Loanword accentuation in Yanbian Korean: A weighted-constraints analysis

Chiyuki Ito

Keywords: loanword phonology; Korean; accentuation; weighted constraints
Abstract: This paper analyzes the factors that determine the assignment of accent to Western (primarily English) and Japanese loanwords in the Yanbian dialect of Korean. The study is based on a corpus of 1,737 words. The major findings are as follows. In Yanbian loanwords, the accent is basically located in a two-syllable window at the right edge of the word. The accent pattern differs between disyllabic and longer words. Penultimate is the strong default accent in disyllabic loanwords, and syllable weight affects the distribution gradiently. On the other hand, the default accent in Yanbian native words is final. Statistical analysis shows that the different accent distributions between the native words and loanwords is attributed to the lexical class difference. The discrepancy between native words and loanwords is supported by a wug test. Our hypothesis is that Yanbian loanword accentuation results from the grammar of the source language and lexical statistics, along with some adjustments by Yanbian native grammar. By comparing the three different loanword categories in Yanbian that derive from different source languages with different prosodic types (English—stress, Japanese—pitch accent, Mandarin—tone), we show statistically that each has its own accentual adaptation system. We propose a loanword adaptation model in which the loanword adaptation is understood as an induction process from a faithfulness constraint to the source language into relevant markedness constraints. Through a learning process, the original faithfulness constraints to the source language are demoted below relevant markedness constraints. These markedness constraints are weighted by the learning algorithm so that the weight hierarchy can achieve a more or less "faithful adaptation" of the source language. Under this view, each separate sublexicon can have a different weight hierarchy of markedness constraints.

1. Introduction

In the study of loanword phonology, segmental adaptation has been discussed in previous literature but suprasegmental adaptation is relatively understudied. Following the common typology of three different suprasegmental categories (i.e. stress, pitch accent and tone), theoretically, nine possible adaptation patterns exist: (i) stress → stress, (ii) stress → pitch accent, (iii) stress → tone, (iv) pitch accent → stress, (v) pitch accent → pitch accent, (vi) pitch accent → tone, (vii) tone → stress, (viii) tone → pitch accent, (ix) tone → tone. Among them, some adaptation patterns are fairly well studied (e.g. stress → tone, such as English to Cantonese (Silverman 1992; Yip 2006), to Hausa (Leben 1996), to Yoruba (Kenstowicz 2004), and Thai (Kenstowicz and Suchato 2006)), but others are virtually unknown (e.g. tone → stress).

In her recent overview of suprasegmental adaptation in loanwords, Kang (2010) observed a puzzling asymmetry between segmental and accentual loanword adaptation. As many tone and pitch accent languages show, the prominence in the source language is often not respected even if the accentual system or the distributional patterns in the recipient language can allow it to

accommodate the position of the prominence in the input. This is quite different from the tendency in segmental adaptation where segments from the source language are typically respected more or less faithfully if they agree with phonotactic restrictions of the recipient language.

As to this specific property of suprasegmental adaptation, four possible factors are discussed in the previous literature. First, in their study of North Kyengsang Korean, Kenstowicz and Sohn (2001) called attention to the surprising lack of accent correspondence in such loanwords as *al.la.tín* 'Aladdin' (σόσ) and *ti.tsi.t^{h}ál* 'digital' (όσσ). They found that for a substantial subset of the loanwords, the accent was assigned by a rule seeking out a heavy syllable in a final two-syllable window. This in turn raised a learnability puzzle as to the source of such a rule. They suggested that it arises as a default that emerges from Universal Grammar. Parallel questions appear in Japanese loanword accent. Shinohara (1997a, 1997b, 2000) found that the accent in French and English loanwords was unfaithful to its source and fell under a type of Latin stress rule, as observed earlier by McCawley (1968). She suggested that the rule emerged as a default from Universal Grammar.

Kubozono (2006) challenged this conclusion. In a statistical study of the native Japanese lexicon, he found that the Latin stress rule actually holds for a substantial portion of the class of native accented nouns and thus proposed that in adapting a Western loanword, Japanese speakers perceive the accent (forcing the loanword into the accented class) but assign its location probabilistically based on the statistics of the native lexicon. Thus, he concluded that Universal Grammar is not necessary to understand Western loanwords in Japanese and that both phonetic and phonological factors are involved.

Another proposed source of loanword accentuation, originally suggested by Kenstowicz and Sohn (2001) for North Kyengsang Korean and further developed by J-S. Kim (2009) for the same dialect and by Lee (2008) for South Kyengsang Korean, is the covert grammar of the native accent system. According to this proposal, loanword accentuation is assigned by markedness constraints that are ranked higher than faithfulness to the source accent, whereas these markedness constraints play a minor role in native words since the faithfulness constraints to the lexical tone are ranked higher than the markedness constraints. If the adapter fails to equate the location of accent in the source language with the same accent location in the native language, then the native default that is normally hidden by faithfulness will emerge.

Finally, as Hsieh and Kenstowicz (2008) demonstrated for English and Mandarin loanwords in Lhasa Tibetan, universal phonetic considerations can be an additional factor in loanword accentuation. In Lhasa Tibetan, the F0 contours of the source languages (Mandarin and English) are ignored, and loanword accentuation reflects the voicing contrast in the onset of the source language: if it is voiced then it is assigned low tone, and if it is voiceless then it is assigned high tone. The statistical distribution of tones in the native lexicon plays no role in this process.

In this paper, based on our analysis of loanword accentuation in Yanbian Korean, spoken in Yanbian Korean Autonomous Prefecture in China, we offer a new perspective on the origin of loanword accentuation.

Yanbian Korean has a pitch accent system in which exactly one syllable in every lexical item is the locus of a pitch peak. For example, in trisyllabic nouns, four accent patterns are possible: HLL, LHL, LLH, and LLL (H = high, L = low), e.g. *tó.k*ɛ.pi* 'goblin', *tsok.tsé.pi* 'weasel',

kɛ.na.rí ‘forsythia’, *ma.tsi.mak* ‘the end’.[1] The unaccented class (LLL) appears with a final accent (LLH) in isolation forms, and unaccentedness appears in inflectional forms, such as LLL-H (*ma.tsi.mák* in isolation vs. *ma.tsi.ma.-kí* in the nominative form). In Yanbian native words, the majority falls within the final accent class. The Yanbian accent patterns and system are similar to the Hamkyeng dialect in North Korea and are rather different from Kyengsang Korean (Ramsey 1978; Umeda 1993; Park 2001; Che 2004; Miyashita 2007), but they tend to correspond with the latter by regular sound change. As to be mentioned shortly below, comparison of the loanword accentuation between Yanbian and Kyengsang Korean is important, because the majority of native nouns in Kyengsang do not carry final accent but penultimate accent due to a historical tonal retraction from the location in Middle Korean, where the majority of native nouns were in the final accent class (Ramsey 1978, e.g. Middle Korean *mə.rí*, Yanbian *mə.rí*, Kyengsang *mə́.ri* ‘head’). Despite the different default accent locations in native words of these two Korean dialects, penultimate is the default accent class for loanwords in both dialects. This suggests that loanword accentuation cannot simply be the reflection of the native grammar.

The Yanbian accent system is also similar to Tokyo Japanese. The loanword accentuation in Kyengsang and Japanese has been discussed in previous literature, as mentioned above (Kubozono 1996, 2006, 2008; Shinohara 2000; Kenstowicz and Sohn 2001; Lee 2006, 2008; Jun 2006, and others). However, as far as we know, there are not many previous investigations of loanword accentuation in Yanbian, and most of them are about the loanwords from Mandarin (Chi 2008; Ito and Kenstowicz 2009; Shen and Takeuchi 2011 for Mandarin loanwords; Shen 2011 for non-Mandarin loanwords).

In this paper we examine Yanbian loanword accentuation, based on data composed of Western (mainly English) and Japanese loanwords, and compare it with native simplex words and Sino-Korean words. The loanword data was collected from eight native speakers (five female and three male, all of them in their 20–40s) in 2007–2011.[2] Similarly, the native simplex and Sino-Korean data were collected from six speakers and one speaker respectively, who also served as the consultants for our loanword study. The original corpus for this study (Western and Japanese loanwords, native words, Sino-Korean words) was prepared by the author through consulting various dictionaries (Kadowaki, Matsuo, Takashima, and Yutani 1993; M-S. Kim 1997; Kwuklip kwuke yenkwuwen 1999) and previous literature (Ramsey 1978; Ito 2000; Kenstowicz and Sohn 2001; Ito et al. 2006; Jun 2006; Lee 2006, 2008). Each item was checked with all of our

[1] The transcription system for the examples in this paper is as follows. Some symbols such as ㄹ, ㅇ are transcribed differently depending on environment. H is marked by the acute accent and L is unmarked. Other Korean words in the body of the paper follow the Yale Romanization system, except for some authors’ names which follow the customary spelling.

ㄱ	ㄲ	ㄴ	ㄷ	ㄸ	ㄹ	ㅁ	ㅂ	ㅃ	ㅅ	ㅆ	ㅇ	ㅈ	ㅉ	ㅊ	ㅋ	ㅌ	ㅍ	ㅎ
k	k*	n	t	t*	r/l	m	p	p*	s	s*	Ø/ŋ	ts	ts*	ts^h	k^h	t^h	p^h	h

ㅏ	ㅐ	ㅑ	ㅓ	ㅔ	ㅕ	ㅖ	ㅗ	ㅘ	ㅙ	ㅚ	ㅛ	ㅜ	ㅝ	ㅞ	ㅟ	ㅠ	ㅡ	ㅢ	ㅣ
a	ɛ	ja	ə	e	jə	je/e	o	wa	wɛ	we	jo	u	wə	we	wi	ju	ɨ	ɨj/i	i

[2] Three speakers are from Yanji, another three are from Longjing, and the other two are from Antu. The accent of the Antu dialect may be slightly different from the other two dialects in its surface realization, but we believe this is not relevant to the loanword accentuation.

consultants to determine the accentuation since accent is not marked in Korean dictionaries in general, and only the words that at least one speaker reported to use were included in the data. Many words that do not appear in the original corpus were also added, based on consultation with the speakers.

The size of the loanword corpus is 1,737 words. Since there are individual variations for the accent patterns, the data from all speakers were simply aggregated for each word so that the general patterns of loanword accentuation are clarified. For example, *am.mo.ni.a* 'ammonia' appears with a penultimate accent (*am.mo.ní.a*) for three speakers and with a final accent (*am.mo.ni.á*) for five speakers. The count for penultimate/final accent class is thus three and five respectively for this word. The gross total of the loanword data aggregated in this way is 9,156 words. The same procedure was applied for native simplex words as well: there are 1,865 native words and the aggregated total from six speakers is 8,882.

Major findings of this paper are summarized as follows.

a) Given that every Yanbian word has a pitch peak, loanwords must be assigned an accent to be phonologically valid words of the language. Since the location of accent is lexically contrastive in the native grammar, one expects the accent of the source word to be respected in much the same way that phonological segments are faithfully adapted. But as in other tone or pitch accent languages, the actual adaptation is not a mechanical process that faithfully adapts the prominence of the source language, e.g. *in.t^{h}ə́.pju* 'interview' (σσὸ), *kí.t^{h}a* 'guitar' (σσ́). Yanbian loanword accent is basically located in a two-syllable window at the right edge of the word (penultimate or final). This bias to the right edge of the word is due to two high-weighted constraints (*LAPSE-RIGHT and *UNACCENTED) that are at work in the native grammar. On the other hand, Yanbian loanword accentuation is not a simple reflection of the statistics of the native lexicon. The most evident difference between the loanword and the native accent patterns is that the default accent class is different between loanwords (penultimate) and native words (final). If Kubozono's (2006) finding that loanword accentuation reflects the statistical tendencies in native Japanese words was generalized to Yanbian, then Yanbian loanwords should appear in the final accent class more frequently. But this pattern is not reflected in the data. In this respect Yanbian is crucially more informative than Kyengsang, where the majority of both loanwords and native nouns have penultimate accent; one might conclude that penultimate accent in loanwords simply reflects the native grammar bias. Yanbian shows that this inference is not necessarily valid.
b) This different accentuation between loanwords and native words in Yanbian is also not due to the covert grammar of the native system. In order to confirm this, we conducted a wug test whose items include unfamiliar native words and loanwords, e.g. *tɨl.me* 'string to tie a shoe and a foot', *pa.sun* 'bassoon'. The test results show that Yanbian speakers apparently distinguished the two lexical classes (native vs. loan) and assigned a different accent accordingly. (The final accent class was predominant in native nouns, whereas the penultimate accent class was predominant in loanwords.) Since these wug test words are not associated with a faithfulness constraint to the lexical accent but still show the established division between the two lexical classes, we can infer that the covert grammar of the native system is not relevant here.
c) Furthermore, the analysis of the three different suprasegmental adaptation patterns found in

Yanbian (stress → pitch accent (English), pitch accent → pitch accent (Japanese), tone → pitch accent (Mandarin)) shows that Yanbian loanwords have different patterns of adaptation depending on the source language (English, Japanese, Mandarin) that are statistically significant, e.g. English: *jú.mə* 'humor', Japanese: *ta.má,* tama, 'ball', Mandarin: *hó.ma*, hàomǎ 号码 'number'. Given this fact, our hypothesis is that loanword accentuation results from the grammar of the source language, while admitting some adjustments by native grammar. Universal Grammar or universal phonetic considerations do not play any significant role.

d) Given this analysis, we propose a loanword adaptation model. Loanword adaptation is understood as an induction process originating from a faithfulness constraint to the source language and resulting in several relevant markedness constraints. Through a learning process, the original faithfulness constraints to the source language are demoted below relevant markedness constraints. These markedness constraints are weighted higher by the learning algorithm so that the weight hierarchy can achieve a more or less "faithful adaptation" of the source language. The resulting "faithful adaptation" is thus based on this generalization process. Under this view, each separate sublexicon can have different rankings or weight hierarchies of markedness constraints. To test this proposal, we conducted a simulation by using Jäger's (2007) Stochastic Gradient Ascent learning algorithm and obtained a "proof of concept" that the loanword accent system or distribution can differ from one source language to another.

This paper is organized as follows. In Section 2, we introduce the Yanbian accent system and the distribution of each accent class in native words and loanwords. In Section 3, we describe the Yanbian loanword accentuation and clarify the various factors underlying its assignment: a syllable weight effect, default penultimate accent, epenthesis, and others. In Section 4, we discuss where the Yanbian loanword accent patterns come from (Universal Grammar default, native grammar, source language) by comparing the loanword accentuation with native accent distribution. In Section 5, we examine subcategories of Yanbian loanwords (English, Japanese, Mandarin) in detail and show that loanword accentuation comes from the source language as a rule, with some adjustments by native grammar. In Section 6, we propose a loanword accentuation model based on our Yanbian data. Section 7 is a summary and conclusion.

2. Accent class and distribution

In this section, we compare the accent distribution of native nouns with that of loanwords. The loanwords discussed here include nouns borrowed directly from Western languages (mainly English) and Japanese, as well as hybrid loanwords between Western languages and Japanese.[3] Loanwords that are compound nouns are excluded from the data as a rule.

As mentioned above, Yanbian has a pitch accent system in which one syllable in every lexical

[3] Hybrid loanwords are loanwords passed through Japanese that have been subsequently modified in certain respects in order to conform to the current direct transmission correspondences between English and Korean. See Kang et al. (2008) for the details. For example, for 'skirt', a direct loanword from English is *si.k^{h}ə́.t^{h}ɨ*, direct loanwords from Japanese are *si.k*á.to, si.k*á.t*o*, and hybrid loanwords are *si.k^{h}á.t^{h}ɨ, si.k*á.tɨ.*

item is the locus of a pitch peak. The Yanbian accent classes in native simplex nouns are summarized in (1). H and L indicate high and low tones, and [ ] indicates the accent of the following suffix.

(1) Accent classes in Yanbian native simplex nouns
Monosyllabic: H[L], L[H]
Disyllabic: HL[L], LH[L], LL[H]
Trisyllabic: HLL[L], LHL[L], LLH[L], LLL[H]
Quadrisyllabic: LLHL[L], LLLH[L], (HLLL[L], LHLL[L], LLLL[H])

Each accent class is abbreviated to H, HL, etc., ignoring the accent of the following suffix in this paper. Quadrisyllabic nouns can have five accent classes (HLLL, LHLL, LLHL, LLLH, LLLL), but in actuality only LLHL and LLLH normally occur, while HLLL/LHLL/LLLL are extremely marginal. Underlying unaccented classes (LL, LLL, LLLL) merge with the final accent classes (LH, LLH, LLLH) in isolation citation forms (Park 2001; Ito 2008a). (2) shows examples.

(2) Examples of native simplex words. The acute accent represents H, and its absence indicates L. Variations in accent are indicated with "~".

a. H: kíl 'road', k*úm 'dream', mál 'language', móm 'body', nún 'eye', pám 'night', páp 'boiled rice', pɛ́m 'snake', sám 'hemp', sól 'pine', t*áŋ 'land', ts^{h}úm 'saliva'
b. L: ap^{h} 'front', jəp^{h} 'side', mal 'horse', moks 'share', nats^{h} 'face', pak* 'outside', pat^{h} 'field', pits 'debt', talk 'chicken'
c. HL: án.kɛ 'fog', jə́m.tsi 'leek', kól.ts^{h}i 'head', kú.rɨm 'cloud', kɨ́.ne 'swing', k*á.t^{h}si 'magpie', ná.i 'age', sá.ram 'person', tsán.ts^{h}i 'party'
d. LH: ə.k*é 'shoulder', ha.nɨ́l 'sky', ka.sɛ́ 'scissors', ki.rɨ́m 'oil', ma.nɨ́l 'garlic', mɛ.mí 'cicada', pə.sɨ́s 'mushroom', p*u.rí 'root', toŋ.mú 'friend'
e. LL: ən.tək ~ ən.tək 'hill', ka.sɨl ~ ka.sɨl 'autumn', kɨn.sím ~ kɨn.sim 'worries', mu.rɨp^{h} 'knee', pa.k*at^{h} 'outside', pa.rám ~ pa.ram 'wind', sɛ.pjək 'dawn', so.kóm ~ so.kom 'salt'
f. HLL: ə́.rɨ.sin ~ ə.rɨ.sin 'esteemed elder', í.ja.ki 'talk', kə́.ma.ri 'leech', mú.tsi.kɛ 'rainbow', sá.ma.kwi ~ sa.ma.kwí 'mole; praying mantis', tó.k*ɛ.pi 'goblin'
g. LHL: ə.tsə́.k*e 'yesterday', ja.tɨ́.re 'eight days', mjə.nɨ́.ri 'daughter-in-law', na.kɨ́.ne 'traveler', tsok.tsé.pi 'weasel'
h. LLH: ko.sa.rí 'bracken', nol.ka.tsí 'deer', sa.t^{h}u.rí 'dialect', to.ka.ní 'melting pot', tsə.ko.rí 'Korean jacket', tsi.rəŋ.í 'earthworm'
i. LLL: ol.li.mák ~ ol.li.mak 'uprise', kjə.tɨ.ráŋ ~ kjə.tɨ.raŋ 'armpit', ma.tsi.mak 'last', mu.rɨ.p^{h}ák ~ mu.rɨ.p^{h}ak 'knee', pu.sɨ.rəm ~ pu.sɨ.rəm 'ulcer', sim.pu.rɨm 'errand'
j. LLHL: a.tsɨ.má.i 'aunt', ho.rɨ.rɛ́.ki 'whistle', kɛ.ku.rák.tsi 'frog', ko.sɨm.tó.ts^{h}i 'hedgehog', me.ts^{h}u.rá.ki 'quail', mi.k*u.rá.tsi 'loach', su.su.k*é.k*i 'puzzle'
k. LLLH: a.tsi.raŋ.í 'heat haze', k*o.rak.sə.ní 'condition', pu.tsi.k*ɛŋ.í 'poker', sa.t^{h}a.ku.ní 'crotch', tsi.nɨ.rə.mí 'fin', tsu.tuŋ.a.rí 'beak'

The distribution of accent classes in loanwords is different from native simplex nouns, as seen in Table 1. Monosyllabic loanwords appear with H without exception in loanwords, whereas

native simplex words appear with L quite frequently. In polysyllabic loanwords, most words appear with either penultimate accent (HL, LHL, LLHL...) or final accent (LH, LLH, LLLH...), with a higher ratio of the former than the latter. There are very few examples of the unaccented class. On the other hand, in polysyllabic native words, the size of the final accent class is larger than the penultimate class in general, and the unaccented class is not that rare, especially in disyllabic words. The pre-antepenultimate accent class rarely appears in both native words and loanwords, whereas the antepenultimate accent class can partially appear, which is mostly observed in trisyllabic words.

Table 1

Accent distribution of Yanbian native words and loanwords. Parentheses indicate the aggregated number of examples within each accent class.

