The third language acquisition of Japanese pitch accent

Junyu Wu, Martin Desmarais, and John Archibald
Department of Linguistics, University of Victoria

L3 Workshop on Multilingual Language Acquisition, Processing and Use
Chuo University
Tokyo, Japan
May 2024

University of Victoria

University of Victoria

Outline

1. Aim of our research
2. Background
3. Hypotheses
4. Methods
5. Results
6. Discussion
7. Conclusion

Question time

University of Victoria

1 Aim of our research

We are investigating the L2 (L1 English-L2 Japanese) and L3 (L1 Mandarin-L2 English-L3 Japanese) perception of Japanese pitch accent (Tokyo dialect).

Japanese is a pitch accent language (Kawahara, 2015) where we find minimal pairs distinguished by the placement of accent as shown in (1) and (2).

(1)

/ame+ga/ (unaccented) ‘candy+nominative’

/áme+ga/ (initial accented) ‘rain+nominative’

(2)

/haná+ga/ (final accented) ‘flower+ nominative’

/hana+ga/ (unaccented) ‘nose+ nominative’

University of Victoria

2 Background

2.1 Background literature

- There have been several studies on L2A of pitch accent (Goss, 2020; Goss & Tamako, 2019; Muradás-Taylor, 2022)

- Previous work has shown that L1 English listeners can have difficulty phonetically discriminating Japanese pitch accent when the task focusses on phonetic F0 discrimination (Goss, 2020; Shport, 2016).

- However, with exposure L2 Japanese phonetic discrimination ability improves (Wu, Kawase & Wang, 2017)

- Goss & Tamako (2019) showed that lexical knowledge played a larger role than domain-general factors such as auditory processing ability in accounting for accuracy.

Phonological Processing

- Shport (2016) used an AX discrimination task with an inter-stimulus interval (ISI) of 1000ms (which primes phonological discrimination);
  - L1 Japanese listeners achieved 85% accuracy while L1 English listeners 72% accuracy.
- Hirano-Cook (2011) also showed high accuracy (87%) on an AX task with a 7 second ISI.

2.2. Linguistic properties of the three groups

Typology

Our languages of interest vary typologically and prosodically:

Mandarin: A Sino-Tibetan tone language.

English: A Germanic stress language.

Japanese: An Altaic pitch accent language.

It seems unlikely that typology will be the explanatory factor of our behavioural results.

2.2. Linguistic properties of the three groups

Prosody

Mandarin Tone to Weight Principle
T3 syllables are Light (monomoraic) while T1, T2 and T4 are Heavy (bimoraic) (Qu, 2013) .

English assigns stress with weight-sensitive trochaic feet (Dresher & Kaye, 1990); heavy syllables attract stress

Japanese assigns pitch accent via weight-sensitive trochaic feet (Kawahara, 2015); heavy syllables attract pitch accent

2.2. Linguistic properties of the three groups

Mandarin Tones

Tone 1: High (μμ)
Tone 2: Low High (μμ)
Tone 3: Low (μ)
Tone 4: High Low (μμ)

Tone 0: neutral tone; phonetic realization determined by preceding tone

2.2. Linguistic properties of the three groups

English Stress Rules

Heavy penults attract stress.

aróma
agénda
horízon

Otherwise antepenultimate

cínema
cábinet
vénison

"English constructs binary iterative moraic trochees starting from the right edge of a PWd, with final syllable extrametricality set to Yes and End-Rule set to Right" (Özçelik, 2021)

This is basically the Latin Stress Rule (Hayes, 1995).

University of Victoria

2.2. Linguistic properties of the three groups

Japanese Pitch Accent Rules

Heavy penults attract pitch accent in loanwords:

/o.re’n.zi/ ‘orange’
/ro’n.don/ ‘London’
/o.ha’i.o/ ‘Ohio’

Otherwise antepenultimate:

