second language research

Article

Second Language Research
2019, Vol. 35(3) 351–376

DOI: 10.1177/0267658318775143
journals.sagepub.com/home/slr
SAGE

Lexical accent perception in highly-proficient L2 Japanese learners: The roles of language-specific experience and domain-general resources

Seth J Goss
Emory University, USA

Katsuo Tamaoka
Nagoya University, Japan

Abstract

This article reports empirical findings on the roles of domain-general resources and language-specific experience in the second language (L2) acquisition of Japanese lexical pitch accent. Sixty-one advanced-proficiency L2 Japanese learners from two first languages (L1s), Mandarin Chinese and Korean, identified and categorized Japanese nouns embedded in short sentences in two aurally-presented tasks. Mixed effects models showed that although the tonal-language Chinese group outperformed non-tonal Korean speakers, L2 lexical knowledge, but not overall proficiency or learning experience, predicted performance on both perception tasks regardless of L1, suggesting that long-term knowledge of L2 phonological structure facilitates perception of lexical-level prosody. Domain-general resources, however, played no predictive role in advanced learners' accent perception. A decision-tree analysis then revealed further divergence in perception accuracy by accent pattern, L1, and task type. Taken together, the results establish a close connection between language learning experience and L2 speech perception at the advanced level, and highlight the complexity inherent in the learning of non-native prosodic categories.

Keywords

cross-linguistic perception, individual differences, Japanese, lexical accent

Corresponding author:
Seth J Goss, Department of Russian and East Asian Languages and Cultures, Emory University, Modern Languages Building, 532 Kilgo Circle, Atlanta, GA 30322, USA
Email: sethjgoss@emory.edu

I Introduction

Second language (L2) learners often struggle to accurately perceive non-native word-level prosodic features such as pitch, tone, and stress (e.g. Strange and Shafer, 2008; Wong and Perrachione, 2007). In fact, low accuracy and a wide degree of individual variation in prosodic perception ability are commonly reported even among non-novice learners who are otherwise skilled listeners (Shibata and Hurtig, 2008; Taylor, 2011). One example of a perceptually challenging prosodic characteristic is the accentual system of Tokyo Japanese, in which individual words carry a pitch pattern as part of their phonological structure. Word-level accent is an important linguistic property in Japanese because it not only marks prominent syllables and delineates prosodic groupings (Beckman and Pierrehumbert, 1986), but is also exploited by native (L1) listeners in the access and retrieval of words from the mental lexicon (Ōtake and Cutler, 1999). Yet despite its phonological status, numerous studies have shown that L2 Japanese learners find lexical pitch accent difficult both to perceive and produce, irrespective of their proficiency level, native language, or learning environment (e.g. Lee et al., 2006; Shibata and Hurtig, 2008; Shport, 2016; Taylor, 2011).

Theories of L2 speech perception have focused heavily on segmental features, and attribute performance variation to perceptual biases resulting from the relative (dis)similarity in sound categories of the L1 and L2 (e.g. Best and Tyler, 2007; Strange, 2011). That is, cross-linguistic perception models typically take specific sound categories as a central focus of comparison, and often overlook non-phonetic sources of individual variation. While it is certainly true that such analyses have shed much light on how one's L1 constrains perception of L2 sound categories (e.g. Bent et al., 2006; So and Best, 2010), little work has looked at the role that individual differences play in attaining a high level of prosodic perception ability.

L2 learners bring differences in both domain-general resources (e.g. Engel de Abreu and Gathercole, 2012; Wong and Perrachione, 2007) and language-specific experience (e.g. Andringa et al. 2012; Martin and Ellis, 2012) to the task of language acquisition. Yet the relative contribution of these differences to one's ability to perceive L2 prosodic categories is unclear. In the present study, we argue that differences in both learner-internal resources and experience-based variables can account for some of the difficulty, and wide degree of variation, reported in L2 Japanese learners' perception of lexical accent. To test this prediction, we identified four variables that potentially support lexical accent perception: two domain-general capacities, phonological short-term memory (PM) and auditory processing ability; and two variables we classified as experience-based, L2 lexical knowledge and L1 experience with lexical prosody. We first closely matched two groups of advanced L2 Japanese learners, lexical-tone Mandarin Chinese speakers and non-tone Korean speakers, on their Japanese proficiency and learning experience. The four variables were subsequently used as predictors of learners' performance on two lexical accent perception tasks. The following questions guided this study:

1. Do PM and auditory processing ability account for perceptual variation at the advanced level?
2. What role does experience with L2 phonological patterning, as indexed by L2 lexical knowledge, play in learners' perception of lexical accent?

3. Does L1 tone experience aid perception of lexical accent relative to a proficiency-matched non-tone L1 group?

II Background

1 Japanese lexical accent system

In Japanese, pitch accent is specified in the lexicon as a part of word form (Beckman and Pierrehumbert, 1986). In acoustic terms, a fall in fundamental frequency (F0) on a given mora – the timing unit that determines many phonological processes in Japanese – serves to mark a word's accent pattern as distinctive from other possible patterns (Kubozono, 2008). Acoustic analyses indicate that accented words consist of a lexically-linked H(igh) pitch followed by a drop to L(ow) pitch (Sugito, 1982). In accent-bearing words in Tokyo Japanese, this F0 fall does not vary with phonetic context, as does pragmatic focus-marking in English. For example, a three-mora Japanese noun can have *n*+1 pitch patterns, which are realized over the duration of a word, with the H tone extending into the following postposition in the unaccented pattern; see (1d) below. When presented in isolation, that is, without a postposition, patterns 1c and 1d are identical because the pitch fall is not realized until the postposition in (1c) (accented moras in bold; examples from Tanaka and Kubozono, 2012).

(1) a. *megane ga* (**ME**gane-ga; HLL-L) 'glasses + SUBJECT marker'
b. *tamago ga* (ta**MA**go-ga; LHL-L) 'egg + SUBJ'
c. *otoko ga* (oTO**KO**-ga; LHH-L) 'man + SUBJ'
d. *sakana ga* (saKANA-GA; LHH-H) 'fish + SUBJ'

Pitch is phonemic in Japanese, and it is estimated that between 10 to 14% of Japanese homophones can be distinguished by accent pattern alone (Kitahara, 2001; Shibata and Shibata, 1990). However, in this study, we are concerned with L2 learners' use of pitch as a perceptual cue to identify and categorize words by their accent pattern, rather than its narrower role in differentiating words that contrast minimally in pitch (Kubozono, 2008).

The lexically-linked status of pitch accent in Japanese has led researchers to assume that accent patterns are processed and stored as part of a word's phonological form in the mental lexicon (e.g. Beckman and Pierrehumbert, 1986; Ōtake and Cutler, 1999). 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.[1] Because pitch-accent patterns are obligatory and accent location unpredictable in Japanese, felicitous pitch patterns aid word recognition in L1 Japanese speakers (Ōtake and Cutler, 1999; Sekiguchi and Nakajima, 1999). In fact, findings from correctness judgment tasks support this hypothesis, with L1 listeners[2] reporting accuracies greater than 90% (e.g. Goss and Tamaoka, 2015; Sakamoto, 2010; Shibata and Hurtig, 2008), although wide variation has been reported in pattern categorization tasks (Shport, 2015).

2 Accent acquisition by L2 Japanese learners

As their proficiency level increases, Japanese learners make steady gains in the perception of certain linguistic features, such as phonemic vowel and consonant length contrasts (e.g. Hardison and Saigo Motohashi Saigo, 2010; Shibata and Hurtig, 2008). Yet, current findings indicate no such developmental trajectory for lexical accent perception (Ayusawa, 2003). For example, in a cross-sectional study of three proficiency groups of L1 English-speaking Japanese learners in the United States, Shibata and Hurtig (2008) found no difference between the groups on an accent perception task (Novice = 47%, Intermediate = 49%, Advanced = 56%). In contrast, the same learners' identification of phonemic consonant length contrasts (e.g. *kata* 'shoulder' vs. *katta* 'won') clearly improved by level (Novice = 45%, Intermediate = 62%, Advanced = 75%), despite this contrast not sharing phonemic status in the participants' L1. Shibata and Hurtig examined L1 English speakers, but the difficulty in accent acquisition does not appear to be limited to speakers of a non-tonal language. For example, Lee et al. (2006) employed a longitudinal design to measure the accent production ability of three Cantonese speakers over a two-year period. They found that although participants spent a year in Japan and attained a high level of Japanese proficiency, they showed no improvement in their production of lexical accent, even for high-frequency words.

Despite the reported difficulty in perceiving and producing lexical accent, some learners appear to acquire accent more easily than others. Shport (2011) noted a large degree of individual variation among Japanese-naive L1 English speakers, in her study on the initial state of accent perception. In an earlier study by Nishinuma et al. (1996), L1 English speakers who had studied Japanese for two years displayed wide variation on an accent perception task, with the lower third of learners correctly identifying accent patterns at an average accuracy of 42%, while the top third averaged 73%. Finally, Taylor (2011) examined the production ability of Japanese lexical accent in two learner groups, described as less versus more experienced in Japanese, and again found a high degree of variation, leading her to conclude that the acquisition of pitch accent by L1 English speakers is essentially a random process. Although researchers (e.g. Shport, 2011; Strange and Shafer, 2008) have suggested that factors beyond L1–L2 sound category mismatches are a likely source of individual variation in prosodic acquisition, these have not been explored systematically.