Native

σ	Pre-antepenultimate	Antepenultimate	Penultimate	Final	Unaccented	Totals
1				H (1,906)	L (351)	2,257
2			HL (1,119)	LH (3,241)	LL (354)	4,714
3		HLL (84)	LHL (241)	LLH (1,262)	LLL (36)	1,623
4	HLLL (2)	LHLL (1)	LLHL (156)	LLLH (122)	LLLL (7)	288
Totals	2	85	1,516	6,531	748	8,882

Loan

σ	Pre-antepenultimate	Antepenultimate	Penultimate	Final	Unaccented	Totals
1				H (386)	L (0)	386
2			HL (3,023)	LH (614)	LL (24)	3,661
3		HLL (476)	LHL (1,662)	LLH (1,137)	LLL (7)	3,282
4	HLLL (21)	LHLL (52)	LLHL (978)	LLLH (383)	LLLL (6)	1,440
5	LHLLL (0)	LLHLL (47)	LLLHL (161)	LLLLH (86)	LLLLL (2)	296
6	LLHLLL (0)	LLLHLL (14)	LLLLHL (43)	LLLLLH (27)	LLLLLL (0)	84
7	LLLHLLL (0)	LLLLHLL (0)	LLLLLHL (7)	LLLLLLH (0)	LLLLLLL (0)	7
Totals	21	589	5,874	2,633	39	9,156

(3) Examples of loanwords

a. H: k*ɔ́m ‘gum’, pél ‘bell’, p*í ‘B’, p^{h}ék ‘pack’, tó ‘do (solmization)’, t^{h}íp ‘tip’

b. HL: én.tsin ‘engine’, hó.t^{h}el ‘hotel’, nín.tsin ‘carrot’ (< Japanese niNdʒiN), pél.t^{h}ɨ ‘belt’, pén.ts^{h}i ‘bench’
LH: a.ráp ‘Arab’, k^{h}e.ts^{h}áp ‘ketchup’, sɨ.k^{h}í ‘ski’, ne.p^{h}ál ‘Nepal’, pol.ljúm ‘volume’

c. HLL: má.sɨ.k^{h}ɨ ‘mask’, tí.sɨ.k^{h}o ‘disco’, wí.sɨ.k*i ‘whisky’, p^{h}ó.sɨ.t^{h}ə ‘poster’
LHL: o.mé.ka ‘omega’, k^{h}a.né.sjon ‘carnation’, in.t^{h}ə́.pju ‘interview’, pe.rán.ta ‘veranda’
LLH: an.t^{h}e.ná ‘antenna’, in.t^{h}ə.nés ‘internet’, me.mo.rí ‘memory’, ma.ra.t^{h}ón ‘marathon’

d. HLLL: ré.sɨ.t^{h}o.raŋ ~ re.sɨ.t^{h}o.ráŋ ‘restaurant’, ák.se.sa.ri ~ ak.sé.sa.ri ~ ak.se.sa.rí ‘accessory’
LHLL: t*ɨ.rák.t*o.rɨ ‘tractor’, k^{h}u.wé.i.t^{h}ɨ ~ k^{h}u.we.í.t^{h}ɨ ‘Kuwait’
LLHL: e.p^{h}il.ló.kɨ ‘epilogue’, el.le.pé.t^{h}a ‘elevator’, k*wa.t*e.mál.la ‘Guatemala’, sɨ.ri.ráŋ.k^{h}a

‘Sri Lanka’, ta.ma.né.ki ‘onion’ (< Japanese tamanegi)
LLLH: mo.sɨ.kʰɨ.pá ‘Moscow’ (< Russian Moskva), pʰɨ.ro.kɨ.rɛ́m ‘program’, pʰe.ni.sil.lín ‘penicillin’, ma.kɨ.ne.sjúm ‘magnesium’, pʰɨl.lɛ.tʰɨ.hóm ‘platform’

e. LLHLL: pʰi.a.ní.sɨ.tʰɨ ~ pʰi.a.ni.sɨ́.tʰɨ ‘pianist’, o.sɨ.tʰɨ́.ri.a ~ o.sɨ.tʰɨ.ri.á ‘Austria’, ruk.sem.pɨ́.rɨ.kɨ ~ ruk.sem.pɨ.rɨ́.kɨ ‘Luxemburg’
LLLHL: so.pʰɨ.tʰɨ.wé.ə ‘software’, el.le.pe.í.tʰə ‘elevator’, kʰɨ.ri.sɨ.má.sɨ ‘Christmas’
LLLLH: a.rɨ.me.ni.á ‘Armenia’, kʰa.si.o.pʰe.á ‘Cassiopeia’

f. LLLHLL: t*o.sɨ.t*o.jép.sɨ.k*i ‘Dostoevsky’, kʰə.mju.ni.kʰé.i.sjən ~ kʰə.mju.ni.kʰe.í.sjən ‘communication’
LLLLHL: p*e.tʰe.rɨ.pɨ.rɨ́.kɨ ‘Petersburg’ (< Russian Peterburg), ma.ta.k*a.sɨ.k*á.rɨ ‘Madagascar’, u.tsɨ.pe.kʰi.sɨ́.tʰan ~ u.tsɨ.pe.kʰi.sɨ.tʰán ‘Uzbekistan’
LLLLLH: me.so.pʰo.tʰa.mi.á ‘Mesopotamia’, ju.ko.sɨl.la.pí.a ~ ju.ko.sɨl.la.pi.á ‘Yugoslavia’

g. LLLLLHL: ul.la.tsi.o.sɨ.t*ó.kʰɨ, ul.la.tsi.po.sɨ.t*ó.kʰɨ ‘Vladivostok’

Given this distribution, it is safe to say that Yanbian loanword accent is basically located in a two-syllable window at the right edge of the word (penultimate or final). In order to confirm this point (following a very helpful suggestion from one of the reviewers), we focus on longer words where it is possible for the accent to appear earlier than on the final two syllables. Table 2 shows the observed (O), expected (E), and observed/expected (O/E) numbers in longer loanwords (the unaccented class is excluded here). If accent had been assigned randomly to one of the three syllables in trisyllabic words, we would have expected a third of these loanwords (E = 1,092) to have antepenultimate, penultimate and final accent respectively. Similarly, if accent had been assigned randomly in quadrisyllabic words, then half of the loanwords would have appeared with antepenultimate and pre-antepenultimate accent (E = 717), while penultimate and final accent would have each appeared in one fourth of the loanwords (E = 359).

As shown in Table 2, the higher O/E numbers in penultimate and final accent classes support a strong preference for accent to fall on one of the final two syllables. The fact that penultimate has higher O/E numbers than final also indicates a preference for penultimate over final accent.

Table 2
Observed, expected, and observed/expected numbers in longer loanwords

	Earlier than penultimate			Penultimate			Final			
σ	O	E	O/E	O	E	O/E	O	E	O/E	Totals
3	476	1,092	0.44	1,662	1,092	1.52	1,137	1,092	1.04	3,275
4	73	717	0.10	978	359	2.72	383	359	1.07	1,434
5	47	176	0.27	161	59	2.73	86	59	1.46	294
6	14	56	0.25	43	14	3.07	27	14	1.93	84
7	0	5	0.00	7	1	7.00	0	1	0.00	7

In the next section, we discuss what factors affect the accentual assignment in these loanwords. We return to the distributional differences between native words and loanwords in 4.

3. Description of Yanbian loanword accentuation

In this section, we describe the factors working in Yanbian loanword accentuation. In 3.1 we examine how syllable weight is correlated with Yanbian loanword accentuation. In 3.2 we point out the different accentual patterns observed between disyllabic loanwords and three-or-more syllable loanwords. In 3.3 we discuss epenthesis, which tends to disprefer having a pitch peak on the epenthetic vowel. In 3.4 we show the accent distribution of words with l-gemination that corresponds with an onset /l/ in the source language and offer a preliminary phonetic analysis. 3.5 is a summary.

3.1 Syllable weight

As in Kyengsang Korean and Japanese, Yanbian loanword accentuation correlates with syllable weight (WEIGHT-TO-STRESS, Prince and Smolensky 1993; STRESS-TO-WEIGHT, Kager 1999). More precisely, a heavy syllable in a final two-syllable window tends to attract a pitch peak. Heavy syllables contain a coda consonant, while light syllables are open. There are no long vowels in Yanbian and no onset or coda consonant clusters on the surface. Table 3 shows the correlations. The data here do not contain examples with epenthesis and l-gemination, which are examined separately in 3.3 and 3.4 respectively.

Table 3
Accent and syllable weight in loanwords. The "Weight" column lists the weight of the final two-syllable window. "pA, A, P, F, U" indicate pre-antepenultimate accent, antepenultimate accent, penultimate accent, final accent and the unaccented class, respectively. "A%, P%, F%" are the percentages of antepenultimate, penultimate and final accent classes in each syllable weight structure.

Weight	pA	A	P	F	U	Totals	A%	P%	F%
Heavy-Heavy			269	87	5	361	0%	75%	24%
Heavy-Light			586	18	1	605	0%	97%	3%
Light-Heavy	11	33	654	652	22	1,372	2%	48%	48%
Light-Light	6	209	1,766	928	3	2,912	7%	61%	32%
Totals	17	242	3,275	1,685	31	5,250	5%	62%	32%

(4) Examples of accent and syllable weight

a. [Heavy-Heavy] píl.tiŋ 'building', sjám.p^{h}aŋ ~ sjam.p^{h}áŋ 'champagne', móŋ.kol ~ moŋ.kól 'Mongolia', al.k^{h}ól 'alcohol', wa.síŋ.t^{h}on 'Washington', ol.lim.p^{h}ík 'Olympic', pɛ.tɨ.mín.t^{h}on 'badminton'

b. [Heavy-Light] ál.p^{h}a 'alpha', k*óŋ.ko 'Congo', rúm.pa 'rumba', t^{h}ék.si ~ t^{h}ɛk.sí 'taxi', p^{h}én.t^{h}i 'panty', ma.tón.na 'Madonna', u.kán.ta 'Uganda', hel.li.k*óp.t^{h}ə 'helicopter', sɨ.ri.ráŋ.k^{h}a 'Sri Lanka'

c. [Light-Heavy] á.men ~ a.mén ‘amen’, kʰó.mik ~ kʰo.mík ‘comic’, mjú.tsik ‘music’, ré.mon ‘lemon’, pʰe.tál ‘pedal’, ri.póŋ ‘ribbon’ (< Japanese riboN), tʰá.i.miŋ ~ tʰa.i.míŋ ‘timing’, kʰɛ́.tsju.əl ~ kʰɛ.tsjú.əl ~ kʰɛ.tsju.ə́l ‘casual’, ré.sɨ.tʰo.raŋ ~ re.sɨ.tʰo.ráŋ ‘restaurant’
d. [Light-Light] mé.mo ‘memo’, pʰá.tʰi ‘party’, kʰá.re ~ kʰa.ré ‘curry’, há.mo.ni ~ ha.mo.ní ‘harmony’, tsjó.tsi.a ~ tsjo.tsí.a ~ tsjo.tsi.á ‘Georgia’, kʰá.me.ra ~ kʰa.me.rá ‘camera’, si.kʰá.ko ~ si.kʰa.kó ‘Chicago’, a.pʰɨ.rí.kʰa ‘Africa’, in.tʰe.rí.ə ~ in.tʰe.ri.ə́ ‘interior’, mí.sɨ.tʰe.ri ~ mi.sɨ.tʰé.ri ~ mi.sɨ.tʰe.rí ‘mystery’

As shown in Table 3, the syllable weight effect is gradient: it is extremely strong in Heavy-Light, where penultimate accent amounts to 97%, while it is much weaker in Light-Heavy, where only 48% belong to final accent. Heavy-Heavy and Light-Light show a somewhat intermediate tendency between Heavy-Light and Light-Heavy with a higher percentage of penultimate accent than Light-Heavy (75% and 61% respectively). Thus, in the Heavy-Heavy and Light-Light cases, where syllable weight is not decisive, there is a bias to penultimate accent. The effect of syllable weight intensifies the penultimate bias in Heavy-Light and retards it in Light-Heavy. Based on this distribution, we can conclude that penultimate accent is the default accent class in Yanbian loanwords and that the syllable weight effect is strongest in the most preferred syllable structure (Heavy-Light) and is gradiently weakened in Heavy-Heavy/Light-Light and Light-Heavy, as a result of a conflict between the default accent class (penultimate) and the syllable weight effect.

Is there a syllable weight effect in the antepenultimate syllable? In order to see this point, we concentrate on three-syllable words now. Table 4 shows the syllable structure and accent patterns. Among the eight possible patterns of syllable weight, Heavy-Heavy-Heavy does not appear in our corpus.

Table 4
Accent and syllable weight in trisyllabic loanwords

Weight	A	P	F	U	Totals	A%	P%	F%
Heavy-Heavy-Light		34	1		35	0%	97%	3%
Heavy-Light-Heavy	10	19	43	4	76	13%	25%	57%
Heavy-Light-Light	23	141	87		251	9%	56%	35%
Light-Heavy-Heavy		14	5		19	0%	74%	26%
Light-Heavy-Light		88			88	0%	100%	0%
Light-Light-Heavy	20	72	244	1	337	6%	21%	72%
Light-Light-Light	138	309	361		808	17%	38%	45%
Totals	191	677	741	5	1,614	12%	42%	46%

First, as seen in Table 4, if the penultimate syllable is heavy, then there is no word with antepenultimate accent. This confirms the strong syllable weight effect in the final two-syllable window and the default penultimate accent. Next, if we compare Heavy-Light-X and Light-Light-X (X denotes any syllable weight), then it is clear that a heavy syllable in the antepenultimate syllable does not attract the accent more strongly than a light syllable does (the

percentage of antepenultimate accent class in Heavy-Light-X is 10%, whereas that in Light-Light-X is 14%, as shown in Table 5).

Table 5
Accent and syllable weight in trisyllabic loanwords of Heavy-Light-X and Light-Light-X

Weight	A	P	F	U	Totals	A%	P%	F%
Heavy-Light-X	33	160	130	4	327	10%	49%	40%
Light-Light-X	158	381	605	1	1,145	14%	33%	53%

Thus, we can conclude that there is no clear syllable weight effect outside of the final two-syllable window. We return to the exceptional adaptation pattern of the antepenultimate accent class in 3.3.

3.2 Syllable number effect

In actuality, syllable number also affects the accent distribution. As shown in Table 6, disyllabic words in Light-Heavy and Light-Light structures tend to appear with penultimate accent more frequently than three-or-more syllable words.

Table 6
Syllable number effect in loanwords

Weight	σ	pA	A	P	F	U	Totals	A%	P%	F%
Heavy-Heavy	2			218	68	5	291		75%	23%
	3~			51	19		70	0%	73%	27%
Heavy-Light	2			388	16	1	405		96%	4%
	3~			198	2		200	0%	99%	1%
Light-Heavy	2			450	238	13	701		64%	34%
	3~	11	33	204	414	9	671	5%	30%	62%
Light-Light	2			834	123		957		87%	13%
	3~	6	209	932	805	3	1,955	11%	48%	41%

A mixed effects logistic regression model with accent location (penultimate/final, final is the baseline) as the dependent variable and item and subject as random factors was conducted by using the lmer function in the lme4 package (Bates, Maechler, and Bolker 2013) in R software (R Development Core Team 2011). The result shows that all factors (syllable weight in penultimate syllables, syllable weight in final syllables, and syllable number—disyllabic vs. three-or-more syllables) are strongly significant: a heavy penult is positively associated with penultimate accent, whereas a heavy final and a larger syllable number are negatively associated with penultimate accent.

Table 7
Results of a mixed effects logistic regression model: syllable number effect

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	4.7000	0.3438	13.671	< 2e-16	***
Weight-Penult-Heavy	2.9297	0.4295	6.822	8.98e-12	***
Weight-Final-Heavy	-3.3363	0.3065	-10.885	< 2e-16	***
Syllable number-three-or-more syllables	-3.6974	0.3115	-11.869	< 2e-16	***

Significance codes: '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

Thus it is clearer in disyllabic loanwords that penultimate accent is the default accent class. The syllable weight effect in disyllabic loanwords is gradient in that it is strongest in Heavy-Light forms, where almost all words appear with the penultimate accent class, followed by Heavy-Heavy/Light-Light and Light-Heavy forms. As argued below (section 5.2), this is probably because disyllabic words with penultimate accent can faithfully realize the most frequent stress pattern (stressed-unstressed, σ́σ) in English disyllabic words. Thus penultimate accent in Yanbian disyllabic words coincides with the locus of this canonical English stress pattern. In longer words, English stress position is not biased to the penultimate syllable, and hence the stronger default to penultimate accent is potentially blocked.

In the accent distribution of three-or-more syllable words with final Heavy-Light or Light-Heavy, the syllable weight effect emerges more clearly: Heavy-Light almost always appears with penultimate accent, whereas Light-Heavy appears with final accent more frequently than in disyllables. Still, the syllable weight effect is stronger in Heavy-Light than in Light-Heavy, given the fact that Heavy-Light appears with the penultimate accent class almost without exception.

In Light-Light, where syllable weight is not decisive, the frequencies of the penultimate and the final accent classes are more or less evenly distributed in three-or-more syllable words (48% vs. 41%), with a slight bias to penultimate accent. Also Heavy-Heavy in three-or-more syllable words appears with the penultimate accent class more frequently (73%). These facts and the exclusive penultimate accent in Heavy-Light structure suggest that the default accent is still penultimate in three-or-more syllable words as well, while the syllable weight effect acts more strongly here than in disyllabic words.

In conclusion, we can restate the accentuation in Yanbian loanwords as follows: in Yanbian loanwords, penultimate accent is the default accent class; in disyllabic loanwords this default accent appears more strongly and the syllable weight effect is working gradiently; in three-or-more syllable words, the syllable weight effect emerges more clearly and accent polarization (Heavy-Light → penultimate, Light-Heavy → final) is observed, with the effect stronger in Heavy-Light than in Light-Heavy due to the default penultimate accent.

3.3 Epenthesis

In Korean loanwords, sonorant codas are adapted as codas whereas obstruent codas and consonant clusters are mostly adapted with an epenthetic vowel (/-ɨ/ as a rule, but /-i/ after coronal alveolars, Kenstowicz and Sohn 2001). The adaptation of coda /-r/ is different depending on the source language: English /-r/ is usually adapted as Ø, while /-r/ from other languages, such as French, is often adapted as /rɨ/.

An unstressed final syllable in English words such as *action* and *medal*, which are transcribed with a syllabic sonorant in the Oxford dictionary, is also adapted with an epenthetic (less sonorous) vowel (/ɨ/, /ə/) or with other vowels reflecting the spelling of the source word, e.g. *ék.sjən* 'action', *hɛ́ntɨl* 'handle' for the former case; *mó.tel* 'model', *me.tál* 'medal' for the latter. Based on these segmental adaptation patterns, we treat the final syllable of the former case as a heavy syllable with epenthesis and the final syllable of the latter case as a genuine heavy syllable, for the sake of descriptive convenience.[4]

In general, less sonorous syllables and epenthetic vowels disprefer having a pitch peak (Kenstowicz 1997; Shinohara 2000; Lee 2008). Epenthetic vowels tend to disprefer accent in Yanbian loanwords as well. First, we examine the case when the final syllable contains an epenthetic vowel and the penultimate syllable does not.

Table 8

Accent when the final syllable contains an epenthetic vowel. (e) indicates a syllable with epenthesis.