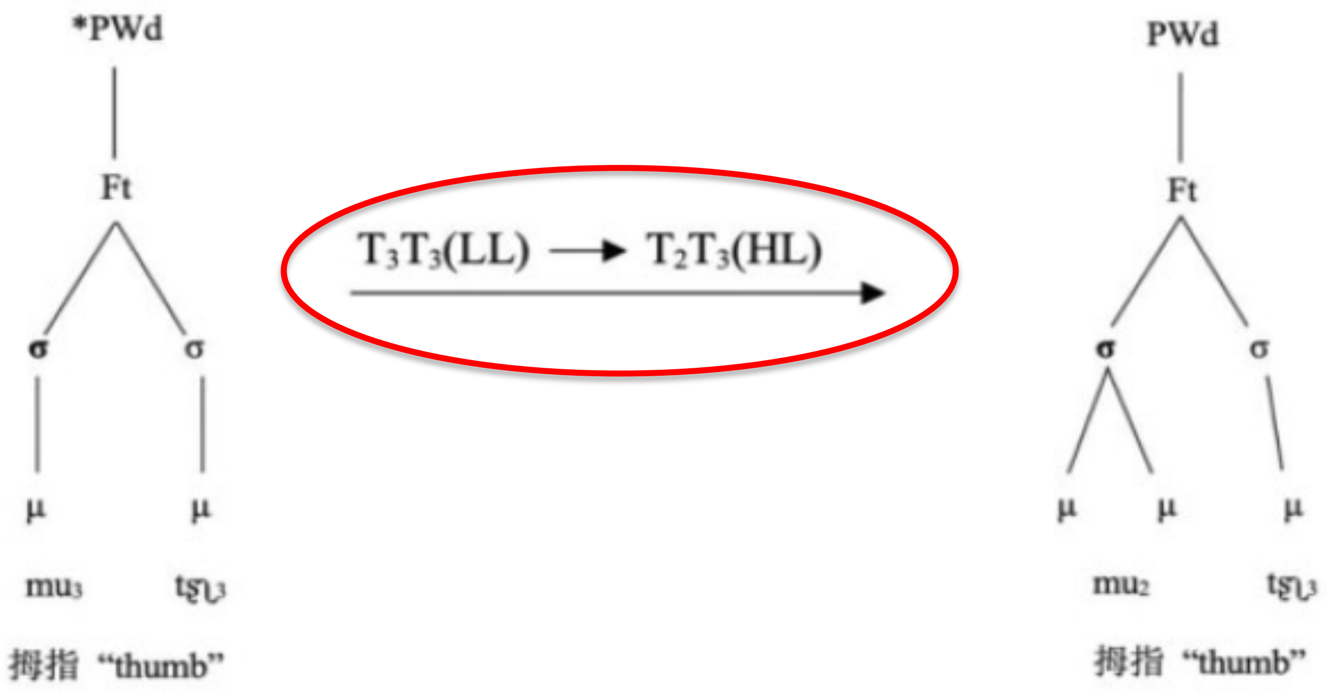
/do’.ra.gon/ ˜ /do.ra’.gon/ ‘dragon’
/ma’.zi.syan/ ˜ /ma.zi’.syan/ ‘magician’
/a’.ma.zon/ ‘Amazon’

This is basically the Latin Stress Rule (Hayes, 1995; Kawahara, 2015)

2.3. Weight and Moras

All three languages have moras which determine syllabic weight. English and Japanese weight is determined by segmental/syllabic properties while Mandarin weight is determined by tone.

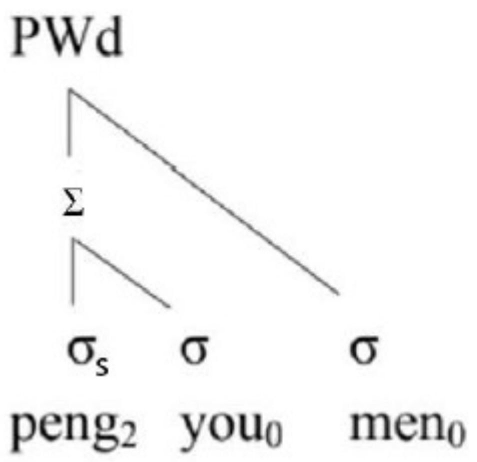
Mandarin also has an allophonic sandhi process in which a T3 T3 sequence (*LL) turns the first T3 into a Heavy T2 syllable thus producing the optimal trochaic form HL.



(*Note*. L stands for light, and H stands for heavy.)

University of Victoria

Mandarin feet are strong on the left (marked with an *s* on the stronger), making them trochaic as shown in the structure for the word ‘friends’.



Note the unfooted final syllable is an enclitic adjoined to the Prosodic Word. Note also that it is the weightless Tone 0 (Depuydt, 2022).

2.4 What is to be acquired in target Japanese?

Initial Accent	*Final Accent*	*Unaccented*
ká t a–g a H L	k a t á–g a L H L	a me g a L H

Phonetically, Japanese pitch accent is realized by a sharp pitch fall (Wiener & Goss, 2018). Accented words have an abrupt F0 fall; unaccented words don’t have the F0 fall.

3 Hypotheses

Hypothesis 1:

Mandarin speakers (L1 Mandarin-L2 English- L3 Japanese) can parse a Japanese initial-accent word with Tone 4 (HL), and an unaccented word with Tone 2 (LH), but they would have to use a Tone 3 *sandhi* pattern (an allophonic representation) (LHL) to parse the Japanese final-accent words, which would diminish accuracy of the L3A group on the final-accented words.