3 Learner variables in L2 speech perception

L2 speech perception involves an interplay of domain-general capacities (O'Brien et al., 2006; Wong and Perrachione, 2007), long-term knowledge of L2 phonological regularities (Speciale et al., 2004), and experience with L1 sound categories (So and Best, 2010). Although previous research has made a case for the involvement of each of these factors in L2 perception, their relative contribution to the perception of lexical prosody is inadequately characterized. We next discuss our domain-general predictors: phonological short-term memory and auditory processing ability, along with the higher-order, language-specific factors of L2 lexical knowledge and L1 tone experience.

a Phonological short-term memory. The phonological loop was proposed by Baddeley (1986; elaborated in Baddeley et al., 1998) as the component of working memory (WM) responsible for the short-term storage and maintenance of verbal input in language processing. Its link to language learning is invoked by the fact that those efficient at the temporary storage and processing of phonological input are better at ultimately transferring this input to long-term memory: a crucial step in mapping novel sounds to concepts (e.g. Cheung, 1996; Hummel, 2009; O'Brien et al., 2006; Speciale et al., 2004).

Although the link between the memory system and lexical prosody has not been examined, previous research on the role of PM in L2 acquisition provides evidence of a relationship. Speciale et al. (2004) found that PM capacity positively correlated with word learning in the early stages of L2 Spanish vocabulary acquisition, but as learning progressed, long-term knowledge of sound regularities became the main factor facilitating word learning. Additionally, Martin and Ellis (2012) reported that PM, as measured by non-word repetition performance, predicted 14% of the variance in learners' ability to acquire vocabulary in an artificial language.

In the present study, we examined the role of PM as a mediator of lexical accent perception, and a potential source of individual variation in L2 speech perception. Despite evidence for a link between PM capacity and L2 word learning ability, none of the languages examined thus far have featured pitch as a lexical property, and it remains to be seen whether prosodic perception is mediated to any extent by a learner's PM capacity. Given that pitch is a lexical property in Japanese, it is plausible that the short-term store is invoked in the brief retention of both pitch and segmental information in a perception task that requires phonological form-based judgments on lexical accent. Specifically, we assumed that PM would be implicated in a two-stage accent perception task, in which learners first judge accent correctness, then following a short pause, categorize spoken words according to visual representations of pitch contours. Since listeners must categorize words based on a memory trace of the input – with the phonological store acting as a buffer where this processing takes place – variation in performance on the delayed categorization task in particular may be modulated by a listener's PM capacity. Alternately, given evidence that PM is intimately tied to language experience, in that it may develop as a function of proficiency level (e.g. Andringa et al., 2012), it may be the case that for the highly-experienced learners in the current study, this construct is inseparable from language-specific predictors like lexical knowledge.

b Auditory processing ability. Lexical accent in Japanese is realized through variations in a single acoustic parameter, fundamental frequency (F0), and thus shares this physical property with non-linguistic or musical pitch. It has thus been argued that the same domain-general auditory processing capacity involved in the perception of non-linguistic pitch may also support the perception of lexical prosody (e.g. Asaridou and McQueen, 2013; Deutsch et al. 2009; Wayland et al., 2010; Wong and Perrachione, 2007). For instance, Wong and Perrachione (2007) suggested that L2 learners acquire lexical prosody through a bottom-up process, which first implicates domain-general auditory resources in the analysis of the phonetic properties of the target feature. They reported that L1 English listeners who were initially more sensitive to non-linguistic pitch variations (i.e. trained musicians) were better at acquiring Mandarin Chinese tones than non-musicians. In fact, initial pitch sensitivity predicted approximately half of the variance in level of attainment after a series of training sessions. Somewhat surprisingly, Goss and Tamaoka (2015) found that among L1 Japanese listeners, the ability to distinguish pairs of pure tones predicted performance on real-word accent perception tasks, suggesting that even L1 speakers are relying on auditory processing resources to categorize words by their lexical accent pattern.

The above findings suggest a link between general auditory processing and the perception and learning of the prosodic features of language. However, this runs counter to the notion that the mechanisms for auditory and speech processing are typically considered to be separate in the human perceptual system (Strange and Shafer, 2008). In general, speech perception models predict a dissociation of acoustic perception from the relative ease with which listeners perceive speech sounds in a known language (e.g. Bent et al., 2006). For example, Wayland et al. (2010) reported that although trained musicians were initially more accurate at identifying Mandarin tones, non-musicians generally narrowed this gap following a series of training tasks, suggesting that the advantage for pitch-sensitive learners may decrease inversely with experience in the target language. Yet, such conflicting findings – particularly those showing a relationship between domain-general auditory processing and the perception of lexical accent in L1 listeners – warrant further investigation in the L2 context. In the current study, we assume that F0 discrimination involves domain-general auditory processing mechanisms, and given that F0 information is relevant to lexical accent perception (e.g. Asaridou and McQueen, 2013; Wong and Perrachione, 2007), aim to measure the relationship between the two.

c L2 lexical knowledge. Lexical knowledge, or the accumulated representations of word forms in the long-term store, has been shown to underpin L2 grammar learning (Williams and Lovatt, 2003), facilitate the automaticity of lexical decisions (Segalowitz et al., 1998), and provide a basis for further L2 lexical learning (Martin and Ellis, 2012). In short, researchers widely acknowledge that L2 learners with larger vocabularies perform better on a range of tasks (e.g. Meara, 1996; Nation, 2010).

As discussed earlier, pitch accent is a lexical property present in the speech input to L2 Japanese learners, but also one that presents ongoing perceptual difficulty. If we consider the two abovementioned domain-general capacities as potentially supporting lexical processing, but still somewhat language-independent, then we can further posit that the long-term lexical store is involved in speech perception to an even greater extent. In terms of Japanese phonology, L2 lexical knowledge encompasses an understanding of the phonological regularities of the language, including segmental structure and lexical accent patterns. Accordingly, accuracy on form-based judgments, such as determining accent correctness, would entail that a word's accent pattern has been acquired and retrieved from long-term memory. Thus, if we take a correctness judgment task to be an indicator of whether a word's accent pattern has been established in long-term memory, then a measure of the size of the L2 lexicon should closely relate to the ability to judge the correctness of individual words. Put differently, lexicon size, or 'breadth' of vocabulary knowledge, is predicted to relate to lexical accent judgments, which can be taken as one measure of the 'depth' of lexical knowledge (Nation, 2010). As some have pointed out, certain measures of lexical depth may only be possible for learners who possess a substantial L2 lexicon (Pignot-Shahov, 2012). The advanced learners in the current study are assumed to know the target words well in a general sense, but insufficient depth of knowledge is likely to be found in judging and categorizing accent patterns, and incorporating L2 lexicon size as a predictor may help account for these gaps.

d L1 experience and accent perception. Learners of Japanese from both L1 Chinese and Korean backgrounds have been noted to attain a high level of Japanese proficiency in a similar length of study (Tamaoka, 2014), and can thus be matched closely on their L2 proficiency. By doing so, we can compare the influence of the prosodic systems of these languages: Mandarin Chinese being a tonal language, while the standard dialect of Korean (Seoul Korean) features a non-tonal, stress-like system of accentuation (Jun, 1998).

In both Mandarin and Japanese, tone and pitch, respectively, have phonemic status, as is illustrated by the presence of minimally contrastive pairs.[3] The frequent use of tone at the syllable/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. Seoul Korean, on the other hand, is a non-tonal language, in that pitch is not used contrastively, nor are pitch variations alone used to mark prominent syllables at the word level (Jun, 1998; see also Silva, 2016). Therefore, in the current study, we make the crucial assumption that pitch is not represented in L1 Korean speakers' mental lexicon as a part of phonological form in the same way as pitch accents are a part of L1 Japanese speakers' lexical knowledge. Namely, there is a difference in phonemic status of lexical pitch between standard Japanese and Seoul Korean, and this difference may result in decreased perceptual salience of word-level accent for Korean learners of Japanese, resulting in greater difficulty in perceiving accent correctness and location.

In the current study, rather than comparing the influence of specific L1 prosodic categories with Japanese pitch patterns, we consider L1 experience with tone as a general contributor to the acquisition of Japanese lexical accent. The presence of phonemic tone in Mandarin may enhance learners' attention to lexical accent in Japanese, making pitch variations more salient to this group (So and Best, 2010), than to L1 Korean speakers.

III Method

1 Participants

Sixty-one advanced learners of Japanese, 31 L1 Mandarin Chinese speakers (Age: M = 24.9 years, SD = 2.6) and 30 L1 Korean speakers (M = 25.6, SD = 3.8) participated in this study. All participants were recruited from the undergraduate and graduate communities at a large research university in Japan.

