ACCENT PLACEMENT IN JAPANESE BLENDS

Rachel Broad

A thesis submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Arts in the Department of Linguistics

Chapel Hill
2015

Approved by:

Jennifer L. Smith

Elliott Moreton

Katya Pertsova

ABSTRACT

Rachel Broad: Accent Placement in Japanese Blends
(Under the direction of Jennifer L. Smith)

This thesis investigates the factors influencing accent placement in Japanese lexical blends (e.g. *go*zira 'Godzilla,' *gorira* 'gorilla' + kuzira 'whale'). Previous studies on English blends have claimed that the relative contributions of source words are influenced by factors such as their linear order (Bat-El & Cohen 2012, Gries 2004a,b) or the privileged position of the head (Shaw 2013). I present accent data collected for a corpus of Japanese blends to show that their accent placement is determined by head faithfulness rather than linear order or patterns found in other word formation processes. This work has several implications. First, it provides support for the claim that the position of the head is relevant in blend formation. It also demonstrates that factors influencing blends are cross-linguistically relevant. Finally, it provides evidence that blending is a distinct process and that the internal morphological structure of blends is unlike that of compounds and reduced compounds.

ACKNOWLEDGEMENTS

This thesis could not have been completed without help from many different people. First, I am immensely grateful to my advisor, Jen Smith, for providing patient advice and support every step of the way. I'm also grateful to my thesis committee members for providing helpful comments and advice. In particular, I'm grateful to Elliott Moreton for encouraging me to consider problems from multiple angles and to Katya Pertsova for early advice in focusing the project more clearly. Thanks to the Sporklab group as a whole, including Brandon Prickett, for being my research family and for providing me with so many opportunities to learn and grow. Thanks to P-Side members for comments and discussion and to Chris Wiesen of the Odum Institute for statistical consultation. Thanks also goes to Yasushi and Katsumi Yoshimoto for their generous help with Japanese translation and survey stimuli and to Shigeto Kawahara for assistance distributing the surveys. This research was supported by NSF grant #CNS 1318520, "Towards pronounceable authentication strings."

I'm also grateful to all of my family, friends, and colleagues who supported me throughout this process. Thanks to all those who participated in my surveys and provided helpful feedback. A special thanks goes to Makiko Mizuta for invaluable assistance finding new examples of blends and the willingness to painstakingly provide a large number of accent judgments. I'm especially grateful to Sarah Andrabi for providing a constant source of inspiration. I'm also grateful to my parents for their unfailing encouragement and support.

TABLE OF CONTENTS

1. INTRODUCTION

Much recent work has been done to investigate the factors influencing the segmental and prosodic formation of blends in English (Bat-El & Cohen 2012; Arndt-Lappe & Plag 2013; Shaw 2013). However, relatively little has been done to investigate this word formation process in other languages. In this thesis, I present a study on Japanese blends, focusing specifically on accent placement. I show that accent is determined by head faithfulness, a factor that has previously been shown to be relevant for blends in English (Shaw 2013). These findings reveal that factors affecting blend formation are cross-linguistically relevant and also provide a better understanding of the morphological structure of blends.

Lexical blends are words formed through a process of truncation and overlap. More specifically, they take material from the beginning of the first source word and the end of the second source word. Some examples of blends in Japanese are given in (1) below.

(1) Examples of Japanese blends

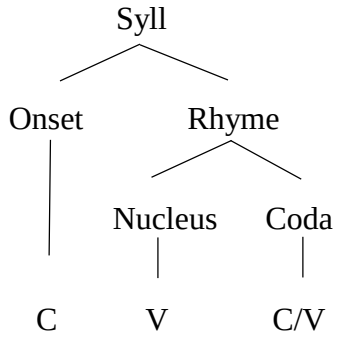
Blend	Source Word 1	Source Word 2
*potet*oruneedo ‘a curly potato snack’	*poteto* ‘potato’	toruneedo ‘tornado’
*go*zira ‘Godzilla’	*gorira* ‘gorilla’	toruneedo ‘tornado’
*ne*mozii ‘hungry and sleepy’	*nemui* ‘sleepy’	kuzira ‘whale’
*h*ine ‘a hie-rice hybrid’	*hie* ‘barnyard millet’	ine ‘rice-plant’

This construction distinguishes blends from other combinatory word formation processes in Japanese such as compounding and reduced compounding (to be discussed further in section

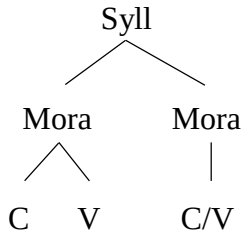
2). While some work has been done on blends in Japanese (Kubozono 1990), none has focused on accent placement as of yet.

In order to understand the proposed analysis, some of the assumptions about prosodic structure in Japanese will need to be explained. Firstly, the syllable structure shown in (2) may be familiar as the traditional syllable structure. This structure contains both an onset and a rhyme and groups together the syllable nucleus and coda as constituents. However, Kubozono (1989) states that there is no evidence for such a syllable constituent in Japanese. Instead, he claims that evidence from speech errors indicates that there is a "cohesiveness" between the onset and the nucleus and he argues for a syllable structure like that in (3). This structure contains two moras, the first of which groups together the onset and the coda.

(2) Traditional syllable model



(3) Japanese syllable model



Kubozono (1999) further elaborates on the importance of both the mora and the syllable in Japanese. He argues that the mora is an essential unit in the language since it is used in a number of ways including as a unit for temporal regulation, segmentation, and measuring phonological distance. He also argues that Japanese cannot just be considered a "mora language" and that the syllable unit is also critical for explaining certain phenomena such as accent placement and minimal word formation rules. As the linguistic evidence points to a syllable structure like the one in (3), this will be the assumed structure for this thesis.

Accent in Japanese is characterized by a pitch fall and it is the mora preceding this fall that carries the accent of a word (Kawahara 2013). The Tokyo dialect has a contrastive pitch accent system and minimal pairs may be distinguished by accent alone. Accent is culminative, occurring only once per word, but is not obligatory. As a result of this, a large number of words in Japanese are unaccented. Some examples of contrastive pitch accent are given in (4) below.

(4) Examples of contrastive pitch accent in Japanese

Initial:	Final:	Unaccented:
há.si(ga)	ha.sí(ga)	ha.si(ga)¯
'chopstick'	'bridge'	'edge'

As these examples demonstrate, accent may fall on any syllable in the word[1] or the word may be unaccented. While there does appear to be a default accentual pattern for Japanese that will be discussed further in the next section, words with alternative accent placements are considered to have this accent specified in the lexicon. Accent will be represented in this paper with an acute accent on the vowel of the relevant mora and a ' ¯ ' mark at the end of the word to represent unaccentedness.

In the next section, I present several different possibilities for the analysis of accent in Japanese blends. In 2.1 I discuss the possibility that accent will model that of simplex nouns in the language. In sections 2.2 and 2.3 I consider what the pattern would look like if blend accent was similar to other word formation processes in Japanese, namely reduced compounds and compounds. In sections 2.4 and 2.5 I examine analyses for linear order effects and head faithfulness—analyses that have been shown to be relevant for blends in English in the past. In section 3 I discuss a corpus study conducted to obtain accent judgments for Japanese blends

[1]Some exceptions exist. For words with five or more moras, initial and final accent are extremely rare (Kubozono 2008).

and evaluate the response conformity for each analysis discussed in the previous section. Finally, in section 4 I conclude.

2. PREDICTIONS

2.1 Simplex Accent Pattern

The first logical hypothesis for accent placement in Japanese blends is that it will follow the same patterns as the simplex nouns of the language. Under this hypothesis (the 'Simplex hypothesis') the accent of the two source words has no influence over the accent of the blend. Similarly, this assumes that, should internal morpheme boundaries exist, they have no effect on accent. To better understand the predictions of this hypothesis we must first discuss the existing the accentual patterns for simplex nouns in Japanese. In this section I will discuss both the default accent of Japanese and the pattern of unaccentedness.

While the accentual system at first appeared to be quite complicated, recent work has shown that it does largely conform to certain predictable patterns (Kubozono 2006, 2008; Ito & Mester 2012). Kubozono (2008) argues that accent in simplex nouns generally follows a default antepenultimate pattern. Nouns that conform to this pattern place accent on the syllable containing the antepenultimate mora. Some examples of this are shown in (5).

(5) Antepenultimate accent (examples from Kubozono 2008)

a. ku.ri.sú.ma.su 'Christmas'	d. yoo.róp.pa 'Europe'
b. su.tó.re.su 'stress'	e. pai.náp.pu.ru 'pineapple'
c. san.do.ít.ti 'sandwich'	f. ba.do.mín.ton 'badminton'

However, this is not an exceptionless pattern and many nouns do not have antepenultimate accent. The role of the default in Japanese is to provide an accent placement for nouns that do not already have lexically specified accent. For instance, nonce words and sequences of sounds

such as *kakikúkeko* (k-column in the syllabary) consistently follow this antepenultimate pattern (Ito & Mester 2012). A large number of the words that conform to this pattern are also loanwords. According to Kubozono (2008) 96% of accented three-mora loanwords have antepenultimate accent as compared to only 59% of native words. This can be explained since loanwords are less likely to have lexically specified accent.

Ito & Mester (2012) account for the antepenultimate pattern in Optimality Theory (Prince & Smolensky 1993/2004) with three high-ranking constraints: NONFINALITY, RIGHTMOST, and NOLAPSE (collectively termed the ANTEPENULT constraints for the remainder of this thesis). Satisfaction of these three constraints as a group occurs when the foot that contains the accent is one syllable away from the end of the word. This is demonstrated in the tableau in (7).