A Mandarin Parse of a Japanese Initial Pitch Accent

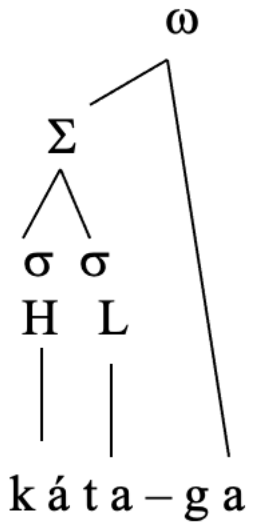
Input: HLL on *kataga*

L1 Match: T4 (HL)

Initial Parse: H L

ka ta ga

Prosodification



Assumption: the unfooted syllable is parsed as Tone 0, adjoined to the Prosodic Word; the preceding tone spreads.

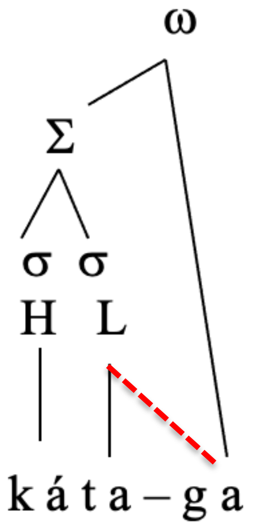
A Mandarin Parse of a Japanese Initial Pitch Accent

Input: HLL on *kataga*

L1 Match: T4 (HL)

Initial Parse:

Prosodification



Assumption: the unfooted syllable is parsed as Tone 0, adjoined to the Prosodic Word; the preceding tone spreads.

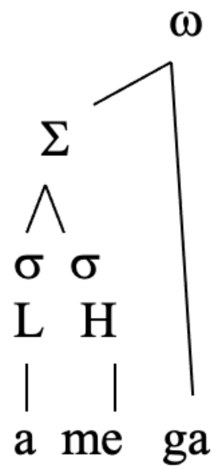
A Mandarin Parse of a Japanese Unaccented Pitch Accent

Input: LHH on *amega*

L1 Match: T2 (LH)

Initial Parse:

Prosodification



Assumption: the unfooted syllable is parsed as Tone 0, adjoined to the Prosodic Word; the preceding tone spreads.

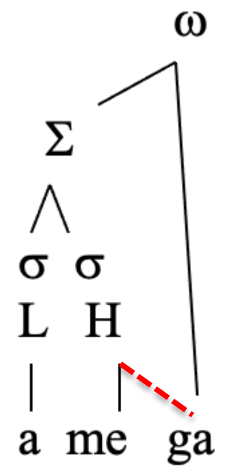
A Mandarin Parse of a Japanese Unaccented Pitch Accent

Input: LHH on *amega*

L1 Match: T2 (LH)

Initial Parse:

Prosodification



Assumption: the unfooted syllable is parsed as Tone 0, adjoined to the Prosodic Word; the preceding tone spreads.

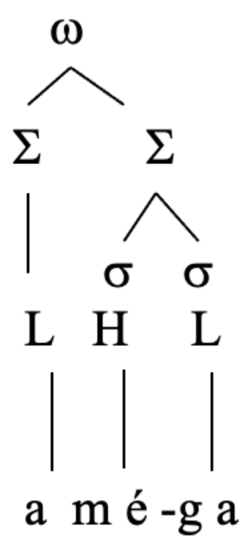
A Mandarin Parse of a Japanese Final Pitch Accent

Input: LHL on *amega*

L1 Match: T3 sandhi

Initial Parse:

Prosodification



Assumption: the initial Low syllable is parsed as a degenerate foot.

The significantly diminished accuracy of the L3A group on the final-accented words is because they would have to use a Tone3 *sandhi* pattern (LHL) to parse the Japanese input .

See Farris Trimble and Tessier, (2019), who show that derived forms (such as Canadian Raising in words like ‘write’ and ‘out’) are slower to be processed than transparent forms without Raising.

‘write’ /rajt/ ⟶ [rʌjt] versus ‘ride’ /rajd/ ⟶ [rajd]

Since both a LH and a HL parse would fail to account for the input, the learners resort to the sandhi (LHL) parse.

University of Victoria

Hypothesis 2:

English speakers (L1 English- L2 Japanese) will have difficulty with both tasks (initial vs unaccented; final vs unaccented) because of an English grammar inability to parse the Japanese unaccented forms.

English allows the following stress patterns:

´ ˘ ˘ as in *cinema* — The first matches initial pitch accent.

˘ ´ ˘ as in *agenda* — The second matches final pitch accent.

But English has no forms of the type ˘ ´ ´

So the Japanese unaccented pattern (L H H) will not be parsed by English stress parameters.

