Updating the analysis of Japanese compound accent

John Alderete, Simon Fraser University; alderete@sfu.ca

Investigating the predictions of a theory requires both a strong purchase of its consequences and the rigorous pursuit of these consequences in the domains pertaining to them (Prince 2007). In this squib, we attempt to emulate the rigorous pursuit of theory so characteristic of Alan Prince's work by investigating some of the consequences of a recent approach to Japanese pitch accent proposed in Itô & Mester 2015/to appear.

In this approach, pitch accent is a property of the head of a moraic trochee that follows a Latin-like pattern, where, roughly speaking, accent falls on the syllable containing the antepenultimate mora. For example, pitch accent falls on the antepenultimate light syllable in loanwords like *barusérona* 'Barcelona' (2b). Words like *amerika* are unaccented in this analysis (1b) because it is impossible to produce the antepenultimate pattern without violating other high-ranking constraints in the language. The patterns predicted by this approach overlap with many of the patterns found in noun-noun compounds, supporting the conjecture that the analysis of the two patterns are unified in some sense (see Kubozono 1988/1993: 17 for a particularly clear statement of this). However, certain recalcitrant facts prevent a direct carry-over of Itô & Mester's analysis to compounds. First, when the second member is of a compound is "long" (=three or more moras), accent regularly falls on the first syllable of the second member (1a), but this regular accent is not found in non-compound words like *amerika* (1b). Second, when the second member of a compound is "short" (1-2 moras), accent falls on the last syllable of the first member (2a). While on the surface (2a) resembles *barusérona* in having antepenultimate accent, for a variety of reasons, prior research assumes that accent falls on the head of a penultimate foot in this case, and that this foot is iambic. These assumptions require foot flipping and do not accord with Itô & Mester's "close to final" moraic trochee analysis.

	Compound noun	*Non-compound words*	
(1)	… (L'L)(LL)	(L L)(L L)	*Accentedness problem*
	a. minami-(áme)(rika)	b. (ame)(rika)	
(2)	L(LL')-(LL)	(L L)(L' L) L	*Foot-flipping problem*
	a. ni(sikí)-(hebi)	b. (baru)(séro)na	

The two problems above seem to suggest that compound and non-compound nouns have certain irreconcilable differences in terms of the existence and placement of feet, and that the possibility of a single internally consistent analysis is limited. I will argue instead that a unified analysis is possible that both maintains a formal commitment to Itô & Mester's approach to antepenultimacy, and realizes Kubozono's contention that the two systems are in fact related by the antepenultimacy generalization.

Let us review the core facts of noun-noun compounds in more detail before delving into analysis. These facts, and many analytical assumptions for organizing them, are guided by prior research, including Kubozono 1988/1993, Poser 1990, and the recent OT analyses of Kubozono 1995, 1997, Alderete 2001, Tanaka 2001, and Itô & Mester 2007. The data below illustrate the three basic patterns in compounds with short second members (N2's henceforth). While lexically specified accent (shown with underlining) can bring about surface accent on a heavy or a non-final light (3b), the default pattern in these forms is for accent on the last syllable of the first conjunct (3a), a pattern that can even trump lexical accent, e.g., *huransú-pan*. Each of these syllabic profiles can also surface unaccented (3c), and

this is indeed very common in some forms. For example, approximately 70% of words with LL-LL profiles are unaccented. Tanaka (2001) states that Japanese appears to be moving away from […σ-L'L] patterns, towards […σ'-LL], supporting the default pattern as a core pattern in the system. Also, (3b) may be more characteristic of foreign words, and some nativized loans have shifted to the (3a) pattern, which again supports regularization to the default.

(3) Compound nouns with short N2 (1 or 2 μ)

a. Default final in N1	b. Lexical μ' in N2	c. Unaccented compound
… σ'-L nagoyá-si 'Nagoya City'	(…σ-L' *rare*)	… σ-L nihon-go 'Japanese language'
… σ'-L L* nisikí-hebi 'reticulating python'	… σ-L'L niwaka-áme 'shower'	… σ-LL kinu-ito 'silk thread'
… σ'-L L kabutó-musi 'beetle'		μ μ -LL mizu-bana 'running nose'
… σ'-H huransú-pan 'french bread'	… σ-H' sunakku-báa 'snack bar'	… σ-H efferu-tou 'The Eiffel Tower'

In compound nouns with long N2's, we again have the emergence of non-final lexical accent (4b) and a default pattern (4a). However, in this context, default accent falls on the first syllable of the second conjunct. Support for the default comes in two forms: (i) it surfaces when the input nouns have no accent, e.g., *minami-ámerika*, and (ii) the fact that some forms alternate between lexical antepenultimate accent and the default, e.g., *neri-hámigaki ~ neri-hamígaki*.