Weight	pA	A	P	F	U	Totals	A%	P%	F%
Heavy-Heavy(e)			80			80	0%	100%	0%
Heavy-Light(e)		15	643			658	2%	98%	0%
Light-Heavy(e)		16	191	33	1	241	7%	79%	14%
Light-Light(e)	3	107	1,297	13		1,420	8%	91%	1%

(5) Examples of accent and syllable weight when the final syllable contains an epenthetic vowel

a. [Heavy-Heavy(e)] ék.sjən 'action', hɛ́ntɨl 'handle', s*ém.p^{h}ɨl 'sample', síŋ.kɨl 'single', sɨ.k^{h}én.tɨl 'scandal', ts^{h}én.nəl 'channel'

b. [Heavy-Light(e)] ín.ts^{h}i 'inch', kól.p^{h}ɨ 'golf', rám.p^{h}ɨ 'lamp', p*óm.p*ɨ 'pump', k^{h}on.sén.t^{h}ɨ 'consent', t^{h}ɨ.rə́m.p^{h}ɨ 'trump', és.s*en.sɨ ~ es.s*én.sɨ 'essence', ta.i.a.món.tɨ 'diamond'

c. [Light-Heavy(e)] k^{h}ə́.p^{h}ɨl 'couple', ó.p^{h}in 'open', p^{h}é.sjən 'fashion', rú.pɨl 'ruble', t*ə́.pɨl ~ t*ə.pə́l 'double', ra.í.pəl ~ ra.i.pə́l 'rival', si.ɛ́.t^{h}ɨl ~ si.ɛ.t^{h}ɨ́l 'Seattle', k^{h}ən.tí.sjən 'condition', t^{h}ə́.mi.nəl ~ t^{h}ə.mi.nə́l 'terminal'

d. [Light-Light(e)] ké.kɨ 'gag', k^{h}í.sɨ 'kiss', p^{h}ó.tsɨ 'pose', ó.p^{h}i.sɨ ~ o.p^{h}í.sɨ 'office', t^{h}é.ni.sɨ ~ t^{h}e.ní.sɨ 'tennis', i.mí.tsi ~ i.mi.tsí 'image', i.rá.k^{h}ɨ 'Iraq', ol.lí.pɨ 'olive', p^{h}ɨl.lá.s*ɨ 'plus', sɨ.wí.ts^{h}i 'switch', mas.sá.tsi ~ mas.sa.tsí 'massage', hó.sɨ.t^{h}i.sɨ ~ ho.sɨ.t^{h}í.sɨ 'hostess'

In this case, most words tend to appear with penultimate accent regardless of the syllable structure. The percentage of penultimate accent is slightly lower in Light-Heavy(e) (79%), which

[4] As pointed out by one of the reviewers, an alternative interpretation in terms of contextually light closed syllables is also possible (Chung 2002; J-S. Kim 2009).

is probably because the syllable weight effect is simultaneously at work.

Next is the case when the penultimate syllable contains an epenthetic vowel and the final syllable does not. There are no forms in our data that have Heavy(e) in the penultimate syllable.

Table 9

Accent when the penultimate syllable contains an epenthetic vowel

Weight	pA	A	P	F	U	Totals	A%	P%	F%
Light(e)-Heavy		27	61	233	2	323	8%	19%	72%
Light(e)-Light		82	66	77		225	36%	29%	34%

(6) Examples of accent and syllable weight when the penultimate syllable contains an epenthetic vowel

a. [Light(e)-Heavy] kɨ.rám ‘gram’, sɨ.k^{h}ín ‘skin’, sɨ́.p^{h}un ~ sɨ.p^{h}ún ‘spoon’, p^{h}ɨ́.raŋ ~ p^{h}ɨ.ráŋ ‘franc’, t^{h}ɨ.rə́k ‘truck’, ke.rɨ.mán ‘German’, ho.rɨ.món ‘hormone’, pó.sɨ.t^{h}on ~ po.sɨ.t^{h}ón ‘Boston’, p^{h}ɨ.ro.kɨ.ré̞m ‘program’, an.te.rɨ́.s*en ‘Andersen’, k*a.tsa.hɨ.sɨ́.t^{h}an ~ k*a.tsa.hɨ.sɨ.t^{h}án ‘Kazakhstan’

b. [Light(e)-Light] sɨ.k^{h}í ‘ski’, sɨ́.t^{h}a ~ sɨ.t^{h}á ‘star’, p^{h}ɨ́.ro ~ p^{h}ɨ.ró ‘program’, tí.sɨ.k^{h}o ‘disco’, wí.sɨ.k*i ‘whisky’, pó.sɨ.t^{h}ə ‘poster’, ko.rí.k*i ‘Gorki’, p^{h}á.t^{h}ɨ.nə ~ p^{h}a.t^{h}ɨ.nə́ ‘partner’, e.ne.rɨ́.ki ‘energy’, al.le.rɨ́.ki ‘allergy’, tsja.k*a.rɨ́.t^{h}a ‘Jakarta’, mo.sɨ.k^{h}ɨ.pá ‘Moscow’ (< Russian Moskva), al.la.sɨ́.k^{h}a ~ al.la.sɨ.k^{h}á ‘Alaska’

In Light(e)-Heavy, most words appear with final accent.[5] Compared to cases when only the final, and not the penultimate, syllable contains an epenthetic vowel, the bias towards having final accent in order to avoid accenting an epenthetic vowel is not as strong. This fact can be understood given the default penultimate accent in Yanbian loanwords.

In Light(e)-Light, theoretically the final accent class should be preferred to both antepenultimate and penultimate accent classes, given the fact that in Yanbian loanwords the accent is basically located in a two-syllable window at the right edge of the word, and the fact that the penultimate syllable contains an epenthetic vowel. However, in actuality, the three accent classes appear more or less at the same rate. This can be interpreted as the result of the interaction of several differently weighted constraints, such as *LAPSE-RIGHT, NON-FINALITY, WEIGHT-TO-STRESS, *ACCENTED EPENTHESIS, etc., which is discussed in section 4.[6] The fact that antepenultimate accent is relatively frequent not only in Light(e)-Light but also in Light-Light and Light-Light(e) can be explained in a similar way.

Finally, we see the case when both penultimate and final syllables contain an epenthetic vowel.

[5] According to Kubozono (1996, 2001, 2006, 2008) and Tanaka (1996), a similar pattern is observed in Tokyo Japanese as well: almost all the instances of the Light-Heavy pattern with final accent have an epenthetic vowel in the initial light syllable, and almost all Light-Heavy disyllables with an initial epenthetic vowel take final accent.

[6] An alternative way of formalizing exceptions (non-uniform vocabulary) with lexically indexed constraints (Pater 2009) is also possible.

Table 10
Accent when both the penultimate and final syllables contain an epenthetic vowel

Weight	pA	A	P	F	U	Totals	A%	P%	F%
Light(e)-Light(e)	1	76	63			140	54%	45%	0%

(7) Examples of accent and syllable weight when both the penultimate and final syllables contain an epenthetic vowel
ál.p^{h}ɨ.sɨ ~ al.p^{h}ɨ.sɨ ‘Alps’, má.sɨ.k^{h}ɨ ‘mask’, pé.sɨ.t^{h}ɨ ‘best’, só.p^{h}ɨ.t^{h}ɨ ‘soft’, hé.rɨ.ts*ɨ ~ he.rɨ.ts*ɨ ‘hertz’, má.rɨ.k^{h}ɨ ~ ma.rɨ.k^{h}ɨ ‘mark’, ham.pu.rɨ.kɨ ‘Hamburg’, mo.ts*a.rɨ.t^{h}ɨ ‘Mozart’, tan.ma.rɨ.k^{h}ɨ ‘Denmark’, sɨ.p^{h}iŋ.k^{h}ɨ.sɨ ‘sphinx’, t^{h}ɨ.wí.sɨ.t^{h}ɨ ~ t^{h}ɨ.wi.sɨ.t^{h}ɨ ‘twist’

Both antepenultimate and penultimate accents are observed, and there is no example of final accent. As in the cases discussed above, the avoidance of accented epenthesis motivates the assignment of antepenultimate accent, whereas the default accent in Yanbian loanwords seems to be relevant in the assignment of penultimate accent. Thus again we see the gradient (weighted) effect of several constraints.

In sum, loanword syllables with an epenthetic vowel tend to resist the accent even when they are closed. This is one of several factors relevant in accent assignment, suggesting that multiple constraints are interacting with one another to produce the variations observed in our data.

3.4 l-gemination

/l/ in the source language is often adapted as l-gemination in Korean, and Kyengsang Korean consistently treats the apparent heavy syllable which contains the first part of l-gemination (CVl) as light (Kenstowicz and Sohn 2001). Yanbian data do not contradict this observation: CVl-Heavy tends to appear with final accent (68%) whereas CVl-Light tends to appear with penultimate accent (80%), although the percentages are not exactly the same as Light-Heavy structure (penultimate: 48%, final: 48%) or Light-Light structure (penultimate: 61%, final: 32%). This may partially be due to an Island of Reliability effect (Albright 2002). For example, most words of CVl-Heavy structure derive from words ending in the syllables *-lin*, *-ljum*, as in *in.sjul.**lín*** ‘insulin’, *ka.sol.**lín*** ‘gasoline’, *pol.**ljúm*** ‘volume’, *sel.**ljúm*** ‘cerium’, which could accidentally increase the percentage of final accent, due to a strong association between these syllable shapes and the accent type.

Table 11
Accent and l-gemination

Weight	pA	A	P	F	U	Totals	A%	P%	F%
CVl-Heavy		2	44	109	5	160	1%	28%	68%
CVl-Light		2	148	34		184	1%	80%	18%
Totals	0	4	192	143	5	344	1%	56%	42%

(8) Examples of accent and l-gemination

a. [CVl-Heavy] hél.ljum ~ hel.ljúm 'helium', pol.ljúm 'volume', sel.ljúm 'cerium', pól.liŋ ~ pol.líŋ 'bowling', sál.loŋ ~ sal.lóŋ 'salon', ts^{h}jo.k^{h}ol.lés 'chocolate', in.sjul.lín 'insulin', ka.sol.lín 'gasoline', na.il.lón 'nylon', pa.pil.lón 'Babylon', sá.hal.lin ~ sa.hal.lín 'Sakhalin', s*ɨ.t*ál.lin ~ s*ɨ.t*al.lín 'Stalin', p^{h}e.ni.sil.lín 'penicillin'

b. [CVl-Light] ál.la 'Allah', ts^{h}él.lo 'cello', ts^{h}íl.le ~ ts^{h}il.lé 'Chile', k^{h}íl.lə 'killer', k^{h}ól.la 'Cola', mál.li 'Mali', mél.lo 'melodrama', s*ól.lo 'solo', t*ál.la 'dollar', á.p^{h}ol.lo ~ a.p^{h}ól.lo 'Apollo', al.k^{h}al.lí 'alkali', ko.ríl.la ~ ko.ril.lá 'gorilla', ma.níl.la ~ ma.nil.lá 'Manila', na.p*ól.li ~ na.p*ol.lí 'Napoli', ho.nol.lúl.lu 'Honolulu', k*wa.t*e.mál.la 'Guatemala', s*aŋ.p^{h}a.úl.lo 'Sao Paulo', mi.k^{h}el.lan.tsél.lo 'Michelangelo'

Concerning the treatment of l-gemination, Lee (2008) studies the durational difference between geminate /-ll-/ and /-nn-/ in South Kyengsang data taken from three speakers, and reports that /-ll-/ is consistently pronounced with shorter duration than /-nn-/. That is, the treatment of /CVl$_1$/ in /CVl$_1$.l$_2$V(C)/ as a light syllable is due to or reflected in the phonetic grammar of Kyengsang Korean. For comparison, we measured the duration of geminate /-ll-/, /-mm-/ and /-nn-/ taken from five Yanbian speakers. In order to see whether syllable weight structure and accent as well as the geminated consonant can affect the duration of gemination, we recorded words with the following 12 possible combinations.

Table 12

The list of test words. The words with " † " have individual variations in the accent class. Only one example was available for /-mm-/ of Heavy-Light structure with LH accent.

Consonant	Weight structure	Accent	Examples
/-ll-/	Heavy-Heavy	HL	múl.lon 'of course', nól.lam 'surprise'
/-ll-/	Heavy-Heavy	LH	mil.lím 'thick forest', sal.lím 'life'
/-ll-/	Heavy-Light	HL	hwál.lo 'means of survival', t^{h}ál.lo 'escape road'
/-ll-/	Heavy-Light	LH	pəl.lé 'insect', kal.lɛ́ 'divergence'
/-mm-/	Heavy-Heavy	HL	kə́m.mun 'inspection', kɨ́m.mul 'taboo'
/-mm-/	Heavy-Heavy	LH	nam.mún 'south gate', † ts^{h}im.mól 'sinking'
/-mm-/	Heavy-Light	HL	ə́m.ma 'mother', † nám.mɛ 'brother and sister'
/-mm-/	Heavy-Light	LH	mam.má 'meal'
/-nn-/	Heavy-Heavy	HL	pón.niŋ 'instinct', hwán.nan 'sufferings'
/-nn-/	Heavy-Heavy	LH	on.nán 'mildness', man.nám 'encounter'
/-nn-/	Heavy-Light	HL	ə́n.ni 'older sister', pún.no 'anger'
/-nn-/	Heavy-Light	LH	an.nɛ́ 'guidance', njən.nɛ́ 'within the year'

The five subjects were asked to read the same randomly sorted list three times. Table 13 shows the mean values taken from all subjects based on each conditioning factor: consonant (/-ll-/, /-mm-/, /-nn-/), syllable weight structure (Heavy-Heavy, Heavy-Light) and accent (HL, LH).

Table 13

Duration (ms) of gemination based on consonant, syllable weight and accent

Subject	/-ll-/	/-mm-/	/-nn-/	Heavy-Heavy	Heavy- Light	HL	LH
MH	225	245	245	244	231	240	236
YH	174	172	169	172	172	177	166
MK	197	190	197	198	191	190	200
SL	177	185	186	181	183	183	182
CK	218	223	210	214	220	216	218

As seen in Table 13, we cannot say that /-ll-/ consistently has a shorter duration than /-mm-/ or /-nn-/. The duration of each consonant is different depending on the subject. For example, /-ll-/ has the shortest duration in subjects MH and SL while /-mm-/ is shortest in subject MK, and /-nn-/ is shortest in subjects YH and CK. The same is true for syllable weight and accent. Some subjects pronounce one subcategory longer than the other, but other subjects pronounce the other category longer. Paired T-tests done by subject show that there are no significant differences between the three consonants (/-ll-/ and /-mm-/, /-ll-/ and /-nn-/, /-mm-/ and /-nn-/), the two syllable weight structures (Heavy-Heavy and Heavy-Light) and the two accent types (HL and LH).

Table 14
Results of paired T-tests: l-gemination

Data	Mean of the differences	95 percent confidence interval (Lower)	(Upper)	t	df	p
/-ll-/, /-mm-/	-4.896	-17.87	8.078	-1.048	4	0.354
/-mm-/, /-nn-/	1.698	-7.41	10.806	0.518	4	0.632
/-ll-/, /-nn-/	-3.198	-17.266	10.87	-0.631	4	0.562
Heavy-Heavy, Heavy-Light	2.584	-6.762	11.93	0.768	4	0.486
HL, LH	1.168	-8.599	10.935	0.332	4	0.757

Thus, as far as our limited data is concerned, no systematic correlation between the duration of gemination and consonant, syllable weight, and accent was observed. Given the fact that with a df = 4 the power of the test is very low, the results showing no difference should be interpreted cautiously. But we can provisionally conclude that the status of /CVl$_1$/ in /CVl$_1$.l$_2$V(C)/ as a light syllable in Yanbian loanwords is not based on the actual phonetic realization patterns in the Yanbian native system. Why then is /CVl$_1$/ treated as a light syllable? A possible answer is simply because /-l$_1$.l$_2$-/ corresponds with just the onset singleton consonant /l/ in the source language, and there is no corresponding coda which makes a heavy syllable. That is, /CVl$_1$/ corresponds with /CV/ in the source language, so it is naturally treated as light in Yanbian loanwords as well. In sum, at least in Yanbian, the accentual behavior of geminate /-ll-/ in loanwords is not from the native phonetic grammar but from the source language.

CVl containing an epenthetic vowel disprefers having accent, following the general tendency that an epenthetic vowel tends to disprefer accent in Yanbian loanwords. Most cases appear with the final accent class, but there are some cases with the antepenultimate accent class.

Table 15
Accent and l-gemination with epenthesis

Weight	pA	A	P	F	U	Totals	A%	P%	F%
CVl(e)-Heavy		4	2	40		46	9%	4%	87%
CVl(e)-Light		16	4	23		43	37%	9%	53%
Totals	0	20	6	63	0	89	22%	7%	71%

(9) Examples of accent and l-gemination with epenthesis

a. [CVl(e)-Heavy] k^{h}ɨl.ləp ~ k^{h}ɨl.ləp ‘club’, ts^{h}já.p^{h}ɨl.lin ~ ts^{h}ja.p^{h}ɨl.lín ‘Chaplin’, i.sɨl.lám ‘Islam’, mu.sɨl.lím ‘Muslim’, pe.rɨl.lín ‘Berlin’

b. [CVl(e)-Light] hi.t^{h}ɨl.lə ~ hi.t^{h}ɨl.ləˊ ‘Hitler’, me.tɨl.lí ‘medley’, t*ə́.bɨl.lju ~ t*ə.bɨl.ljú ‘W’, ín.p^{h}ɨl.le ~ in.p^{h}ɨl.lé ‘inflation’

3.5 Summary

The descriptive generalizations underlying Yanbian loanword accentuation can be summarized as follows:

(10) Yanbian loanword accentuation

a. In Yanbian loanwords, the accent is basically located in a two-syllable window at the right edge of the word; antepenultimate accent can appear when the penultimate syllable is light (especially with an epenthetic vowel).
b. A discrepancy is observed between disyllabic and longer words. In disyllabic loanwords, there is a preference for penultimate accent when the syllables in the two-syllable window are equivalent in weight (Heavy-Heavy, Light-Light). Heavy-Light structures increase the penultimate bias while Light-Heavy decrease it.
c. In three-or-more syllable words, there is still a preference for penultimate accent, but the syllable weight effect emerges more clearly and accent polarization (Heavy-Light → penultimate, Light-Heavy → final) is observed.
d. Epenthetic vowels disprefer having a pitch peak.
e. A syllable with l-gemination is treated as light (as in Kyengsang Korean), and this accentual behavior is not due to the native phonetic grammar but originates from the source language.
f. As a whole, Yanbian loanword accentuation can be interpreted with the interaction of several relevant constraints weighted differently.

In the next section, we examine how these descriptive generalizations and the gradient distribution can be understood with weighted constraints by employing a learning algorithm. We also compare the accent distribution of loanwords with that of native and Sino-Korean, and discuss where the loanword accentuation comes from.

4. Analysis

One of the major questions in loanword phonology is where loanword adaptation patterns come from (the emergence question, Kang 2011). For example, the tendency listed in (10a)—"Yanbian loanword accent is located in a two-syllable window at the right edge of the word"—seems to reflect Yanbian native grammar, since as mentioned above, at least in native simplex longer words, the possible accent classes are only penultimate and final as a rule. On the other hand, (10d)—"Epenthetic vowels disprefer having a pitch peak"—is probably not from Yanbian native grammar.[7] This can result from either English phonology/phonetics (the segment to which a syllable with an epenthetic vowel corresponds does not have any stress) or general phonetics/Universal Grammar (less sonorous syllables disprefer having a pitch peak).

Thus Yanbian loanword accentuation raises the following questions: do the loanword

[7] For example, an epenthetic vowel /ɨ/ that is inserted before some suffixes such as *-mjən* (conditional), *-ni* (sequential), etc. is assigned a high pitch accent when it follows an unaccented verbal stem in native grammar, e.g. *mək-* 'eat' *mə.k-ɨ́.-mjən, mə.k-ɨ́.-ni*.

adaptation patterns result from the grammar of the recipient language or from the source language? Is the adaptation based on phonology or phonetics or both? Can Universal Grammar be working in the adaptation? In this section, we examine how Yanbian loanword accentuation is explained from these perspectives.

4.1 Constraint weighting

As mentioned above, Yanbian loanword accentuation can be interpreted with the interaction of several relevant weighted constraints. We first propose an accent assignment model for Yanbian with weighted constraints. We employ Jäger's (2007) Stochastic Gradient Ascent learning algorithm, using a Perl script implemented by Albright (2008).[8] In this learning algorithm, faithfulness constraints start at 0 and markedness constraints at 10. The learning data is fed into the learning program. The data contains the input forms, output candidates, the actual frequency distribution of each output candidate (e.g. frequency of each accent class) and the constraints that are assumed to be relevant. Based on the existing frequencies, the learning algorithm assigns different weights to each constraint. The goodness of a form is the result of the sum of weighted constraint violations, as in stochastic OT, maximum entropy models, and Harmonic Grammar. We ran the simulation 50,000 times. As far as we know, this is the first study to try to find a gradient constraint weighting in loanword adaptation patterns. (11)–(13) show the inputs, outputs, and constraints that are employed in the learning data. For simplicity, the structures with [l] are excluded in this simulation. We use standard constraints from metrical phonology except for INITIAL, which is included to see the difference between the disyllabic and longer loanwords.