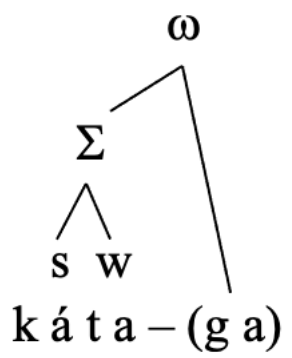
An English Parse of a Japanese Initial Pitch Accent

Input: HLL on *kataga*

L1 Match: all light syllables (= cínema)

Initial Parse: s w
| |
kataga

Prosodification



The final syllable will be extrametrical.

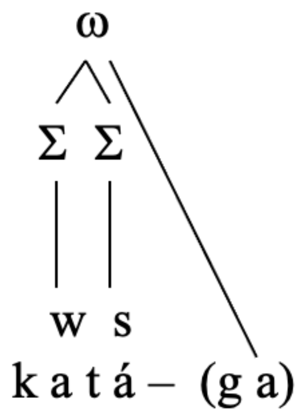
An English Parse of a Japanese Final Pitch Accent

Input: LHL on *kataga*

L1 Match: heavy penult (= agénda)

Initial Parse: s
|
kataga

Prosodification



The final syllable will be extrametrical.

University of Victoria

An English Parse of a Japanese Unaccented Pitch Accent

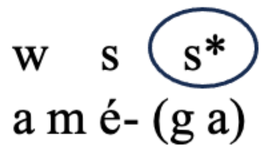
Input: LHH on *amega*

L1 Match: None

Initial Parse: s
|
amega

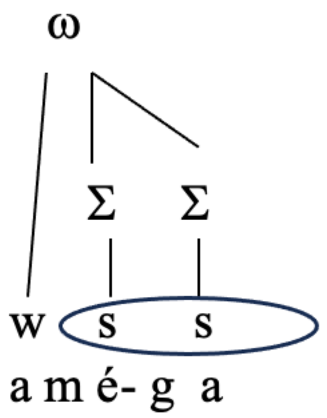
Prosodification

Option 1



L1 does not allow stressing an extrametrical syllable.

Option 2



L1 does not allow stress clash violations.

University of Victoria

Hypothesis 3: If English parsing (in the trilinguals) predominates then initial- and final-accented items should be inaccurate because of inability to parse unaccented words. If Mandarin parsing (in the trilinguals) predominates then initial pitch accent should be more accurate than final pitch accent.

4 Method and data collection

The ABX discrimination task was conducted by arranging initial vs unaccented words or final vs unaccented words in the orders of ABB, BAA, ABA, BAB . i.e. A. amega B. ámega X?

The selection of words is based on previous studies (Sugiyama, 2006; Muradás-Taylor, 2022) and discussion with Japanese instructors at UVic.

Stimuli were recorded by two Japanese speakers in a sound-isolated recording booth. Recordings were normalized for amplitude and spliced into separate sound files using Praat.This yielded a total of 100 trials.

The inter-stimulus interval (ISI) is 1500ms, triggering a phonemic level of perception (Werker & Logan, 1985).

The word list

Unaccented	Initial accented
/ame+ga/ candy+NOM	/a’me+ga/ rain+NOM
/sake+ga/ alcohol+NOM	/sa’ke+ga/ salmon+NOM
/kaki+ga/ persimmon+NOM	/ka’ki+ga/oyster+NOM
/aki+ga/ availability+NOM	/a’ki+ga/ autumn+NOM
/kami+ga/ paper+NOM	ka’mi+ga/ god+NOM
/ashi+ga/ foot+NOM	/a’shi+ga/ reed+NOM
/niwa+ga/ yard+NOM	ni’wa+ga/ second chapter+NOM
/kaen+ga/ flame+NOM	/ka’en+ga/ orchard+NOM
/eigo+ga/ English+NOM	e’igo+ga/ guarding+NOM
/hon+ga/ translation+NOM	ho’n+ga/ book+NOM

Unaccented	Final accented
/hasi + ga / edge + NOM	/hasi’ + ga/ bridge + NOM
/hana + ga/ nose + NOM	/hana’ +ga/ flower + NOM
/yuki + ga/ going + NOM	/yuki’ + ga/ snow+ NOM
/hane + ga/ feather + NOM	/ hane’ + ga/ jump + NOM
/shitu + ga/ quality + NOM	/ shitu’ + ga/ room+ NOM
/hachi + ga/ bee + NOM	/hachi’ + ga/ eight + NOM
/nori + ga/ ride + NOM	/nori’ + ga/ seaweed + NOM
/osu’ + ga / male + NOM	/osu + ga/ vinegar + NOM
/take + ga/ bamboo + NOM	/take’ + ga / length + NOM
/shita + ga/ below + NOM	/shita’ + ga/ tongue + NOM

We included the –*ga* suffix because final accent is only realized before certain morphemes, and the subject marker –*ga* is one of them (Kawahara, 2013).

