013). Phonological short-term and working memory in bilinguals' native and second language. *Applied Psycholinguistics*, 34, 1005-1037.
- Kitahara, M. (2001). Category structure and function of pitch accent in Tokyo Japanese. Doctoral dissertation. Indiana University, Bloomington.
- Kormos, J. & Safar, A. (2008). Phonological short-term memory, working memory and foreign language performance in intensive language learning. *Bilingualism: Language and Cognition*, 11, 261-271.
- Lee, W., Murashima, K., & Shirai, Y. (2006). Prosodic development in the acquisition of Japanese: A longitudinal study of three native speakers of Cantonese. *Japan Journal*, 10, 38-51.
- Levelt, W. J. M, Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1-75.
- Liu, F., Patel, A., Fourcin, A., & Stewart, L. (2010). Intonation processing in congenital amusia: Discrimination, identification and imitation. *Brain*, 133, 1682-1693.
- Logan & Pruitt, (1995). Methodological issues in training listeners to perceive nonnative phonemes: Linguistic experience and the "perceptual magnet effect." In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 121-154). Baltimore: York Press.
- Maekawa, K. (2003). *Corpus of Spontaneous Japanese*. Available from: http://www.ninjal.ac.jp/corpus_center/csj/misc/preliminary/index_e.html

- Mandel, J. (2009). Adaptive pitch test. Available from:
<http://tonometric.com/adaptivepitch/>
- Martin, K. & Ellis, N. (2012). The roles of phonological short-term memory and working memory in L2 grammar and vocabulary learning. *Studies in Second Language Acquisition*, 34, 379-413.
- Masoura, E. & S. E. Gathercole (2005). Contrasting contributions of phonological short-term memory and long-term knowledge to vocabulary learning in a foreign language. *Memory*, 13, 422-429.
- Masuda-Katsuse, I. (2006). Contribution of pitch-accent information to Japanese spoken word recognition. *Acoustical Science and Technology*, 27, 97-103.
- Minematsu, N. & Hirose, K. (1995). Role of prosodic features in the human process of perceiving spoken words and sentences in Japanese. *Journal of the Acoustical Society of Japan*, 16, 311-320.
- Miwa, K., Dijkstra, T., Bolger, P., & Baayen, R. (2013). Reading English with Japanese in mind: Effects of frequency, phonology and meaning in different-script bilinguals. *Bilingualism: Language and Cognition*, 17, 445-463.
- Miyaoka Y, Tamaoka, K. & Sakai, H. (2011) Nihongo goi tesuto no kaiatsu to shinraisei: Chugokugo o bogo tosuru nihongo gakushusha no deta niyoru tesuto hyoka [Developing Japanese vocabulary test and its reliability: Test evaluation by data taken from native Chinese speakers learning Japanese] *Research Journal of Hiroshima University of Economics*, 34, 1-18.
- Miyaoka Y, Tamaoka, K. & Sakai, H. (in press, 2014) Nihongo bunpo noryoku tesuto no kaiatsu to shinraisei: Nihongo gakushusha no deta niyoru tesuto hyoka [Developing Japanese grammar test and its reliability: Test evaluation by data taken from Japanese learners] *Research Journal of Hiroshima University of Economics*, 36.
- Morett, L. & Chang, L. (2014). Emphasizing sound and meaning: Pitch gestures enhance Mandarin lexical tone acquisition. *Language, Cognition and Neuroscience*, 98, 773-794.
- Nan, Y., Sun, Y., & Peretz, I. (2010). Congenital amusia in speakers of a tone language: Association with lexical tone agnosia. *Brain*, 133, 2635-2642.
- Nishinuma, Y., Arai, M., & Ayusawa, T. (1996). Perception of tonal accent by Americans learning Japanese. In *Proceedings of the 4th International Conference on Spoken Language Processing*, Philadelphia, PA.