(4) Compound nouns with long N2 (3+μ)

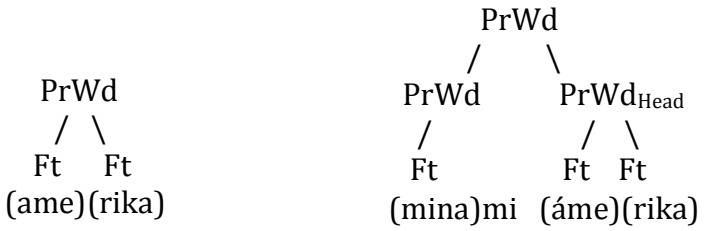
a. Default initial in N2	b. Non-initial lexical accent
… σ-L'LL onna-gókoro 'women's heart'	*In general, non-final lexical accent can cause non-initial accent in N2*
… σ-L'LLL neri-hámigaki 'toothpaste'	… σ-LL'LL neri-hamígaki
… σ-L'LLL minami-ámerika 'South America'	

Any analysis of accent in compound nouns will have to reckon with (i) the emergence of lexical prosody, (ii) unaccented compounds with short N2s, and (iii) the two default patterns. I follow prior research (Kubozono 1995, 1997, Alderete 2001, Tanaka 2001) in assuming that the constraints that predict the defaults interact with faithfulness constraints, allowing lexical prosody to surface in non-final syllables. It seems that some measure of lexical specification is necessary to account for both realization of lexical prosody and its suppression by the default, and presumably faithfulness constraints can be co-indexed with distinct correspondence relations in a way that accounts for both classes (Itô & Mester 1999). I will also not analyze unaccented words here, but just note that many of the ideas put forth in Itô

& Mester (2015/to appear) seem to straightforwardly extend to some of these cases. Several of these words are tri- and quadramoraic, which mirrors these patterns in non-compound words.

I focus instead on the default patterns and the problems they raise mentioned above. The accentedness problem is that compounds have regular accent, but some words that are accented in compounds are unaccented in non-compound words, e.g., *minami-ámerika*, cf. *amerika*. A key insight into this problem comes from the idea that compounds are headed structures, and that headed constituents have special properties (Alderete 2001, Itô & Mester 2007). In the typology of compounds developed in Itô & Mester (2007), based on many original insights in Kubozono (1988/1993), the noun-noun compounds under discussion here are 'word compounds' that involve a recursive prosodic word. In Alderete (2001), these recursive prosodic words are right-headed, as shown below. Given these assumptions, the difference between unaccented *amerika* and *minami-ámerika* can be explained by assuming the head PrWd of the larger compound must have a peak prominence. *Amerika* does not have a head PrWd, so this constraints is vacuously satisfied in non-compounds.

(5) Accentedness to the head PrWd



We may formalize this intuition with the constraint PRWDHEADACCENT, cf. Itô & Mester's MINWORDACCENT, which is similar in spirit. Inserting this constraint at the top of the hierarchy resolves the inconsistency between these two classes of examples (7).

(6) PRWDHEADACCENT (PRWDHEADACC)
The head PrWd of a prosodic word compound must have a peak prominence.

(7) Inconsistency resolved by identifying $PrWd_{Head}$

Winner ~ Loser	PRWDHDACCT	RIGHTMOST	WORDACCENT
a. minami-(áme)(rika) ~ minami-(ame)(rika)	W	L	W
b. (ame)(rika) ~ (áme)(rika)		W	L