University of Victoria

Participants

(1) L1 Mandarin; L2 English; L3 Japanese (n=23)
(2) L1 English; L2 Japanese (n=20)
(3) Japanese natives (NS) (n=21)

The two learner groups were at intermediate to advanced levels of Japanese proficiency as measured by a self-rated background questionnaire (their course level, JLPT scores)

The Mandarin speakers' L2 English proficiency level was measured by IELTS (average score 7.0).

Statistical analysis

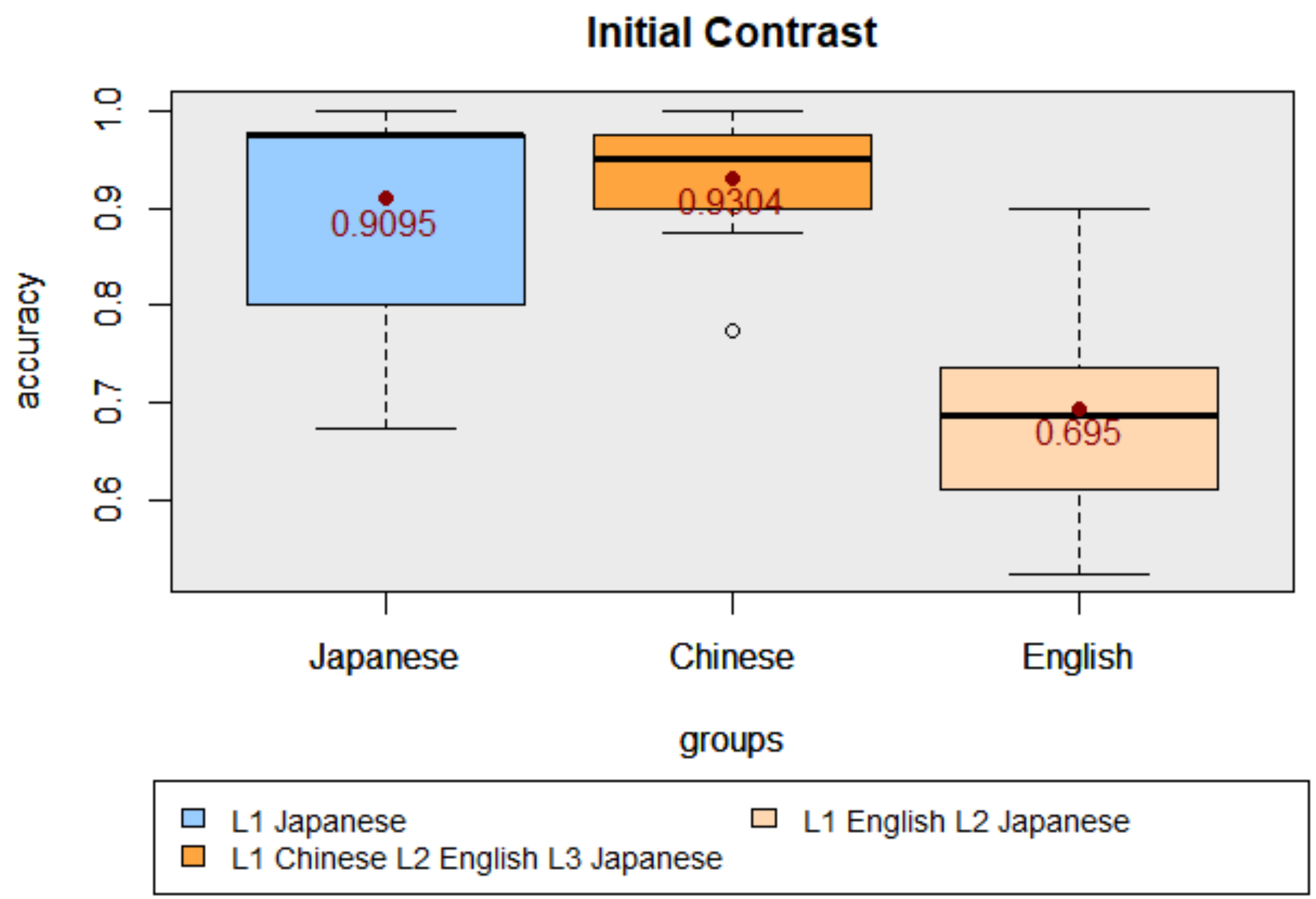
A linear mixed effect model was run with ‘subject’ (participants) as a random effect and ‘pitch accent placement (initial vs final)’ and 'language background' as fixed factors.

Under this model, the emmeans function was used to conduct a post-hoc test for pairwise comparison that drives the significant effects in the data.