(11) Input (33 inputs based on syllable weight. X indicates any weight.)
a. Monosyllabic (2): Heavy, Light
b. Disyllabic (10): Heavy-Heavy, Heavy-Light, Light-Heavy, Light-Light, Heavy-Heavy(e), Heavy-Light(e), Light-Heavy(e), Light-Light(e), Light(e)-Heavy, Light(e)-Light
c. Trisyllabic (11): X-Heavy-Heavy, X-Heavy-Light, X-Light-Heavy, X-Light-Light, X-Heavy-Heavy(e), X-Heavy-Light(e), X-Light-Heavy(e), X-Light-Light(e), X-Light(e)-Light(e), X-Light(e)-Heavy, X-Light(e)-Light
d. Quadrisyllabic (10): X-X-Heavy-Heavy, X-X-Heavy-Light, X-X-Light-Heavy, X-X-Light-Light, X-X-Heavy-Light(e), X-X-Light-Heavy(e), X-X-Light-Light(e), X-X-Light(e)-Light(e), X-X-Light(e)-Heavy, X-X-Light(e)-Light

(12) Output
a. Monosyllabic: Final, Unaccented
b. Disyllabic: Penultimate, Final, Unaccented
c. Trisyllabic: Antepenultimate, Penultimate, Final, Unaccented

[8] Other models of weighted constraints such as stochastic OT (Boersma and Hayes 2001), the Maxent Grammar Tool (http://www.linguistics.ucla.edu/people/hayes/MaxentGrammarTool/, Hayes and Wilson 2008), or noisy Harmonic Grammar (Potts, Pater, Jesney, Bhatt, and Becker 2010) could also be investigated. This paper is not a contribution to learning theory but uses the tools made available by recent learning theory for descriptive analytic purposes.

d. Quadrisyllabic: Pre-antepenultimate, Antepenultimate, Penultimate, Final, Unaccented

(13) Constraints (all markedness constraints)[9]

a. WEIGHT-TO-STRESS (WtoS): a heavy syllable cannot appear in nonprominent position. (Prince and Smolensky 1993)
b. STRESS-TO-WEIGHT (StoW): a light syllable cannot appear in prominent position. (Kager 1999)
c. *UNACCENTED (*U): every word must have a pitch peak. (Cf. Culminativity, Hayes 1995:24-25)
d. NON-FINALITY (NF): a final syllable is extra-metrical (Hung 1994).
e. INITIAL: a word-initial accent is favored.
f. *LAPSE-RIGHT (*LAPSE-R): a maximum of one unstressed syllable separates the rightmost stress from the right edge of a stress domain. (No more than one syllable separates the rightmost syllable with a level 1 grid mark from the right edge of the word. Cf. Green and Kenstowicz 1995; Gordon 2002)
g. *ACCENTED EPENTHESIS (*ə́): an epenthetic vowel cannot have a pitch peak. (Cf. Shinohara 1997; Alderete 1999)

Table 16 shows the example of the learning data. "Input" column indicates the syllable structure of the learning data, "Output" column indicates the candidate accent class, and "Output #" is the number of words that appear with each accent class. "WtoS"—"*ə́" are the relevant constraints.

[9] We have avoided constraints that explicitly recognize metrical grouping, since we lack the evidence to claim that there is such a hierarchical grouping in Yanbian or more generally in contemporary Korean. To the best of our knowledge, the language lacks the prosodic morphology effects found in languages such as English, Arabic, Japanese, etc.: CV light syllables are possible words (no minimality), the segmental phonology does not refer to foot boundaries (as in English flapping). The only segmental effect is the preference for tense/aspirated and high tone—a relation that refers directly to F0 with no need for metrical grouping. The Yanbian accent has a single and uniform phonetic correlate (high pitch) and so it seems sufficient to propose constraints that govern its distribution directly. Also as a reviewer points out, *ACCENTED EPENTHESIS is normally interpreted as a faithfulness constraint: every accented vowel in the output must have a correspondent in the input. We have classified this as a markedness constraint for analytic convenience.

Table 16
Example of the learning data (pA = pre-antepenultimate, A = antepenultimate, P = penultimate, F = final, U = unaccented)

Input	Output	Output #	WtoS	StoW	*U	NF	INITIAL	*LAPSE-R	*ə́
Heavy	F	256				1			
	U		1		1		1		
Light	F	130		1		1			
	U				1		1		
Heavy-Heavy	P	218	1						
	F	68	1			1	1		
	U	5	2		1		1		
X-Heavy-Heavy	A		2					1	
	P	33	1				1		
	F	18	1			1	1		
	U		2		1		1		
X-X-Heavy-Heavy	pA		2					2	
	A		2				1	1	
	P	18	1				1		
	F	1	1			1	1		
	U		2		1		1		
Heavy-Heavy(e)	P	72	1						
	F		1			1	1		1
	U		2		1		1		
X-Heavy-Heavy(e)	A		2					1	
	P	6	1				1		
	F		1			1	1		1
	U		2		1		1		

The obtained weights are as follows.

Table 17
Obtained weights

*U	9.02
*LAPSE-R	4.57
*ə́	2.66
INITIAL	1.56
StoW	1.55
NF	0.63
WtoS	0.02

The existing patterns are well reflected in the obtained weights. *U is expectedly the strongest constraint since the unaccented class does not appear in Yanbian loanwords as a general rule.

Similarly *LAPSE-R was assigned a higher weight, reflecting the relative rarity of antepenultimate accent. INITIAL is weaker than *LAPSE-R, thus resulting in the biased initial (= penultimate) accent in disyllabic loanwords while minimizing the initial (= antepenultimate/pre-antepenultimate) accent in trisyllabic/quadrisyllabic loanwords. NF results in the default penultimate accent as a whole.

This weight hierarchy also explains the discrepancy between the disyllabic and three-or-more syllable words discussed above. Since INITIAL is weighted higher than syllable weight constraints (in particular WtoS), penultimate accent appears strongly in disyllabic words; thus the syllable weight effect works gradiently. In trisyllabic/quadrisyllabic loanwords, on the other hand, since *LAPSE-R is ranked much higher than INITIAL, antepenultimate/pre-antepenultimate accent is extremely disfavored. As a result, the effect of syllable weight constraints emerges clearly in these cases, since INITIAL is essentially inactive due to the higher weight of *LAPSE-R. On the other hand, NF is ranked lower than at least one of the syllable weight constraints (StoW), and so the default is not strongly biased to penultimate accent.

Table 18 shows the application of the weights in the Heavy-Heavy structure. (We referred to Hayes and Wilson 2008:383–385, for this calculation.) The obtained weights are listed under each constraint. "Score" indicates the weighted sum of the candidate's constraint violations. The Maxent value is calculated based on the negated score. The output probability is determined by its share in the summation of maxent values of all candidates (in the case of Table 18, the summation of maxent values of all candidates is 0.980 + 0.110 + 2.44e-05 = 1.090). Thus we see that the output probability of the penultimate accent class is highest (0.899), followed by final accent class (0.101) and unaccented class (0.000) in the Heavy-Heavy structure.

Table 18

Application of the weights in Heavy-Heavy structure

	WtoS 0.02	StoW 1.55	*U 9.02	NF 0.63	INITIAL 1.56	*LAPSE-R 4.57	*ə 2.66	Score	Maxent value	Output probability
P	1							(0.02·1) = 0.02	exp(-0.02) = 0.980	0.980/1.090 = 0.899
F	1			1	1			(0.02·1) + (0.63·1) + (1.56·1) = 2.21	exp(-2.21) = 0.110	0.110/1.090 = 0.101
U	2		1		1			(0.02·2) + (9.02·1) + (1.56·1) = 10.62	exp(-10.62) = 2.44e-05	2.44e-05/1.090 = 0.000

Table 19 shows the given/predicted distribution of Heavy-Heavy, Heavy-Light, Light-Heavy, and Light-Light structures in the simulation based on Jäger (2007). The discrepancy is more or less well reflected in this simulation. (The percentage of final accent in X-X-Heavy-Heavy is lower in the predicted distribution than in the given distribution, but this may be an accidental result since the original given data with this structure is very small (only 19).)

Table 19
Given/predicted distribution based on Jäger (2007). Most frequent adaptations are in bold.

Weight	Cand	Predict	Given	Weight	Cand	Predict	Given
Heavy-Heavy	P	**0.899**	**0.749**	Light-Heavy	P	**0.650**	**0.642**
	F	0.101	0.234		F	0.350	0.340
	U	0.000	0.017		U	0.000	0.019
X-Heavy-Heavy	A	0.031	0.000	X-Light-Heavy	A	0.061	0.058
	P	**0.633**	**0.647**		P	0.264	0.280
	F	0.337	0.353		F	**0.675**	**0.651**
	U	0.000	0.000		U	0.000	0.012
X-X-Heavy-Heavy	pA	0.000	0.000	X-X-Light-Heavy	pA	0.001	0.080
	A	0.007	0.000		A	0.014	0.014
	P	**0.648**	**0.947**		P	0.277	0.391
	F	0.345	0.053		F	**0.709**	**0.500**
	U	0.000	0.000		U	0.000	0.014
Heavy-Light	P	**0.977**	**0.958**	Light-Light	P	**0.899**	**0.871**
	F	0.023	0.040		F	0.101	0.129
	U	0.000	0.002		U	0.000	0.000
X-Heavy-Light	A	0.042	0.000	X-Light-Light	A	0.132	0.141
	P	**0.863**	**0.993**		P	**0.566**	**0.439**
	F	0.096	0.007		F	0.302	0.420
	U	0.000	0.000		U	0.000	0.000
X-X-Heavy-Light	pA	0.000	0.000	X-X-Light-Light	pA	0.002	0.010
	A	0.009	0.000		A	0.031	0.030
	P	**0.892**	**0.980**		P	**0.631**	**0.596**
	F	0.099	0.020		F	0.336	0.361
	U	0.000	0.000		U	0.000	0.003

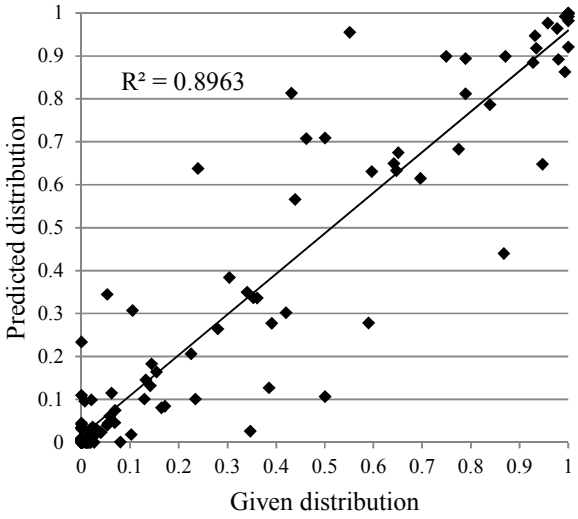
A regression analysis using glm (general linear model) between the given and predicted values shows a significant result. Figure 1 shows this correlation.

Table 20
Results of a regression analysis using glm: given and predicted values

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	0.01428	0.01247	1.146	0.254	
Given	0.94464	0.02862	33.008	< 2e-16	***

Significance codes: '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

Figure 1
Correlation between the given and predicted values



Thus, we tentatively conclude that the gradient patterns underlying Yanbian loanword accent can be expressed with a grammar employing weighted constraints. We now turn to the question of the origin of the weight hierarchy of these constraints. The null hypothesis is that these reflect biases in the native grammar.

4.2 Comparison with native and Sino-Korean words

4.2.1 Accent distribution in native and Sino-Korean words

Let us compare the accent class distributions in loanwords with native simplex nouns on the one hand and with Sino-Korean nouns on the other. Sino-Korean words are composed of morphemes borrowed much earlier (9–10th c.) from Middle Chinese (Kōno 1968; Ito 2007).

Table 21 shows the accent distribution in native simplex nouns, where the data are aggregated from our six speakers (the same data as Table 1).

Table 21
Accent distribution in native simplex nouns

Monosyllabic		
Accent	Numbers	Ratios
H	1,906	84%
L	351	16%
Totals	2,257	

Disyllabic		
Accent	Numbers	Ratios
LH	3,241	69%
HL	1,119	24%
LL	354	8%
Totals	4,714	

Trisyllabic		
Accent	Numbers	Ratios
LLH	1,262	78%
LHL	241	15%
HLL	84	5%
LLL	36	2%
Totals	1,623	

Quadrisyllabic		
Accent	Numbers	Ratios
LLHL	156	54%
LLLH	122	42%
LLLL	7	2%
HLLL	2	1%
LHLL	1	0%
Totals	288	

In native simplex nouns, the final accent classes (H, LH, LLH) are the dominant classes in monosyllabic/disyllabic/trisyllabic words. In quadrisyllabic words, the ratio of the penultimate accent class is only slightly higher than the final accent class, but in actuality the accent patterns are quite strongly correlated with the segmental sequences in the final two-syllable window (Island of Reliability effect, Albright 2002). For example, the words that end with *-ŋ.i* tend to appear with final accent (*a.tsi.raŋ.í* 'heat haze', *na.pu.rɛŋ.í* 'scraps'), whereas the words that end with *-ra.ki* or *-rɛ.ki* tend to appear with penultimate accent (*hɛ.o.rá.ki* 'white egret', *ho.rɨ.ré.ki* 'whistle'). Given this correlation between the segmental composition and the accent patterns and the relatively small number of examples of quadrisyllabic words compared with monosyllabic/disyllabic/trisyllabic words, we can assume that the frequency distribution in quadrisyllabic words is somewhat accidental and that final accent is the default accent class in native simplex nouns as a whole. That is, the initial (or NF) constraint is not active (weighted low) in native words. This is a different tendency from the loanword accentuation where penultimate accent is the default class. It is worth noting that Yanbian differs from Kyengsang Korean in this respect. The majority of native nouns in Middle Korean had final accent (Ramsey 1978). Kyengsang retracted the accent by one syllable so the majority of native nouns are in the penultimate accent class in Kyengsang, where penultimate accent is the default in loanwords as well.

On the other hand, the longer the syllable count is, the lower the ratio of the unaccented class (L, LL, LLL, LLLL). LLLL was observed in only one speaker, and other speakers pronounced the same words with final accent. These facts coincide with the highest weight of *U (9.02) in Yanbian loanword accentuation. Similarly, antepenultimate accent shows a very low ratio in trisyllabic native nouns (HLL, 5%), and in quadrisyllabic native nouns there is virtually no example of pre-antepenultimate/antepenultimate accent (HLLL/LHLL, 1% and 0% respectively).

This tendency again mirrors the higher weight of *LAPSE-R (4.57) in loanword accentuation.

Now we compare the relation between syllable weight and accent class in native simplex words with that in loanwords. We first concentrate on the disyllabic words, where the difference between the two lexical classes is most evident. The unaccented class is aggregated with the final accent class, since: a) the unaccented class merges to the final accent class in isolation (unsuffixed citation form); b) the unaccented class does not appear in loanwords as a rule; c) the number of words in the unaccented class is relatively small in general; d) individual variations for the unaccented class are substantial and in many cases the words with the unaccented class for one speaker are pronounced with the final accent class by other speakers.

Table 22

Comparison of the relation between syllable weight and accent class in disyllabic native simplex nouns and disyllabic loanwords

	Native					Loan				
Weight \ Accent	P	F	Totals	P%	F%	P	F	Totals	P%	F%
Heavy-Heavy	118	410	528	22%	78%	218	73	291	75%	25%
Heavy-Light	327	677	1,004	33%	67%	388	17	405	96%	4%
Light-Heavy	212	1,396	1,608	13%	87%	450	251	701	64%	36%
Light-Light	462	1,112	1,574	29%	71%	834	123	957	87%	13%
Totals	1,119	3,595	4,714	24%	76%	1,890	464	2,354	80%	20%

In native words where the final accent class is a default, the percentage of final accent is highest in Light-Heavy (87%), which is the most preferred syllable structure for this accent pattern. Similarly the percentage of final accent is lowest in Heavy-Light (67%), which is the least preferred syllable structure for this accent pattern, although the difference between Heavy-Light and Light-Light is not large (67% vs. 71%). On the other hand, in loanwords where the penultimate accent class is the default, the percentage of penultimate accent is highest in Heavy-Light (96%) and lowest in Light-Heavy (64%), which again reflects the correlation between the preferred syllable structure and the default accent class. Given this distributional difference, we cannot say that loanword accentuation simply reflects the accent patterns in native words. We ran a mixed effects logistic regression model with accent location (penultimate/final) as the dependent variable, based on the data from six subjects who participated in the investigation of both native words and loanwords. A random intercept was set for subjects.[11] The result shows that all factors (syllable weight of penultimate/final syllables, lexical class—native vs. loan) are strongly significant: a heavy penult and a loanword (as opposed to a native word) are positively associated with penultimate accent, whereas a heavy final is negatively associated with penultimate accent.

[11] When a random intercept was set for both items and subjects, the problem of false convergence occurred, and hence the random item factor was excluded. The same is true for other statistical analyses in this paper that have only a random subject factor.

Table 23
Results of a mixed effects logistic regression model: lexical class difference between disyllabic native words and disyllabic loanwords

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-0.91728	0.07964	-11.52	< 2e-16	***
Weight-Penult-Heavy	0.36129	0.06472	5.58	2.37e-08	***
Weight-Final-Heavy	-0.98039	0.06476	-15.14	< 2e-16	***
Lexical class-Loan	2.74228	0.07346	37.33	< 2e-16	***

Significance codes: '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

This suggests that the accentual assignment in native words and loanwords in Yanbian follows different rules, which are internally defined depending on each lexical class. In this respect, Yanbian is crucially more informative than Kyengsang. As observed above, the majority of native Kyengsang nouns have penultimate accent due to the historical accent shift. One might thus conclude that penultimate accent in loanwords simply reflects the native grammar bias. Yanbian shows that this inference is not necessarily valid.

Next we compare trisyllabic native simplex words with trisyllabic loanwords. The unaccented class is aggregated with the final accent class.

Table 24
Comparison of the relation between syllable weight and accent class in trisyllabic native simplex nouns and trisyllabic loanwords

	Native							Loan						
Weight \ Accent	A	P	F	Totals	A%	P%	F%	A	P	F	Totals	A%	P%	F%
Heavy-Heavy			7	7	0%	0%	100%		33	18	51	0%	65%	35%
Heavy-Light	13	41	307	361	4%	11%	85%		140	1	141	0%	99%	1%
Light-Heavy	5	1	104	110	5%	1%	95%	30	146	345	521	6%	28%	66%
Light-Light	66	199	880	1,145	6%	17%	77%	165	513	491	1,169	14%	44%	42%
Totals	84	241	1,298	1,623	5%	15%	80%	195	832	855	1,882	10%	44%	45%

As seen in Table 24, the accent is strongly biased to the final syllable in native words (80% as a whole). A syllable weight effect is observed in that penultimate accent can appear relatively frequently in Heavy-Light (11%) and Light-Light (17%) but essentially not at all in Light-Heavy (1%). Still, the syllable weight effect in native words is not as strong as that in loanwords, where 99% of words appear with penultimate accent in Heavy-Light while 66% of words appear with final accent in Light-Heavy. Thus, the distribution is clearly different between trisyllabic native words and trisyllabic loanwords. We ran a mixed effects logistic regression model (excluding the small antepenultimate accent class), again based on the data from six subjects. A random intercept was set for items and subjects. It shows a similar result to the disyllabic words. Weight in the penultimate as well as the final syllables and lexical class are significant predictors: a heavy penult and a loanword (as opposed to a native word) are positively associated with

penultimate accent, whereas a heavy final is negatively associated with penultimate accent. Given this result, we can conclude that the accent distribution in loanwords does not result from the statistics of native words.

Table 25
Results of a mixed effects logistic regression model: lexical class difference between trisyllabic native words and trisyllabic loanwords

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-6.4406	0.5288	-12.180	< 2e-16	***
Weight-Penult-Heavy	1.5739	0.7806	2.016	0.0438	*
Weight-Final-Heavy	-5.0117	0.7177	-6.983	2.90e-12	***
Lexical class-Loan	8.0117	0.6178	12.968	< 2e-16	***

Significance codes: '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

Table 26 shows the accent distribution in Sino-Korean nouns taken from one speaker. As for the Sino-Korean nouns, only monosyllabic and disyllabic words are examined, since trisyllabic and longer Sino-Korean words are as a rule compounds. Precisely speaking, disyllabic Sino-Korean words are composed of two Sino-Korean morphemes, but in general most Sino-Korean morphemes are bound forms and hence we treat disyllabic Sino-Korean words as a kind of simplex word.

Table 26
Accent distribution in Sino-Korean nouns

Monosyllabic		
Accent	Numbers	Ratios
H	263	82%
L	58	18%
Totals	321	

Disyllabic		
Accent	Numbers	Ratios
HL	5,346	67%
LH	2,284	29%
LL	322	4%
Totals	7,952	

(14) Examples of Sino-Korean nouns
a. H: hjə́ŋ 刑 'punishment', káŋ 江 'river', mák 膜 'membrane', pjə́ŋ 病 'sickness', ts^{h}ɛ́k 冊 'book'
b. L: hjəŋ 兄 'older brother', kɨm 金 'gold', tsən 前 'before', ts^{h}ən 千 'thousand', ts^{h}im 鍼 'needle'
c. HL: án.ma 按摩 'massage', ó.il 五日 'five days', hjə́n.pəp 憲法 'constitution', kó.a 孤兒 'orphan', tóŋ.saŋ 銅像 'bronze statue'
d. LH: jən.kí 煙氣 'smoke', hɛ.án 海岸 'seashore', koŋ.hák 工學 'engineering', njən.tɛ́ 年代 'age', ts^{h}a.í 差異 'difference'
e. LL: in.sam 人蔘 'ginseng', o.tsən 午前 'forenoon', koŋ.wən 公園 'park', pjəŋ.wən 病院 'hospital', toŋ.nam 東南 'southeast'

Monosyllabic Sino-Korean words show a similar distribution to monosyllabic native words, but disyllabic Sino-Korean words have a totally different distribution than disyllabic native words: HL is the largest accent class in Sino-Korean, whereas LH is the largest class in native words. Sino-Korean words may be similar to loanwords in this respect, but in actuality they show different patterns with respect to the syllable weight effect. Table 27 compares the relation between syllable weight and accent class in disyllabic Sino-Korean words and loanwords. The final accent class and the unaccented class are aggregated.