(6) ANTEPENULT Constraints (Ito & Mester 2012)

NONFIN(ALITY):	assign a violation when accent falls on the word-final prosodic unit
RIGHT(MOST):	assign a violation when a foot falls between the accented foot and the rightmost edge of the prosodic word
NOLAPSE:	assign a violation when there are two consecutive low-toned daughters of PrWd

Additional Constraints:

PARSE- σ:	assign a violation for each syllable that is not contained by a foot
FTFORM:	assign a violation when accent in a foot does not fall on the first Syllable

(7)

		ANTEPENULT			
/baruserona/ 'Barcelona'	FTFORM	NOLAPSE	RIGHT	NONFIN	PARSE-σ
☞ a. (baru)(séro)na					*
b. ba(ruse)(róna)				*!	*
c. ba(rúse)(rona)			*!		*
d. (báru)(sero)na			*!		*
e. ba(rúse)rona		*!			***
f. (baru)(seró)na	*!				*

As this tableau shows, NONFINALITY prevents accent from being too close to the right edge of the word. Candidate b) with penultimate accent loses because it violates this constraint by placing accent on the word-final foot. On the other hand, the RIGHTMOST and NOLAPSE constraints prevent accent from being too far from the right edge of the word. Candidates c) and d) violate RIGHTMOST by placing accent on the non-rightmost foot. Since tone on moras following accent in Japanese is always low, candidate e) violates NOLAPSE by having two low-toned, unparsed syllables. It should also be noted that these ANTEPENULT constraints do not actually force the accent to be antepenultimate. An additional high-ranking constraint, termed FTFORM, is necessary to prevent candidates where the accent is on the second syllable of the head foot, such as like f), from winning. With the combination of these constraints, candidate a) with antepenultimate accent emerges as the winner.

While Kubozono's (2008) descriptive generalization accounts for the accent placement of a large number of words in Japanese, several studies have noted another pattern of unaccentedness (Kubozono 2006; Ito & Mester 2012). As Kubozono (2006) discusses, four-mora words are much more likely to be unaccented than words of other lengths. This is especially true of four-mora words ending in two light syllables.

(8) Unaccented words with four moras (from Kawahara 2013)

a. a.me.ri.ka¯ ‘America’
b. i.ta.ri.a¯ ‘Italy’
c. kon.so.me¯ ‘consommé’
d. an.te.na¯ ‘antenna’

Ito & Mester (2012) attempt to explain this pattern of unaccentedness by proposing the ranking of the INITIALFT and ANTEPENULT constraints over WORDPROM, a constraint that is violated when a word does not have an accent. Crucially, the highly-ranked constraint INITIALFT causes four-mora words to be exhaustively footed. Any possible accent placement would then violate the ANTEPENULT constraints which are ranked higher than WORDPROM. This is demonstrated in the tableau in (10).

(9) INITIALFT : assign a violation when a prosodic word does not begin with a foot
WORDPROM: assign a violation when a prosodic word does not contain a prominence peak

(10) Four-mora word ending in two light syllables (Ito & Mester 2012)

		ANTEPENULT				
/amerika/ ‘America’	INITIALFT	NOLAPSE	NONFIN	RIGHT	WORDPROM	PARSE- σ
☞a. (ame)(rika)¯					*	
b. (áme)(rika)			*!	*!		
c. (ame)(ríka)			*!			
d. (áme)rika		*!				**
e. a(méri)ka	*!					**

In this tableau, the accented candidates b), c), and d) violate RIGHTMOST, NONFINALITY, and INITIALFT respectively while the unaccented candidate a) only violates the low-ranking WORDPROM constraint. It is also important to note that this constraint ranking does not result in unaccentedness for three or five-mora words since satisfaction of INITIALFT will not cause violations of the ANTEPENULT constraints.

Crucially, this ranking does not incorrectly cause four mora words not ending in two light syllables to become unaccented. These words are not exhaustively footed and therefore an accented candidate does not violate any of the ANTEPENULT constraints. An example of this is shown in the tableau below.

(11) Four mora word not ending in two light syllables (Ito & Mester 2012)

		ANTEPENULT				
/takusii/ 'taxi'	INITIALFT	NOLAPSE	NONFIN	RIGHT	WORDPROM	PARSE- σ
☞a. (tá.ku)sii						*
b. (ta.ku)(sii)					*!	
c. (tá.ku)(sii)				*!		
d. (ta.ku)(síi)			*!			
e. ta.kú(sii)	*!					**

In this tableau, candidate b) which is exhaustively footed and unaccented loses to candidate a) which is accented. Since the final syllable of the candidate a) is heavy, accentedness does not lead to a violation of the NOLAPSE constraint. Similarly, as this candidate does not foot the final syllable, it does not violate the NONFIN constraint. It is also important to note that the specific prediction of initial accent for words of this form ([LLH]) differs from the prediction of the antepenultimate rule. However, for the sake of this thesis we will consider this prediction to be the correct one.

In summary, if blends conform to the predictions of the Simplex hypothesis, then they will be unaccented if they are four moras long and end in a sequence of light syllables. Otherwise, they will follow the default antepenultimate pattern. If this hypothesis is shown to be correct, then it will indicate that blending is unique from other word formation processes in Japanese that have accent influenced by other morphological factors. It would strongly indicate that

blends are analyzed as simplex rather than morphologically complex. In the next few sections I will discuss alternative hypotheses that involve patterns from other word formation processes.

2.2 Reduced Compounds

The next hypothesis that I will consider is that blends will follow the accentual pattern of reduced compounds, a common word formation process in Japanese that is very similar to blending. These reduced (or 'truncated') compounds are like blends in that they take segments from two source words (rather than the whole words as in regular compounds) and combine them in order to construct a new word. Reduced compounds are crucially distinct from blends in that they never involve overlap and are formed when the *beginnings* of each source word are concatenated. Examples demonstrating the difference between these two types of words are shown in (12).

(12)

REDUCED COMPOUND	BLEND
(deji)taru 'digital' + (kame)ra 'camera'	(pera)pera 'fluent' + pu(raido) 'pride'
dejikame 'digital camera'	peraido 'pride in being fluent'

What does the accentual pattern of reduced compounds look like? While it has not been seriously addressed until recently, at least one study has noted that they tend to be unaccented regardless of their length (Ito & Mester 1992). The hypothesis that blends will follow this descriptive generalization (termed 'Reduced Compounds I') predicts that they should all be unaccented. This is a simple prediction that will be very easy to test.

However, if we want to argue that blends actually have the same internal morphological structure as reduced compounds then it is useful to consider a formal analysis of reduced compound unaccentedness. This hypothesis will be termed the 'Reduced Compounds II' hypothesis. In their recent study, Ito & Mester (2012) also attempt to account for this pattern

in OT. They propose a highly ranked LEXFT constraint which requires every lexical morpheme to minimally project its own foot. Since each source word in a reduced compound always contributes either one or two moras of material, reduced compounds will always be exhaustively footed. This causes them to become unaccented in the same way as the four-mora words discussed in the previous section. An example of this is shown in the tableau in (14).

(13) LEXFT: assign a violation for every lexical morpheme that does not minimally projects its own foot (Ito & Mester 2012) (to be redefined below)

(14) Three-mora reduced compound (Ito & Mester 2012)

/seku-me/ 'section mate'	LEXFT	ANTEPENULT	WORDPROM	PARSE- σ
☞ a. (se.ku)(me) ¯			*	
b. (sé.ku)(me)		*! W	L	
c. (sé.ku)me	*! W		L	* W

In the tableau above, the accented candidate c) violates the high-ranking LEXFT constraint. An attempt to satisfy both LEXFT and WORDPROM results in a violation of the ANTEPENULT constraints as in candidate b). Therefore, the unaccented candidate a) emerges as the winner despite the fact that it is only three moras long and would likely be accented if it were simplex.

In order to understand what this analysis would look like for blends we must first define what a lexical morpheme is for blends. Specifically, we need to determine which segments we should consider part of what morphemes in each blend and where the morpheme boundaries are. This is a slightly more complicated question than it may at first appear because of the existence of overlapping segments in some blends. Consider the examples in 0.

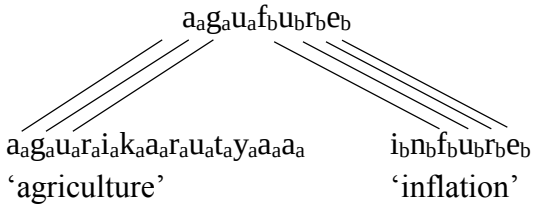
(15) Blends with varying overlap

a. *k**o**tona 'child-adult' (*ko)domo* 'child' + (otona) 'adult'
b. *h**omo**dati 'gay friend' (*homo)* 'gay' + t(omodati) 'friend'
c. *gu**suri**ipu 'good sleep' (*gussuri*) 'sound asleep' + (suriipu) 'sleep'

Each of these examples contains one or more segments of overlap. It is not immediately clear where the morpheme boundaries are and it is entirely possible to imagine the overlapping segments belonging to either or both of the morphemes in these blends.

As a solution to this problem, we will make reference to the principle of *Consistency of Exponence*, first proposed by McCarthy & Prince (1993, 1994). According to this principle, a restriction is placed on the function Gen so that it cannot generate forms where the lexical specifications (including segments) of a morpheme have been changed. For blends, then, we would expect that each segment taken from a particular source word will still inherit the lexical specification of that source word. An example representation is provided below in (16) where morpheme specification is indicated by indices.