Participants' Japanese learning background was controlled as follows. First, the L1 groups were matched for proficiency using tests of Japanese lexical and grammatical knowledge (Miyaoka et al., 2011, 2014; details are provided below). No significant differences were found between the L1 groups on the measures of lexical (L1 Chinese: M = 37.9, SD = 4.93; L1 Korean: M = 38.7, SD = 5.87; $t(59) = 0.55$, n.s.) or grammatical (L1 Chinese: M = 31.9, SD = 3.05; L1 Korean: M = 32.6, SD = 2.67; $t(59) = 0.90$, n.s.)

knowledge. 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. Length of residence (LoR) was approximately the same as well, with Chinese participants having lived in Japan for an average of 27 months, and Korean participants for 28 months. Self-reported Japanese usage per week (hours spent speaking and listening) also indicated comparability, with L1 Chinese reporting an average of 10.6 hours compared to Koreans' 10.4 hours. The proficiency tests and self-reported measures clearly indicated the similarity of the two groups' Japanese learning experience.

2 Materials

The following tasks were presented to all participants in this study. The first three measures – F0 discrimination, serial non-word recognition, and L2 lexical knowledge – comprised the predictor variables. The accent pattern correctness judgment (PitchID), reaction time on these judgments (PitchID.RT), and the categorization (PitchCAT) task were the dependent variables.

a F0 discrimination. An adaptive pitch test, which increased in difficulty based on performance, was used to measure participants' just noticeable difference (JND) in the F0 of paired tones. This task was web-based and similar in format to commonly used AX discrimination tasks, in that the first pitch stimulus (A) remained constant, while the second stimulus (X) varied by F0 height. In this task specifically, the first tone of the two-item pair was a 500 Hz pure tone, and the second tone differed by a predetermined interval of hertz (96, 48, 24, 12, 6, 3, 1.5 Hz, etc.), with the between-stimulus Hz difference either increasing or decreasing based on the accuracy of participants' responses (see Mandel, 2009). 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. The resulting score represented the distance in Hz at which a listener could discriminate the paired tones, with a lower score indicating a lower JND for pure-tone discrimination (Jongman et al., 2017).[4] Stimuli were presented through headphones at a loudness of approximately 65 dB, a level comfortable for any normally hearing participant.

b Phonological short-term memory (PM). A serial non-word recognition (SNWR) task was chosen as the measure of phonological memory capacity. The SNWR task closely followed the design of the English syllable-based test used by O'Brien et al. (2006). In contrast with the more common non-word repetition tasks measuring PM capacity, the SNWR task requires no vocalization of the non-word stimuli, thereby removing the articulatory burden arising from speech production.

The task was presented aurally to participants and was composed of Japanese mora-based non-words adhering to the Japanese phonotactic structure. PM measures featuring non-words that are more word-like in the target language are purported to be better predictors of L2 vocabulary acquisition (Martin and Ellis, 2012). We selected an L2-based memory measure primarily because our participants lacked a common L1, but also since we assumed that our advanced-level participants were capable of handling an aurally-presented L2 memory task.

Non-word stimuli were recorded by a speaker of Tokyo Japanese in a soundproof recording booth at a sampling rate of 44.1 kHz. All of the non-words were spoken with a Low-High pitch accent pattern, so that variability in pitch would not influence task performance. We recorded the stimuli with an existent accent pattern on the assumption that pitch-carrying non-words sound more like speech than non-speech, and thus may better predict real word task performance. Audio files for the PM task were then compiled in sound-editing software as follows. For each trial, two lists consisting of the same number of non-words were created with a 1,500 ms pause separating the lists; see (2) below. Non-word lists increased from 4 items in the practice phase to 5, 6, and 7 non-words for the test phase. The inter-stimulus interval (ISI) between the non-words in each list was set to 750 ms. There were 8 trials at each of the 3 list lengths (i.e. 5, 6, and 7 non-words), yielding a total of 24 trials.

Two types of trials were then created from these lists, ‘same’ trials, meaning that the non-words in both lists were presented in an identical order, and ‘different’ trials in which the order of two of the non-words 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. The same/different sets were randomized within each set length. 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 non-words in order to make this decision. Note that participants were not required to use pitch accent to make judgments on the non-word stimuli. A weighted scoring measure was used, so that the longer the set length, the more points that were awarded for a correct response (see O’Brien et al., 2007).

(2) Two samples of 5 non-word-list length:
(Same) gohe zuka imyo hezi baro <1.5 s> gohe zuka imyo hezi baro
(Different) tida ***ruge hami*** zare kebu <1.5 s> tida ***hami ruge*** zare kebu

c Lexical and grammatical knowledge tests. Participants’ L2 lexical and grammatical knowledge were measured with two multiple-choice tests, a 48-item lexical knowledge test (Miyaoka et al., 2011) and a 36-item grammar test (Miyaoka et al., 2014). Both tests were created with vocabulary and grammar items selected from the highest two levels of the Japanese Language Proficiency Test (JLPT; Japan Foundation, 2002), and were validated in the two studies above, which used these tests to control for L2 proficiency. The current study, likewise, used the tests first as a control for proficiency level, but more importantly, used the L2 lexical knowledge test score as a predictor of lexical accent perception.[5]

d Pitch accent identification and categorization. Forty-eight nouns were selected as stimuli, and the length of the carrier sentences was set to 6–7 moras 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 moras for 4-mora targets as in 地下鉄に乗る / *tikatetu ni noru* / ‘(I) ride on the subway’ (M = 6.70 moras; SD = 0.52).[6] Half of the 48 stimuli were 3-mora nouns, and the other half 4-mora nouns. These were further divided into four accent patterns for both word lengths, yielding the following patterns for 3-mora nouns (+L/H indicates the monomoraic particle that followed each

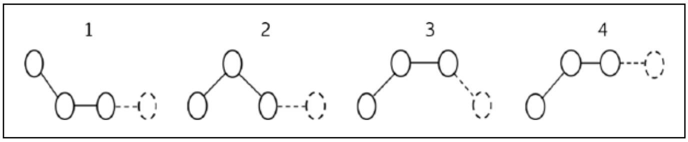

Figure 1. In the PitchCAT task, listeners categorized the noun-plus-postposition (shown in dotted lines) into one of four pitch contours by selecting a visual representation of the pitch pattern.

word): Pattern 1: HLL+L (initial-accent), Pattern 2: LHL+L (2nd-mora accent), Pattern 3: LHH+L (ultimate accent), Pattern 4: LHH+H (unaccented); and likewise for 4-mora nouns: HLLL+L, LHLL+L, LHHH+L, LHHH+H. We omitted the possible LHHL+L (penultimate) accent pattern from the tests to limit the 4-mora nouns to a total of four patterns. Additionally, 24 of the noun stimuli were spoken with the correct pitch accent and 24 with an incorrect accent pattern. The number of correctly and incorrectly accented items was balanced to control for potential response bias. Stimuli were recorded by a male native speaker of Tokyo Japanese with extensive training in accent production, and were sampled at 44.1 kHz and presented to participants with no phonetic manipulation.

Stimuli were examined for frequency for each of the four accent patterns using the NTT corpus of approximately 300 million words (Amano and Kondo, 2000). Given the participants' high proficiency level and LoS (>5 years), we assumed that lexical frequency and familiarity effects would approximate those of L1 Japanese speakers (Crossley et al., 2014). Total mean normative frequency for all stimuli was 30.3 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). Word familiarity, a subjective measure which has been shown to influence both recognition accuracy and latency (e.g. Ueno et al., 2014), was also checked for each of the stimuli. Amano and Kondo (2000) measured the combined lexical familiarity for a word's spoken and written forms on a Likert scale from 1 (not familiar at all) to 7 (very familiar). The mean familiarity rating for all stimuli was 6.19 (SD = 0.29), indicating the words were subjectively very familiar to L1 Japanese speakers. Refer to the Appendix 1 for a complete list of target words.

The PitchID and PitchCAT tasks were combined into one task using shared stimuli as follows. First, participants had to judge the correctness of a target word's accent pattern on the PitchID task, where reaction time was measured. Then, for the correctly-accented items only (n = 24), immediately following the PitchID judgment, participants categorized the noun-plus-postposition into one of four pitch contours by selecting a graph which corresponded to the pitch contour (Figure 1). Target sentences were only heard once by the listener, and importantly, participants were never required to categorize a word that was spoken with the incorrect pitch accent.[7]

3 Procedure

Data for the PitchID task were structured as 2 word lengths (3 and 4-moras) × 4 accent patterns × 6 words × 61 listeners (n = 2928), with half the number of words (i.e. correctly

Table 1. Descriptive statistics for all test scores for the L1 Chinese and Korean groups.