- Norris, D., McQueen, J., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences*, 23, 299-370.
- O'Brien, I., Segalowitz, N., Collentine, J., & Freed, B. (2006). Phonological memory and lexical narrative, and grammatical skills in second language oral production by adult learners. *Applied Psycholinguistics*, 27, 377-402.
- O'Brien, I., Segalowitz, N., Freed, B. & Collentine, J. (2007). Phonological memory predicts second language oral fluency gains. *Studies in Second Language Acquisition*, 29, 557-581.
- Otake, T., & Cutler, A. (1999). Perception of suprasegmental structure in a nonnative dialect. *Journal of Phonetics*, 27, 229-253.
- Perrachione, T., Lee, J., Ha, L., & Wong, P. (2011). Learning a novel phonological contrast depends on interactions between individual differences and training paradigm design. *Journal of the Acoustical Society of America*, 130, 461-472.
- Pierrehumbert, J. (1980). The phonology and phonetics of English intonation. Doctoral dissertation, MIT.
- Plonsky, L. (2013). Study quality in SLA: An assessment of designs, analyses, and reporting practices in quantitative L2 research. *Studies in Second Language Acquisition*, 35, 655-687.
- Roy, P., & Chiat, S. (2004). A prosodically controlled word and nonword repetition task for 2- to 4-year-olds: Evidence from typically developing children. *Journal of Speech, Language, and Hearing Research*, 47, 223-234.
- Sadakata, M. & McQueen, J. (2014). Individual aptitude in Mandarin lexical tone perception predicts effectiveness of high-variability training. *Frontiers in Psychology*, 5, 1-15.
- Sakamoto, E. (2010). An investigation of factors behind foreign accent in the L2 acquisition of Japanese lexical pitch accent by adult English speakers. Doctoral dissertation. University of Edinburgh.
- Segalowitz S, Segalowitz, N., & Wood, A. (1998). Assessing the development of automaticity in second language word recognition. *Applied Psycholinguistics* 19: 53-67.
- Sekiguchi, T., & Nakajima, Y. (1999). The use of lexical prosody for lexical access of the Japanese language. *Journal of Psycholinguistics*, 28, 439-454.

- Semal, C., Demany, L., Ueda, K., & Halle, P.A. (1996). Speech versus nonspeech in pitch memory. *Journal of the Acoustic Society of America*, 100, 1132-1140.
- Service, E. & V. Kohonen (1995). Is the relation between phonological memory and foreign language learning accounted for by vocabulary acquisition? *Applied Psycholinguistics* 16, 155–172.
- Shibata, T., & Hurtig, R. (2008). Prosody acquisition by Japanese learners. In Z. Han. (Ed.), *Understanding second language process* (pp. 176-204). Clevedon, UK: Multilingual Matters.
- Shibata, T. & Shibata, R. (1990) Akusento wa doo'ongo o donoteido benbetsu shiuruno ka?: Nihongo, eigo, chuugokugo no baai [How much can accent distinguish homophones?: cases of Japanese, English and Chinese]. *Mathematical Linguistics*, 17, 317-327.
- Shibatani, M. (1990). *The Languages of Japan*. Cambridge: Cambridge University Press.
- Short, G., Hirose, K., & Minematsu, N. (2013). Japanese lexical accent recognition for a CALL system by deriving classification equations with perceptual experiments. *Speech Communication*, 55, 1064-1080.
- Shport, I. (2008). Acquisition of Japanese pitch accent by American learners. In P. Heinrich & Y. Sugita (Eds.), *Japanese as foreign language in the age of globalization* (pp. 165-187). München: Iudicium Verlag.
- Shport, I. (2011) Cross-linguistic perception and learning of Japanese lexical prosody by English listeners. Doctoral dissertation. University of Oregon, Eugene.
- Skehan, P. (1989). *Individual differences in second-language learning*. London: Edward Arnold.
- So, C. & Best, C. (2010). Cross-language perception of non-native tonal contrasts: Effects of native phonological and phonetic influences. *Language and Speech*, 53, 273-293.
- Speciale, G., Ellis, N., & Bywater, T. (2004). Phonological sequence learning and short-term store capacity determine second language vocabulary acquisition. *Applied Psycholinguistics*, 25, 293-321.
- Strange, W. (2011). Automatic selective perception (ASP) of first and second language speech: A working model. *Journal of Phonetics*, 39, 456-466.