University of Victoria

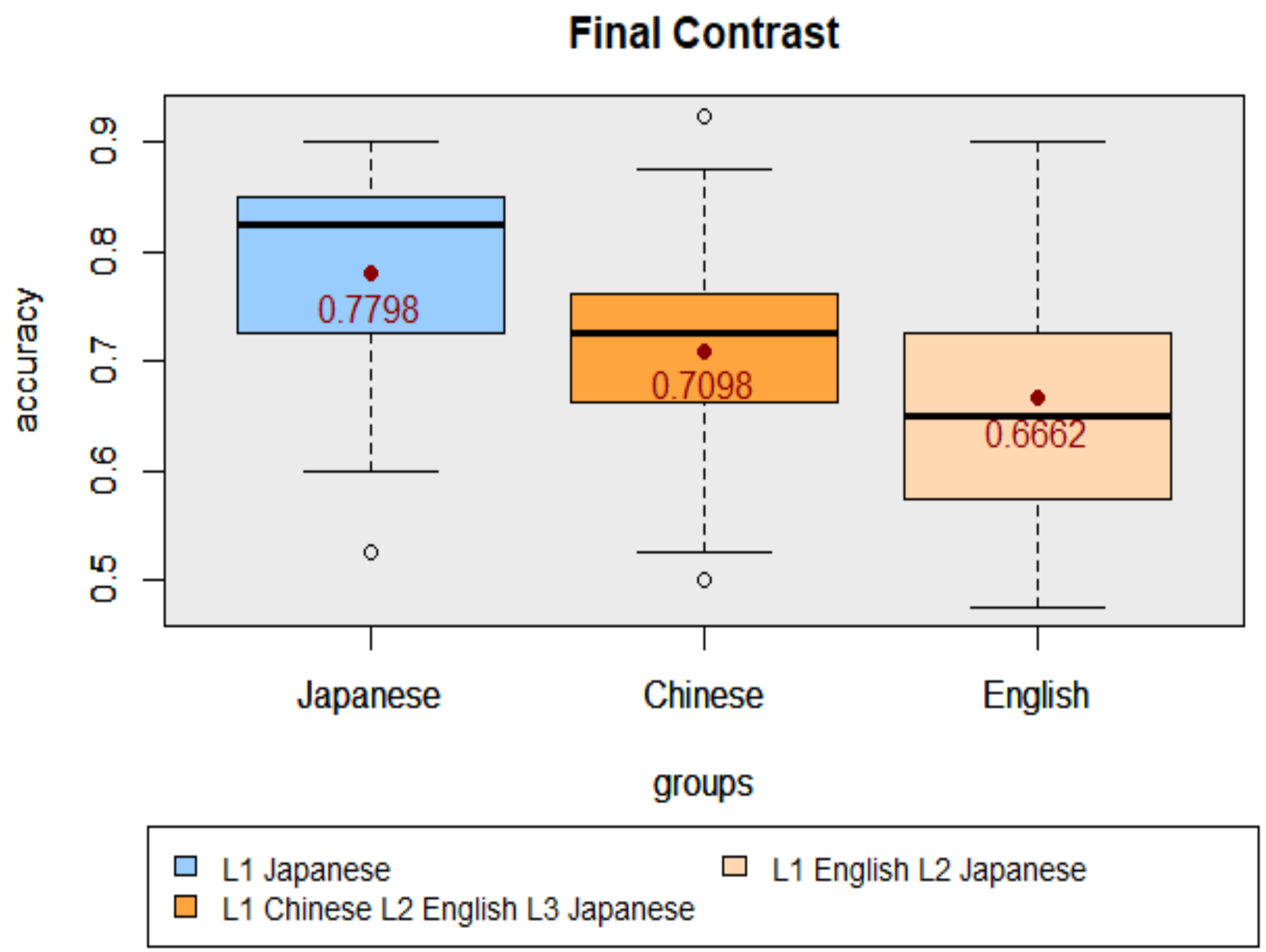
5 Results

5.1 Initial accented vs unaccented words



University of Victoria

5.2 Final accented vs unaccented words



University of Victoria

Table 1 Accuracy scores and p-values

Participants	Initial	Final	p-value
L1 Mandarin-L2 English-L3Japanese	93% (37.2/40)	71% (27.8/40)	$p < 0.001$*
L1 English-L2 Japanese	70% (28.4/40)	67% (26.7/40)	$p=0.176$
Japanese natives	91% (36.4/40)	78% (31.2/40)	$p < 0.001$*

Table 2 Pairwise Comparison (p-value <0.05)

Group comparison	Initial	Final
L3 Japanese vs L2 Japanese	$p < 0.001$*	p=0. 35
L3 Japanese vs Japanese natives	p=0.74	p= 0.07
L2 Japanese vs Japanese natives	$p < 0.001$*	$p < 0.002$*

University of Victoria

Summary of results

1. The L3A group was not significantly different from the L1A group on initial-accent words.
2. The L3A group was significantly more accurate on initial-accent compared to final-accent.
3. By contrast, the L2A group treated initial- and final-accented forms indistinguishably, and they were significantly different from the L1A group on both initial- and final- accented forms.