Variable	L1 Chinese (n = 31)			L1 Korean (n = 30)		
	Max	*M*	*SD*	Max	*M*	*SD*
F0 discrimination	–	24.66	19.65	–	9.83	11.65
PM	144	88.19	17.50	144	74.57	21.28
Lexical test	48	37.94	4.93	48	38.70	5.87
Grammar test	36	31.94	3.05	36	32.60	2.67
PitchID	48	37.16	3.03	48	28.40	5.49
PitchID.RT	–	3,545	1,442	–	3,320	1,306
PitchCAT	24	14.45	4.93	24	11.83	3.48

Notes. Max = maximum possible score; Pitch sensitivity is measured in Hz, lower mean score indicates greater sensitivity; PitchID.RT is measured in milliseconds.

accented items only) for PitchCAT (n = 1464). All tasks were presented to participants on a laptop computer through headphones in a soundproofed room, except for the two written proficiency tests. Stimuli for the PM task and the PitchID/PitchCAT perception 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 measured for the PitchID task only. The PitchCAT task employed a 4-button response in which each button corresponded to one of the pitch patterns in Figure 1. Experimental tasks were presented in the following order, with a short break following the PM task: (1) F0 discrimination, (2) PM, (3) PitchID, and (4) PitchCAT. Each task began with a brief practice phase to familiarize participants with the procedures. Due to the complexity of the PitchCAT task, participants were familiarized with sample pitch contour graphs similar to those in Figure 1 during the practice phase. All participants confirmed they were clear with the categorization procedure prior to starting the main experiment. Upon completing the tasks, the lexical and grammar knowledge tests were administered to all participants. All inclusive, the task participation time was about 1 hour 15 minutes per individual. After the test completion, participants filled out a background questionnaire and received a small payment for their participation.

IV Results

1 Comparison of L1 Chinese and Korean groups

Descriptive results for all variables are presented in Table 1. On the F0 discrimination task, the L1 Korean speakers were significantly more sensitive than L1 Chinese speakers to variations in F0 height ($t(59) = 3.25$, $p < .01$). However, L1 Chinese displayed a greater PM capacity than L1 Koreans ($t(59) = 2.74$, $p < .01$). As mentioned previously, no difference was found between L1 Chinese and L1 Korean speakers on the L2 lexical knowledge text. These three variables – F0 discrimination, PM, and L2 lexical knowledge – were subsequently examined, along with L1 background, as predictors of accent perception ability.

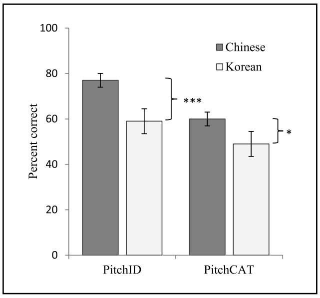

Figure 2. Mean accuracy scores on PitchID and PitchCAT tasks by L1 group.
Notes. $n = 61$; * $p < .05$; *** $p < .001$; Bars indicate standard deviations.

Figure 2 displays the accuracy scores on the PitchID and PitchCAT perception tasks separated by L1 group. Here we see a clear difference in accuracy rates between the groups on both tasks. On the PitchID task, L1 Chinese attained a mean accuracy of 77.5%, while L1 Korean displayed a mean of 59.2% ($t(59) = 7.74$, $p < .001$). The same trend was found on the PitchCAT task – in which learners matched the correctly-accented stimuli from the PitchID task with visual representations of pitch contours – with the L1 Chinese performing at a 60.2% mean accuracy and L1 Koreans at 49.3% ($t(59) = 2.39$, $p < .05$). Bear in mind that the categorization task required a 1-out-of-4 answer choice, so we consider the lower scores to be partially an outcome of task design.

2 Learner-variables and accent perception

In order to explore the relationship between our predictor variables and accent perception, we next examined correlation coefficients for the combined L1 groups ($n = 61$). We found that Japanese lexical knowledge correlated positively with both PitchID ($r = .32$; $p < .05$) and PitchCAT scores ($r = .34$; $p < .01$). Taking into account the differences in perception accuracy between the L1 groups – given that we also considered L1 knowledge an experienced-based predictor – these results suggest a relationship between the experience-based predictors and lexical accent perception. By contrast, neither PM nor F0 discrimination, the domain-general predictors, yielded significant correlations on either of the accent perception tasks. Our analysis also revealed that listeners who could accurately judge accent correctness on the PitchID task showed a strong tendency to perform well at categorization ($r = .52$; $p < .01$). Reaction time on the PitchID task yielded the only significant correlation with PitchCAT accuracy ($r = -.26$; $p < .05$), indicating that listeners who made rapid decisions on accent correctness were also better at pattern categorization.

These results suggest that along with the apparent advantage for the L1 Chinese participants, who have rich experience using tone phonemically, the size of a learner's L2 lexicon is also related to their ability to perceive lexical accent. To further explore the relation of the predictors to L2 accent perception, we next submitted the data for each of the dependent variables (PitchID, PitchID.RT, and PitchCAT) to mixed-effects analysis in R (version 3.2.0) using the lme4 package (Bates et al., 2015).

Table 2. Summary of mixed effects models for PitchID and PitchID.RT.

	Predictor	Est.	*SE*	*z*
PitchID	(Intercept)	−5.24	2.08	−2.52*
	L1	−1.70	0.29	−5.82***
	Lexical knowledge	5.99	1.14	5.21***
	PM	−1.63	0.62	−2.60**
	F0 discrimination	0.03	0.04	0.80
PitchID.RT	(Intercept)	4,449.02	3,067.23	1.45
	L1	−183.44	234.01	−0.79
	Lexical knowledge	−1931.76	1,654.35	−1.15
	PM	1,183.99	917.52	1.29
	F0 discrimination	−58.81	67.02	−0.87

Notes. $n = 61$; ***$p < .001$; **$p < .01$; *$p < .05$; Model formula: PitchID ~ L1 + LexKnow + PM + F0 + (1 | Subject) + (L1 | Item).

First, a logistic mixed-effects regression model was conducted for the PitchID data (Jaeger, 2008). The model included fixed-effect predictors for L1, Lexical knowledge, PM, and F0 discrimination. Interactions between fixed effects were also entered into the model (Barr et al., 2013). The random effects structure included random intercepts for subjects and items, and by-item random slopes for L1. Significant main effects were found for the predictors L1, Lexical knowledge, and PM. We then compared the full model, with interactions included, to one without, and found that removing the interactions yielded a better fit, as assessed by Bayesian information criteria (full model: BIC = 2,696, without interactions: BIC = 2,674; lower number indicates a preferred model). We must note that because of the correlation between the two domain-general predictors, PM ($r = .33$) and F0 discrimination ($r = .39$), and L1, these fixed effects and their interactions failed to account for meaningful variance in the model. The multicollinearity of the two predictors may also explain why PM reached significance in the model: the higher PM L1 Chinese outperformed the L1 Korean group on the PitchID task. Next, to determine if the two experience-related predictors, L1 and Lexical knowledge, improved model fit, we then compared the model with L1 removed to the full model, and found that its addition yielded a significantly better fit (χ^2 (1) = 53.741, $p < 0.001$). This was repeated with Lexical knowledge, the inclusion of which also improved the model fit (χ^2 (1) = 22.843, $p < 0.001$). Finally, marginal R^2 was calculated for the combined predictors, which accounted for 58.5% ($R^2 = .585$) of variance in correctness judgment accuracy. The results of the best-fitting model are shown in Table 2.

A second regression model was performed on the reaction times on the PitchID task (Table 2), with an identical fixed and random effects structure as above. However, this model failed to reveal any significant predictors of learners' decision latencies on the correctness judgments, suggesting the task's processing demands were heavy, and that reaction time was an unsuitable index of accent knowledge.

Finally, a logistic regression model was fitted to the PitchCAT data, with the same fixed and random effects structure as in the PitchID model. Significant main effects were found for the predictors L1 and Lexical knowledge. As with the PitchID model, the

Table 3. Summary of mixed effects model for PitchCAT.

Predictor	Est.	*SE*	*z*
(Intercept)	−7.04	3.08	−2.28*
L1	−0.63	0.28	−2.24**
Lexical knowledge	5.18	1.66	3.10**
PM	−0.43	0.93	−0.47
F0 discrimination	−0.02	0.06	−0.30

Notes. $n = 61$; ***$p < .001$; **$p < .01$; *$p < .05$; Model formula: PitchCAT ~ L1 + LexKnow + PM + F0 + (1 | Subject) + (L1 | Item).

inclusion of interactions did not improve model fit (full model: BIC = 1972, without interactions: BIC = 1953), likely due to the intercorrelation of predictors stated earlier. Comparisons again showed that models with L1 (χ^2 (1) = 4.897, $p < 0.05$) and Lexical knowledge (χ^2 (1) = 9.086, $p < 0.01$) added significantly improved fit over the reduced models without these respective predictors. Marginal R^2 for the fixed effects accounted for 22% ($R^2 = .221$) of the variance in the categorization task. The final model is shown in Table 3. To summarize, both L1 and L2 experience were significant predictors in the models that we fitted onto the accuracy data for L2 Japanese learners' accent judgment and categorization.