- Strange, W., & Shafer, V. L. (2008). Speech perception in second language learners: The re-education of selective perception. In J. G. Hanson Edwards & M. L. Zampini (Eds.), *Phonology and second language acquisition* (pp. 193-218). Amsterdam: John Benjamins.
- Sunderman, G. & Kroll, J. (2009). When study-abroad experience fails to deliver: The internal resource threshold effect. *Applied Psycholinguistics*, 30, 79-99.
- Szmalc A, Brysbaert, M., & Duyck, W. (2012). Working memory and (second) language processing. In J. Altarriba & L. Isurin (Eds.) *Memory, language, and bilingualism: Theoretical and applied approaches*. Cambridge: Cambridge University Press, pp. 74-94.
- Tamaoka, K., & Ikeda, F. (2010). Whiskey or Bhiskey?: Influence of first element and dialect region on sequential voicing of shoochuu. *Gengo Kenkyu*, 137, 65-79.
- Tamaoka, K., Saito, N., Kiyama, S., Timmer, K., & Verdonschot, R. (2014). Is pitch accent necessary for comprehension by native Japanese speakers? An ERP investigation. *Journal of Neurolinguistics*, 27, 31-40.
- Tamaoka, K. (2014). The Japanese writing system and lexical understanding. *Japanese Language and Literature*, 48, 431-471.
- Tanaka, S. & Kubozono, H. (2012). *Nihongo no hatsuon kyooshitsu: Riron to renshuu*. [Introduction to Japanese Pronunciation: Theory and Practice]. Tokyo: Kurosio.
- Taylor, B. (2011). Variability and systematicity in individual learners' Japanese accent. *Pozanan Studies in Contemporary Linguistics*, 47, 146-158.
- Toda, T. (2001). Hatsuo shidoo ga akusento no chikaku ni ataeru eikyoo. [The effects of pronunciation training on accent perception]. *Waseda University Center for Japanese Language Studies*, 14, 67-88.
- Tsujimura, N. (2014) *An Introduction to Japanese Linguistics*, 3rd edition. Oxford: Blackwell.
- Tsukada K (2012) Non-native Japanese listeners' perception of vowel length contrasts in Japanese and Modern Standard Arabic. *Second Language Research*, 28, 151-168.
- Ueno, T., Saito, S., Saito, A., Tanida, Y., Patterson, K., & Lambon Ralph, M. (2014). Not lost in translation: Generalization of the primary systems hypothesis to Japanese-specific language processes. *Journal of Cognitive Neuroscience*, 26, 433-446.

- Ueyama, M. (2000). Prosodic transfer: An acoustic study of L2 English vs. L2 Japanese. Doctoral dissertation. University of California, Los Angeles.
- Vance, T. (2008). *The sounds of Japanese*. Cambridge: Cambridge University Press.
- Venditti, J. (2005). The J_ToBI Model of Japanese Intonation In S.-A. Jun (Ed.), *Prosodic Typology: The phonology and intonation of phrasing* (pp.172-200). Oxford: Oxford University Press.
- Wang, X. (2013). Perception of Mandarin tones: The effect of L1 background and training. *The Modern Language Journal*, 97, 144-160.
- Warner, N., Otake, T., & Arai, T. (2010). Intonational structure as a word-boundary cue in Tokyo Japanese. *Language and Speech*, 53, 107-131.
- Warrier, C., & Zatorre, R. (2004). Right temporal cortex is critical for utilization of melodic contextual cues in a pitch constancy task. *Brain*, 127, 1616-1625.
- Wayland, R., Herrera, E., & Kaan, E. (2010). Effects of musical experience and training on pitch contour perception. *Journal of Phonetics*, 38, 654-662.
- Werker, J., & Curtin, S. (2005). PRIMIR: A developmental model of speech processing. *Language Learning and Development*, 1, 197-234.
- Williams, J., & Lovatt, P. (2003). Phonological memory and rule learning. *Language Learning*, 53, 67-121.
- Wong, P., & Perrachione, T. (2007). Learning pitch patterns in lexical identification by native English-speaking adults. *Applied Psycholinguistics*, 28, 565-585.
- Yuzawa, M. (2002). Effects of pitch accent on short-term retention of words in young children. *Japanese Journal of Psychology*, 73, 258-263. (In Japanese with English abstract)
- Yuzawa, M., & Saito, S. (2006). The role of prosody and long-term phonological knowledge in Japanese children's nonword repetition performance. *Cognitive Development*, 21, 146-157.

Appendix A – Experiment 1 test stimuli

Target words ($N = 32$) and postpositions (underlined) with carrier sentences used in the PitchID and PitchCAT tasks in Experiment 1.