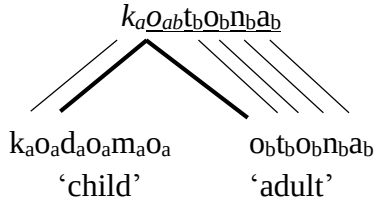
(16) Example representation of blend morpheme specification (without overlap)

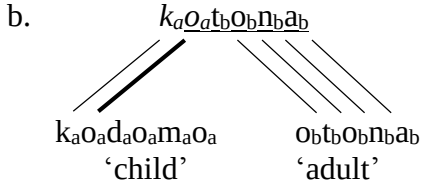


In this diagram, the segments taken from the first source word are part of one morpheme and the segments taken from the second source word are part of another morpheme. Thus, in blends such as this with no overlap, there are two clear and separate morphemes.

For blends with overlap, the picture is slightly more complicated. Some possible representations of such a blend are shown in (17).

(17) Possible representations of blend morpheme specification (with overlap)

a. 

b. 

In a), the overlapping segment, [o], is taken from each source word and crucially has the lexical specification of *both* morphemes. An alternative representation is shown in b) where the segment is only taken from one source word. While this type of representation is indeed possible, it is considered unlikely due to the prevalence of overlap in blends. It is likely that this overlap is desirable in part because of fact that it results in more segments being preserved from each word. Therefore, we would expect that an analysis like a) would be preferred since it preserves more segments. For this analysis we will assume a representation like that of a).

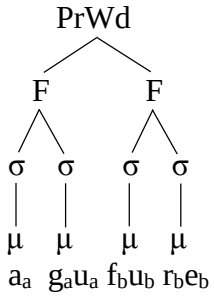
Now that we have a decent understanding of how the segments of the morphemes are determined, the next question to be addressed is how a high-ranking LEXFT constraint will affect accent placement in certain blend types. Due to the differences in segmentation between blends and reduced compounds, it is necessary to somewhat revise Ito & Mester’s (2012) definition of LEXFT to be more explicit. This new definition is given in (18).

(18) LEXFT: for any two morphemes x and y where some segments in morpheme x occur before those in morpheme y, assign a violation if there does not exist a foot that contains segments from morpheme x that occurs before a foot that contains segments from morpheme y.

As this definition specifies, each morpheme must project a separate foot but the segments in this foot need not all be from the same morpheme. Alternatives to this definition are certainly possible, but in order to restrict the range of possibilities we will only consider this one.

To see how this constraint operates for blends, let us examine a few examples. For blends where each morpheme is one or two moras the analysis is clear. An example of the prosodic structure of such a blend is given in (19).

(19) Prosodic structure of a blend (with only two moras in each morpheme)



agufure ‘agriculture inflation’ = agurikarutyaa ‘agriculture’ + infure ‘inflation’

This structure satisfies LEXFT by having each lexical morpheme minimally project its own foot. A tableau showing the predicted winning accent placement of this blend is given below.

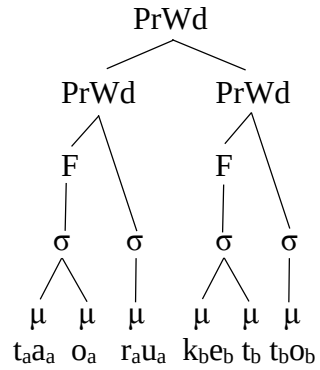
(20)

/ $a_a g_a u_a r_a i_a k_a á_a r_a u_a t_a y_a a_a a_a$+ $i_b n_b f_b u_b r_b e_b$¯/ ‘agriculture inflation’	LEXFT	ANTEPENULT	WORDPROM
☞ a. $(a_a g_a u_a)(f_b u_b r_b e_b)$¯			*
b. $(á_a g_a u_a)(f_b u_b r_b e_b)$		*!	
c. $(á_a g_a u_a) f_b u_b r_b e_b$	*!		

As this tableau demonstrates, the winning candidate is the unaccented one, as it is for reduced compounds. This is because under footing that satisfies LEXFT, any accent placement that is chosen would then violate the ANTEPENULT constraints. As this example illustrates, we then predict that any blend where each morpheme is one or two moras long should be unaccented.

For blends where the source word contributions are longer than two moras, the result is somewhat different. The predicted prosodic structure of one such example is given in (21) below.

(21) Prosodic structure of a blend (with more than 2 moras in each morpheme)



toruketto ‘towel-blanket’ = taoru ‘towel’ + buranketto ‘blanket’

For this blend, the separate morphemes once again project one foot each. However, it is important to note that they are not exhaustively footed. The tableau for this blend is given below.

(22)

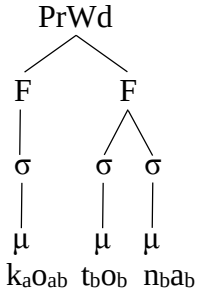
/t_aá$_a o_a r_a u_a$ + $b_b u_b r_b a_b n_b k_b$é$_b t_b t_b o_b$/ ‘towel-blanket’	LEXFT	ANTEPENULT	WORDPROM
☞ a. ($t_a a_a o_a$)$r_a u_a$(k_bé$_b t_b$)$t_b o_b$			
b. ($t_a a_a o_a$)$r_a u_a$($k_b e_b t_b$)$t_b o_b$ ¯			*! W

As this tableau demonstrates, the winning candidate is actually not the unaccented one, but is instead the candidate with antepenultimate accent. This is because the footing structure allows an accent placement that does not violate ANTEPENULT. Under the Reduced Compounds II hypothesis, then, we expect blends to have antepenultimate accent if the second morpheme is more than two moras long and where the first morpheme can form a foot without crossing a boundary (more on this later). It is important to note that this prediction differs from the surface pattern of unaccentedness. If blends are like reduced compounds in that they are subject to the LEXFT constraint, then we will see this difference being borne out.

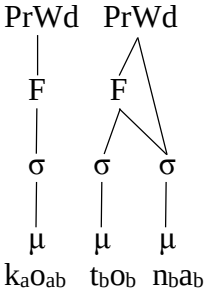
Where at first the predictions for blends with overlap appeared to be more complicated, we can see that with the addition of the consistency of exponence, the predictions are much like those for blends without overlap. The important point is that to satisfy LEXFT, each morpheme (whether some of those segments are actually part of more than one morpheme or not) must project its own foot. An example of a proposed prosodic structure for a blend with overlap is given below.

(23) Possible prosodic structures for a blend (with overlap)

a.



b. *Violates proper bracketing



In a) above, the first two segments which form a morpheme have projected their own foot despite the fact that one of the segments in the first foot also belongs to the second morpheme. Thus, in this structure LEXFT is not violated. An alternative prosodic structure is presented in b) where the overlapped segment is included in the feet from both morphemes. Such a structure is rejected, however, because it violates proper bracketing. No one syllable can be part of two feet. The tableau for this blend is shown in (25) where the unaccented candidate is the winner as a result of the exhaustive footing. BRACKETING is shown here as a violable constraint, although it may in fact be a restriction on Gen.

(24) BRACKETING: assign a violation each time a syllable is part of more than one foot.

(25)

/$k_a o_a d_a o_a m_a o_a$‾ + $o_b t_b o_b n_b a_b$‾/ ‘Child-adult’	BRACKETING	LEXFT	ANTEPENULT	WORDPROM
☞a. ($k_a o_{ab}$)($t_b o_b n_b a_b$)‾				*
b. (k_aó$_{ab} t_b o_b$)$n_b a_b$		*!		
c. (k_a($ó_{ab}$)$t_b o_b$)$n_b a_b$	*!			

While the example above merely illustrated an example where a foot contained a segment that was part of two morphemes, there are also blends for which, in order to satisfy the LEXFT constraint, the foot must contain segments that are not a part of the morpheme that is projecting it. However, this should be allowed under the revised definition of LEXFT as long as each morpheme projects its own foot. Some example blends where satisfying LEXFT would result in this are given in (26) below.

(26) Satisfies LEXFT Violates LEXFT

a. ($m_a u_b k_b$)($k_b u_b$)‾ vs. *m_aú$_b k_b$($k_b u_b$) $m_a a_a g_a a_a z_a i_a n_a$ ‘magazine’+$b_b u_b k_b k_b u_b$ ‘book’

b. $r_a e_a$($t_a a_{ab} k_b$)($k_b u_b s_b u_b$)‾ vs. *$r_a e_a t_a$á$_{ab} k_b$ ($k_b u_b s_b u_b$) $r_a e_a t_a a_a a_a$ ‘letter’+$f_b a_b k_b k_b u_b s_b u_b$ ‘fax’

c. ($f_a a_b$)$b_b u_b$(r_bé$_b t_b$)$t_b o_b$ vs. *$f_a a_b b_b u_b$(r_bé$_b t_b$)$t_b o_b$ $f_a o_a n_a$ ‘phone’+$t_b a_b b_b u_b r_b e_b t_b t_b o_b$ ‘tablet’

In the first two examples, satisfaction of LEXFT leads to unaccentedness. This is not always the case, however, and as c) shows, if the length and structure of the blend allows then antepenultimate accent will be predicted.

In summary, an investigation of whether accent placement in blends is like that of reduced compounds will give us insight into the internal morphological structure of Japanese blends. If we find that they are like reduced compounds then this would indicate that they are indeed constructed with segments from separate source words being part of two different morphemes, each of which projects its own foot. This would also provide further evidence for the influence of a LEXFT constraint in Japanese. Since Ito & Mester (2012) also attempt to use this constraint to account for the prevalence of unaccentedness in native words (as they are more likely to be morphologically complex), confirmation of

the utility of this constraint could have consequences for the analysis of accent in all Japanese nouns and not just blends. In the next section I consider an analysis for accent in blends based off of another word formation process—that of compounding.

2.3 Compound Accent

Another possibility for accent placement in Japanese blends is that they will follow the accent pattern of compounds. Compounding is yet another type of word formation process that involves the combination of two source words. However, unlike blends they do not truncate or overlap any material. Examples of each of these two types of words are given in (27) below.