Table 27

Comparison of the relation between syllable weight and accent class in disyllabic Sino-Korean nouns and loanwords. Syllable weight in Sino-Korean is based on the underlying weight of each morpheme and a resyllabification across the two morphemes is not taken into account. The loanword data is the same as in Table 22.

	Sino-Korean					Loan				
Weight \ Accent	P	F	Totals	P%	F%	P	F	Totals	P%	F%
Heavy-Heavy	1,956	1,057	3,013	65%	35%	218	73	291	75%	25%
Heavy-Light	1,243	534	1,777	70%	30%	388	17	405	96%	4%
Light-Heavy	1,333	728	2,061	65%	35%	450	251	701	64%	36%
Light-Light	814	287	1,101	74%	26%	834	123	957	87%	13%
Totals	5,346	2,606	7,952	67%	33%	1,890	464	2,354	80%	20%

Although the frequency of the penultimate accent class as a whole is much higher in Sino-Korean (67%) than in native words (24%), it is lower than in loanwords (80%). Similarly the syllable weight effect in Heavy-Light is not as strongly observed in Sino-Korean, where only 70% of words appear with penultimate accent, whereas 96% of words in this structure appear with penultimate accent in loanwords. A logistic regression model using glm function (Table 28) based on the data from one subject who participated in the investigation of both Sino-Korean words and loanwords shows that the weight of the final syllable and the lexical class (Sino-Korean words and loanwords) are highly significant: a heavy penult and a loanword (as opposed to a Sino-Korean word) are positively associated with penultimate accent, whereas a heavy final is negatively associated with penultimate accent.

Table 28
Results of a logistic regression model: lexical class difference between disyllabic Sino-Korean words and disyllabic loanwords

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	1.01297	0.04897	20.684	< 2e-16	***
Weight-Penult-Heavy	0.01562	0.04601	0.339	0.734	
Weight-Final-Heavy	-0.46230	0.04731	-9.771	< 2e-16	***
Lexical class-Loan	0.62161	0.06739	9.224	< 2e-16	***

Significance codes: '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

Thus, we cannot attribute the loanword accentuation pattern to Sino-Korean words, which appear to show a similar accent distribution as loanwords. The bias to HL in disyllabic Sino-Korean words is rather due to the historical development of this lexical class and the regular correspondence between Middle Korean accent and Middle Chinese tones (Ito 2008b). That is, Sino-Korean word accent basically reflects the Middle Chinese tonal distribution and constraints such as INITIAL/NF that seem to be active in loanwords are irrelevant here. In sum, we conclude that the correlation between syllable weight and accent class in loanwords derives from neither native nor Sino-Korean word accentuations.

As an experiment, we performed a simulation by again using Jäger (2007) to test if different lexical classes can have a different hierarchy of constraint weights. For simplicity, we treated only disyllabic native words and loanwords with Heavy-Heavy, Heavy-Light, Light-Heavy and Light-Light structures, using the constraints NF, WtoS, StoW, and ALIGN-RIGHT (ALIGN-R, a high pitch is assigned at the right edge of the word). INITIAL is aggregated with NF since these two have essentially the same effect in disyllabic words. Given the distributions of disyllabic native words and loanwords discussed above, we expect ALIGN-R to get the greatest weight in native words, whereas NF should get the greatest weight in loanwords; also we expect that the syllable weight constraints (WtoS, StoW) should obtain relatively smaller weights in native than in loanwords. In fact, this is what we obtained. Tables 29 and 30 show the obtained weights and given/predicted distributions. Thus, we conclude that each lexical class can have different weighted constraint hierarchies.

Table 29
Obtained weights

Native		Loan	
ALIGN-R	10.27	NF	10.59
NF	9.18	ALIGN-R	8.98
WtoS	0.41	WtoS	0.72
StoW	0.26	StoW	0.57

Table 30
Given/predicted distributions

		Native		Loan	
Weight	Cand	Predict	Given	Predict	Given
Heavy-Heavy	P	0.252	0.223	0.833	0.749
	F	0.748	0.777	0.167	0.251
Heavy-Light	P	0.397	0.326	0.948	0.958
	F	0.603	0.674	0.052	0.042
Light-Heavy	P	0.147	0.132	0.579	0.642
	F	0.853	0.868	0.421	0.358
Light-Light	P	0.252	0.294	0.833	0.871
	F	0.748	0.706	0.167	0.129

Still one may think that the different accent distributions between native words and loanwords are due to a faithfulness constraint to a lexical accent that exists in native words but not in loanwords (or is weighted much lower than relevant markedness constraints in loanwords) and that the actual weight hierarchy of markedness constraints is the same or similar between these two classes (see the analysis by J-S. Kim 2009 for North Kyengsang Korean and Lee 2008 for South Kyengsang Korean for analyses along these lines). It is difficult to judge whether or not this covert grammar is operative in Yanbian, since we examine only the existing data. In order to investigate this point further, we need to conduct a wug test with Yanbian speakers: if Yanbian speakers would tend to assign different accentual patterns to a wug test word that looks like a native word versus a wug test word that looks like a loanword, then we can conclude that these two lexical classes have different weight hierarchies and thus exclude the possibility of covert grammar, since wug test words (by definition) lack an underlying lexical accent that a faithfulness constraint could refer to. In fact, this is the result observed in our wug test, which is discussed in the next section.

4.2.2 Wug test

The tendency discussed in 4.2.1 suggests that Yanbian speakers have different accent assignment rules for each lexical class (native, Sino-Korean, loanword). In order to confirm this point, we conducted a wug test with three Yanbian native speakers who were also consultants of this study. The subjects were instructed to read a list of 180 words consisting of extremely unfamiliar (infrequent) native words and loanwords taken from contemporary Korean dictionaries or websites, whose syllable structures are either Heavy-Light, Light-Heavy, or Light-Light (e.g. *tɨl.me* ‘string to tie a shoe and a foot’, *ko.sak* ‘supporting wood’, *nə.ri* ‘pyorrhea’ for native, *p*^{h}*ɛn.tsi* ‘pansy’, *pa.sun* ‘bassoon’, *ho.mə* ‘Homer’ for loans). Note that in Yanbian many fewer loanwords are used than in South Korean. For the entire list of these wug test words, see the appendix. We used disyllabic wug test words because the difference between the two lexical classes is most evident in this condition. Subjects were asked to read the tested words (sorted in

random order) by putting them in the following sentence frame: *ikəsɨn* _____ *imnida*, 'This is ____'. The tested words, while assumed to be infrequent words (the rarity was judged by the author), are actually attested in dictionaries/websites, and so the subjects were asked to indicate if they knew the word. The words that subjects could recognize were excluded from our analysis (4% on average were excluded). Our expectation is that even though the subjects do not recognize the words, they can still classify them into the two lexical classes (native and loan) and assign accent classes differently. Table 31 shows the results.

Table 31
Wug test results

Subject YH	Native						Loan					
Weight \ Accent	P	F	U	Totals	P%	F%	P	F	U	Totals	P%	F%
Heavy-Light	15	14		29	52%	48%	29	1		30	97%	3%
Light-Heavy	1	28		29	3%	97%	20	10		30	67%	33%
Light-Light	11	17		28	39%	61%	30			30	100%	0%
Totals	27	59	0	86	31%	69%	79	11	0	90	88%	12%

Subject SL	Native						Loan					
Weight \ Accent	P	F	U	Totals	P%	F%	P	F	U	Totals	P%	F%
Heavy-Light	11	16		27	41%	59%	17	11		28	61%	39%
Light-Heavy	1	24	1	26	4%	92%	11	17	2	30	37%	57%
Light-Light	6	19		25	24%	76%	28	2		30	93%	7%
Totals	18	59	1	78	23%	76%	56	30	2	88	64%	34%

Subject MH	Native						Loan					
Weight \ Accent	P	F	U	Totals	P%	F%	P	F	U	Totals	P%	F%
Heavy-Light	18	12		30	60%	40%	27	2		29	93%	7%
Light-Heavy	8	23		31	26%	74%	20	10		30	67%	33%
Light-Light	18	12		30	60%	40%	29			29	100%	0%
Totals	44	47	0	91	48%	52%	76	12	0	88	86%	14%

As seen in Table 31, subject YH apparently distinguished the two lexical classes and showed a different accentual assignment accordingly.[12] As a whole, final accent was predominant in native words (69%), whereas penultimate accent was predominant in loanwords (88%). The syllable weight effect was also different between the two lexical classes, in that Light-Heavy in native words received a very high percentage of final accent (97%), whereas Heavy-Light in loanwords appeared with penultimate accent exclusively (97%). The other two subjects SL and MH showed a similar tendency as speaker YH, with a different degree of bias to each accent pattern in two

[12] It is assumed that Yanbian speakers distinguish native words and loanwords based on different phonotactic patterns that are observed in existing words. For example, vowel harmony or the ban against having two aspirated onsets in one word appears more strongly in native words than in loanwords; CV.Cə rarely appears in native words but it is very frequent in loanwords.

lexical classes.[13]

A mixed effects logistic regression model with accent location (penultimate/final) as the dependent variable (random intercepts set for items and subjects) demonstrates that weight in the final syllable and lexical class are strongly significant factors (Table 32): a final heavy is negatively associated with penultimate accent, while a loanword (as opposed to a native word) are positively associated with penultimate accent. Given this result, we conclude that there are different strategies at work for accentual assignment in native words vs. loanwords in Yanbian, which in turn supports the hypothesis that the loanword accentuation does not directly derive from Yanbian native grammar.

Table 32
Results of a mixed effects logistic regression model: wug test

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-0.04851	0.56947	-0.085	0.932	
Weight-Penult-Heavy	-0.40315	0.41953	-0.961	0.337	
Weight-Final-Heavy	-2.91433	0.44942	-6.485	8.90e-11	***
Lexical class-Loan	3.44086	0.37182	9.254	< 2e-16	***

Significance codes: '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

The results of the wug test are also important in that these tested words (by definition) do not have a faithfulness constraint to the lexical accent but still showed a similar division between the two lexical classes. That is, the different accentuation patterns between native words and loanwords are not due to the lexical accent but due to the lexical class itself. Thus, we conclude that the covert grammar of the native system is not relevant in the loanword accentuation in Yanbian.

4.3 Where does the loanword accentuation derive from?

Where then does the loanword accentuation pattern come from? One possibility is that it results from Universal Grammar. Kenstowicz and Sohn (2001) suggest that the Universal Grammar default setting plays a role in accentual adaptations in Kyengsang Korean. Similar examples of "emergence of the unmarked" effects (McCarthy and Prince 1994) in Japanese loanword adaptations are discussed in this manner by Shinohara (1997a, b, 2000). Following this proposal, one might say that the default penultimate accent in Yanbian loanwords comes from Universal

[13] The fact that some native words appear with the penultimate accent class is considered to be a "nonce word effect": in Yanbian, it is assumed that the words which do not look like any existing native/Sino-Korean/loanwords are treated as perfect "nonce words" and that the accentual assignment in nonce words tends to be similar to loanwords, based on their shared feature, "non-nativeness" (= foreignness). For example in the case of speaker MH, she probably classified at least some native test words as nonce words, and hence these words appeared with penultimate accent. Thus, some differences among the subjects are attributed to different treatments of unfamiliar words: to what extent each speaker treats unfamiliar words as nonce words or not.

Grammar, although this is speculative.

On the other hand, Kubozono (2006, 2008) proposes that loanword accentuation in Japanese is closely related to native accentuation. His major findings are summarized as follows.

(15) Loanword accentuation in Japanese (Kubozono 2006, 2008)

a. In Tokyo Japanese, loanwords favor the accented pattern whereas native and Sino-Japanese words prefer the unaccented pattern. These apparent differences in the accent distributions of the three word classes are attributed to distributional differences in syllable structure among them. (For example, native words contain more words composed of only light syllables, whereas loanwords are abundant in heavy syllables and epenthetic vowels.)
b. Accented nouns tend to follow the same accent rule (the "antepenultimate rule" of McCawley 1968), which is fundamentally similar to the accent rule of Latin and English, regardless of the word class.
c. Preference for the accented class in loanwords reflects the abrupt pitch fall found in the major stressed syllable of English words when they are pronounced in isolation (faithfulness to accentedness in English source).
d. The precise accent locus in loanwords is determined by the default accent rule for "accented" nouns in Japanese (= antepenultimate rule).

Thus, Kubozono (2006, 2008) points out that at least for Japanese it is not necessary to assume an accent rule specifically for loanwords, but rather that loanwords basically follow the most productive accent pattern taken by accented native words.

Can Yanbian loanword accentuation be explained in a similar fashion? Such an analysis encounters several problems. First, if Kubozono's (2006) finding that loanword accentuation reflects the statistical tendencies in native words is correct, Yanbian loanwords should appear with the final accent class more frequently; but this is not consistent with our data.

Second, a Latin-type accent rule (weight-sensitive accent system) is not as strong a factor in the accent distribution of Yanbian native and Sino-Korean words as in loanwords, as discussed above. Rather, the default accent is fixed separately for each lexical class, based on such factors as the earlier stage of the language, Middle Chinese tone, etc., while syllable weight affects the degree of bias towards this default accent type.

Third, the distribution of each syllable weight structure in the three word classes does not align with the general tendencies constituting the default accent class (native: final, Sino-Korean and loanword: penultimate in disyllables). Table 33 shows the ratios of each syllable weight structure in the three word classes. Native words and loanwords have somewhat similar distributions in that Light-Light and Light-Heavy are relatively larger classes, while Heavy-Heavy is smallest. However, Sino-Korean words have the opposite profile in that Heavy-Heavy is the largest class. In spite of this, it is Sino-Korean words and loanwords, not native words and loanwords, that show similar accent distribution.

Table 33
Distribution of weight structure of final two syllables in three lexical classes (disyllabic words)

Weight \ Word class	Native		Sino-Korean		Loan	
Heavy-Heavy	120	(12%)	3,013	(38%)	59	(13%)
Heavy-Light	225	(22%)	1,777	(22%)	81	(18%)
Light-Heavy	329	(32%)	2,061	(26%)	134	(30%)
Light-Light	354	(34%)	1,101	(14%)	176	(39%)
Totals	1,028		7,952		450	

Similarly, it is difficult to relate the distribution in trisyllabic native words and loanwords to the default final accent in native words and the accentuation patterns following syllable weight in loanwords.

Table 34
Distribution of weight structure of final two syllables in two lexical classes (trisyllabic words)

Weight \ Word class	Native		Loan	
Heavy-Heavy	1	(0%)	11	(3%)
Heavy-Light	71	(21%)	30	(8%)
Light-Heavy	20	(6%)	105	(28%)
Light-Light	246	(73%)	230	(61%)
Totals	338		376	

Thus, the distribution of syllable weight is not a factor which explains the different accent patterns among the three lexical classes. The apparent differences in accent distributions of native and Sino-Korean are rather attributed to differences in the default accent class which resulted from the historical background: the preference of the HL class in Sino-Korean disyllabic words is originally due to the regular correspondence between the Sino-Korean accent in Middle Korean and the Middle Chinese tone, which was enhanced in the evolution from Middle Korean to Yanbian (Ito 2008b).

Fourth, the preference for the accented (as opposed to the unaccented) class in Yanbian loanwords may be based on the pitch falls/accentedness in every English word, as in Japanese, but it could also be explained by the *UNACCENTED constraint which appears in the native grammar.

Fifth, it may be problematic to attribute the preference for penultimate accent in Yanbian loanwords to the adjustments by the native/Sino-Korean grammar, since in Yanbian the default accent classes differ between native and Sino-Korean words. It is not reasonable to assume that Sino-Korean words, but not native words, induced the default penultimate accent to loanword accentuation, given the different distributions of syllable weight in Sino-Korean vs. loanwords, as mentioned above. Thus it is difficult to explain the Yanbian loanword accentuation pattern only based on the native/Sino-Korean grammar.

Our hypothesis is that Yanbian loanword accentuation results from the grammar of the source

language and lexical statistics, along with some adjustments by Yanbian native grammar. If this is correct, different accentuation patterns should appear depending on the source language. This prediction is observed in our Yanbian data, as we discuss in the next section.

5. Difference among English, Japanese and Mandarin loanwords

So far we have discussed various kinds of loanwords as a single group, but in actuality, different patterns are observed depending on the source language. In this section, we examine Yanbian loanword accentuation from the major two source languages (English and Japanese) and compare them with Mandarin loanwords reported by Ito and Kenstowicz (2009).

We will first focus on the details of English loanwords. The examined data is restricted to direct English loanwords. Hybrid loanwords between English and other languages are excluded. Also, loanwords which seem to be based on the spelling of English source words are not included here, e.g. *a.tam* 'Adam' (English /æ/ not adapted with Yanbian /ɛ/, English /ə/ not adapted with Yanbian /ə/), *k^ho.k^ha.in* 'cocaine' (English /ei/ not adapted with Yanbian /ei/). For some words, it is difficult to judge the source; we included such words as far as there is a possibility that they are direct loanwords from English.[14] Since Japanese is a CV language and thus Japanese loanwords in Yanbian do not have an epenthetic vowel, the discussion here concentrates on the adaptation patterns in Heavy-Heavy, Heavy-Light, Light-Heavy and Light-Light structures in the final two-syllable window.

5.1 General tendencies

First, let us compare the accent distributions in the Yanbian loanwords from English and Japanese, shown in Table 35.

[14] English /ɔ/ is adapted with Yanbian /o/, except for two words *k^hə.p^hi* 'coffee' and *wə.siŋ.t^hon* 'Washington'. The exceptional adaptation pattern of *k^hə.p^hi* 'coffee' is probably the result of avoiding homophony with *k^ho.p^hi* 'copy', and we assume this word (*k^hə.p^hi*) is a direct English loanword. The adaptation with *wə* in *wə.siŋ.t^hon* 'Washington' is assumed to be due to a phonotactic restriction (*wo) in Korean. (Still this word is not treated as a direct English loanword since English /ə/ is adapted with Yanbian /o/ in *wə.siŋ.$\underline{t^h on}$*, based on its spelling.) On the other hand, British English /ɒ/ that corresponds with general American English /ɑ/ and Eastern American English /ɔ/, is mostly adapted with Yanbian /o/, except for a few words that adapt it with /a/. The adaptation with /a/ is probably a relatively recent pattern based on American English: e.g. *si.t^hop* 'stop', *sjo.k^hi* 'shock', *t*al.la* 'dollar', *hal.li.u.ti* 'Hollywood'. Both adaptation patterns are included here as direct English loanwords.

Table 35

Accent distributions in the Yanbian loanwords from English and Japanese

English

Weight	A	P	F	U	Totals	A%	P%	F%
Heavy-Heavy		107	18		125	0%	86%	14%
Heavy-Light		134	9		143	0%	94%	6%
Light-Heavy	6	187	91	2	286	2%	65%	32%
Light-Light	13	263	34		310	4%	85%	11%
Totals	19	691	152	2	864	2%	80%	18%

Japanese

Weight	A	P	F	U	Totals	A%	P%	F%
Heavy-Heavy		17	3		20	0%	85%	15%
Heavy-Light		169	1	1	171	0%	99%	1%
Light-Heavy	2	13	103	2	120	2%	11%	86%
Light-Light	41	381	299		721	6%	53%	41%
Totals	43	580	406	3	1,032	4%	56%	39%

As a whole, the percentage of penultimate accent is higher in English loanwords (80% vs. 56%), whereas the percentage of final accent is higher in Japanese loanwords (18% vs. 39%). The difference between the two types of loanwords is most evident in Light-Heavy structures: in English loanwords only 32% of words of this structure appear with final accent, while in Japanese loanwords this figure reaches 86%. Given that the Heavy-Light structure appears with penultimate accent in a high percentage in both loanword types (94% for English loans vs. 99% for Japanese loans), we may assume that Japanese loanwords are more sensitive to syllable weight. We ran a mixed effects logistic regression model with accent location (penultimate/final) as the dependent variable. A random intercept was set for subjects. The result shows that all three factors (syllable weight in penultimate and final syllables, source language—English vs. Japanese) are strongly significant, affecting the differences in accent distribution: a penult heavy is positively associated with penultimate accent, while a final heavy and Japanese loanwords (as opposed to English loanwords) are negatively associated with penultimate accent.