3 Analysis of individual accent patterns

a PitchID. Considering that L1 was a significant language-experience predictor of accuracy on both perception tasks, we next evaluated the response data by accent pattern, to further examine if L1-related differences were reflected in the identification accuracy of individual patterns. Previous research has shown perception accuracy to differ widely by accent pattern (e.g. Nishinuma et al., 1996; Toda, 2001), but cross-linguistic comparisons are lacking.

To this purpose, we conducted a decision-tree analysis (e.g. Tamaoka and Ikeda, 2010) to explore how PitchID responses diverged by both accent pattern and L1 group. Although more common in business and decision science, decision tree analyses are well-suited for analysing categorical response data and can be insightful in second language studies. In this analysis, we mapped participants' response data onto a multi-tiered chart, which is separated into branches, or nodes, according to the variables of interest. Chi-squared tests of independence were used to divide the responses into these nodes, with statistically-different variables (i.e. accent pattern or L1 group) being assigned to separate nodes. It is important to note that variables which do not differ significantly are combined into a single node in the tree, enabling easy visualization of where response data diverged.

Figure 3 shows in the uppermost node (Node #0) the total number of correct and incorrect responses made by the combined groups ($n = 61$). The next tier down displays responses by accent pattern in their order of accuracy, from left to right. Here we observe that the accent patterns were separated into three nodes, yielding the order: Pattern 4

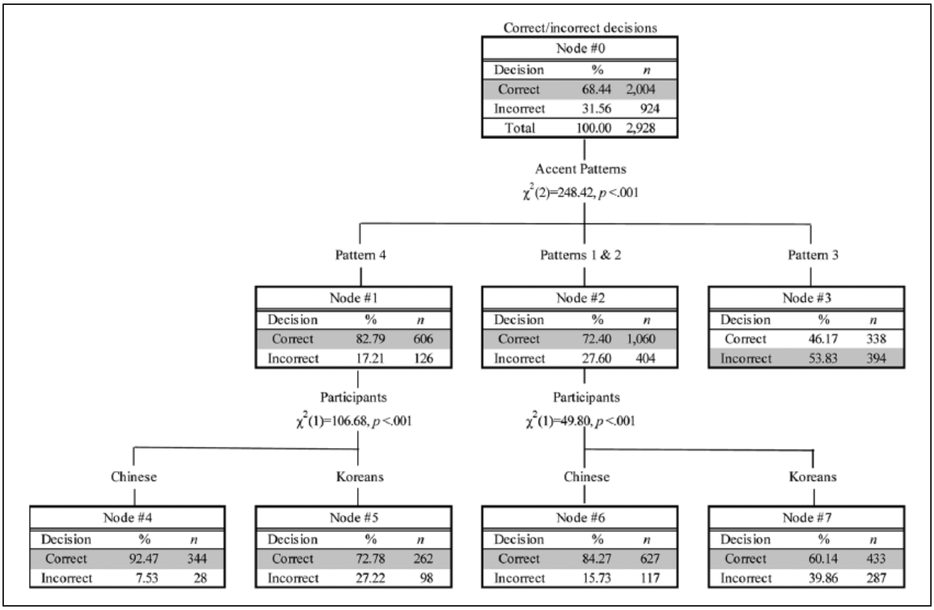

Figure 3. Decision tree analysis for the PitchID task by the L1 Chinese and Korean groups.
Notes. Pattern 1 = initially accented; Pattern 2 = 2nd-mora accent; Pattern 3 = ultimate accent; Pattern 4 = unaccented.

(82.79%) > Pattern 1 = Pattern 2 (72.40%) > Pattern 3 (46.17%). In instances where no accuracy differences were detected between patterns, data were combined as in Patterns 1 and 2 (Node #2). In the lowest tier of branches, responses are further divided by L1 group. We can thus interpret the split below Pattern 4 (Nodes #4 and #5) as reflecting a significant difference between Chinese (92.47% correct) and Korean speakers (72.78% correct) on their judgments of unaccented words, as well as in their accuracy on Patterns 1 and 2 (84.27% vs. 60.14% correct) (Nodes #6 and #7). Finally, the lack of a split below Node #3 reveals that Pattern 3 words were equally difficult for both groups.

b PitchCAT. The results of the 4-choice PitchCAT task for the L1 Chinese group are presented as a confusion matrix in Table 4. The 'Target' column on the left indicates the actual lexical pitch pattern heard by listeners, while the 'Response Percentage' row at the top displays the response listeners selected from among the four patterns. For example, on Pattern 1 targets, L1 Chinese correctly categorized 69.35% of these as Pattern 1, while only 10.22% were mistaken for Pattern 2. Contrast this with Pattern 3 targets, in which listeners only selected the correct pattern on 42.47% of the targets.

Here we see a tendency for Chinese speakers to misidentify Pattern 3 as Pattern 4 (42.47% vs. 40.32%), with a nearly equal percentage of responses made between these accent types when categorizing Pattern 3 stimuli. However, note that for Pattern 4 targets, this group was relatively accurate (69.89%). Table 5 displays L1 Korean speakers' accent categorization results. Here we observe that this group was accurate in their

Table 4. Confusion matrix for accent categorizations by the L1 Chinese group (percentages).

Target	Response percentage			
	Pattern 1	Pattern 2	Pattern 3	Pattern 4
Pattern 1	69.35	10.22	7.53	12.90
Pattern 2	5.38	59.14	19.35	16.13
Pattern 3	5.38	11.83	42.47	40.32
Pattern 4	6.99	6.99	16.13	69.89

Notes. *n* = 186 total responses per target. Shaded cells indicate correct responses.

Table 5. Confusion matrix for accent categorizations by the L1 Korean group (percentages).

Target	Response percentage			
	Pattern 1	Pattern 2	Pattern 3	Pattern 4
Pattern 1	72.78	11.67	9.44	6.11
Pattern 2	8.33	52.22	26.67	12.78
Pattern 3	9.44	23.89	46.67	20.00
Pattern 4	17.22	21.67	34.44	26.67

Notes. *n* = 180 total responses per pattern. Shaded cells indicate correct responses.

categorization of Pattern 1 stimuli, but for the remainder of the accent patterns, there was a gradual decrease in accuracy, with the lowest scores on Pattern 4.

We then created a decision tree for the categorization data from the combined Chinese and Korean groups (Figure 4). The uppermost node (#0) indicates the total correct and incorrect choices on this task. In the next tier down, the tree separates accent patterns by accuracy rates, revealing an accuracy order of Pattern 1 (71.04%) > Pattern 2 (55.74%) > Pattern 3 = Pattern 4 (46.58%). The split in Nodes 4 and 5 shows that the L1 groups differed only in their categorization of Patterns 3 and 4. The bottom-most nodes (#6–9) then reveal that Chinese listeners displayed much higher accuracy on Pattern 4 stimuli than Pattern 3, while Korean listeners were more accurate on Pattern 3 stimuli than Pattern 4, with the latter group's performance at nearly random-chance for Pattern 4 words (26.67%).

V Discussion

The present study examined how individual variation in two broad areas – domain-general resources and language-specific experience – influenced the L2 perception of Japanese lexical accent. The primary motivation for this research was to account for the wide degree of hereto unexplained variability in accent perception ability across a range of L2 Japanese learners. Our findings paint a complex picture of the variables involved in highly-proficient Japanese learners' perception of lexically-accented speech, with four

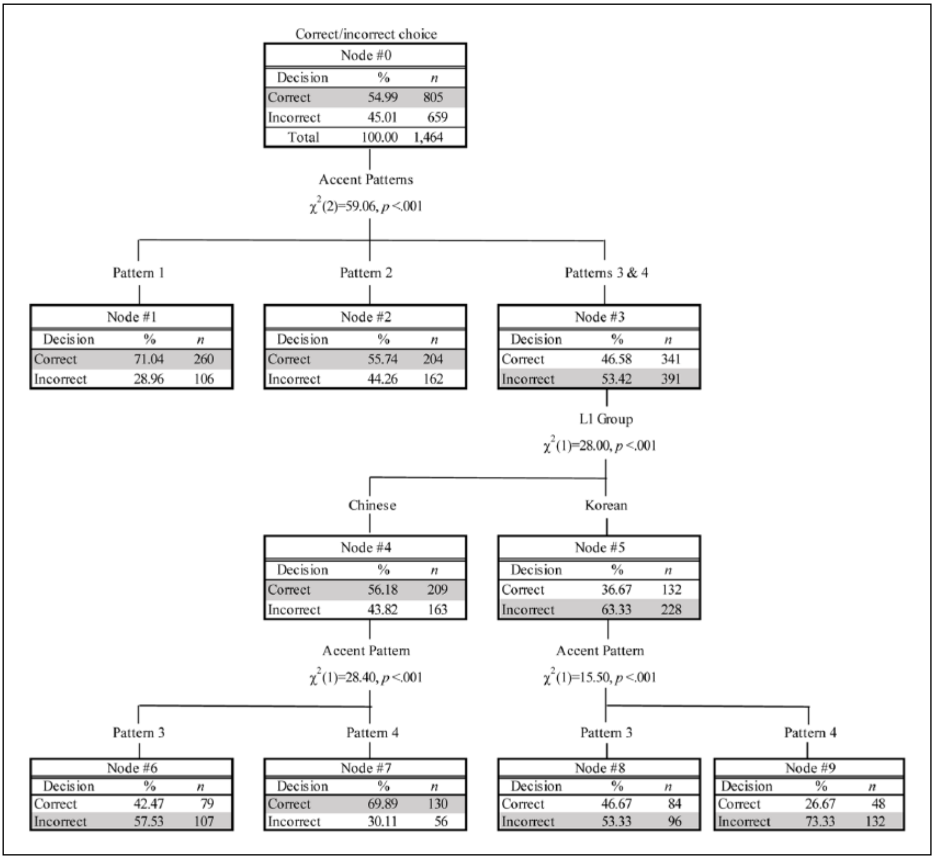

Figure 4. Decision tree analysis for the PitchCAT task by the L1 Chinese and Korean groups.