Discussion

1. Mandarin redeployment vs English redeployment

We argue that the L3A initial-form accuracy is because the Mandarin grammar can parse a Japanese initial-accent word with Tone 4 (HL), and an unaccented word with Tone 2 (LH). The significantly diminished accuracy of the L3A group on the final-accented words is because they would have to use the allophonic Tone 3 *sandhi* pattern (LHL) to parse the Japanese input.

The L2A difficulty in parsing initial (and final) vs unaccented forms is because of the English grammar inability to parse Japanese unaccented forms: The Japanese unaccented pattern (L H H) cannot be parsed by English stress parameters.

University of Victoria

2. L1 group's low accuracy on final-accent words

Japanese native speakers' lower accuracy on final-accent forms stems from the fact that in isolation the pronunciation of finally accented words and unaccented words is phonetically very similar (Vance, 1995; Kawahara, 2013).

The –*ga* clitic is phonologically neutral so sometimes it's High and sometimes it's Low so this would be a more difficult task.

Both forms start with an LH sequence, and the listeners' predictions/expectations as to what comes next may be variable.

University of Victoria

3. Functional load

Input effects

English pitch cues carry a relatively low functional load (Van Lancker, 1980) in that they are informative for the identification of less than 1% of English words (Shibata & Shibata, 1990).

The four discrete Mandarin pitch patterns or tones are informative for the identification of roughly 70% of Mandarin words (Shibata & Shibata, 1990), causing Mandarin tones to carry a functional load as high as that of Mandarin vowels (Surendran & Levow, 2004).

In Japanese about 20% of the lexical items differ by pitch accent alone (Weiner & Goss, 2019)

Thus, the input cues for the learners are not all that frequent, and for the English leaners the functional load of stress minimal pairs is quite low as well.

Both these factors would lead us to expect that learning this phenomenon may take time.

Limitations

1. Sample size
2. Online experiment

Conclusion

1. Initial pitch accent is discriminated more accurately than final pitch accent for all groups.

2. This implies that Mandarin parsing (in the trilinguals) predominates because if English parsing predominated, then initial- and final-accented items should be perceived equally. This may well be a dominance effect in our trilinguals (see Cabrelli et al., 2023).

3. The L3A accuracy emerges because the Mandarin grammar can parse a Japanese initial-accent word with Tone 4 (HL), and an unaccented word with Tone 2 (LH). Parsing with L1 allophonic patterns (T3 sandhi) is more difficult.

4. The L2A difficulty in parsing initial and final accents stems from an English grammar's inability to parse the Japanese *unaccented* forms which were involved in both of our discrimination tasks. Future research should probe discriminating initial- from final-accent.

University of Victoria

Thank you.

junyuwu@uvic.ca
johnarch@uvic.ca

University of Victoria

Acknowledgment

We would like to express our appreciation to many kind and supportive people.

We would like to express our gratitude to Min Zeng for assisting us with the recruitment process at Waseda University. Additionally, we want to thank Dr. Tim Iles, Dr. Mamoru Hatakeyama. and Mariko Yuguchi sensei for their help with Japanese pitch accent and recruitment at UVic. We would also like to extend our thanks to Yi Yan for his unwavering support in statistics.

University of Victoria

References

Cabrelli, J., Pichan, C., Ward, J., Rothman, J., & Serratrice, L. (2023). Factors that moderate global similarity in initial L3 transfer: Intervocalic voiced stops in heritage Spanish/English bilinguals' L3 Italian. *Linguistic Approaches to Bilingualism*, 13(5).

Depuydt, L. (2022). Neutral tone in Chinese: a comprehensive theory bridging east and west. *Open Journal of Modern Linguistics* 12: 768-789.

Dresher, B. E., & Kaye, J. D. (1990). A computational learning model for metrical phonology. *Cognition*, 34(2), 137-195.

Farris-Trimble, A., & Tessier, A. M. (2019). The effect of allophonic processes on word recognition: Eye-tracking evidence from Canadian raising. *Language*, 95(1), e136-e160.

Goss, S. (2020). Exploring variation in nonnative Japanese learners' perception of lexical pitch accent: The roles of processing resources and learning context. *Applied Psycholinguistics*, 41(1), 25-49.