<u>医学</u> を学ぶ <i>igaku o manabu</i> ‘(I) study medicine’	<u>家内</u> に話す <i>kanai ni hanasu</i> ‘(I) speak to my wife’
<u>めがね</u> をかける <i>megane o kakeru</i> ‘(I) put on glasses’	<u>技術</u> をみがく <i>gijutsu o migaku</i> ‘(I) polish my skills’
<u>フォーク</u> で食べる <i>fooiku de taberu</i> ‘(I) eat with a fork’	<u>荷物</u> を運ぶ <i>nimotsu o hakobu</i> ‘(I) carry luggage’
<u>おもちゃ</u> で遊ぶ <i>omocha de asobu</i> ‘(I) play with a toy’	<u>あなた</u> がいる <i>anata ga iru</i> ‘You are there’
<u>八時</u> に起きる <i>hachiji ni okiru</i> ‘(I) wake up at 8 o’clock’	<u>ハサミ</u> がある <i>hasami ga aru</i> ‘There are scissors’
<u>中身</u> を見る <i>nakami o miru</i> ‘(I) look at the contents’	<u>刺身</u> を食べる <i>sashimi o taberu</i> ‘(I) eat sashimi’
<u>ななめ</u> にする <i>naname ni suru</i> ‘(I) turn it sideways’	<u>昼間</u> に行く <i>hiruma ni iku</i> ‘(I) go in the afternoon’
<u>言葉</u> にする <i>kotoba ni suru</i> ‘(I) put it into words’	<u>娘</u> がいる <i>musume ga iru</i> ‘(I) have a daughter’
<u>夜中</u> に起きる <i>yonaka ni okiru</i> ‘(I) wake up at midnight’	<u>男</u> がいる <i>otoko ga iru</i> ‘There is a man’
<u>いなか</u> に住む <i>inaka ni sumu</i> ‘(I) live in the countryside’	<u>左</u> に曲がる <i>hidari ni magaru</i> ‘(I) turn left’
<u>ハガキ</u> を出す <i>hagaki o dasu</i> ‘(I) send a postcard’	<u>昔</u> に戻る <i>mukashi ni modoru</i> ‘(I) return to the past’
<u>手前</u> に引く <i>temae ni hiku</i> ‘(I) pull it toward me’	<u>仕事</u> を探す <i>shigoto o sagasu</i> ‘(I) look for a job’
<u>来月</u> に行く <i>raigetsu ni iku</i> ‘(I) go next month’	<u>タクシー</u> に乗る <i>takushii ni noru</i> ‘(I) get in a taxi’
<u>あさって</u> まで待つ <i>asatte made matsu</i> ‘(I) wait until the day after tomorrow’	<u>飛行機</u> に乗る <i>hikooki ni noru</i> ‘(I) ride in an airplane’
<u>九日</u> に行く <i>kokonoka ni iku</i> ‘(I) go on the 9th’	<u>妹</u> がいる <i>imooto ga iru</i> ‘(I) have a younger sister’
<u>夕方</u> になる <i>yuugata ni naru</i> ‘It becomes evening’	<u>地下鉄</u> に乗る <i>chikatetsu ni noru</i> ‘(I) ride the subway’

Appendix B – Experiment 2 test stimuli

Target words and postpositions (underlined) with carrier sentences used in the PitchID and PitchCAT tasks in Experiment 2.