(27)

COMPOUND	BLEND
(maneki) 'beckoning' + (neko) 'cat'	(pera)pera 'fluent' + pu(raido) 'pride'
maneki-neko 'lucky cat'	peraido 'pride in being fluent'

While the accentual pattern of compound nouns in Japanese is one that has been well studied, the generalizations provided in the literature are somewhat contradictory. In the interest of accounting for all possibilities, I will discuss both of the generalizations in this section.

The first account that I will discuss was put forward by Kubozono (1997, 2008). While older accounts of the phenomena treated compounds with long versus short second members separately (Akinaga 1985; McCawley 1968; Poser 1990), Kubozono (2008) argues that their accent patterns can be described together. As he describes, compounds in Japanese keep the accent of the second word *unless* it is on the last syllable. If the accent is on the last syllable (or if the word is unaccented) then the compound will accent the rightmost, nonfinal foot. Within that foot it will accent the "syllable that is closer to the

word-internal morpheme boundary." The hypothesis that blends will follow this descriptive generalization will be termed the 'Compound Kubozono' hypothesis.

Kubozono (2008) proposes that this generalization can be accounted for in OT with the constraints MAX(ACC), and ALIGN in addition to the previous ANTEPENULT constraints. Since Kubozono (2008) does not provide tableaus, I have attempted to account for his generalizations in my own tableaus below. I use FAITH(ACC)-RT to include both Kubozono's proposed MAX(ACC) constraint as well as a NOFLOP(ACC) constraint (but crucially not a DEP(ACC) constraint since this would prevent compounds with unaccented source words from becoming accented). I have also added a NONFINALITY-SYL constraint which he mentions in his paper, but does not include in his proposal. The tableaus shown below in (29) through (32) motivate the ranking provided in (33).

(28) NONFIN(ALITY)-SYL: assign a violation when accent falls on a word-final syllable
ALIGN: assign a violation when accent does not fall on a syllable at a morpheme boundary

(29) ALIGN >> FTFORM

/hirosima + sí/ 'Hirosima city'	NONFIN-SYL	FAITH(ACC)-RT	ANTE	ALIGN	FTFORM
☞a. hiro(simá)-si		*			*
b. hiro(síma)-si		*		*! W	L
c. hirosi(má-si)		*	*! W		L
c. hirosima-(sí)	*! W	L	* W		L

In the tableau above, the second word is only one mora long and thus the internal morpheme boundary is very close to the end of the compound. Candidates which place accent in the rightmost, non-final foot win over the alternatives. However, candidate b) with antepenultimate accent crucially violates the ALIGN constraint and loses to candidate

a) which places accent closer to the morpheme boundary. This indicates that ALIGN outranks FTFORM.

(30) ANTEPENULT >> ALIGN

/nyúu + karedonia¯/ 'New Caledonia'	NONFIN-SYL	FAITH(ACC)-RT	ANTE	ALIGN	FTFORM
☞a. nyuu-kare(dóni)a				*	
☞b. nyuu-(kare)(doni)a¯				*	
c. nyuu-(káre)(doni)a			*! W	L	

Next, the second word in the compound above is five moras long, placing the internal morpheme boundary of the compound very far from the right edge. In this case, the candidate that places the accent at the boundary (thus satisfying ALIGN) ends up violating the RIGHTMOST constraint from ANTEPENULT. The candidates a), which contains antepenultimate accent, and b), which is unaccented, are both possible winners (as stated by Kubozono 2008). Thus, the ANTEPENULT constraints outranks ALIGN.

(31) FAITH(ACC)-RT >> ANTEPENULT

/maneki + néko/ 'Lucky cat'	NONFIN-SYL	FAITH(ACC)-RT	ANTE	ALIGN	FTFORM
☞ a. maneki-(néko)			* NONFIN		
b. mane(kí-ne)ko		*! W	L		

In the tableau above, the second word in the compound contains penultimate accent. As is consistent with the Kubozono (2008) analysis, candidate a) which preserves the accent of the second word (and satisfies FAITH(ACC)-RT) is the winner over candidate b) which satisfies the ANTEPENULT constraints. Therefore, FAITH(ACC)-RT outranks ANTEPENULT.

(32) NONFINALITY-SYL >> FAITH(ACC)-RT

/ten.zyoo + kawá/ ‘raised-bed river’	NONFIN-SYL	FAITH(ACC)-RT	ANTE	ALIGN	FTFORM
☞a. ten(zyóo)-kawa		*			
b. tenzyoo-(káwa)		*	*! W		
c. tenzyoo-(kawá)	*! W	L	* W	*W	*W

Lastly, the tableau above contains a compound whose second word has final accent. In this case, candidate c), which preserves this accent, violates the NONFIN-SYL constraint and loses to candidate a), which does not preserve the accent. Thus, NONFIN-SYL outranks FAITH(ACC)-RT. A summary of the proposed ranking for Kubozono (2008) is given below.

(33) Proposed ranking for analysis in Kubozono (2008)
NONFINALITY-SYL >> FAITH(ACC)-RT >> ANTEPENULT >> ALIGN >> FTFORM

The most important aspect to consider here is that the FAITH(ACC)-RT constraint is ranked higher than the ALIGN constraint.

The second descriptive generalization that I will discuss is provided by Ito & Mester (2012). According to their account, the accent (or unaccentedness) of the second member is preserved when the second member is more than four moras long. If it contains four or fewer moras, then the accent falls on the syllable in the non-final foot that is closest to the morpheme boundary. Some examples of this accent pattern are given below in (34).

(34) Examples of compound accent in Ito & Mester (2012)

a.	inú	akitá-(inu)	‘Akita dog’
b.	mushi¯	kaabutó-(musi)	‘beetle’
c.	sakana¯	nama-(záka)na	‘raw fish’
d.	kamisóri	denki-(kámi)(sori)	‘electric razor’
e.	kariforunia¯	minami-kariforunia¯	‘Southern California’
f.	kaoáwase	hatsu-kao(áwa)se	‘first face-to-face encounter’

Some of the examples given by Ito & Mester (2012) are blatantly contradictory to the generalizations made by Kubozono (2008), as shown in (35). The authors of both papers

do acknowledge that there are cases of compounds that follow alternative accent patterns (which incidentally appear to be the same patterns proposed by each other's papers). However, both state that these are minority patterns and that their rules should account for the majority of compounds. Since an investigation into which generalization truly does account for more compounds is outside of the scope of this thesis, I will simply include both analyses.

(35) Contradictory predictions for compound accent

		Ito & Mester (2012)	Kubozono (2008)	
a.	éki	kyootó-eki	kyooto-éki	'Kyoto station'
b.	kokóro	onna-gókoro	onna-gokóro	'a woman's heart'
c.	saibansyó	chihoo-saibansyó	chihoo-saibánsyo	'district court'

In terms of constraints, Ito & Mester (2012) propose that compounds must follow the junctural accent requirement (assigned by an ALIGN constraint). However, if the second word is too long (as in greater than four-mora words) and violates the NOLAPSE(TONE) constraint, then it will preserve the accent structure of the second word (FAITH(ACC)-RT). It is also important to note that this FAITH constraint includes DEP as well as MAX and NOFLOP, although both analyses have DEP being low-ranked. The tableaus in (36) and (37) motivate the ranking provided in (38).

(36) ALIGN >> FAITH(ACC)-RT

/onna + kokóro/ 'a woman's heart'	NOLAPSE(TONE)	ALIGN	FAITH(ACC)-RT
☞a. onna-(góko)ro			*
b. onna-go(kóro)		*!W	L

As the tableau above demonstrates, when the second word is short (less than five moras), the candidate that violates the ALIGN constraint loses even if it satisfies the FAITH(ACC)-RT constraint. This motivates the ranking of ALIGN >>FAITH(ACC)-RT.

(37) NoLapse(tone) >> Align

/chihoo + saibansyó/ ‘district court’	NoLapse(tone)	Align	Faith(acc)-RT
☞a. chihoo-saibansyó		*	
b. chihoo-sai(bán)syo		*	*!W
c. chihoo-(sái)bansyo	*!W	L	*W

On the other hand, when the second word is longer than four moras, as is the case in the tableau above, the candidate that satisfies the Align constraint ends up violating the NoLapse(tone) constraint. This causes the faithful candidate a) to be the winner. Thus NoLapse(tone) outranks Align. A summary of the proposed ranking is given below.

(38) Proposed ranking for analysis derived from Ito & Mester (2007)
NoLapse(tone) >> Align >> Faith(acc)-RT

The main difference between their two accounts appears to be a difference in the ranking of the Faith constraints (specifically Max and NoFlop) and the Align constraint. On the one hand, Kubozono (2008) claims that the accent of the second word will be preserved unless it is on the last syllable. On the other hand, Ito & Mester (2012) claim that the accent is only preserved if the second word is five moras or longer. Since it is impossible to determine which account is correct without obtaining accent judgments for the compounds, I will consider the implications of both analyses for accent placement in blends.

To understand the predictions of these accounts for blends we must understand where the morpheme boundaries are and how this affects the Align constraint. To do this, we can make reference to the definition of a morpheme for blends given in the previous section. Recall that under consistency of exponence, the OT candidates must not make changes to the lexical specification of the morpheme segments. We propose, then, that there is a

morpheme boundary at any location where one morpheme ends. Some examples of morpheme boundaries in blends are given below in (39).