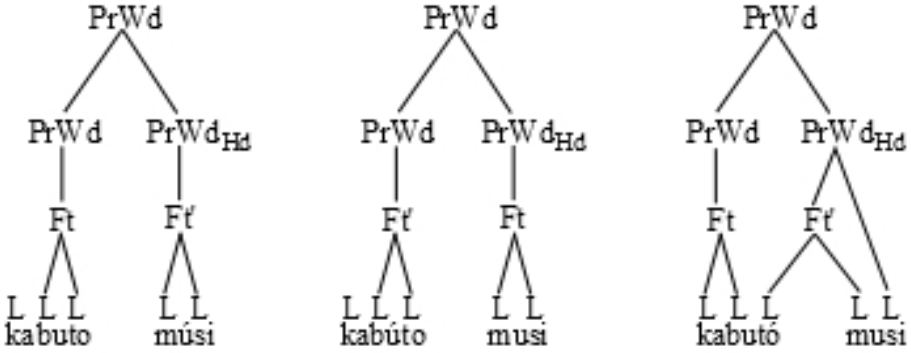
I assume that the relevant constraint requiring accentedness in long N2's is prosodic in nature because, while we focus on right-headed compound nouns morpho-syntactically, Sino-Japanese compounds behave in similar respects to these compounds, but they are left-headed (Kageyama 1982).

It seems self-evident that compounds should have heads and that they should be demarcated somehow through prominence. In a sense, therefore, a constraint like PRWDHEADACCENT is the null hypothesis for why compounds have regular accent on long second conjuncts. Interestingly, this same constraint paves the way to solving the foot-flipping problem, as shown below. The 'normal' way of satisfying PRWDHEADACCENT would be to accent the final LL, as in candidate (9a) (illustrated in (8)), but this puts accent on a final foot, which is outlawed by Itô & Mester's NONFINAL(FT'). Another sensible thing to do would be to accent the penultimate foot, as with candidate (9c), but this does not satisfy PRWDHDACCENT. Candidate (9b) can thus be viewed as the optimal way of satisfying this constraint. It is not perfect, because the head foot straddles the morpheme boundary (violating MATCH(STEM, PRWD) of Alderete and MacMillan (2015)), but it is otherwise perfect prosodically.

Finally, since mismatches are allowed, I note that stretching the $PrWd_{Head}$ two syllables inward, as in (9d), is not okay for the same reason it is not okay elsewhere: this leads to a violation of RIGHTMOST.

(8) Attempts at satisfying PRWDHDACC in short N2's

Candidate (9a) *Candidate (9c)* *Candidate (b)*



(9) Antepenultimacy in compounds with short N2 (PrWd boundaries shown with [ ])

/kabuto + musi/	PRWDHDACCT	RIGHTMOST	NONFINAL(FT')	MATCH(STEM, PRWD)
a. [ka(buto)]-[(músi)]			*!	
b. → [(kabu)][(tó-mu)si]				*
c. [ka(búto)]-[(musi)]	*!			
d. [ka][(búto)-(musi)]		*!		*

In this analysis, PRWDHEADACCENT works in concert with all of the constraints involved in predicting the antepenultimate pattern in Itô & Mester's system, even their INITIALFOOT (as demonstrated in the associated OTWorkPlace files available from the author's webpage). There is thus strong theoretical motivation for a footing like that given in (9b). But what other evidence might support (9b), which is, after all, a little odd given its misalignment of prosodic and morpho-syntactic structure?

Further evidence may come from another well-known morpho-phonological process in compounds, rendaku (Itô & Mester 1986; McCawley 1968; Vance 1987). Recent work in Rosen (2001, 2003) has examined how rendaku, the voicing of initial obstruents in compounds, can be blocked in semi-regular ways. In particular, Rosen documents a loose correlation between irregular accent in compounds with short N2s and a higher incidence of rendaku blocking. While there are no absolute regularities between accent pattern and rendaku blocking, Rosen's data support an interesting conjecture about voicing and footing. Suppose the foot straddling the two conjuncts in (9b) creates a kind of phonological cohesion that sets this structure apart from other compounds. In such a context, we might expect words with greater cohesion to be more word-like, and thus subject more often to idiosyncractic exceptions, like the resistance to rendaku. To check this, I have examined the raw data from Rosen's study (thanks to Eric Rosen for sharing this on short notice) and found an interesting trend in LL-LL and LL-L compounds (10). Compounds with straddled feet, including all LL-L cases and LL'-LL compounds with antepenultimate accent, seem to have a greater propensity for rendaku blocking than unaccented LL-LL words, which must be parsed as (LL)-(LL) under Itô & Mester's analysis. These facts also fit with a pattern of allomorphy in island name constructions in which pre-accented stems (which have straddling feet) correlate with rendaku blocking, e.g., *(itu)(kú-si)ma*, but unaccented words do not, as in *(saku)ra-(zima)* (Tanaka 2005).