医学を学ぶ <i>igaku o manabu</i> ‘(I) study <u>medicine</u> ’	家内に話す <i>kanai ni hanasu</i> ‘(I) speak to my <u>wife</u> ’
めがねをかける <i>megane o kakeru</i> ‘(I) put on <u>glasses</u> ’	技術をみがく <i>gijutsu o migaku</i> ‘(I) polish my <u>skills</u> ’
フォークで食べる <i>fooku de taberu</i> ‘(I) eat with a <u>fork</u> ’	荷物を運ぶ <i>nimotsu o hakobu</i> ‘(I) carry <u>luggage</u> ’
おもちゃで遊ぶ <i>omocha de asobu</i> ‘(I) play with a <u>toy</u> ’	あなたがいる <i>anata ga iru</i> ‘ <u>You</u> are there’
八時に起きる <i>hachiji ni okiru</i> ‘(I) wake up at <u>8 o’clock</u> ’	ハサミがある <i>hasami ga aru</i> ‘There are <u>scissors</u> ’
中身を見る <i>nakami o miru</i> ‘(I) look at the <u>contents</u> ’	刺身を食べる <i>sashimi o taberu</i> ‘(I) eat <u>sashimi</u> ’
ななめにする <i>naname ni suru</i> ‘(I) turn it <u>sideways</u> ’	屋間に行く <i>hiruma ni iku</i> ‘(I) go in the <u>afternoon</u> ’
言葉にする <i>kotoba ni suru</i> ‘(I) put it into <u>words</u> ’	娘がいる <i>musume ga iru</i> ‘(I) have a <u>daughter</u> ’
夜中に起きる <i>yonaka ni okiru</i> ‘(I) wake up at <u>midnight</u> ’	男がいる <i>otoko ga iru</i> ‘There is a <u>man</u> ’
いなかに住む <i>inaka ni sumu</i> ‘(I) live in the <u>countryside</u> ’	左に曲がる <i>hidari ni magaru</i> ‘(I) turn <u>left</u> ’
ハガキを出す <i>hagaki o dasu</i> ‘(I) send a <u>postcard</u> ’	昔に戻る <i>mukashi ni modoru</i> ‘(I) return to the <u>past</u> ’
手前に引く <i>temae ni hiku</i> ‘(I) pull it <u>toward</u> me’	仕事を探す <i>shigoto o sagasu</i> ‘(I) look for a <u>job</u> ’
来月に行く <i>raigetsu ni iku</i> ‘(I) go <u>next month</u> ’	太陽がのぼる <i>taiyoo ga noboru</i> ‘The <u>sun</u> rises’
湖でおよぐ <i>mizuumi de oyogu</i> ‘(I) swim in the <u>lake</u> ’	タクシーに乗る <i>takushii ni noru</i> ‘(I) get in a <u>taxi</u> ’
将来を考える <i>shoorai o kangaeru</i> ‘(I) think of the <u>future</u> ’	マンションに住む <i>manshon ni sumu</i> ‘(I) live in an <u>apartment</u> ’
あさってまで待つ <i>asatte made matsu</i> ‘(I) wait until the <u>day after tomorrow</u> ’	黒板に書く <i>kokuban ni kaku</i> ‘(I) write on the <u>blackboard</u> ’
八月になる <i>hachigatsu ni naru</i> ‘It becomes <u>August</u> ’	飛行機に乗る <i>hikooki ni noru</i> ‘(I) ride in an <u>airplane</u> ’
オレンジを食べる <i>orenji o taberu</i> ‘(I) eat an <u>orange</u> ’	てぶくろを買う <i>tebukuro o kau</i> ‘(I) buy <u>gloves</u> ’
九日に行く <i>kokonoka ni iku</i> ‘(I) go on the <u>9th</u> ’	料金を払う <i>ryookin o harau</i> ‘(I) pay the <u>fee</u> ’

成績をつける *seiseki o tsukeru* ‘(I) assign grades’

正月を楽しむ *shoogatsu o tanoshimu*
‘(I) look forward to New Year’s Day’

地下鉄に乗る *chikatetsu ni noru* ‘(I) ride the subway’

専門を学ぶ *senmon o manabu* ‘(I) study for my major’

土曜日に行く *doyoobi ni iku* ‘(I) go on Saturday’

妹がいる *imooto ga iru* ‘(I) have a younger sister’

二日目になる *futsukame ni naru* ‘It becomes the 2nd’

六月になる *rokugatsu ni naru* ‘It becomes June’

夕方になる *yuugata ni naru* ‘It becomes evening’

独身になる *dokushin ni naru* ‘(I) become single’

Appendix C – Experiment 3 test stimuli

Targets (underlined) with carrier sentences used in the PitchID/PitchCAT tasks in Experiment 3.