(39) Examples of morpheme boundaries in blends

a. $a_a g_a u_a | f_b u_b r_b e_b$ = $a_a g_a u_a r_a i_a k_a a_a r_a u_a t_a y_a a_a a_a$ + $i_b n_b f_b u_b r_b e_b$

b. $p_a o_a t_a e_a | t_{ab} o_{ab} | r_b u_b n_b e_b e_b d_b o_b$ = $p_a o_a t_a a_a t_a o_a$ + $t_b o_b r_b u_b n_b e_b e_b d_b o_b$

In the case of blends without overlap, as in a), it is clear where the morpheme boundary should be. Thus, ALIGN will be satisfied with an accent placement on either side of that boundary. For blends with overlap, as in b), this means that there are two word-internal morpheme boundaries despite the fact that there are only two morphemes in the word. Specifically, the boundary precedes the leftmost segment in the second morpheme and follows the rightmost segment of the first morpheme. With this boundary placement, ALIGN will be satisfied if the accent falls on any of the three moras straddling the boundaries.

Now that we have an understanding of the morpheme boundaries for blends, we will examine the predictions of each account. First, we consider the predictions of the Kubozono (2008) account for compounds. Under this account, we expect blends to always follow the accent of the second source word unless it is on the final syllable or the second source word is unaccented. Thus, as the tableau below demonstrates, the prediction for those blends with non-final accent is a straightforward one. The candidate that preserves the accent of the second source word will always be the winner.

(40)

/$y_a a_a k_a i_a$ + $t_b o_b o_b m_b ó_b r_b o_b k_b o_b s_b i_b$/ (a corn snack)	NONFIN-SYL	FAITH(ACC)-RT	ANTE	ALIGN	FTFORM
☞ a. $y_a a_a k_a i_a$ ($m_b ó_b r_b o_b$) ($k_b o_b s_b i_b$)			*		
b. $y_a a_a k_a i_a m_b o_b$ ($r_b ó_b k_b o_b$) $s_b i_b$		*!		*	

For those blends with second source words that have final accent or are unaccented, the accent placement will almost always be on the syllable containing the antepenultimate mora. The only situations where the low-ranked ALIGN would influence the placement of the accent would be when the segments from the second source word only constitute one mora. Due to other factors involved in blend construction, the only blends that have one mora from the second source word are those that are two syllables long. In those blends, ALIGN will cause the accent to fall on the first mora, as shown in (41).

(41)

/$o_a o_a m_a o_a r_a i_a^{\_}$ + $k_b a_b m_b a_b t_b a_b^{\_}$/ 'Ota' (placename)	NONFIN-SYL	FAITH(ACC)-RT	ANTE	ALIGN	FTFORM
☞a. $ó_a o_a t_b a_b$					
b. $o_a o_a t_b a_b$				*!	

Under the Ito & Mester (2012) account, however, the ALIGN constraint plays a more central role. Here, blends must always place accent on a morpheme boundary unless that would cause the candidate to violate the NOLAPSE(TONE) constraint. As blends are highly unlikely to have five moras of material contributed from the second source word alone (there are no such blends in the database), NOLAPSE(TONE) only comes into play if the second source word contributes at least four moras ending in a heavy syllable followed by a light syllable (…HL]), as shown in the tableau below.

(42)

/$p_a o_a t_a é_a t_a o_a$ + $t_b o_b r_b u_b n_b é_b e_b d_b o_b$/ (a potato snack)	NOLAPSE(TONE)	ALIGN	FAITH(ACCENT)-RT
☞ a. $p_a o_a t_a e_a t_{ab} o_{ab} r_b u_b (n_b é_b e_b) d_b o_b$		*	
b. $p_a o_a t_a e_a t_{ab} o_{ab} r_b ú_b (n_b e_b e_b) d_b o_b$	*!		*

In this tableau, the accent placement that satisfies ALIGN causes there to be two daughters of the phonological word after the accent thus violating NOLAPSE(TONE). This causes the winner to be the candidate with faithful accent placement.

For all other blends, the winning candidate will be one that satisfies the ALIGN constraint. Interestingly, the effect of overlap on ALIGN is that there are simply more accent placements that would satisfy ALIGN. Under this account, blends are highly likely to have antepenultimate accent. This is, of course, not the case for all blends. Some, like the one in (43), are predicted to have accent fall on the mora closer to the boundary but not necessarily in the antepenultimate position.

(43)

/$y_a a_a k_a i_a$ + $t_b o_b o_b m_b ó_b r_b o_b k_b o_b s_b i_b$/ (a corn snack)	NOLAPSE(TONE)	ALIGN	FAITH(ACC)-RT
☞ a. $y_a a_a k_a i_a$ ($m_b ó_b r_b o_b$) ($k_b o_b s_b i_b$)			
b. $y_a a_a k_a i_a m_b o_b$($r_b ó_b k_b o_b$)$s_b i_b$		*!	*

In summary, an investigation into whether blend accent placement is similar to that of compound accent will also reveal information about the internal morphological structure of a blend. If accent is like that of the Ito & Mester (2012) account, then it will show that morpheme boundaries do exist in blends and that the ALIGN constraint is relevant for accent placement. On the other hand, if accent is like the Kubozono (2008) account, then it will show that the formation of blends is highly similar to that of compounds and that faithfulness to the right hand source word is an important factor. In the next two sections I will consider hypotheses than have been shown to be relevant for blends in previous research.

2.4 Linear Order

A slightly more straightforward hypothesis for the accent placement of blends is that they will be faithful to the accent of the right-hand source word. This type of 'linear order' claim has been made in several studies on English blends (Bat-El & Cohen 2012, Gries 2004a,b) as well as one on Japanese blends (Kubozono 1990). In Kubozono's (1990) study, he claimed that in Japanese and English the second source word determines the phonological length of the blend. He claimed that this was evidence for the Right-hand Head Rule (RHR) as proposed by Williams (1981), which states that the head of a morphologically complex word is the right-hand member. If the RHR is relevant for Japanese blends as Kubozono (1990) suggests, then it is entirely plausible that the accent of the blend could be determined by the second source word as well.

The predictions made by this hypothesis are very similar to the predictions of the Kubozono (2008) version of the compound accent hypothesis. One important difference is that according to this hypothesis, even if the second source word has final accent, then the blend will as well. The relevant constraints are defined in (44) and (45) below with faithfulness either to the right or left source word.

(44) MAX(ACCENT)-R: assign a violation for an accent in the right source word that does not have a correspondent in the blend

NOFLOP(ACCENT)-R: assign a violation for an accent in the right source word that where the corresponding accent in the blend has shifted

DEP(ACCENT)-R: assign a violation for an accent in the blend that does not have a correspondent in the right source word

(45) MAX(ACCENT)-L: assign a violation for an accent in the left source word that does not have a correspondent in the blend

NOFLOP(ACCENT)-L: assign a violation for an accent in the left source word that where the corresponding accent in the blend has shifted

DEP(ACCENT)-L: assign a violation for an accent in the blend that does not have a correspondent in the left source word

According to this prediction, all of the faithfulness constraints for the right source word will outrank the faithfulness constraints for the left source word. This is shown in the tableau in (46) where the right-faithful candidate is the clear winner.

(46)

/báka + ahó/ 'an idiot'	MAX-R	NOFLOP-R	DEP-R	MAX-L	NOFLOP-L	DEP-L
☞a. bahó				*		
b. báho		*!		*		
c. baho⁻	*!			*		

It is also important to note that this faithfulness to the accent of the right source word does not necessarily mean that the blend will preserve unaccentedness as well as accent placement. We have already seen in the Kubozono (2008) analysis that compounds are not faithful to unaccentedness and therefore it seems plausible that blends may not exhibit this faithfulness either. Thus, there is a strong and a weak version of this hypothesis. In the strong version, the DEP(ACCENT)-R constraint is highly ranked and prevents a new accent from being inserted. This is demonstrated in (47) below.

(47) Strong Version:

/erotíkku + burogu⁻/ 'erotic blog'	MAX-R	NOFLOP-R	DEP-R	ANTE/ ALIGN	MAX-L	NOFLOP-L	DEP-L
☞a. erogu⁻				*			
b. érogu			*!				

As the tableau above shows, this ranking results in the blend being faithful to even an unaccented right source word. In the weak version, the DEP(ACC)-R constraint is outranked

by another constraint, possibly ANTEPENULT or ALIGN, thus preventing the blend from remaining unaccented. This is shown in (48) below.

(48) Weak Version:

/erotíkku + burogu⁻/ 'erotic blog'	MAX-R	NOFLOP-R	ANTE/ ALIGN	DEP-R	MAX-L	NOFLOP-L	DEP-L
☞a. érogu				*			
b. erogu⁻			*!				

If the data is shown to be consistent with this hypothesis then it will provide further evidence that the Right-hand Head Rule is at work in Japanese and that it is relevant for blends. It would also show that Japanese is consistent with the past studies of blends in English that argue for a linear order effect (Bat-El & Cohen 2012, Gries 2004a,b). The behavior of the blend when the right source word is unaccented will also show us whether DEP(ACC) is also high-ranking. If the weak version is shown to be correct, then this could provide further evidence for the Kubozono (2008) version of compound analysis. In the next section, I will consider an alternative influencing factor for blends that has been proposed recently—head faithfulness.

2.5 Head Faithfulness

The last hypothesis for accent placement in Japanese blends that I will consider here is that they will be faithful to the accent of the head of the blend. This is consistent with the claim by Shaw (2013) who argued that blends are subject to head faithfulness effects. Shaw (2013) proposed this analysis in opposition to the linear order analyses discussed in the previous section. She argued that the reason a right-hand effect was found in those studies was that a majority of blends in English are right-headed despite. However, left-headed and coordinating blends exist as well and according to Shaw's findings, these blends

exhibit different behavior. While right-headed blends have their stress determined by the second source word, coordinating blends followed the default stress. Shaw's analysis is superior to the linear order accounts in that it does not need to make reference to concepts like "left" and "right" that are not privileged positions in other domains. Instead, it makes reference to the morphologically privileged position of the head (Revithiadou 1999, Roon 2006).