(10) Rendaku blocking in LL-LL and LL-L compounds

	Rendaku	Rendaku blocked	Totals
L(L-L), L(L'-L)L	146	98 (40%)	244
(LL)-(LL)	297	148 (33%)	445

While the trend in (10) is not statistically significant, together with the allomorphy pattern, it lends support to the foot straddling analysis, and in turn the larger goal here of unifying the analysis of compound and non-compound words. Furthermore, the analysis proposed here really involves a small modification of Kubozono and Tanaka's original OT analyses. In essence, it motivates a role for PrWd heads in requiring compounds to have a peak prominence, even if this leads to a slight mismatch between the prosody and morpho-syntax (9b). This single assumption has a big impact, which can be appreciated by comparing the analyses in (7) and (9) with the foot-flipping approach in (1-2). The latter approach essentially involves a hybrid system of accenting the penultimate foot: the penultimate foot is a moraic trochee in words like *minami-(áme)(rika)*, but an iamb in *ka(butó)-(musi)*. This system is in two ways distinct from non-compound accent: it allows iambs in a system that is otherwise trochaic, and it assigns accent to the penultimate foot, not the final foot. Both approaches have odd foot parses, but the approach advocated here unifies both compound and non-compound systems with foot straddling. The iambic analysis in words like *ka(butó)-(musi)* accounts for the antepenultimate pattern, but in a way rather distinct from non-compound accent.

References

Alderete, John. 2001. Morphologically governed accent in Optimality Theory New York: Routledge.

Alderete, John & Kayleigh MacMillan. 2015. Reduplication in Hawaiian: Variations on a theme of minimal word. Natural Language and Linguistic Theory 33.1-45.

Itô, Junko & Armin Mester. 1986. The phonology of voicing in Japanese: Theoretical consequences for morphological accessibility. Linguistic Inquiry 17.49-73.

—. 1999. The structure of the phonological lexicon. The handbook of Japanese linguistics, ed. by N. Tsujimura, 62-100. Malden, MA: Blackwell Publishers.

—. 2007. Prosodic adjunction and Japanese compounds. MIT Working Papers in Linguistics. Formal Approaches to Japanese Linguistics 4, 97-111.

—. 2015/to appear. Unaccentedness in Japanese. Linguistic Inquiry.

Kageyama, Taro. 1982. Word formation in Japanese. Lingua 57.215-58.

Kubozono, Haruo. 1988/1993. The Organization of Japanese Prosody: Kurosio Publishers [revised 1988 doctoral dissertation from the University of Edinburgh].

—. 1995. Constraint interaction in Japanese phonology: Evidence from compound accent. Phonology at Santa Cruz [PASC], ed. by R. Walker, O. Lorentz & H. Kubozono, 21-38. Santa Cruz: Linguistics Research Center, UC Santa Cruz.

—. 1997. Lexical markedness and variation: A nonderivational account of Japanese compound accent. Proceedings from the West Coast Conference on Formal Linguistics 15, ed. by B. Agbayani & S.-W. Tang, 273-87.

McCawley, James D. 1968. The phonological component of a grammar of Japanese The Hague, Paris: Mouton.

Poser, William J. 1990. Evidence for foot structure in Japanese. Language 66.78-105.

Prince, Alan. 2007. The pursuit of theory. The Cambridge handbook of phonology, ed. by P. de Lacy, 33-60: Cambridge University Press.

Rosen, Eric. 2001 Phonological processes interacting with the lexicon: Variable and non-regular effects in Japanese phonology: University of British Columbia.

—. 2003. Systematic irregularity in Japanese Rendaku: How the grammar mediates patterned lexical exceptions. Canadian Journal of Linguistics 48.1-37.

Tanaka, Shin-ichi. 2001. The emergence of the 'unaccented': Possible patterns and variation in Japanese compound accentuation. Issues in Japanese phonology and morphology, ed. by J. van ed Weijer, 159-92. Dordrecht: Mouton de Gruyter.

—. 2005. Where voicing and accent meet: (In) completeness in the harmonic scale of phonological prominence. Voicing in Japanese, ed. by J. van de Weijer, K. Nanjo & Nishihara, 261-78. Berlin: Mouton de Gruyter.

Vance, Timothy. 1987. An Introduction to Japanese Phonology New York: SUNY Press.