Table 36

Results of a mixed effects logistic regression model: source language difference

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	2.0794	0.1482	14.03	< 2e-16	***
Weight-Penult-Heavy	2.1412	0.2019	10.61	< 2e-16	***
Weight-Final-Heavy	-1.6194	0.1538	-10.53	< 2e-16	***
Source language-Japanese	-1.8054	0.1502	-12.02	< 2e-16	***

Significance codes: '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

This suggests that for both English and Japanese loanwords, a syllable weight effect is observed in accent assignment. But the way the effect works and its degree differ between the two source languages. How do these differences arise? In the following sections, we examine each loanword type separately and try to identify the causes of the adaptation patterns.

5.2 English loanwords

In this section, we examine English stress patterns in order to determine whether there are any correlations between English stress patterns and the accentual assignment in Yanbian loanwords, while taking into account the syllable weight effect. The data here is composed of loanwords that are borrowed from English words with the same syllable number, e.g. disyllabic → disyllabic: *én.tsin* 'engine', trisyllabic → trisyllabic: *ti.rék.t^{h}ə* 'director'. Since Yanbian loanword accent is basically located in a two-syllable window at the right edge of the word (penultimate or final), we compare the English stress patterns in penultimate and final syllables with the Yanbian accentual assignment.

One view of English stress is that if the syllable has a full (tense) vowel then it is stressed (primary or secondary). Since lax vowels such as ɪ, ɛ, ʊ are banned at the end of the word in English, dictionaries such as Merriam-Webster transcribe stressed and unstressed alternatives for words ending in a tense vowel such as 'window', 'kiwi'. In this paper, we assume that the vowel that is ambiguous due to this kind of phonotactic constraint is unstressed. Also for simplicity, we aggregate the primary and secondary stresses as 'stressed (S)', contrasting with 'unstressed (U)'. Thus, the English stress patterns are classified into four categories: stressed-unstressed (SU) [< σ́σ, σσ́, σ̀σσ́, etc.], stressed-stressed (SS) [< σ́σ́, σ̀σ́, σ́σ̀], unstressed-stressed (US) [< σσ́, σσ̀σ́, σ̀σσ́], and unstressed-unstressed (UU) [< σσσ, σσσσ, σ̀σσσ]. The English stress patterns in our data were confirmed by a native English speaker.

Table 37 shows the relations between English stress patterns and syllable weight structures on the one hand and the accentual adaptation patterns on the other. Syllable weight structures are based on the Yanbian loanwords (output forms), not on the English input.

Table 37
Adaptation patterns of English loanwords. S = stressed, U = unstressed.

	Heavy-Heavy		Heavy-Light		Light-Heavy		Light-Light		
English stress	P	F	P	F	P	F	P	F	Totals
SS	15	9	9		71	20			124
SU	92	9	114	9	104	27	231	12	598
US					17	37	16	4	74
UU			11		10	14	16	18	69
Totals	107	18	134	9	202	98	263	34	865

(16) Examples of English loanwords
a. SS: in.p^{h}ús 'input' (σ́σ̀), ín.t^{h}ən ~ in.t^{h}ə́n 'intern' (σ́σ̀), mján.ma 'Myanmar' (σ̀σ́), hó.t^{h}el

‘hotel’ (ȍ̀), ó.tson ~ o.tsón ‘ozone’ (ó̀ȍ), sú.tan ~ su.tán ‘Sudan’ (ȍ̀)

b. SU: kʰép.tʰin ‘captain’ (óσ), pʰéŋ.kwin ‘penguin’ (óσ), ro.mɛn.tʰík ‘romantic’ (σóσ~ȍ̀óσ), píŋ.ko ‘bingo’ (óσ), tʰɛ́k.si ~ tʰɛk.sí ‘taxi’ (óσ), ti.rék.tʰə ‘director’ (σóσ), wé.tiŋ ‘wedding’ (óσ), tsʰí.kʰin~ tsʰi.kʰín ‘chicken’ (óσ), jú.mə ‘humor’ (óσ), té.mo ‘demo’ (óσ), kʰəm.pʰjú.tʰə ‘computer’ (σóσ), si.kʰá.ko ~ si.kʰa.kó ‘Chicago’ (σóσ)

c. US: tsʰim.pʰén.tsi ‘chimpanzee’ (ȍ̀σó or ȍ̀ó̀ȍ), í.ran ~ i.rán ‘Iran’ (σó), mə.sín ‘machine’ (σó), hə.ní.mun ~ hə.ni.mún ‘honeymoon’ (óσȍ̀), pʰil.lí.pʰin ~ pʰil.li.pʰín ‘Philippines’ (óσȍ̀), u.ni.pʰóm ‘uniform’ (óσȍ̀), kí.tʰa ‘guitar’ (σó), in.tʰə́.pju ‘interview’ (óσȍ̀), sé.mi.na ~ se.mi.ná ‘seminar’ (óσȍ̀)

d. UU: me.sín.tsə ‘messenger’ (óσσ), he.ró.in ~ he.ro.ín ‘heroin’ (óσσ), tsʰém.pʰi.ən ~ tsʰɛm.pʰí.ən ~ tsʰɛm.pʰi.án ‘champion’ (óσσ), e.nə́.tsi ~ e.nə.tsí ‘energy’ (óσσ), mɛ.ní.tsə ‘manager’ (óσσ)

As can be seen in Table 37, the English stress patterns as well as syllable weight seem to be relevant to the accentual assignment in Yanbian. In order to confirm this, a mixed effects logistic regression model with accent location (penultimate/final) as the dependent variable was run, with English stress pattern (stressed or unstressed in penultimate and final syllables) and syllable weight in Yanbian loanwords (heavy or light in penultimate and final syllables) as independent variables. A random intercept was set for items and subjects. The result is that not only syllable weight in the final syllable but also the English stress pattern in the penultimate syllable are significant: a penultimate stress is positively associated with penultimate accent, whereas a final heavy is negatively associated with penultimate accent.

Table 38

Results of a mixed effects logistic regression model: English loanwords

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	1.5019	0.7389	2.033	0.04208	*
Weight-Penult-Heavy	0.4311	0.7211	0.598	0.54994	
Weight-Final-Heavy	-2.3288	0.7155	-3.255	0.00114	**
Stress-Penult-Stressed	3.7032	0.7533	4.916	8.83e-07	***
Stress-Final-Stressed	-0.5831	0.5984	-0.974	0.32983	

Significance codes: ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1

This tendency for the English stress pattern only in the penultimate syllable to be taken into account in the accent assignment is understandable, given that English stress (especially the primary stress) rarely appears in the final syllable in nouns due to a strong NON-FINALITY constraint. That is, the final syllable, being extrametrical, is frequently unstressed. Thus, it is a locus that Yanbian speakers do not have to pay much attention to whether it is stressed or not, since there are very few cases of stress contrast, if any. This marginal status of the stress contrast in the final syllable results in a “stress-deafness” in this position of English loanwords.

Thus, we assume that Yanbian loanword accentuation actually results from two statistical

patterns in the English source words: a) the correlation between syllable weight and stressed syllables, and b) the statistical bias of stressed syllables to the penultimate/antepenultimate positions in English lexical items. In other words, the accentuation patterns in Yanbian loanwords were adapted from English grammar, not from Universal Grammar, and the default penultimate accent reflects the lexical statistics of English words, not native words. In the initial stage of adaptation, Yanbian speakers tried to adapt the English stress patterns faithfully. While doing so, they discovered a generalization (= Latin stress rule) in English grammar supported by lexical frequency and started to generalize the rule to English loanwords as a whole.

Table 39 shows the frequency of the stress patterns in the English words which are adapted as loanwords in Yanbian: Table 39a indicates the number of English words that have the same syllable number as Yanbian loanwords, whereas Table 39b indicates the number of English words as a whole that are borrowed into Yanbian.[15] The distributions are more or less the same between Table 39a and Table 39b. On the other hand, Table 40 shows the accent distribution of English loanwords in Yanbian: Table 40a is based on the loanwords that have the same syllable number as English words, and Table 40b is based on all the English loanwords in our corpus. Note that Table 40a, b are the aggregated total of Yanbian loanword accent from all speakers and so these numbers do not coincide with Table 39a, b, where the number of English words is counted.

Table 39

Stress patterns in English lexical items (final two syllables) which are adapted as loanwords in Yanbian. S = stressed, U = unstressed.

	(a)					(b)				
Syllable \ Stress	S	SU	SS	US	UU	S	SU	SS	US	UU
Monosyllabic	56					279				
Disyllabic		102	24	6			217	54	17	
Tri/Quadrisyllabic		10	1	7	14		28	1	18	23

Table 40

Accentual distribution of English loanwords in Yanbian (aggregated for all speakers)

	(a)		(b)	
Syllable \ Accent	Penult	Final	Penult	Final
Monosyllabic		332		332
Disyllabic	595	88	1,627	224
Tri/Quadrisyllabic	96	64	789	238

There are many similarities between the stress patterns in English lexical items and the accent distribution of English loanwords in Yanbian. First, the fact that all monosyllabic words are stressed (Table 39a: 56, b: 279) supports the exclusive final accent (not unaccented) in monosyllabic loanwords in Yanbian. Second, the fact that in disyllabic words the most frequent

[15] The loanwords in Table 39b but not in Table 39a arise primarily from the addition of epenthetic vowels that adjust English consonants and consonant clusters to Yanbian syllable structure.

pattern is SU (Table 39a: 102, b: 217), which is much larger than US (Table 39a: 6, b: 17), coincides with the Yanbian loanword accentuation in that penultimate accent, not final accent, is a strong default class in disyllable words (INITIAL).

On the other hand, in trisyllabic/quadrisyllabic English words, UU (= antepenultimate primary stress patterns (σσσ), Table 39a: 14, b: 23) and SU (Table 39a: 10, b: 28) are slightly more frequent than US (Table 39a: 7, b: 18). It is assumed that due to the restricted distribution of the antepenultimate and unaccented classes in Yanbian native grammar (*LAPSE-R and *UNACCENTED), the frequency distribution in trisyllabic/quadrisyllabic English words could not be reflected in Yanbian loanwords as is. Instead, Yanbian speakers made use of more general rules working in English (Latin stress rule) in their accent assignments in this case. Table 41 shows the data from Table 37, reclassifying the accent distribution by taking into account syllable weight. (The totals do not agree between Tables 41 and 37, since in Table 37 the variant stress patterns that a given word can have are counted separately.) Unlike disyllabic words, where most words appear with penultimate accent, the syllable weight effects emerge more clearly in trisyllabic words: Heavy-Light and Light-Heavy vs. Heavy-Heavy and Light-Light. A mixed effects logistic regression model shows that syllable weight in the final syllable and syllable number are highly significant factors: a final heavy and trisyllabic words are negatively associated with penultimate accent.

Table 41

Accentual distribution of English loanwords in Yanbian (reclassified; there are no examples of quadrisyllabic words)

weight, accent / syllable	Heavy-Heavy		Heavy-Light		Light-Heavy		Light-Light	
	P	F	P	F	P	F	P	F
Disyllabic	106	14	117	9	174	57	198	8
Trisyllabic	1	4	17		13	34	65	26

Table 42

Results of a mixed effects logistic regression model: syllable number effect in English loanwords

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	5.6198	0.7931	7.086	1.38e-12	***
Weight-Penult-Heavy	0.7094	0.7063	1.004	0.315	
Weight-Final-Heavy	-3.3060	0.7655	-4.319	1.57e-05	***
Syllable number-Trisyllabic	-3.7485	0.7875	-4.760	1.93e-06	***

Significance codes: '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

Still, penultimate accent is expected to be preferred to final accent in Yanbian loanwords, given the frequency difference in English. This is observed in the fact that Heavy-Light appears with penultimate accent almost without exception, whereas Light-Heavy is not biased to the final accent class so strongly.

Thus, we conclude that the English loanword accentuation in Yanbian is not from Universal Grammar but basically from English native grammar and lexical statistics, along with some adjustments by Yanbian native grammar.

5.3 Japanese loanwords

Standard Japanese has a pitch accent system and so we might expect that Japanese loanwords in Yanbian reflect the pitch accent patterns of each Japanese source word. On the other hand, as we have seen so far, Yanbian loanwords as a whole tend to correlate with syllable weight structures. In Japanese loanwords, which factor is decisive in accentual assignment: either one of them or both? In order to see this point, we concentrate on Yanbian loanwords from Japanese here (hybrid loanwords between Western languages and Japanese are excluded), and compare their accentuation with that of the loanwords from English. (As to the segmental adaptation patterns of Japanese loanwords into Korean, see Ito et al. 2006.)

Table 43 shows the correspondence between standard Tokyo Japanese accent classes (based on Hirayama 1960 and NHK hōsōbunka kenkyūjo 1998) and Yanbian accent classes in Japanese loanwords. As seen in Table 43, most loanwords appear with either penultimate or final accent. The Japanese penultimate accent class tends to appear with Yanbian penultimate accent more frequently (78%), whereas Japanese final accent tends to appear with Yanbian final accent (81%). On the other hand, Japanese pre-antepenultimate and antepenultimate accent classes rarely correspond with Yanbian pre-antepenultimate and antepenultimate accent classes. Similarly, the Japanese unaccented class, which is expected to be adapted as the unaccented class or the final accent class in Yanbian since both accent classes appear with final accent in isolation forms in Yanbian, appears with Yanbian penultimate accent with a slightly higher percentage.

Table 43

Comparison between Tokyo Japanese accent and the accent of Yanbian loanwords.[16] The leftmost column indicates Japanese accent, and the topmost row indicates the Yanbian accent. Japanese accent is syllable-based.[17] pA = pre-antepenultimate accent, A = antepenultimate accent, P = penultimate accent, F = final accent, U = unaccented class.

Japanese accent \ Yanbian accent	pA	A	P	F	U	Totals	Regular correspondence
pA			8			8	0%
A		35	78	101		214	16%
P		2	382	104	1	489	78%
F			17	72		89	81%
U		8	196	179	2	385	1%
Totals		45	681	456	3	1,185	41%

We may thus assume that when Japanese source words appear with penultimate or final accent, which are the two major accent patterns in Yanbian, the Japanese accent is relatively faithfully adapted into Yanbian; elsewhere the original Tokyo pitch accent information is disregarded, and the accentual assignment depends on some other factors. Based on this, we aggregate Japanese pre-antepenultimate accent, antepenultimate accent and unaccented classes into one category and examine the correlations with the syllable weight structures in the final two-syllable window and the Yanbian two major accent classes (penultimate and final). Table 44 shows these correlations.[18]

[16] Monosyllabic loanwords, which are excluded from this table, appear with final accent: *p*áŋ* 'bread' < J. *paN*, *én* 'Yen' < J. *eN*, *hwá* 'fa, solmization' < J. *ɸa*.

[17] Note that the position of the Japanese pitch accent in a syllable-based analysis can differ from the position in a mora-based analysis. For example, the accent on the antepenultimate mora is equal to the accent in the penultimate syllable when the penultimate mora is a geminate, e.g. *bákku* 'back': the accent is on the antepenultimate mora and in the penultimate syllable.

[18] Following the suggestion from one of the reviewers, we simplified the data by ignoring the length distinction in Japanese that is reflected in accentual assignment: the long syllables in Japanese, CVV and CVO (the final consonant constitutes the first part of the gemination), which are adapted as an apparent light syllable in Yanbian, tend to attract the accent, especially in the penultimate syllable, e.g. *ó.pa* 'overcoat' < J. *óobaa*, *sé.ta* 'sweater' < J. *séetaa*, *mó.ta* 'motor' < J. *móotaa*, *tó.k*jo* 'Tokyo' < J. *tookjoo*, *hó.k*u* 'hook, snap' < J. *hókku*, *kó.p*u* 'glass' < J. *koppu*, *p*á.k*u* 'back' < J. *bákku*, *p*á.ma* 'permanent wave' < J. *páama*, *réru* 'rail' < J. *réeru/reeru*, *ts*o.k*í* 'vest' < J. *tʃokki*, *tsó.ro* 'watering can' < J. *dʒooro*, *si̱.k*á.t*o* 'skirt' < J. *sukáato*. The sensitivity to length in Japanese is important because 1) in many cases long syllables are not distinguished from short syllables at the segmental level in Yanbian loanwords; and 2) loanwords from English do not have a comparable distinction. Similar adaptation strategies are observed in Taiwanese loanwords from Japanese (Hsieh 2006), where Japanese CVV/CVN syllables are in general differentiated from CV/CVO syllables by means of tone.

Table 44
Correlation between Japanese accent and syllable weight and accent of Yanbian loanwords

Yanbian weight, accent / Japanese accent	Heavy-Heavy		Heavy-Light		Light-Heavy		Light-Light		
	P	F	P	F	P	F	P	F	Totals
pA, A, U	8	1	79	1	2	65	193	213	562
P	11	3	126	1	12	38	233	62	486
F	2	1	8			7	7	64	89

(17) Examples of Japanese loanwords. Japanese words are shown with the pitch accent (´) of the Tokyo dialect. Unaccented words in Tokyo Japanese have no diacritic.

a. [Heavy-Heavy] ts*ám.p*oŋ 'Japanese-style noodle' < tʃáNpoN, nín.tsin 'carrot' < niNdʒiN, tsjáŋ.k*en 'the game of 'scissors-paper-rock'' < dʒaNkeN
b. [Heavy-Light] móm.p*e 'baggy work pants' < móNpe, tén.ts*i 'battery' < déNtʃi/deNtʃi, tán.su 'chest of drawers' < taNsu, p*éŋ.k*i 'paint' < peNki
c. [Light-Heavy] ú.toŋ ~ u.tóŋ 'Japanese-style noodle' < udoN, o.téŋ 'Japanese style stew' < odéN, ri.póŋ 'ribbon' < ríboN, han.ts*i.póŋ 'short pants' < haNdzúboN, wa.ri.k*aŋ 'sharing the expense' < warikaN
d. [Light-Light] ta.má 'ball' < tamá, e.rí 'collar' < erí, sá.ra 'plate' < sara, tá.pi ~ ta.pí 'socks' < tábi, sa.si.mí 'sashimi' < saʃimí, wá.sa.pi ~ wa.sá.pi ~ wa.sa.pí 'Japanese horseradish' < wásabi, o.sá.k*a ~ o.sa.k*á 'Osaka' < oosaka, mi.ts*i.pi.sí ~ mi.ts*i.pí.si ~ mi.ts*i.pi.sí 'Mitsubishi' < mitsúbiʃi, ta.ma.né.ki 'onion' < tamanégi

A mixed effects logistic regression model (random intercepts set for item and subject) shows that syllable weight in penultimate/final syllables and Japanese accentedness in the penultimate syllable play a significant role in the accent assignment: a penult heavy and a Japanese penultimate accent are positively associated with penultimate accent, whereas a final heavy is negatively associated with penultimate accent. However, Japanese accentedness in the final syllable is not significant, which is probably because the final accent class in Japanese is relatively small compared to the other accent classes. This is comparable to English final stress, with respect to the marginal contrastiveness, at least in the words that have been loaned into Yanbian. On the other hand, in Japanese loanwords, the syllable weight in both penultimate and final syllables is strongly significant. This is because 1) Japanese words have the accent correlation with syllable weight regardless of word class (Kubozono 2006, 2008) and 2) the accent classes in nouns are not biased to penultimate accent as strongly as in English. In English loanwords, the bias to the penultimate stress especially in disyllabic words cancels out the syllable weight effect in this position, resulting in the significant syllable weight effect only in the final syllable.

Table 45
Results of a mixed effects logistic regression model: Japanese loanwords

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	0.4546	0.4752	0.957	0.3388	
Weight-Penult-Heavy	8.1219	2.0498	3.962	7.42e-05	***
Weight-Final-Heavy	-6.0511	1.3228	-4.574	4.78e-06	***
Japanese accent-Penult	1.4264	0.5008	2.848	0.0044	**
Japanese accent-Final	-0.5608	0.8987	-0.624	0.5326	

Significance codes: '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

Thus, we conclude that Yanbian speakers are sensitive to Japanese pitch accent and syllable weight in their accent assignment, and display a different tendency from the accent adaptation for English loanwords.