In Shaw's analysis, the head is the source word that either 1) determines the lexical category or 2) is the semantic head. A blend with source words that are the same lexical category may also be coordinating. Examples of each type of headedness can be found in Japanese, as shown in (49) below. This goes against the argument by Kubozono (1990) that all Japanese blends have right-hand heads.

(49) Examples of Headedness in Japanese
   a. Right-Headed
      i. *homo*dati 'homosexual friend' = *homo* 'homosexual' + tomodati 'friend'
      ii. *zya*bitto (Giant's mascot) = *zyaiantu* 'Giant' + rabitto 'rabbit'
   b. Coordinating
      i. *b*aho 'a fool' = *baka* 'idiot' + aho 'idiot'
      ii. *h*ine 'a hie-rice hybrid' = *hie* 'barnyard millet' + ine 'rice-plant'
   c. Left-Headed
      i. *pote*toruneedo 'a potato snack' = *poteto* 'potato' + toruneedo 'tornado'
      ii. *o*ttyen 'girl-like man' = *otoko* 'man' + mettyen (German for 'girl')

The relevant constraints for this analysis are given below. The constraints in (50) penalize any unfaithfulness to the accent of the source words while the constraints in (51) only penalize unfaithfulness to the head. It should also be noted that any violation of the constraints in (51) will lead to violations in the corresponding constraints in (50).

(50) MAX(ACCENT): assign a violation for every accent in the input that does not have a correspondent in the output

NOFLOP(ACCENT): assign a violation for every accent in the input where the corresponding accent in the output has been shifted

DEP(ACCENT): assign a violation for every accent in the output that does not have a corresponding accent in the input

(51) MAX(ACCENT)$_{HEAD}$: assign a violation for every accent in the input head that does not have a correspondent in the output

NOFLOP(ACCENT)$_{HEAD}$: assign a violation for every accent in the input head where the corresponding accent in the output has been shifted

DEP(ACCENT)$_{HEAD}$: assign a violation for every accent in the output that does not have a corresponding accent in the input head

The head faithfulness constraints in this account cannot be ranked in relation to the normal faithfulness constraints. However, blends that violate head faith will lose simply because they are violating more constraints. This is shown in the tableaus in (52) and (53).

(52)

/ása + samuké*Head*/ 'morning chills'	MAX$_{HEAD}$	NOFLOP$_{HEAD}$	DEP$_{HEAD}$	MAX	NOFLOP	DEP
☞a. asamuké				*		
b. asámuke		*!		*	*	
c. asamuke^{-}	*!			**		

(53)

/báka + ahó/ 'idiot'	MAX$_{HEAD}$	NOFLOP$_{HEAD}$	DEP$_{HEAD}$	MAX	NOFLOP	DEP
☞a. bahó				*		
☞b. báho				*		
c. baho^{-}				**!		

As (52) shows, if the blend is headed then the candidate that is faithful to the accent of the head becomes the winner. However, if the blend is coordinating, as in (53), then none of the head faithful constraints will be violated and there may be several possible winners. The actual winner would likely be decided by another constraint that is ranked below the head faith constraints: possibly ANTEPENULT or ALIGN.

Just as in the previous account, faithfulness to the accent of the head does not necessarily mean that the blend will be unaccented if the head is unaccented. Therefore, there are two different version of this hypothesis; a strong version and a weak version. In the strong version, the DEP(ACCENT)$_{HEAD}$ constraint is ranked above any other constraints like ANTEPENULT or ALIGN. As the tableau in (54) shows, this will cause the unaccented candidate to be the winner if the head is also unaccented. In the weak version of this hypothesis, the DEP(ACCENT)$_{HEAD}$ constraint is ranked below ANTEPENULT or ALIGN, which causes the candidate with accent to be the winner. This is shown in (55).

(54) Strong Version:

/erotíkku + burogu^{-}*Head*/ 'erotic blog'	MAX$_{HEAD}$	NOFLOP$_{HEAD}$	DEP$_{HEAD}$	ANTE/ ALIGN	MAX	NOFLOP	DEP
☞a. erogu^{-}				*	*		
b. érogu			*!				*

(55) Weak Version:

/erotíkku + burogu^{-}*Head*/ 'erotic blog'	MAX$_{HEAD}$	NOFLOP$_{HEAD}$	ANTE/ ALIGN	DEP$_{HEAD}$	MAX	NOFLOP	DEP
☞a. érogu				*			
b. erogu^{-}			*!		*		

The implications of Shaw's findings are expanded upon in Shaw et al. (2014) where it is argued that head faithfulness is an *emergent* effect in English. As these authors state, there is no evidence of head faithfulness elsewhere in the language. In fact, English compounds usually preserve the stress of the first element, as in the right-headed compound *blackboard* (Plag 2006), which is in direct conflict with head faithfulness. Since speakers could not have learned these constraints from the ambient language data, this supports the idea that they are part of a universal set. However, if a head faithfulness effect is found in blends in Japanese, it will not be considered emergent because there is some evidence for

this effect in the language already. As discussed previously, both analyses of compounds in Japanese include some form of faithfulness constraint to the right-hand element. All of the compounds discussed until this point have been right-headed and therefore it is possible that this faithfulness could be reinterpreted as head faithfulness. As it turns out, dvandva (or coordinating) compounds behave differently than their headed counterparts and generally preserve the accent and accentedness of the left member (Poser 1984). This evidence shows that a distinction between headed and coordinating structures already exists in the language and that speakers have likely already learned a constraint for head faithfulness. Therefore, we could not then argue that it was emergent in Japanese.

If head faithfulness is found to be relevant for blends in Japanese, then this will be significant for several reasons. Firstly, it will show that morphological heads in Japanese blends are not simply the second source word (in opposition to the claim by Kubozono 1990). It would also show that it is possible to explain blend accentuation without making reference to the concepts of "left" or "right" and that head faithfulness is actually the relevant concept in the formation of blends. Further, it would provide support for Shaw's (2013) analysis of English blends and would provide evidence that there is a fundamental similarity between blends cross-linguistically.

2.6 A Note on Matching Analysis

Several of the hypotheses discussed above (namely Compounds, Linear Order and Head Faith) are dependent upon matching a blend's accent with a source word's accent. However, as Shaw (2013) discusses, there are several options for determining this matching. The first option is to treat accent as a segmental feature such that a blend is only faithful to the accent of the source word if it both preserves the accented segment and that segment still carries

the accent. In this way, *ryóteru* ‘a mix of a Japanese and western hotel’ (*ryokan*¯ + *hóteru*) would be considered to be preserving the accent of the second source word while *ósyabari* ‘chatty and intrusive’ (*osyáberi* + *désyabari*) would not. Shaw (2013) utilizes a second option: determining the matching by alignment. By this procedure, a blend is faithful to the accent of its source word if accent falls on the same number of syllables (or moras) away from the edge. This would mean that both r*yóteru* and *ósyabari* would be considered faithful since the accent in both of the blends and source words were an equal number of syllables from the edge as shown in (56).

(56) Accent matching by alignment

Source word	*hó te ru*	*dé sya ba ri*
	| | |	| | | |
Blend	r*yó te ru*	*ó sya ba ri*

Since this procedure takes into account the clear similarities between these two examples I will consider accent to be determined by alignment rather than identity. For faithfulness to the second source word accent, alignment will be from the right edge. Additionally, for Head Faith matching to the first source word, left-headed blends will be aligned from the left edge.

It is also important to note that the exact procedure for this matching is not immediately clear. It could be determined by matching the accent placement of the source word and blend by either the number of moras or the number of syllables away from the edge. This is demonstrated in (57) below where an attempt to preserve the accent of the source word *zookin* ‘dustcloth’ could result in two different accent placements for the blend *dasukin* (a cleaning company). Since it is impossible to determine in advance which matching is relevant for blends, both possibilities will be discussed in the results section below.

(57) Example of matching ambiguity

	Match by Mora	**Match by Syllable**
Source Word 2	zó.o.ki.n	(zóo)(kin)
Blend	dá.su.ki.n	(da)(sú)(ki.n)

3. CORPUS STUDY

In this section, I test the hypotheses discussed above against accent judgements for a corpus of Japanese blends. As the vast majority of blends in the corpus are not found in dictionaries and their accent is not readily available, it was necessary to obtain data from native speakers. To accomplish this, I conducted an online survey and collected accent judgments for both the blends and their source words. This study follows several other corpus analyses investigating blend prosody (Shaw 2013; Arndt-Lappe & Plag 2013; Gries 2012), but represents the first attempt to conduct an analysis of this type in Japanese.

In section 3.1 I discuss the stimuli that were chosen and in section 3.2 I discuss the survey administration. Sections 3.3 and 3.4 present the participant information and the results of the study respectively while section 3.5 provides a discussion of these results.

3.1 Stimuli

Before determining what the survey stimuli would be, I collected a corpus of 102 Japanese blends (the largest corpus to date). This was created by obtaining blends from various sources including native speakers, other literature on Japanese blends (Kubozono 2008 and Ito 2011), and a variety of online websites. The survey stimuli that were chosen included a total of 40 blends, 31 of which were attested and 9 of which were novel. While the corpus of attested blends was much larger than this, the number of items was restricted

in order to reduce the burden on participants. Therefore, the first task in determining the stimuli was to decide which attested blends to include. I first excluded all blends that did not have the lexical category of noun. I also excluded any items where it was ambiguous whether it was a blend or a reduced compound. For example, the word *tundere* 'hot-cold personality,' which is a combination of the source words *tuntun* 'aloof' and *deredere* 'idling,' could be considered a blend or a reduced compound depending on whether the contribution from the second source word comes from the beginning or the end of the word.