pertinent results emerging. First, Mandarin Chinese speakers' experience with L1 tones greatly facilitated the perception of lexical accent, leading to a large between-group asymmetry in perception ability unrelated to Japanese proficiency measures. Second, L2 lexical knowledge predicted performance independently of L1. Next, we found that the domain-general resources of F0 discrimination and phonological short-term memory were not implicated in the perception of lexical accent in highly-proficient learners of Japanese. Finally, decision tree analyses revealed that accuracy varied in relation to three factors: accent pattern, L1 group, and task type.

The first finding was that L1 Chinese speakers were more accurate than L1 Korean speakers at judging the correctness of accent patterns in Japanese noun-plus-postposition units (PitchID: 77.5% vs. 59.2%). A parallel result was also obtained on the four-choice pattern categorization task (PitchCAT: 60.2% vs. 49.2%). Although the two L1 groups were closely matched on language learning experience, L1 Chinese were significantly more accurate on both perception tasks. We conjecture that this is an instance of experience with an L1 phonological property facilitating L2 perception ability. Mandarin

Chinese contains lexical tone contrasts that indicate phonemic variations; for example, Tone 2 (rising) and Tone 4 (falling) have been noted to approximate the rising (L-H) and falling (H-L) pitch contours found in Japanese (So and Best, 2010). Although these pitch shapes are not phonetically identical in the two languages, and are used over different timing units (i.e. single syllable in Mandarin versus two moras in Japanese), the phonemic status of tone in Mandarin likely heightened awareness of word-level Japanese pitch accent for these speakers. Additionally, the fact that Mandarin speakers focus on multiple phonetic cues – pitch height and movement – when perceiving tones in their L1 potentially aided their perception of the pitch cues in Japanese (Gandour and Harshman, 1978). Furthermore, most of this group's errors were confined to their incorrect rejection and miscategorization of a single accent type (Pattern 3). Pattern 3 words feature an H-L pitch fall, as do Patterns 1 and 2, but it is located on the word-final mora, and its audibility is dependent on the presence of a following postposition. Although knowledge of a specific Mandarin tone (Tone 4) may have facilitated perception of the accent fall in Pattern 1 words, in Tokyo Japanese, 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 an L1 tone contour similar to the Japanese H-L pattern alone may not have aided in the perception of Patterns 2 and 3. Thus, because the presence of pattern-specific facilitation was unclear in the Chinese group, the most plausible account for their higher accuracy on non-Pattern 3 words is that experience with L1 lexical tone draws attention to the L2 accentual system in general, which in turn facilitates acquisition of the contrasting accent patterns.[8]

Although the data indicated a marked advantage for the tonal-L1 group, we found that L2 lexical knowledge was closely tied to perception accuracy in both L1 groups. The mixed-effects regression models showed that L2 lexical knowledge significantly predicted learners' ability to judge the correctness of and categorize lexical accent patterns (Tables 2 and 3). Importantly, even for Korean speakers, individuals estimated to possess a larger Japanese lexicon were more accurate on both accent perception tasks. In other words, individual variation in L2 lexical knowledge predicted accent perception ability separate of the between-group factor L1 experience. If we consider that the two groups displayed no differences in the Japanese experience measures (i.e. L2 lexical and grammatical knowledge, length of study, and length of residence in Japan), thereby ruling out the possibility that performance was a function of general L2 proficiency and exposure, we can conclude that accent perception ability is specifically tied to learners' lexical knowledge. Possessing a large L2 lexicon provides a wealth of exemplars of word form with which to compare spoken input (Nation, 2010), and in Japanese in particular, our results suggest that this predictive relationship also holds for word-level prosodic perception, which in the current tasks can be construed as a measure of depth of lexical knowledge (Meara, 1996). Furthermore, this finding ties in with theories suggesting a phonetic-to-lexical continuum of L2 acquisition, which propose that learners increasingly rely on lexical strategies, as opposed to basic perceptual resources, when processing speech input (e.g. Wong and Perrachione, 2007). This suggests that learners who possess a large L2 lexicon are more likely to have internalized accent patterns, and are thus situated at the lexical end of the acquisition continuum.

This finding also echoes what SLA researchers have said about the role of long-term memory in lexical learning by non-beginner L2 learners. Namely, that stable phonological representations take a much more prominent role in language processing as learners' proficiency develops (e.g. Andringa et al., 2012; Martin and Ellis, 2012; Speciale et al., 2004). Rather than reverting to domain-general resources such as auditory processing or PM, the advanced learners in the present study likely utilized their explicit knowledge of the Japanese accent system, in particular when categorizing words by accent pattern. The results from this experiment support the prediction that learners who possess a larger L2 lexicon, but who are not necessarily more experienced in the language, also have more robust representations of L2 accent categories, enabling greater accuracy on measures of perception ability.

Yet in contrast with the language-specific variables, the domain-general resources of F0 discrimination and PM were not significant predictors of perception ability.[9] With advanced learners, auditory processing in a pure-tone discrimination task did not mediate performance on either accent perception task. In fact, most studies examining auditory processing ability and lexical perception have focused on target-language naive (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 (Wayland et al., 2010; Wong and Perrachione, 2007; see also Goss and Tamaoka, 2015). In these studies, listeners with no knowledge of the target language perhaps had no recourse but to rely on basic perceptual resources when making decisions on Mandarin tones, since these were probably perceived as non-speech. Moreover, the discrimination task in the current study measured listeners' acuity in distinguishing Hz variations in pure tones, which although may have roughly approximated the F0 height parameter in the bitonal (L-H/H-L) pitch patterns found in Japanese, undoubtedly tapped an acoustic, rather than linguistic, mode of perception (Strange and Shafer, 2008). Furthermore, it is possible that the group difference on the F0 discrimination task (L1 Chinese: 25 Hz vs. L1 Korean: 10 Hz) may not have been perceptually relevant to F0 perception in real words, since the relationship between perceptual scales for pitch (e.g. mels) and Hz is non-linear above 500 Hz. Perhaps this can account for its failure to explain any significant variance in the regression models. In short, auditory processing ability as measured in this study was dissociated from lexical accent perception in both tone and non-tone L1 participants.

The main observation regarding memory resources and accent perception is that in highly-proficient learners, phonological short-term store capacity was no longer involved in making judgments on short spoken stimuli. On the categorization task, in particular, which was more metalinguistic in nature, we predicted that learners would rely on a memory trace of a spoken stimulus in order to compare it with schematized accent contours. However, it appears that learners were able to make these decisions by accessing word forms from their mental lexicon, rather than invoking the phonological loop for short-term maintenance of the stimuli. The fact that PM appeared to be negatively predictive in the model for correctness judgments was most likely a result of the imbalance in PM scores between the L1 groups, but may also suggest that this task measured a construct that is only indirectly related to language processing (Andringa et al., 2012). PM is generally considered to be most involved in the early stages of L2 acquisition, where it has been shown to be predictive of not only vocabulary learning, but also the development of grammar and speech fluency (e.g. Kaushanskaya, 2012; Martin and

Ellis, 2012; O'Brien et al., 2007). In future research on PM and L2 acquisition, training studies or longitudinal designs would perhaps be more suitable in evaluating the relationship between the short-term phonological store and the emergent knowledge of L2 prosodic categories.

Finally, decision-tree analyses revealed that perception accuracy differed by accent pattern, L1 group and task type. On the PitchID task, we noted difficulty on Pattern 3 words, which was especially pronounced in the L1 Chinese group, who were quite accurate on the other three patterns. L1 Korean speakers also found Pattern 3 problematic, although their errors were more diffusely spread throughout the other accent types. Why the difficulty with Pattern 3? Pattern 3 words depend on the following postposition (see Section II, Background) 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 (unaccented) words. High accuracy on Pattern 4 has been previously reported (e.g. Nishinuma et al., 1996; Toda, 2001), but only one other recent study has noted such a stark contrast between the final-accented and unaccented patterns (Shport, 2016). Perceiving the difference between these patterns requires listeners to attend to pitch beyond the word boundary, onto the following postposition, with a failure to do so resulting in the inability to distinguish these patterns.[10]

On the pattern categorization task, we again observed comparatively low accuracy on Pattern 3 in both listener groups. Mandarin speakers were again more accurate on the other accent types, as shown in the confusion matrix in Table 4, while Korean speakers made few errors with Pattern 1 categorizations, but showed a gradual decrease in accuracy across Patterns 2, 3, and 4 (Table 5). It appears that for Korean speakers, rather than a specific pattern causing categorization difficulty, any non-initially accented word resulted in increased error rates. We 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 a pitch fall was not detected on the first mora, then the word was randomly categorized as one of the other three patterns, resulting in low accuracy outside of Pattern 1.