Goss, S. J., & Tamaoka, K. (2019). Lexical accent perception in highly-proficient L2 Japanese learners: The roles of language-specific experience and domain-general resources. *Second Language Research*, 35(3), 351-376.

Hayes, B. (1995). *Metrical stress theory: Principles and case studies*. University of Chicago Press.

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).

Kawahara, S. (2013). Testing Japanese loanword devoicing: Addressing task effects. *Linguistics*, 51(6), 1271-1299.

Kawahara, S. (2015). 11 The phonology of Japanese accent. Handbook of Japanese phonetics and phonology, 445-492.

Muradás-Taylor, B. (2022). Accuracy and Stability in English Speakers' Production of Japanese Pitch Accent. *Language and Speech*, 65(2), 377-403.

Özçelik, Ö. (2021). L2 acquisition of a complex stress pattern: UG-constrained learning paths in Khalkha Mongolian. *Frontiers in Psychology*, 12, 627797.

Qu, C. (2013). *Representation and acquisition of the tonal system of Mandarin Chinese.* (PhD thesis). McGill University. Montreal, Canada.

Shport, I. A. (2016). Training English listeners to identify pitch-accent patterns in Tokyo Japanese. *Studies in Second Language Acquisition*, 38(4), 739-769.

Sugiyama, Y. (2008). Production and Perception of Pitch Accent in Japanese. *In Annual Meeting of the Berkeley Linguistics Society* (pp. 317-328).

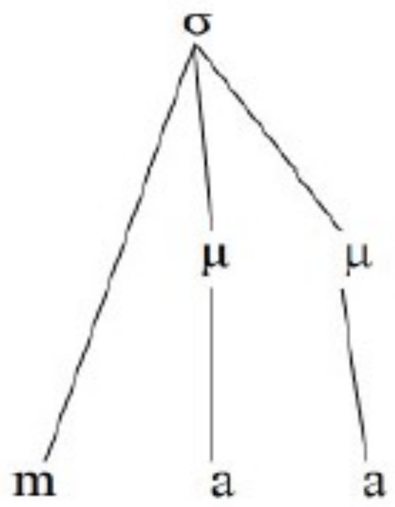
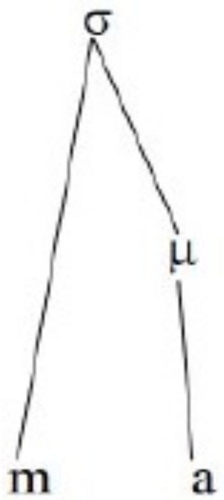
Vance, T. J. (1995). Final accent vs. no accent: Utterance-final neutralization in Tokyo Japanese. *Journal of Phonetics*, 23(4), 487-499.

Werker, J. F., & Logan, J. S. 1985. Cross-language evidence for three factors in speech perception. *Perception & Psychophysics*, 37(1), 35-44.

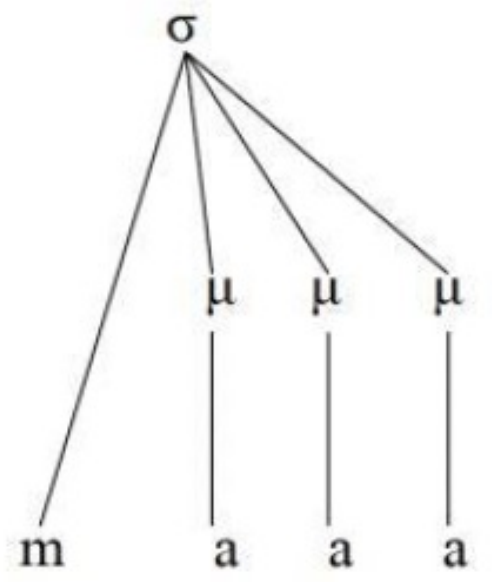
Wiener, S., & Goss, S. (2019). Second and third language learners' sensitivity to Japanese pitch accent is additive: an information-based model of pitch perception. *Studies in Second Language Acquisition*, 41(4), 897-910.

Wu, X., Kawase, S., & Wang, Y. (2017). Effects of acoustic and linguistic experience on Japanese pitch accent processing. *Bilingualism: Language and Cognition*, 20(5), 931-946.

University of Victoria

Heavy ($\sigma_{\mu\mu}$)	$T_{1/2/4}$ syllables	 妈“mother” (T_1 in disyllabic word)
Light (σ_{μ})	T_3 syllables	 马 “horse” (T_3 in disyllabic word)
Weightless (σ)	T_0 syllables	 吗 “question particle” (T_0)

University of Victoria

Syllable Weight	Tone	Internal Structure
Super-heavy ($\sigma_{\mu\mu\mu}$)	$T_{1/2/3/4}$ in isolation	 马 “horse” (T_3 in isolation)