With 77 blends remaining, I next determined whether an adequate number of distinctions could be made between each hypothesis. This would ensure that the study would not inadvertently be unable to differentiate between two hypotheses. To this end, I first consulted with a native Japanese speaker to obtain preliminary accent judgments on each of the source words. The native speaker was unaware of the hypotheses of the study and accent judgments were obtained by both listening to the pronunciations of the words and conferring about the placement of the pitch fall. After obtaining accent judgments for the source words, I then generated the blends' predicted accent placement for each hypothesis based on these accent placements. Any blends which had the same predictions for each hypothesis were then excluded as they would likely not provide any useful data for this study in distinguishing between hypotheses.

The number of distinguishing cases between each hypothesis was also calculated and cases where hypotheses had fewer than 10 blends distinguishing them were noted. In particular, it was found that hypotheses which differed in predictions for blends where the second source word had final or penultimate accent were lacking in distinguishing cases. To correct for this and increase the number of distinctions, I created a set of novel blends.

These blends were created by collecting nouns which had final or penultimate accent as the second source words and matching them with semantically similar or plausible first source words. A native speaker was consulted to help choose the final set of novel blends based on naturalness and plausibility.

After combining the remaining attested blends with the novel blends, I further excluded a number of the items. The predictions for each hypothesis were generated again and I determined which blends had the most distinguishing cases between hypotheses. I then chose to include the 40 blends (including the 9 novel blends) which had the most distinctions and also allowed there to be at least 10 distinctions between each hypothesis according to the judgments provided by the native speaker.

Definitions were also created for each of the 40 blends. Many speakers of Japanese would not necessarily be familiar with the meanings of the blends and thus would not have access to their headedness. To illustrate this point, with blends such as *ottyen*, a combination of *otoko* 'man' and *mettyen* 'girl,' the meaning of the blend is not obvious from the source words. A speaker who is not already familiar with the word could imagine that meaning is something like 'man-like girl' rather than the actual meaning 'girl-like man.' Including definitions helps to resolve this ambiguity for the participants. Definitions for the attested blends were created and double checked for naturalness with help from native speakers. For a full list of the blends and definitions used in this survey, see Appendix A.

3.2 Survey

The survey was administered as a web-based experiment using a modified version of the Experigen software (Becker & Levine 2014). In this survey, participants were asked to

provide their accent judgments for both the blends and their source words. It was important to obtain judgments for the source words since individual judgments for these might differ across participants. This information was crucial for determining conformity in several hypotheses. For each word, participants could select the accent from a list of each possible accent placements for that word, as shown below. The transcriptions were not included in the experimental materials but are provided here for the reader's benefit.

(58) Accent options for *ottyen* 'girl-like man'

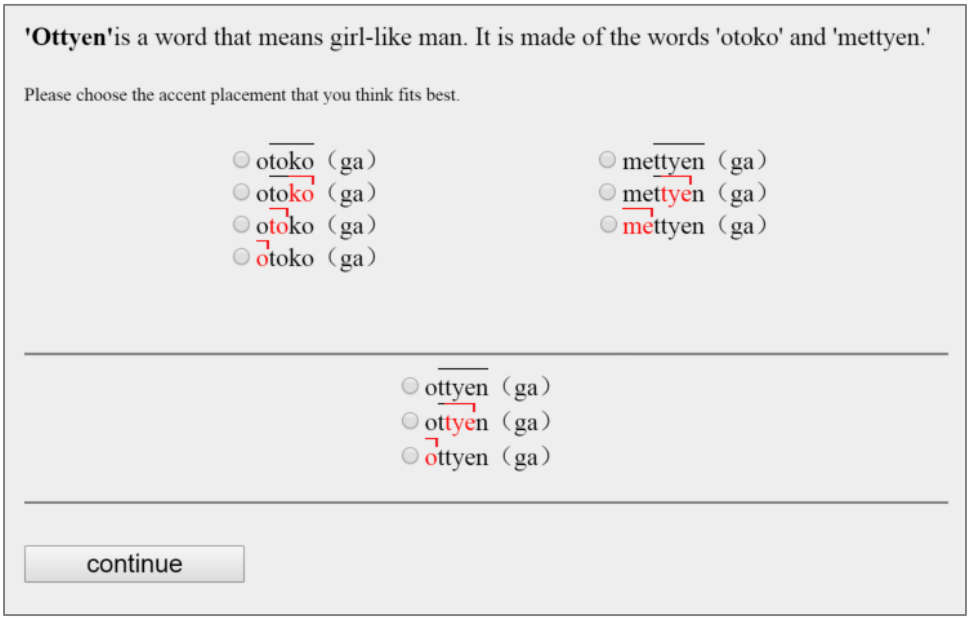
Japanese orthography:		Transcription:
オッチェン	（が）	o t tye n (ga)
オッチェン	（が）	o t tye n (ga)
オッチェン	（が）	o t tye n (ga)

Accent was indicated by a red line over the mora before which the pitch fall occurs with a small notch at the fall. Unaccented words were indicated by containing only a black line from the second mora to the end of the word. Each option was also followed by the subject particle 'ga' in parentheses in order to allow the participant to distinguish between the unaccented and final accent options.

The survey was conducted with instructions in Japanese and consisted of four different sections including instructions, a training page, the test items, and a post-survey questionnaire. The instructions briefly explained what a blend is and that the participants would be asked to look at words and decide which pronunciation is best. They were also given a short explanation of what accent is and how the markings in this survey were used to indicate it. In the training page, they were asked to choose the accent of three Japanese words: *kaze* 'wind,' *hata* 'flag', and *sora* 'sky.' This was also used as a diagnostic to help determine which dialect they might have.

Next, in the testing pages participants were first presented with a blend and its source words in their normal orthography and also given the definition of the blend. They were then asked to choose the accent for each of these words with the relevant options written either in hiragana or katakana. A representative screen shot of one of the testing pages with translated instructions and transcriptions for the items is given below for reference. Actual Japanese instructions are given in Appendix B.

(59) Example test page for *ottyen* with instructions in English (Japanese version provided in Appendix B)



Finally, in the questionnaire section, participants were asked to provide demographic information about what year they were born, their sex, their handedness and their native language. They were also asked what prefecture and country they were born in and whether they spoke a regional dialect or other languages and how well they spoke them.

Four versions of the surveys were created with 10 blends in each. This subset was presented in order to reduce the amount of work for each participant. The blends were

randomly placed in one of these four surveys. Within the surveys, the order of the blends was randomized.

3.3 Participants

Participants for this survey were recruited mainly through social media sites and e-mail. All participants were volunteers and were not compensated. Forty-three participants completed this study with 24 male and 19 female respondents. The age range was between 19 and 75 with the median age being 37 years old. All but one participant reported being right handed with the remaining participant being ambidextrous. All speakers reported Japanese as their native language. A variety of dialects were reported with 11 participants either not responding with a dialect or reporting a Tokyo dialect. Additionally, 14 participants reported a dialect from Kyushu or Okinawa (including Fukuoka, Kitakyushu and Hakata dialects), 2 from Chugoku, 7 from Kansai (including Mie and Osaka), 2 from Chubu, 4 from Kanto (including Ibaraki, Kanagawa and Tochigi), 1 from Tohoku, and 2 from Hokkaido. The participants also reported a variety of education levels from High School degree/GED to PhDs. Only 22 out of 43 participants responded to the diagnostic questions at the beginning of the surveys as expected for a Tokyo speaker.

3.4 Results

Each response for this experiment was coded as conforming or non-conforming to each hypothesis. For some hypotheses (e.g. Linear Order and Head Faith), the conformity depended on the source word accent judgments. In those cases, the conformity was determined by taking into account the accent judgments for the accent of the source word as indicated by the participant (as opposed to, for example, the expected accent for a Tokyo

dialect speaker or the most common judgement for that word). It should also be noted that any given response could potentially conform to multiple hypotheses. For example, in the theoretical responses below, a) conforms to four hypotheses, b) conforms to one and c) doesn't conform to any.

(60) Example response conformity to hypotheses

	Blend	SW2	SLO	SIM	CMK	CMI	RC2
a.	nékama	ókama	✓	✓	✓	✓	×
b.	biniron¯	náiron	×	×	×	×	✓
c.	okinagamé	nágame	×	×	×	×	×

SLO: Strong Linear Order; SIM: Simplex; CMK: Compound Kubozono; CMI: Compound Ito & Mester; RC2: Reduced Compounds II

For the purpose of analyzing the results, the hypotheses were also separated into two groups depending on the number of relevant blends. The hypotheses in the first group, termed "Group A," include only those that provided predictions for each response. The remaining hypotheses placed in "Group B," however, only provided predictions for a subset of the responses. For example, Strong Head Faith did not provide a definitive prediction for coordinating blends and therefore including these responses for this hypothesis in some way would run the risk of artificially inflating or depreciating the percent conforming responses. Further, Weak Linear Order lacked predictions for blends with unaccented second source words and Weak Head Faith lacked predictions for both coordinating blends and those with unaccented second source words.

In the sections below, I first present an analysis of the results in Group A. After eliminating several less promising hypotheses I then discuss the top results from Group B and compare each one to the winner of the previous section. Before presenting these

specific results, however, in section 3.4.1 I discuss the procedure used for determining whether the accent of a blend matched the accent of its source words.

3.4.1 Matching Analysis Tendencies

As discussed in section 2.6, the choice of how to determine whether blend accent matches source-word accent for several of the relevant hypotheses (including Compounds, Linear Order and Head Faith) is not immediately clear. Matching could be determined either by the number of moras or the number of syllables away from the edge. In the data, there were relatively few cases where this distinction was of concern. The chart in (61) gives a summary of the responses patterns for these cases.