5.4 Mandarin loanwords

A particularly striking influence of the source language can be observed in Yanbian loanwords from Mandarin (Ito and Kenstowicz 2009). Note that the Chinese used in Yanbian is basically the same as standard Chinese (Putonghua). The Mandarin loanwords discussed here are different from Sino-Korean words which are much older borrowings. Based on a corpus of about 250 Yanbian loanwords from Mandarin, collected from one of the consultants for this study (Mandarin tonal patterns in our data were also confirmed with the same consultant), Ito and Kenstowicz (2009) show that only penultimate and final accent classes appear, and the choice is predictable on the basis of the tones that occupy the final two syllables of the Mandarin source word, as shown in Table 46.[19] The Mandarin data are transcribed in Pinyin (b, d, g indicate voiceless unaspirated stops) with the numbers that customarily indicate the four tones (Tone 1 = [55] High, Tone 2 = [35] Rise, Tone 3 = [21(4)] Low, Tone 4 = [51] Fall). Tone 0 indicates a toneless syllable. For example, if the penultimate syllable is Tone 1 and the final syllable is Tone 1 as well, then the corresponding Yanbian loanword appears with final accent (e.g. jia^{1}banr^{1} (加班儿) 'overtime work' → tsja.pál); if the penultimate syllable is Tone 1 and the final syllable is Tone 2, then the corresponding Yanbian loanword appears with penultimate accent (e.g. feng^{1}tian^{2} (丰田) 'TOYOTA' →fə̀ŋ.t^{h}ɛn); if the penultimate syllable is Tone 2 and the final syllable is Tone 3, then the corresponding Yanbian loanword appears with penultimate accent (e.g. pi^{2}jiu^{3} (啤酒) 'beer' → p^{h}í.tsju).

[19] Mandarin data have been compiled from previous research (Chi 2008), checked with our bilingual Mandarin/Yanbian consultant, who also added new loanwords.

Table 46
Mandarin tones of penultimate and final syllables and their Yanbian correspondences. Tone 3 + Tone 3 is altered to Tone 2 + Tone 3 by tone sandhi. F = final accent, P = penultimate accent.

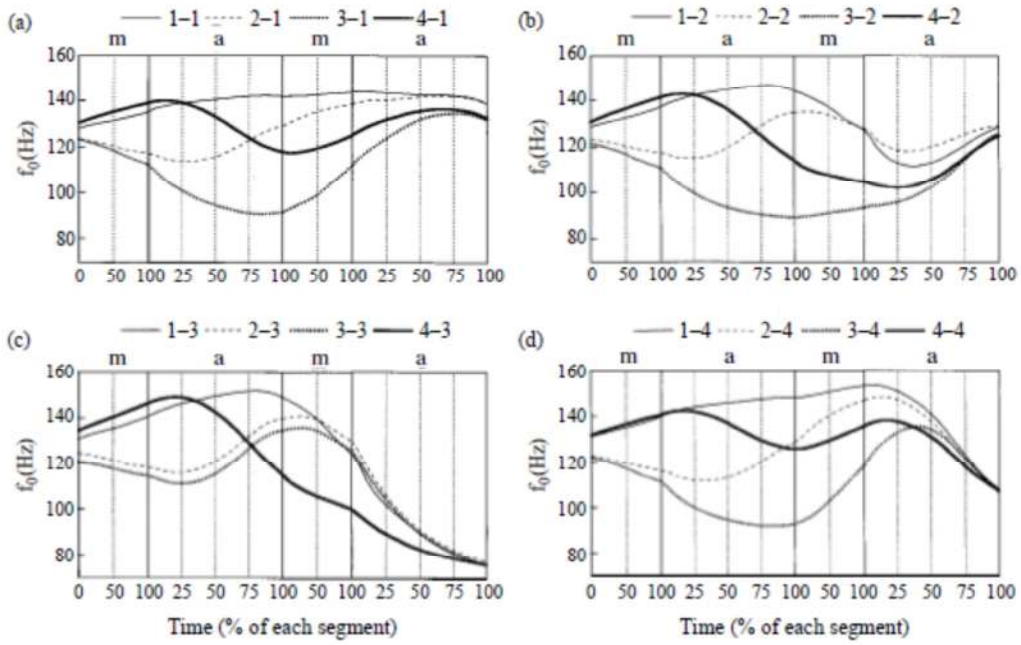
Penultimate \ Final	Tone 1	Tone 2	Tone 3	Tone 4	Tone 0
Tone 1	F	P	P	F	P
Tone 2	F	P	P	F	P
Tone 3	F	F	—	F	F
Tone 4	F	P	P	F	P

(18) Examples of Mandarin loanwords

Mandarin		Yanbian	
ji^{1}chang^{3}	机场	tsí.ts^{h}aŋ	‘airport’
bing^{1}gunr^{4}	冰棍儿	p*iŋ.kól	‘popsicle’
mei^{2}chu^{1}xi^{0}	没出息	mei.ts^{h}ú.si	‘to be not promising’
bai^{2}ganr^{1}	白干儿	p*ɛ.kál	‘spirits; liquor’
cun^{2}zhe^{2}	存折	ts^{h}ún.tsə	‘bankbook’
guo^{2}mao^{4}	国贸	kwə.mó	‘International trade building’
man^{2}tou^{0}	馒头	mán.t^{h}u	‘Chinese-style steamed bread’
er^{3}ji^{1}	耳机	əl.tsí	‘earphone’
lao^{3}tour^{2}	老头儿	no.t^{h}ól	‘old male person’
qi^{3}ma^{3}	起码	ts^{h}í.ma	‘at least’
duan^{3}ku^{4}	短裤	twan.kú	‘short pants’
ling^{3}zi^{0}	领子	liŋ.tsɨ́	‘collar’
da^{4}yi^{1}	大衣	t*a.í	‘overcoat’
bing^{4}du^{2}	病毒	p*íŋ.tu	‘virus’
di^{4}nuan^{3}	地暖	t*í.nwan	‘floor heating’
dian^{4}shi^{4}	电视	t*ɛn.s*ɨ́	‘television’
ci^{4}ji^{0}	刺激	ts^{h}ɨ́.tsi	‘to stimulate’

Ito and Kenstowicz (2009) find that the tonal adaptations in the loanwords from Mandarin are based on the F0 relation between the end of the first syllable and the start of the second, and this relation depends on the actual trans-syllabic F0 contours that are the product of articulatory “smoothing” rather than the abstract phonological categories. Figure 2 below taken from Xu’s (1997) investigation of Mandarin tonal coarticulation shows the normalized contours for all 16 tonal combinations over a dummy disyllabic [mama] string, based on an averaging over 48 utterances (eight subjects and six repetitions) per combination.

Figure 2
Normalized Mandarin carry-over tonal coarticulation (Xu 1997:69)



Xu (1997) finds that the tonal targets for the second syllable—high vs. low in panels (a) vs. (c) and rise vs. fall for panels (b) vs. (d)—are only reached in the second mora. The first mora including the onset is a zone of carryover articulation from the target of the preceding syllable. Focusing on this transition zone, we see that in the first panel (a), where the second syllable is Mandarin Tone 1, all of the transitions have a rising configuration that corresponds to the Yanbian LH (final) adaptation. In the (d) panel [X+4], all the transitions are rising and hence are best matched by Yanbian LH (final). In panel (b) [X+2], all the contours are falling except for [3+2], which is rising, thus all [X+2] combinations are adapted as Yanbian HL (penult) except for [3+2], which is LH (final). Finally, panel (c) [X+3] shows a falling contour for the [1+3] and [4+3] combinations that matches the Yanbian HL (penult) adaptation. The Mandarin [2+3] (and [3+3]) case is more ambiguous in that the onset zone shows a slightly rising contour followed by a sharp fall, which may be reflected in relatively lower rates of regular adaptation: HL (15) vs. LH (5). Thus, the coarticulatory configurations documented by Xu (1997) for his Beijing Mandarin subjects offer striking support for the hypothesis that the Mandarin → Yanbian tonal adaptations are based on the trans-syllabic F0 contour. The recent theoretical literature on borrowing has discovered a number of other cases where phonologically redundant phonetic information plays a crucial role in shaping the loanwords (e.g. Hsieh et al. 2009).

Here syllable weight does not play any role as shown in Table 47, unlike in the loanwords from English or Japanese.

Table 47
Correlation between syllable weight and accent class in Mandarin loanwords

Weight	P	F	U	Totals	P%	F%
Heavy-Heavy	22	26	4	52	42%	50%
Heavy-Light	29	32		61	48%	52%
Light-Heavy	11	21	1	33	33%	64%
Light-Light	16	31		47	34%	66%
Totals	78	110	5	193	40%	57%

A logistic regression model using glm shows that syllable weight structures are not strongly significant in deciding the accentual adaptation patterns in Mandarin loanwords. Thus Mandarin loanwords again indicate that the loanword accentuation can differ depending on the source language.

Table 48
Results of a logistic regression model: Mandarin loanwords

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-0.64085	0.26653	-2.404	0.0162	*
Weight-Penult-Heavy	0.52810	0.30537	1.729	0.0837	.
Weight-Final-Heavy	-0.03606	0.30208	-0.119	0.9050	

Significance codes: '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

5.5 Summary

In this section, we compared three different loanword categories in Yanbian from different source languages with different prosodic types (English: stress, Japanese: pitch accent, Mandarin: tone), and showed that each has its own accentual adaptation system that may reflect original phonological oppositions or phonetic realizations. As shown in Table 49, a syllable weight effect is active in English and Japanese loanwords (final for the former and penultimate/final for the latter) but not in Mandarin loanwords, whereas suprasegmental type (stress/pitch accent/tone) affects the adaptation patterns in Mandarin more than in English and Japanese. These differences are probably due to the general tendencies or the grammar of the original source languages: in English and Japanese, a syllable weight effect is observed to a certain extent, while it is not clearly observed in Mandarin; Mandarin has contrastive tone more strongly than in English and Japanese, where a lexical contrast by stress/accent tends to be marginal in the final syllable; English has a strong bias to the penultimate stress, especially in disyllabic words, compared to Japanese. Thus, we conclude that loanword adaptation is not from Universal Grammar but instead reflects the phonology/phonetics of the original source languages along with some adjustments, based on the native grammar. Also, given that all the relevant factors/constraints for the Yanbian loanword accentuation come from the source language or from the native grammar,

there is no need to assume universal phonetic considerations (at least for our Yanbian data).

Table 49

Differences among three loanword categories. "+/(+)" indicates that each factor (syllable weight, suprasegmental type) plays a (partial) role in the accent adaptation in the Yanbian loanwords, whereas "–" indicates that no strong effect for that factor is observed.

Source	Syllable weight	Suprasegmental type
English	(+)	(+)
Japanese	+	(+)
Mandarin	–	+

This conclusion differs from other recent literature on loanword phonology, where it is found that loanword accentuation systems are the same regardless of the source language. For example, Lee (2008) reports that English and Japanese loanwords in South Kyengsang Korean have the same accentuation rules. Hsieh and Kenstowicz (2008) also find the same tendency for English and Mandarin loanwords in Lhasa Tibetan. However, the failure to find any differences may be because differences between the two loanword categories appear as statistical tendencies with different frequencies, so at first glance it is difficult to recognize an underlying difference. If we take into account a faithfulness constraint to the source language on the one hand and the data of variation on the other, then loanword subcategories could have different weight hierarchies, even though their accentuation rules look similar.

6. Loanword adaptation model

Finally, we propose a loanword adaptation model based on our analysis of Yanbian loanword accentuation. As mentioned above, Kubozono (2006, 2008) points out that native, Sino-Japanese, and Western loanwords in Japanese exhibit very similar accent patterns and preferences, and the apparent differences among them (e.g. more biased distribution of accented class in loanwords) can be attributed to a phonetic factor (faithfulness to the English abrupt pitch fall). Similarly, Itô and Mester (1995, 1999) analyze the phonological lexicon in contemporary Japanese and propose a core-periphery theory where the phonological properties of each separate sublexicon fall into a hierarchy of implicational relations. In this theory, the lexical constraint systems are depicted as the nesting of constraint domains, which entails the existence of a core area (unmarked or native vocabulary), governed by the maximum set of lexical constraints. For the whole lexicon, there is a single markedness constraint ranking and lexical stratification is a consequence of faithfulness constraint reranking. For example, the four different lexical strata (a~d) can be represented as follows:

(19) M1 » Fa » M2 » Fb » M3 » Fc » M4 » Fd (M = markedness, F = faithfulness)

Thus, Itô and Mester (1995, 1999) assume that an invariant ranking of markedness constraints exists for all lexical strata in Japanese and that the differences among the strata result from the ranking of faithfulness constraints.

Can we explain the lexical stratification observed in Yanbian accentuation (native, Sino-Korean, loanwords) in the same way? Following this theory, we may analyze the lexical stratification in Yanbian by postulating the three faithfulness constraints for native, Sino-Korean, and loanwords (= Faith/Native, Faith/Sino-Korean, Faith/Loan) and relevant markedness constraints such as *U, *HH (high tone does not appear more than once in a prosodic word). Since no word appears with HH in Yanbian, and the unaccented class does not appear in loanwords, whereas it does in native and Sino-Korean words, the ranking can be represented as follows:

(20) *HH » Faith/Native, Faith/Sino-Korean » *U » Faith/Loan

However, as we have seen earlier, the accentuation and the accent distribution in each lexical class is not so simple, and it is also impossible to explain all the accent patterns of Yanbian loanwords only based on the faithfulness constraint to stress/accent patterns in the source language. Rather, Yanbian loanword accentuation seems to result from the interaction of gradiently weighted constraints, not from a single fixed ranking of markedness constraints with a faithfulness constraint to the source language. The core-periphery theory, at least as far as our Yanbian data is concerned, simplifies the lexical stratification too much and leaves many exceptions.

One possible analysis is to understand the loanword adaptation as an induction process originating from a faithfulness constraint to the source language and resulting in several relevant markedness constraints:

(a) In the very beginning stages of loanword adaptation, all loanwords are adapted as faithfully as possible to the source language; the output candidates are evaluated by a faithfulness constraint to the source language.
(b) After a certain number of loanwords are adapted, speakers start to analyze the accent location in loanwords phonologically and discover the relevant markedness constraints through a learning process.[20] The generalizations abstracted in this way are then extended to loanwords as a whole.
(c) The original faithfulness constraints are demoted below the relevant markedness constraints, which may or may not be highly ranked in the recipient language. These markedness constraints, along with other constraints which are originally "active" in the recipient language, are weighted by the learning algorithm so that the weight hierarchy can achieve a more or less "faithful adaptation" of the source language.
(d) The resulting "faithful adaptation" is thus based on this generalization, which was originally made by assessing the words that were introduced into the recipient language in the beginning

[20] The assumption in standard OT and Harmonic Grammar is that constraints do not need to be learned, whereas in the Hayes and Wilson Maxent model (2008), both constraints and their weighting are learned. At this point with our data, we cannot choose between these alternatives.

stages of the adaptation. At this later stage, each loanword is no longer faithful to the source language, but instead loanwords tend to be faithful to the general patterns observed in the source language. In this sense, we can still maintain that the source language determines the output forms of the loanwords.

A major implication of this idea is that loanword accentuation can have different adaptation patterns depending on the source language. In order to demonstrate this point, we tried to simulate the adaptation of both English and Japanese loanwords (monosyllabic, disyllabic, trisyllabic) by using Jäger's (2007) Stochastic Gradient Ascent learning algorithm, adjusting faithfulness constraints to start at 10 and markedness constraints at 0. Table 50 shows the obtained weights of two loanword classes and Figures 3 and 4 show the weight change process of some constraints as the algorithm is rerun over the data. The x-axis indicates the number of iterations of the algorithm over the data and the y-axis indicates the weight.

Table 50
Obtained weights of two loanword classes

English		Japanese	
*LAPSE-R	3.66	*U	4.50
*U	3.30	*LAPSE-R	3.12
INITIAL	2.00	WtoS	2.75
WtoS	0.91	Faith	1.29
Faith	0.81	INITIAL	0.93
StoW	0.04	NF	0.61
NF	0.04	StoW	0.14

Figure 3
Simulation of the learning process of English loanwords

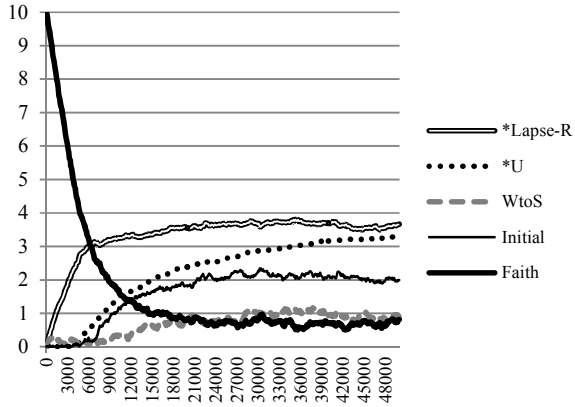
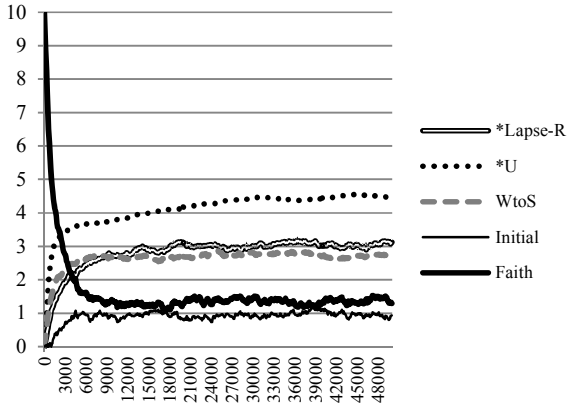


Figure 4
Simulation of the learning process of Japanese loanwords



As seen in Figures 3 and 4, the faithfulness constraints to English stress and Japanese accent are demoted as the learning process proceeds, whereas some markedness constraints such as *LAPSE-R and *U are promoted and weighted higher than the faithfulness constraint. Also, the resulting weight hierarchies differ between English and Japanese loanwords: e.g. INITIAL is weighted relatively higher in English loanwords than in Japanese loanwords, while WtoS is weighted higher in Japanese loanwords than in English loanwords.

This kind of adaptation process probably occurred because most Yanbian speakers do not have many opportunities to consult with native speakers of English/Japanese unlike Mandarin, and because stress/accent does not appear in the orthography of these source languages. Thus, we assume that at some point Yanbian people started to generalize the adaptation patterns based on their limited understanding and available data. The generalization is the process of introducing (or promoting) new markedness constraints into the loanword section of the weight hierarchy.

Under this view, the native/Sino-Korean lexicon and each loanword category can have different constraint sets or weight hierarchies. Also "new" constraints in the loanword section are not due to the "emergence of the unmarked" but originate from the source language, being computed from the existing phonological rules supported by lexical frequencies. There is no need to assume Universal Grammar in this hypothesis.[21]

7. Summary and conclusion

In accentual loanword adaptation, the prominence in the source language is often not respected, even if the accentual system or the distributional patterns in the recipient language could allow it to accommodate the position of the prominence in the input. This contrasts with segmental adaptation, in which segments in the source language tend to be respected more or less faithfully if they agree with phonotactic restrictions of the recipient language. Concerning this puzzling asymmetry, previous literature points out several possible factors: a) Universal Grammar, b) statistical tendencies in the native lexicon, c) the covert grammar of the native accent system, and d) universal phonetic considerations. In this paper, we studied loanword accentuation in Yanbian Korean and offered a new perspective on the origin of the loanword accentuation, through an analysis of the factors affecting the accentual assignment rules.

In Yanbian loanwords, the accent is basically located in a two-syllable window at the right edge of the word. The accent pattern differs between disyllabic and longer words. Penultimate is the strong default accent in disyllabic loanwords, and syllable weight affects the distribution gradiently. In particular, the syllable weight effect is strongest in Heavy-Light, which is the most preferred structure for the default penultimate accent: almost 100% of words with Heavy-Light structure appear with penultimate accent in loanwords. In three-or-more syllable words, the syllable weight effect emerges more clearly and accent polarization (Heavy-Light → penultimate, Light-Heavy → final) is observed, although penultimate is still the default accent.

[21] A problem for this hypothesis is that it does not differentiate between different words when the phonological conditionings are the same, and predicts the same kind of variation or non-variation for every single word. Accounting for how actual individual words pattern is in general a task for future research wherein models try to predict patterns in the lexicon.

On the other hand, the default accent in Yanbian native words is final. A syllable weight effect is observed in native words as well, but it is not as strong as in loanwords. Statistical analysis shows that the different accent distributions between the native words and loanwords are attributed to the lexical class difference itself. The same is true for Sino-Korean words, although the default in disyllabic Sino-Korean words is the penultimate accent class, as in disyllabic loanwords. The distributional difference in syllable structure between the three lexical classes does not explain the different tendencies in the default accent class and accent patterns. Thus, we can assume that each lexical class has different accentual assignment rules, and the default accent class or accent assignment rules are fixed for each category separately. The result of the wug test supports this hypothesis: speakers tended to distinguish the two lexical classes (native vs. loan) and showed a different accentual assignment accordingly. The result of the wug test is also important in that it shows that the different accentuation patterns between native words and loanwords are not due to a faithfulness constraint to a lexical accent that exists in native words but not in loanwords (or weighted much lower than relevant markedness constraints in loanwords) but rather are due to the lexical class itself. Thus, we conclude that the covert grammar of the native system is not relevant in the loanword accentuation in Yanbian.