あちらに行く	<i>atira ni iku</i>	“(I) go <u>over</u> there.”
タクシーに乗る	<i>takusii ni noru</i>	“(I) take a <u>taxi</u> .”
仕事をする	<i>shigoto o suru</i>	“(I) go to <u>work</u> .”
花屋に行く	<i>hanaya ni iku</i>	“(I) go to the <u>flower shop</u> .”
土曜日に行く	<i>doyoobi ni iku</i>	“(I) go on <u>Saturday</u> .”
本屋に行く	<i>honya ni iku</i>	“(I) go to the <u>bookstore</u> .”
あさってからする	<i>asatte kara suru</i>	“(I) will do it the <u>day after tomorrow</u> .”
家族に言う	<i>kazoku ni iu</i>	“(I) tell my <u>family</u> .”
飛行機に乗る	<i>hikooki ni noru</i>	“(I) ride an <u>airplane</u> .”
ビールを飲む	<i>biiru o nomu</i>	“(I) drink <u>beer</u> .”
あなたがいる	<i>anata ga iru</i>	“(You) are <u>here</u> .”
デパートで買う	<i>depaato de kau</i>	“(I) buy it at the <u>department store</u> .”
新聞を読む	<i>sinbun o yomu</i>	“(I) read the <u>newspaper</u> .”
八時に来る	<i>hatizi ni kuru</i>	“(I) come at <u>8 o'clock</u> .”
二十歳になる	<i>hatati ni naru</i>	“(I) turn <u>twenty</u> .”
今晚までいる	<i>konban made iru</i>	“(I) will stay until <u>tonight</u> .”
フランスに行く	<i>furansu ni iku</i>	“(I) go to <u>France</u> .”
手前で止まる	<i>temae de tomaru</i>	“(I) stop in <u>front</u> (of someplace).”
ベーコンを食べる	<i>beekon o taberu</i>	“(I) eat <u>bacon</u> .”
砂糖を使う	<i>satoo o tukau</i>	“(I) use <u>sugar</u> .”
手紙を書く	<i>tegami o kaku</i>	“(I) write a <u>letter</u> .”
赤ちゃんがいる	<i>akatyan ga iru</i>	“(I) have a <u>baby</u> .”
たまごを食べる	<i>tamago o taberu</i>	“(I) eat <u>eggs</u> .”
グリーンもある	<i>guriin mo aru</i>	“There are <u>green</u> (ones) also.”
ホテルに泊まる	<i>hoteru ni tomaru</i>	“(I) stay in a <u>hotel</u> .”
外国に行く	<i>gaikoku ni iku</i>	“(I) go <u>abroad</u> .”
名前を呼ぶ	<i>namae o yobu</i>	“(I) call (someone’s) <u>name</u> .”
図書館に行く	<i>tosyokan ni iku</i>	“(I) go to the <u>library</u> .”
彼女がいる	<i>kanojo ga iru</i>	“(She) is here/(I) have a <u>girlfriend</u> .”
グレーもある	<i>guree mo aru</i>	“There are <u>grey</u> (ones) also.”

両親に会う	<i>ryoosin ni au</i>	“(I) meet my <u>parents</u> .”
左へ曲がる	<i>hidari e magaru</i>	“(I) turn <u>left</u> .”
弁当を買う	<i>bentoo o kau</i>	“(I) buy a <u>boxed lunch</u> .”
毎日を楽しむ	<i>mainiti o tanosimu</i>	“(I) enjoy <u>each day</u> .”
勉強をする	<i>benkyoo o suru</i>	“(I) <u>study</u> .”
地下鉄に乗る	<i>tikatetu ni noru</i>	“(I) ride the <u>subway</u> .”

Appendix D – Background Questionnaire

1. Please indicate your age: _____

2. Please indicate your gender: Male Female

3. What is your first language(s)? _____

4. What is your status at Ohio State? Please indicate the one that applies to you.

Undergraduate Student Graduate Student Other (Specify) _____

5. How many years/months have you studied Japanese? _____

6. What Japanese language class(es) are you taking now (level and name)?

7. Have you been to Japan before? Yes No

If yes, please indicate the total time you spent in Japan and the reason for the stay(s):

8. Besides Japanese, do you know any other foreign languages? Yes No

Which foreign languages and how well do you speak them?

9. How many hours per week do you spend using Japanese outside of class to:

Do homework (JSL audio/videos)	0	1-2	3-4	5-6	7+
Listen to music	0	1-2	3-4	5-6	7+
Watch TV, movies	0	1-2	3-4	5-6	7+
Use internet sites (Youtube, etc.)	0	1-2	3-4	5-6	7+
Talk to friends	0	1-2	3-4	5-6	7+

10. Do you play any musical instruments? Yes No

If yes, which instruments and how long? _____