VI Conclusions

Accurately perceiving L2 prosodic categories is often challenging for language learners even into the advanced proficiency levels. The current study sheds light on why some learners, despite very similar L2 learning experiences, are more successful than others at attuning their perceptual system to Japanese lexical accent. Our results showed that L1 experience with lexical tone afforded Mandarin Chinese speakers a marked advantage over proficiency-matched Korean speakers. When learning Japanese, Mandarin speakers place greater focus on word-level pitch variations, and are therefore more likely to process this prosodic characteristic as an integral part of lexical form. Yet by no means do we take this finding as indicative of the inability of Japanese learners from a non-tone L1 to acquire lexical accent. To the contrary, L2 lexical knowledge predicted learners' ability to accurately perceive lexically-accented stimuli, regardless of the phonological status of tone in their L1. That is, the breadth of one's L2 vocabulary knowledge, and presumably the accompanying knowledge of accent patterns as well, proved to be a critical factor underpinning L2 prosodic perception in advanced-proficiency learners. The fact that differences traceable to variation in lexical knowledge emerged in experienced learners' prosodic perception ability underscores the importance of sustained vocabulary instruction throughout the acquisition process.

Acknowledgements

We thank three anonymous reviewers for their helpful comments on earlier versions of this article. We also thank Xiang Juan Chu, Kyoko Hayakawa, Sunju Park, and Kexin Xiong for assistance with participant recruitment and data collection.

Declaration of Conflicting Interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The first author was supported by a Japan Foundation Dissertation Fellowship while conducting this research.

Notes

1. With verbs and other inflecting forms, however, accent pattern is often predictable based on the type of morphology attached to the verb, such as the negative suffix *-nai* (accented H-L).
2. These listeners were undergraduates from the Tōkai region, a standard dialect area in central Japan (Shibatani, 1990).
3. It must be noted that this lexical-distinguishing function of tone is far more common in Chinese, in which 71% of homophones are distinguished by tone (Shibata and Shibata, 1990).
4. This instrument does not generate data for each individual trial, but rather gives a single score for the entire task. We acknowledge the imprecision of the measure, but at the same time feel the need to expand the toolbox for complex research designs in L2 learning.
5. The test was designed as a measure of receptive vocabulary knowledge and can be considered as an estimate of L2 lexicon size (Nation, 2010). Vocabulary was selected from all speech categories. Distribution of lexical accent patterns and word length of the test items were not controlled. Words from the most difficult level (Level 1) of the JLPT, which constituted approximately half of the items on the current lexical knowledge measure, are from a low frequency band. Estimates suggest that learners capable of achieving a score above 70% (cut-off for passing) on this test possess a lexicon of approximately 10,000 words (Ishizaki et al., 1999).
6. Considering the substantial amount of previously-reported variation in L1 perception of lexical accent, we must note that 32 of the 48 nouns were tested on L1 Japanese speakers in a previous study by Goss and Tamaoka (2015). They found that L1 listeners judged correctness with a mean accuracy of 93%, suggesting little perceptual variation for this set of nouns. The 16 nouns added to the current study were produced by the same talker and are of similar frequency and familiarity to those used in the previous study.
7. As a reviewer pointed out, the PitchID and PitchCAT tasks likely tap different mental processes. That is, the PitchID task is a type of lexical decision task that taps language-specific knowledge, while the PitchCAT task requires more of a metalinguistic judgment, and could in theory be performed by a Japanese-naive listener.

8. Although all participants reported familiarity with the accent system, L1 Chinese who studied Japanese in their home country may have received more form-focused accent instruction than Korean speakers, the importance of which is often emphasized in Japanese courses at Chinese universities (personal communication, X. Chu).
9. We acknowledge that the variation between the L1 groups likely influenced the predictive power of the two domain-general variables. Namely, L1 Chinese showed a much greater mean PM capacity, while L1 Koreans were more sensitive to non-linguistic pitch. This unexpected 'noise' from the participant sample needs to be accounted for in future studies on individual differences.
10. As two reviewers pointed out, advanced learners may be sensitive to frequency distributions of the pattern types themselves, as has been shown for L1 Japanese (see Ueno et al., 2014). If we consider that at the 3- and 4-mora word lengths, Patterns 4 and 1 are the most frequent, then it appears that at least in light of accuracy orders on the PitchID task (Figure 3), this may indeed be the case for advanced L2 speakers as well.

References

Amano S and Kondo K (2000) *Nihongo no goitokusei – Siyoohindo* [Lexical properties of Japanese – Frequency]. Tokyo: Sanseido.

Andringa S, Olsthoorn N, van Beuningen C, Schoonen R, and Hulstijn J (2012) Determinants of success in native and non-native listening comprehension: An individual differences approach. *Language Learning* 62: 49–78.

Asaridou S and McQueen J (2013) Speech and music shape the listening brain: Evidence for shared domain-general mechanisms. *Frontiers in Psychology* 4: 1–14.

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, and Papagno C (1998) The phonological loop as a language learning device. *Psychological Review* 105: 158–73.

Barr D, Levy R, Scheepers C, and Tily H (2013) Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68: 255–278.

Bates D, Maechler M, Bolker B, and Walker S (2015) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67: 1–48.

Beckman M and Pierrehumbert J (1986) Intonational structure in Japanese and English. *Phonology Yearbook* 3: 255–309.

Bent T, Bradlow A, and 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* 23: 97–103.

Best C and Tyler M (2007) Nonnative and second-language speech perception. In: Bohn O and Munro M (eds) *Language experience in second language speech learning* (pp. 13–34). Amsterdam: John Benjamins.

Cheung H (1996) Nonword span as a unique predictor of second-language vocabulary learning. *Developmental Psychology* 32: 867–73.

Crossley S, Salsbury T, Titak A, and McNamara D (2014) Frequency effects and second language acquisition. *International Journal of Corpus Linguistics* 19: 301–32.

Deutsch D, Dooley K, Henthorn T, and 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.

Engel de Abreu P and Gathercole S (2012) Executive and phonological processes in second-language acquisition. *Journal of Educational Psychology* 104: 974–86.

Gandour J and Harshman R (1978) Cross-language differences in tone perception: A multidimensional scaling investigation. *Language and Speech* 21: 1–33.

Goss S and Tamaoka K (2015) Predicting lexical accent perception in native Japanese speakers: An investigation of acoustic pitch sensitivity and working memory. *Japanese Psychological Research* 57: 143–54.

Hardison D and Motohashi Saigo M (2010) Development of perception of second language Japanese geminates: Role of duration, sonority, and segmentation strategy. *Applied Psycholinguistics* 31: 81–99.

Hummel K (2009) Aptitude, phonological memory, and second language proficiency in nonnovice adult learners. *Applied Psycholinguistics* 30: 225–49.

Ishizaki A, Furuichi Y, and Ie M (1999) *Gokaku dekiru! 1-kyuu Nihongo noryoku shaken taisaku* [*You can pass it! Preparing for the Japanese Language Proficiency Test Level One*]. Tokyo: ALC Press.

Jaeger T (2008) Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language* 59: 434–46.

Japan Foundation (2002) *Nihongo nooryoku siken syutudai kizyun - kaiteiban* [The production standard of Japanese proficiency tests – Revised version]. Tokyo: Japan Foundation.

Jongman A, Qin Z, Zhang J, and Sereno J (2017) Just noticeable differences for pitch direction, height, and slope for Mandarin and English listeners. *The Journal of the Acoustical Society of America* 142: EL163.

Jun S (1998) The accentual phrase in the Korean prosodic hierarchy. *Phonology* 15: 189–226.

Kaushanskaya M (2012) Cognitive mechanisms of word learning in bilingual and monolingual adults: The role of phonological memory. *Bilingualism: Language and Cognition* 15: 470–89.

Kitahara M (2001) *Category structure and function of pitch accent in Tokyo Japanese*. Unpublished doctoral dissertation. Indiana University, Bloomington.

Kubozono H (2008) Japanese accent. In: Miyagawa S and Saito M (eds) *The Handbook of Japanese Linguistics* (pp. 165–91). Oxford: Oxford University Press.

Lee W, Murashima K, and Shirai Y (2006) Prosodic development in the acquisition of Japanese: A longitudinal study of three native speakers of Cantonese. *Japan Journal* 10: 38–51.

Mandel J (2009) *Adaptive pitch test*. Available at: http://tonometric.com/adaptivepitch (accessed April 2018).