So where does the Yanbian loanword accentuation come from? Our hypothesis is that Yanbian loanword accentuation results from the grammar of the source language and lexical statistics, along with some adjustments by Yanbian native grammar, as exemplified by *LAPSE-RIGHT or *UNACCENTED. If this is correct, then different accentuation patterns should appear depending on the source language. This difference is in fact observed in our Yanbian data.

We compared the three different loanword categories in Yanbian that are from different source languages with different prosodic types (English: stress, Japanese: pitch accent, Mandarin: tone) and showed statistically that each has its own accentual adaptation system. A syllable weight effect is active in English and Japanese loanwords (final for the former and penultimate/final for the latter) but not in Mandarin loanwords, whereas suprasegmental type (stress/pitch accent/tone) affects the adaptation patterns more in Mandarin than in English and Japanese. Thus, we can conclude that loanword adaptation is not from Universal Grammar or universal phonetic considerations, but basically reflects the phonology/phonetics of the original source languages along with some adjustments based on the native grammar.

We proposed a loanword adaptation model based on our analysis of Yanbian loanword accentuation. The loanword adaptation is understood as an induction process from an initial stage with faithfulness to the source language into a state where several relevant markedness constraints come into play. Through a learning process, the original faithfulness constraints to the source language are demoted below several relevant markedness constraints. These markedness constraints are weighted by the learning algorithm so that the weight hierarchy can achieve a more or less "faithful adaptation" of the source language. The resulting "faithful adaptation" is thus based on this generalization process. Under this view, each separate sublexicon can have a different weight hierarchy of markedness constraints.

It would be interesting to test this hypothesis of loanword adaptation with other loanword data. For example, English loanwords in Japanese or Kyengsang Korean may be reanalyzed by postulating a faithfulness constraint to the stress position in English as well as several markedness constraints such as WEIGHT-TO-STRESS and *ACCENTED EPENTHESIS. Comparison of

different loanword categories within one language should be conducted in more detail as well, by taking into account not only the faithfulness constraint to each source language but also variations (distributional frequency), since variations are the concrete numerical manifestations of the constraint interaction that can be different from one subcategory to another. This is a task for future research.

Appendix

The following is the list of words that were used in the wug test.

Loan						**Native**					
Heavy-Light		Light-Heavy		Light-Light		Heavy-Light		Light-Heavy		Light-Light	
알토	al.t^{h}o	에칭	e.ts^{h}iŋ	에뮤	e.mju	앙짜	aŋ.ts*a	아늠	a.nɨm	애채	ɛ.ts^{h}ɛ
인바	in.pa	하켄	ha.k^{h}en	하키	ha.k^{h}i	얼개	əl.kɛ	아습	a.sɨp	여뀌	jə.k*wi
인디	in.ti	이퀠	i.k^{h}wəl	히피	hi.p^{h}i	꺽지	k*ək.tsi	하릅	ha.rɨp	예새	je.sɛ
칸나	k^{h}an.na	이젤	i.tsel	호머	ho.mə	끙게	k*ɨŋ.ke	으름	ɨ.rɨm	까리	k*a.ri
캔터	k^{h}ɛn.t^{h}ə	가닛	ka.nis	이슈	i.sju	갈개	kal.kɛ	야살	ja.sal	꾸미	k*u.mi
맘바	mam.pa	카불	k^{h}a.pul	게토	ke.t^{h}o	갈미	kal.mi	까락	k*a.rak	개리	kɛ.ri
만나	man.na	캐럴	k^{h}ɛ.rəl	카뮈	k^{h}a.mwi	길미	kil.mi	개암	kɛ.am	겨리	kjə.ri
맥시	mɛk.si	큐섹	k^{h}ju.sek	커마	k^{h}ə.ma	맹이	mɛŋ.i	겨릅	kjə.rɨp	괴끼	kwe.k*i
먼로	mən.ro	고딕	ko.tik	캐디	k^{h}ɛ.ti	목새	mok.sɛ	고콜	ko.k^{h}ol	고누	ko.nu
뮌슈	mwin.sju	마임	ma.im	콰시	k^{h}wa.si	몽니	moŋ.ni	고삭	ko.sak	모끼	mo.k*i
넥타	nɛk.t^{h}a	뮤온	mju.on	오더	o.tə	물미	mul.mi	고둥	ko.tuŋ	모루	mo.ru
벌키	pəl.k^{h}i	모굴	mo.kul	베레	pe.re	날피	nal.p^{h}i	귀얄	kwi.jal	너리	nə.ri
범퍼	pəm.p^{h}ə	나셀	na.sel	버디	pə.ti	는개	nɨn.kɛ	구듭	ku.tɨp	너새	nə.sɛ
밴조	pɛn.tso	노넷	no.nes	버저	pə.tsə	능에	niŋ.e	구죽	ku.tsuk	느치	nɨ.ts^{h}i
팬지	p^{h}ɛn.tsi	오팔	o.p^{h}al	파카	p^{h}a.k^{h}a	올미	ol.mi	미늘	mi.nɨl	노깨	no.k*ɛ
펑키	p^{h}əŋ.k^{h}i	바순	pa.sun	페니	p^{h}e.ni	삘기	p*il.ki	너겁	nə.kəp	바대	pa.tɛ
폴카	p^{h}ol.k^{h}a	베냉	pe.nɛŋ	피케	p^{h}i.k^{h}e	밴대	pɛn.tɛ	너설	nə.səl	버캐	pə.k^{h}ɛ
람다	ram.ta	배럴	pɛ.rəl	푸가	p^{h}u.ka	뱅니	pɛŋ.ni	바림	pa.rim	파개	p^{h}a.kɛ
린네	rin.ne	배팅	pɛ.t^{h}iŋ	라미	ra.mi	핑구	p^{h}iŋ.ku	배동	pɛ.toŋ	벼리	pjə.ri
론도	ron.to	파톤	p^{h}a.t^{h}on	래커	rɛ.k^{h}ə	분디	pun.ti	보꾹	po.k*uk	보늬	po.ni
룽기	ruŋ.ki	패럿	p^{h}ɛ.rəs	리마	ri.ma	살미	sal.mi	부럼	pu.rəm	씨아	s*i.a
센나	sen.na	리봅	ri.pop	리라	ri.ra	살피	sal.p^{h}i	부룩	pu.ruk	사래	sa.rɛ
센서	sen.sə	루틴	ru.t^{h}in	세터	se.t^{h}ə	솔기	sol.ki	서덜	sə.təl	스리	sɨ.ri
델타	tel.t^{h}a	서핑	sə.p^{h}iŋ	데카	te.k^{h}a	떨기	t*əl.ki	뚜깔	t*u.k*al	따비	t*a.pi
덴버	ten.pə	세륨	se.rjum	터부	t^{h}ə.pu	똘기	t*ol.ki	더넘	tə.nəm	따리	t*a.ri
댄서	tɛn.sə	시멘	si.men	테제	t^{h}e.tse	달구	tal.ku	드난	tɨ.nan	더께	tə.k*e
던디	tən.ti	데칸	te.k^{h}an	튜너	t^{h}ju.nə	들메	tɨl.me	주럽	tsu.rəp	드므	tɨ.mɨ
진저	tsin.tsə	데릭	te.rik	디키	ti.k^{h}i	종개	tsoŋ.kɛ	두겁	tu.kəp	되리	twe.ri
웰터	wel.t^{h}ə	자칼	tsa.k^{h}al	제너	tse.nə	둔테	tun.t^{h}e	두멍	tu.məŋ	처네	ts^{h}ə.ne
윈저	win.tsə	왜건	wɛ.kən	조커	tso.k^{h}ə	왁대	wak.tɛ	두릅	tu.rɨp	조새	tso.sɛ

Acknowledgments

I thank NLLT editor Junko Ito and three anonymous reviewers for detailed and helpful comments. I also thank Adam Albright for sharing with me the program for the learning algorithm used in section 4.1, and Shigeto Kawahara for advice about statistics. This work was supported by a Grant-in-Aid for Scientific Research (KAKENHI, NO. 21720136). The data on which this study is based is available from the author.

References

Alderete, John. 1999. Head dependence in stress-epenthesis interaction. In Ben Hermans and Marc van Oostendorp (Eds.), *The derivational residue in phonological Optimality Theory*, 29–50. Amsterdam: John Benjamins. Reprinted in John McCarthy (Ed.), *Optimality Theory in phonology: selected readings*, 215–227, Oxford: Blackwell.

Albright, Adam. 2002. Islands of reliability for regular morphology: evidence from Italian. *Language* 78: 684–709.

Albright, Adam. 2008. Perl script. (personal communication.)

Bates, Douglas, Ben Bolker, and Martin Maechler. 2013. Lme4: linear mixed effects models using S4 classes. R package version 0.999999-2. Online: http://cran.r-project.org/web/packages/lme4/index.html.

Boersma, Paul and Bruce Hayes. 2001. Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32: 45–86.

Che, Xiangchun. 2004. Chōsengo ryūsei hōgen-no meishi-no akusento taikei [The accent system of the Longjing dialect of Korean]. *Tokyo University Linguistic Papers* 23: 1–22.

Chi, Feng-hua. 2008. Enpen chōsengo onshakugo-no goon tokuchō-to akusento patān-ni tsuite. [Segmental features and the accent of Mandarin loanwords in Yanbian Korean]. *Chōsen Gakuhō* 207: 1–38.

Chung, Young-Hee. 2002. Contextually-dependent weight of closed syllables: cases of the North Kyungsang dialect and English. *Enehak* 33: 23–35.

Gordon, Matthew. 2002. A factorial typology of quantity insensitive stress. *Natural Language and Linguistic Theory* 20: 491–552.

Green, Thomas and Michael Kenstowicz. 1995. The lapse constraint. *Formal Linguistics Society of the Midwest* 6: 1–15.

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

Hayes, Bruce and Colin Wilson. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39: 379–440.

Hirayama, Teruo. (Ed.) 1960. *Zenkoku akusento jiten* [National accent dictionary]. Tokyo: Tokyodo Publishing.

Hsieh, Feng-fan. 2006. High infidelity: the non-mapping of Japanese accent onto Taiwanese tone. In Feng-fan Hsieh and Michael Kenstowicz (Eds.), *Studies in Loanword Phonology (MITWPL* 52*)*, 1–27. Cambridge: MITWPL.

Hsieh, Feng-fan and Michael Kenstowicz. 2008. Phonetic knowledge in tonal adaptation: Mandarin and English loanwords in Lhasa Tibetan. *Journal of East Asian Linguistics* 17: 279–297.

Hsieh, Feng-fan, Michael Kenstowicz, and Xiaomin Mou. 2009. Mandarin adaptations of coda nasals

in English loanwords. In Andrea Calabrese and Leo Wetzels (Eds.), *Loanword Phonology: Issues and Models*: 131–154. Amsterdam: John Benjamins.
Hung, Henrietta. 1994. *The rhythmic and prosodic organization of edge constituents*. Waltham, MA: Brandeis University dissertation.
Ito, Chiyuki, Yoonjung Kang, and Michael Kenstowicz. 2006. The adaptation of Japanese loanwords into Korean. In Feng-fan Hsieh and Michael Kenstowicz (Eds.), *Studies in Loanword Phonology (MITWPL 52)*: 65–104, Cambridge: MITWPL.
Ito, Chiyuki. 2000. Chūki chōsengo kanjigo akusento shiryō [Glossary of Sino-Korean words with accentuation]. In Rei Fukui (Ed.), *Kankokugo akusento ronsō* [Studies in Korean accentology], 99–247, Tokyo: ICHEL, The University of Tokyo.
Ito, Chiyuki. 2007. *Chōsen kanjion kenkyū* [Sino-Korean phonology]. Tokyo: Kyūko shoin.
Ito, Chiyuki. 2008a. Historical development and analogical change in Yanbian Korean accent. *Harvard Studies in Korean Linguistics* XII: 165–178.
Ito, Chiyuki. 2008b. Analogical changes in the accent of Sino-Korean words in Yanbian Korean. In Natasha Abner and Jason Bishop (Eds.), *Proceedings of the 27th West Coast Conference on Formal Linguistics*, 238–246, Somerville, MA: Cascadilla Proceedings Project.
Ito, Chiyuki and Michael Kenstowicz. 2009. Mandarin loanwords in Yanbian Korean II: tones. *Language Research* 45.1: 85–109.
Itô, Junko and Armin Mester. 1995. The core-periphery structure of the lexicon and constraints on reranking. In Jill Beckman, Suzanne Urbanczyk, and Laura Walsh (Eds.), *University of Massachusetts Occasional Papers in Linguistics [UMOP] 18: Papers in Optimality Theory*, 181–209, Amherst: GLSA.
Itô, Junko and Armin Mester. 1999. The structure of the phonological lexicon. In Natsuko Tsujimura (Ed.), *The handbook of Japanese linguistics*, 62–100, Malden, MA, and Oxford, U.K.: Blackwell Publishers.
Jäger, Gerhard. 2007. Maximum entropy models and stochastic Optimality Theory. In Annie Zaenen, Jane Simpson, Tracy Holloway King, Jane Grimshaw, Joan Maling, and Chris Manning (Eds.), *Architectures, rules, and preferences: variations on themes by Joan W. Bresnan*, 467–479, Stanford: CSLI publications.
Jun, Ho Kyung. 2006. Factors affecting accentual patterns of loanwords in Pusan Korean. *Harvard Studies in Korean Linguistics* XI: 158–170.
Kadowaki, Seiichi, Isam Matsuo, Yoshiro Takashima, and Yukitoshi Yutani. 1993. *Chōsengo jiten* [Korean-Japanese Dictionary]. Tokyo: Shōgakukan.
Kager, René. 1999. *Optimality theory*. Cambridge: Cambridge University Press.
Kang, Yoonjung. 2010. Tutorial overview: suprasegmental adaptation in loanwords. *Lingua* 120: 2295–2310.
Kang, Yoonjung. 2011. Loanword phonology. In Marc van Oostendorp, Colin J. Ewen, Elizabth Hume, and Keren Rice (Eds.), *The Blackwell companion to phonology*, 2258–2282, Malden, MA: Wily-Blackwell.
Kang, Yoonjung, Michael Kenstowicz, and Chiyuki Ito. 2008. Hybrid loans: a study of English loanwords transmitted to Korean via Japanese. *Journal of East Asian Linguistics* 17.4: 299–316.
Kenstowicz, Michael. 1997. Quality-sensitive stress. *Rivista di Linguistica* 9.1: 157–188.

Kenstowicz, Michael. 2004. Tone loans: the adaptation of English loanwords into Yoruba. Unpublished paper presented at the 35th Annual Conference on African Linguistics, Cambridge, MA.

Kenstowicz, Michael and Hyang-Sook Sohn. 2001. Accentual adaptations in North Kyungsang Korean. In Michael Kenstowicz (Ed.), *Ken Hale: A life in language*, 239–270, Cambridge: MIT Press.

Kenstowicz, Michael and Atiwong Suchato. 2006. Issues in loanword phonology: a case study from Thai. *Lingua* 116-7: 921–949.

Kim, Jungsun. 2009. Double accent in loanwords of North Kyungsang Korean and variable syllable weight. *Language Research* 45-1: 67–83.

Kim, Min-Swu. 1997. *Wulimal ewen sacen* [Korean etymological dictionary]. Seoul: Thayhaksa.

Kōno, Rokurō. 1968. *Chōsen kanjion-no kenkyū* [Study on Sino-Korean]. Nara: Tenri jihōsha. Reprinted in Rokurō Kōno 1979, *Kōno rokurō chosakushū* [*Rokurō Kōno works*] 2: 295–512, Tokyo: Heibonsha.

Kubozono, Haruo. 1996. Syllable and accent in Japanese: evidence from loanword accentuation. *The Bulletin (Phonetic Society of Japan)* 211: 71–82.

Kubozono, Haruo. 2001. Epenthetic vowels and accent in Japanese: facts and paradoxes. In Jeroen van de Weijer and Tetsuo Nishihara (Eds.), *Issues in Japanese phonology and morphology*, 113–142, Berlin, New York: Mouton de Gruyter.

Kubozono, Haruo. 2006. Where does loanword prosody come from? A case study of Japanese loanword accent. *Lingua: Special Issue on Loanword Phonology* 116: 1140–1170.

Kubozono, Haruo. 2008. Japanese accent. In Shigeru Miyagawa and Mamoru Saito (Eds.), *The Oxford Handbook of Japanese Linguistics*, 165–191, Oxford: Oxford University Press.

Kwuklip kwuke yenkwuwen [National Institute of the Korean Language]. (Ed.) 1999. *Phyocwun kwuke taysacen* [Standard Korean dictionary]. Seoul: Tusangtonga.

Leben, Wil. 1996. Tonal feet and the adaptation of English borrowings into Hausa. *Studies in African Linguistics* 25: 139–54.

Lee, Dongmyung. 2006. Weight-sensitive tone patterns in loan words of South Kyungsang Korean. *Harvard Studies in Korean Linguistics* XI: 197-210.

Lee, Dongmyung. 2008. *The Loanword Tonology of South Kyungsang Korean*. Bloomington, IN: Indiana University dissertation.

McCawley, John. 1968. *The phonological component of a grammar of Japanese*. Mouton: The Hague.

McCarthy, John and Alan Prince. 1994. The emergence of the unmarked: optimality in prosodic morphology. In *Proceedings of NELS* 24: 333–379, Amherst: GLSA.

Miyashita, Naoko. 2007. *Gengo sessyoku-to chūgoku chōsengo-no seiritsu* [Language contact and the establishment of Korean in China]. Fukuoka: Kyushu University Press.

NHK hōsōbunka kenkyūjo. (Ed.) 1998. *NHK Nihongo hatsuon akusento jiten* [NHK Japanese pronunciation accent dictionary]. Tokyo: NHK publishing.

Park, Youngmae. 2001. Enpen chōsengo-no akusento-ni kansuru kōsatsu [Accent of Yanbian Korean in China]. *Kyoto University Linguistic Research* 20: 181–195.

Pater, Joe. 2009. Morpheme-specific phonology: constraint indexation and inconsistency resolution. In Steve Parker (Ed.) *Phonological argumentation: essays on evidence and motivation*, 123–

154, London: Equinox.
Potts, Christopher, Joe Pater, Karen Jesney, Rajesh Bhatt, and Michael Becker. 2010. Harmonic Grammar with linear programming: from linear systems to linguistic typology. *Phonology* 27: 77–117.
Prince, Alan and Paul Smolensky. 1993. Optimality Theory: constraint interaction in generative grammar. Technical Report #2, Rutgers Center for Cognitive Science, Rutgers University. [Oxford: Blackwell, 2004].
R Development Core Team. 2011. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. http://www.R-project.org.
Ramsey, S. Robert. 1978. *Accent and morphology in Korean dialects*. Seoul: Tower Press.
Shen, Yinji. 2011. On the accent of non-Mandarin loanwords in Yanbian Korean. Handout from the 6th Phonology Festa, Shiga, Feb 17.
Shen, Yinji and Yasunori Takeuchi. 2011. Treatment of the Mandarin unaspirated consonants adapted in Mandarin loanwords in Yanbian Korean. *Phonological Studies* 14: 51–62, The Phonological Society of Japan.
Shinohara, Shigeko. 1997a. Default accentuation and foot structure in Japanese: analysis of Japanese adaptations of French words. In Benjamin Bruening, Yoonjung Kang, and Martha McGinnis (Eds.), *PF : Papers at the Interface* 30: 263–290, Cambridge: MITWPL.
Shinohara, Shigeko. 1997b. *Analyse phonologique de l'adaptation japonaise de mots étrangers* [Phonological analyses of Japanese adaptation of foreign words]. Paris, FR: University of Paris III dissertation.
Shinohara, Shigeko. 2000. Default accentuation and foot structure in Japanese: evidence from Japanese adaptations of French words. *Journal of East Asian Linguistics* 9: 55–96.
Silverman, Daniel. 1992. Multiple scansions in loanword phonology: evidence from Cantonese. *Phonology* 9: 289–328.
Tanaka, Shinichi. 1996. *Nihongo-no onsetsukōzō-to on'ingenshō* [Syllable structure and phonological phenomena in Japanese]. Osaka, JP: Osaka University of Foreign Studies MA thesis.
Umeda, Hiroyuki. 1993. Enpen chōsengo-no on'in [Phonological study of Yanbian Korean]. In *Studies of Linguistics and Cultural Contacts* 6: 131–145, Tokyo: ILCAA, Tokyo University of Foreign Studies.
Xu, Yi. 1997. Contextual tonal variations in Mandarin. *Journal of Phonetics* 25: 61–83.
Yip, Moira. 2006. The symbiosis between perception and grammar in loanword phonology. *Lingua* 116: 950–75.