Martin K and 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.

Meara P (1996) The dimensions of lexical competence. In: Brown G, Malmkjær K, and Williams J (eds) *Performance and competence in second language acquisition* (pp. 35–52). Cambridge: Cambridge University Press.

Miyaoka Y, Tamaoka K, and Sakai H (2011) Nihongo goi tesuto no kaihatsu 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]. *Hiroshima Keizai Daigaku Kenkyu Ronshu* [*Research Journal of Hiroshima University of Economics*] 34: 1–18.

Miyaoka Y, Tamaoka K, and Sakai H (2014) Nihongo bunpo noryoku tesuto no kaihatsu to shinraisei: Nihongo gakushusha no deta niyoru tesuto hyoka [Developing Japanese grammar test and its reliability: Test evaluation by data taken from Japanese learners]. *Hiroshima Keizai Daigaku Kenkyu Ronshu* [*Research Journal of Hiroshima University of Economics*] 36: 33–46.

Nation I (2010) *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Nishinuma Y, Arai M, and Ayusawa T (1996) *Perception of tonal accent by Americans learning Japanese*. In: *Proceedings of the 4th International Conference on Spoken Language Processing: Volume 1* (pp. 646–49). Philadelphia, PA: IEEE.

O'Brien I, Segalowitz N, Collentine J, and 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, and Collentine J (2007) Phonological memory predicts second language oral fluency gains. *Studies in Second Language Acquisition* 29: 557–581.

Ōtake T and Cutler A (1999) Perception of suprasegmental structure in a nonnative dialect. *Journal of Phonetics* 27: 229–53.

Pignot-Shahov V (2012) Measuring L2 receptive and productive vocabulary knowledge. *University of Reading Language Studies Working Papers* 4: 37–45.

Sakamoto E (2010) An investigation of factors behind foreign accent in the L2 acquisition of Japanese lexical pitch accent by adult English speakers. Unpublished doctoral dissertation, University of Edinburgh, UK.

Segalowitz S, Segalowitz N, and Wood A (1998) Assessing the development of automaticity in second language word recognition. *Applied Psycholinguistics* 19: 53–67.

Sekiguchi T and Nakajima Y (1999) The use of lexical prosody for lexical access of the Japanese language. *Journal of Psycholinguistics* 28: 439–54.

Shibata T and Hurtig R (2008) Prosody acquisition by Japanese learners. In: Han Z (ed.) *Understanding second language process* (pp. 176–204). Clevedon: Multilingual Matters.

Shibata T and 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]. *Keiryoo Kokugo-gaku* [*Mathematical linguistics*] 17: 317–27.

Shibatani M (1990) *The languages of Japan*. Cambridge: Cambridge University Press.

Shport I (2011) Cross-linguistic perception and learning of Japanese lexical prosody by English listeners. Unpublished doctoral dissertation, University of Oregon, Eugene, OR, USA.

Shport I (2015) Perception of acoustic cues to Tokyo Japanese pitch-accent contrasts in native Japanese and naive English listeners. *Journal of the Acoustical Society of America* 138: 307–18.

Shport I (2016) Training English listeners to identify pitch-accent patterns in Tokyo Japanese. *Studies in Second Language Acquisition* 38: 738–68.

Silva D (2016) Acoustic evidence for the emergence of tonal contrast in contemporary Korean. *Phonology* 23: 287–308.

So C and Best C (2010) Cross-language perception of non-native tonal contrasts: Effects of native phonological and phonetic influences. *Language and Speech* 53: 273–93.

Speciale G, Ellis N, and Bywater T (2004) Phonological sequence learning and short-term store capacity determine second language vocabulary acquisition. *Applied Linguistics* 25: 293–321.

Strange W (2011) Automatic selective perception of first and second language speech: A working model. *Journal of Phonetics* 39: 456–66.

Strange W and Shafer V (2008) Speech perception in second language learners: The re-education of selective perception. In: Hanson Edwards JG and Zampini ML (eds) *Phonology and second language acquisition* (pp. 193–218). Amsterdam: John Benjamins.

Sugito M (1982) Tōkyō akusento ni okeru 'hana' to 'hana' no sēsē to chikaku [The production and perception of 'flower' and 'nose' with Tokyo accent]. *Nihongo akusento no kenkyū* [*Studies on Japanese accent*] (pp. 182–201). Tokyo: Sanseido.

Tamaoka K (2014) The Japanese writing system and lexical understanding. *Japanese Language and Literature* 48: 431–71.

Tamaoka K and Ikeda F (2010) Whiskey, or Bhiskey?: Influence of first-element and dialect region on sequential voicing of shoochuu. *Gengo Kenkyuu* 138: 65–80.
Tanaka S and 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: 146–58.
Toda T (2001) Hatsuon 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.
Ueno T, Saito S, Saito A, Tanida Y, Patterson K, and Lambon R (2014) Not lost in translation: Generalization of the Primary Systems Hypothesis to Japanese-specific language processes. *Journal of Cognitive Neuroscience* 26: 433–46.
Wayland R, Herrera E, and Kaan E (2010) Effects of musical experience and training on pitch contour perception. *Journal of Phonetics* 38: 654–62.
Williams J and Lovatt P (2003) Phonological memory and rule learning. *Language Learning* 53: 67–121.
Wong P and Perrachione T (2007) Learning pitch patterns in lexical identification by native English-speaking adults. *Applied Psycholinguistics* 28: 565–85.

Appendix 1. Target words and postpositions (underlined) with carrier sentences used in the PitchID and PitchCAT tasks.

医学を学ぶ	*igaku o manabu*	'(I) study medicine'
めがねをかける	*megane o kakeru*	'(I) put on glasses'
フォークで食べる	*fooku de taberu*	'(I) eat with a fork'
おもちゃで遊ぶ	*omotya de asobu*	'(I) play with a toy'
八時に起きる	*hatizi ni okiru*	'(I) wake up at 8 o'clock'
中身を見る	*nakami o miru*	'(I) look at the contents'
ななめにする	*naname ni suru*	'(I) turn it sideways'
言葉にする	*kotoba ni suru*	'(I) put it into words'
夜中に起きる	*yonaka ni okiru*	'(I) wake up at midnight'
いなかに住む	*inaka ni sumu*	'(I) live in the countryside'
ハガキを出す	*hagaki o dasu*	'(I) send a postcard'
手前に引く	*temae ni hiku*	'(I) pull it toward me'
来月に行く	*raigetu ni iku*	'(I) go next month'
湖でおよぐ	*mizuumi de oyogu*	'(I) swim in the lake'
将来を考える	*syoorai o kangaeru*	'(I) think of the future'
あさってまで待つ	*asatte made matu*	'(I) wait until the day after tomorrow'
八月になる	*hatigatsu ni naru*	'It becomes August'
オレンジを食べる	*orenzi o taberu*	'(I) eat an orange'
九日に行く	*kokonoka ni iku*	'(I) go on the 9th'
成績をつける	*seiseki o tukeru*	'(I) assign grades'
正月を楽しむ	*syoogatu o tanosimu*	'(I) look forward to New Year's Day'
地下鉄に乗る	*tikatetu ni noru*	'(I) ride the subway'
専門を学ぶ	*senmon o manabu*	'(I) study for my major'

(Continued)

Appendix 1. (Continued)

土曜日に行く	*doyoobi ni iku*	'(I) go on Saturday'
家内に話す	*kanai ni hanasu*	'(I) speak to my wife'
技術をみがく	*gijutsu o migaku*	'(I) polish my skills'
荷物を運ぶ	*nimotu o hakobu*	'(I) carry luggage'
あなたがいる	*anata ga iru*	'You are there'
ハサミがある	*hasami ga aru*	'There are scissors'
刺身を食べる	*sasimi o taberu*	'(I) eat sashimi'
昼間に行く	*hiruma ni iku*	'(I) go in the afternoon'
娘がいる	*musume ga iru*	'(I) have a daughter'
男がいる	*otoko ga iru*	'There is a man'
左に曲がる	*hidari ni magaru*	'(I) turn left'
昔に戻る	*mukasi ni modoru*	'(I) return to the past'
仕事を探す	*sigoto o sagasu*	'(I) look for a job'
太陽がのぼる	*taiyoo ga noboru*	'The sun rises'
タクシーに乗る	*takusii ni noru*	'(I) get in a taxi'
マンションに住む	*mansyon ni sumu*	'(I) live in an apartment'
黒板に書く	*kokuban ni kaku*	'(I) write on the blackboard'
飛行機に乗る	*hikooki ni noru*	'(I) ride in an airplane'
てぶくろを買う	*tebukuro o kau*	'(I) buy gloves'
弟がいる	*otooto ga iru*	'(I) have a younger brother'
妹がいる	*imooto ga iru*	'(I) have a younger sister'
二日目になる	*hutukame ni naru*	'It becomes the 2nd'
六月になる	*rokugatu ni naru*	'It becomes June'
夕方になる	*yuugata ni naru*	'It becomes evening'
独身になる	*dokusin ni naru*	'(I) become single'