(61) Response patterns for cases of ambiguity (for Strong Linear Order)

	Preserves SW Accent
Match by Syllable	4 (13%)
Match by Mora	26 (87%)
Total:	30

As this table shows, a much larger proportion of ambiguous cases preserved the accent of the source word by matching the number of moras as compared to the number of syllables. This is not surprising as accent placement in Japanese is closely correlated with the the mora. Of the responses that were matched by syllable, three of them came from *komi*ketto (*komikku* 'comic' + maaketto 'market') and one came from *bini*ron (*biniru* 'vinyl' + nairon 'nylon'). It should also be noted that two of these responses came from the same participant and that there were four other responses for *komiketto* that were matched by mora. Further, there appears to be nothing special about the form of these two blends as *dasukin* and *wasyuretto* have the same prosodic structure as *biniron* and *komiketto* and also had 2 and 7 responses matched by mora respectively. As this pattern appears to be robust, conformity

to Head Faith and Linear Order is determined by mora count for all following sections. Next. I present a discussion of the results for Group A.

3.4.2 Analysis of Group A Hypotheses

As discussed previously, the hypotheses in Group A all provided predictions for each blend and are thus easily analyzed together. Under a null hypothesis that accent judgments were chosen by chance, the percentage of conforming responses should be 20.98%[2]. A summary of the total number of conforming responses obtained for each hypothesis is given in (62) below with the percentage of conforming responses given out of the total number of responses. The values were compared to chance by a logistic regression accounting for multiple observations in subjects using the LOGISTIC procedure from the SAS statistical package. The results from this analysis are also provided below.

(62) Responses for Group A

Hypothesis	**Conforming Responses**	**S.E**	χ^2	**P values**
Strong Linear Order (SLO)	60.7%	0.103	253.33	<.0001
Simplex (SIM)	45.6%	0.111	99.73	<.0001
Compound Kubozono (CKB)	42.6%	0.117	70.04	<.0001
Compound Ito & Mester (CIM)	36.5%	0.096	53.12	<.0001
Reduced Compounds I (RC1)	28.6%	0.132	6.19	0.0129

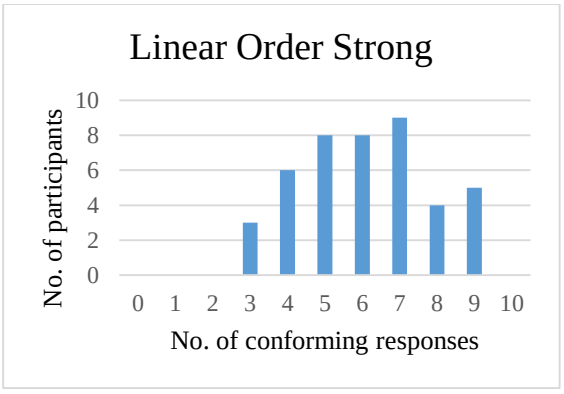
As this data shows, each hypothesis predicts accent significantly better than chance ($p<.0001$ for most). A look at the conforming responses also shows that Strong Linear Order (SLO) is clearly the best predictor from Group A. With just over 60% of the responses conforming to this hypothesis, it beats out the closest competitor from the same

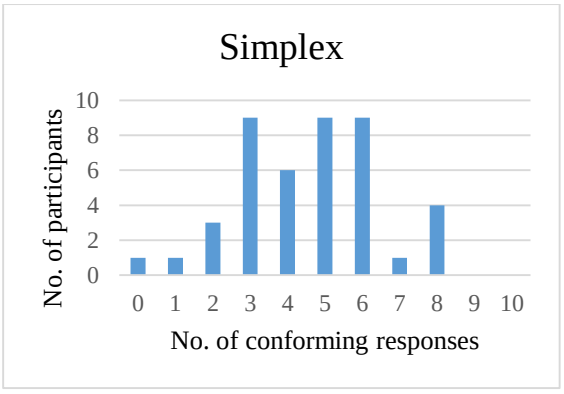
[2]This value was obtained by dividing the number of responses obtained by the number of possible responses (430/2050). The number of possible responses differed by item and included both each possible accent placement and an unaccented option.

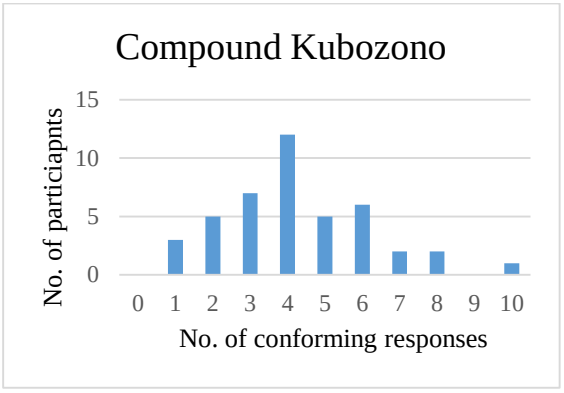
group, Simplex (SIM), by about 15 percentage points. The remaining hypotheses have 42.6%, 36.5%, and 28.6% conformity, all much lower than SLO.

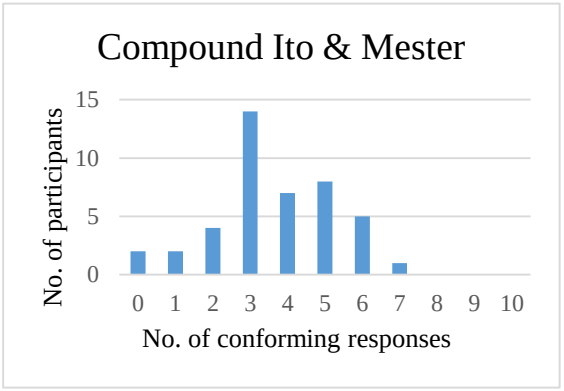
For a more in-depth look at the results from Group A, the histograms below show the distribution of participants that gave between zero and ten conforming responses.

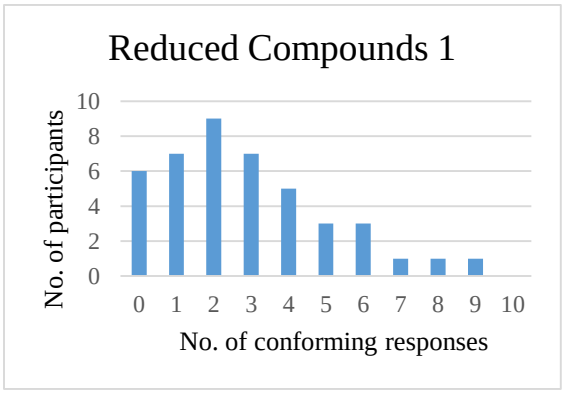
(63) Conforming responses by hypothesis and participant







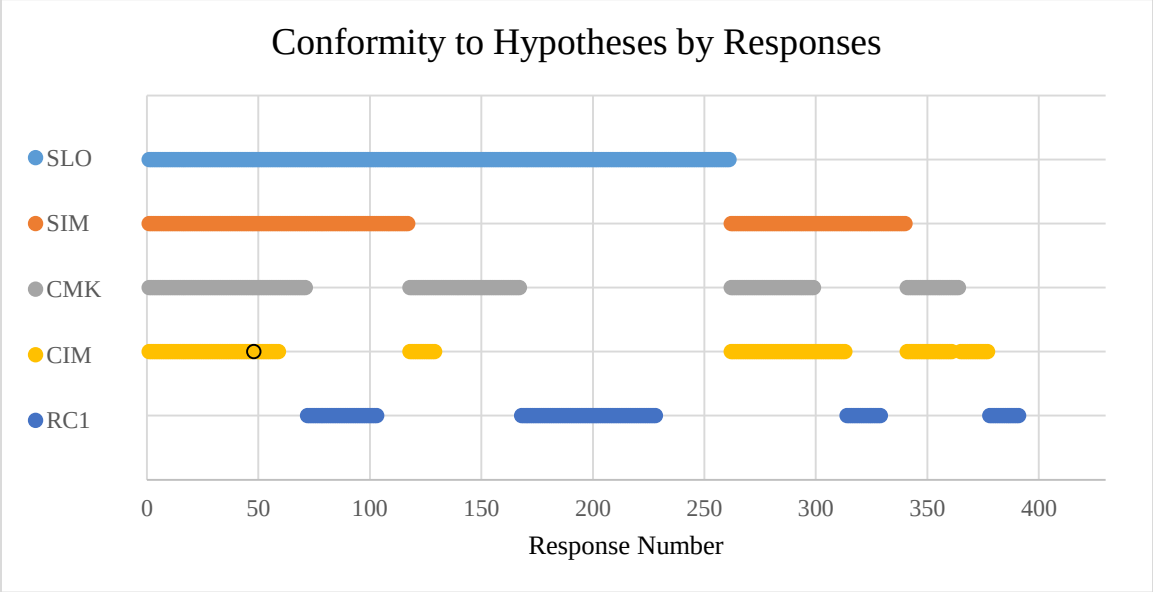




As these histograms demonstrate, a much larger number of participants gave a high number of conforming responses to SLO as compared to the others. About 60% of participants responded with greater than half of responses conforming. This is compared to less than 33% of participants responding with this level of conformity for *all other* hypotheses. Additionally, 5/43 participants had 90% of responses conforming to SLO.

Also, it is possible that one of these hypotheses accounts for the majority of responses and a different hypothesis accounts for the residual cases. Alternatively, the success of some of the hypotheses may be due to their similarity with others. To give an idea of the relative contributions of each hypothesis, the graph in (64) below shows the conformity to each hypothesis (indicated by a circle) for each of the 430 responses.

(64)



The graph above displays each of the 430 responses with conformity to each hypothesis. Circles indicate conformity and appear as solid lines due to close proximity. For example, the dark circle in the CIM row indicates that response #48 was conforming to the CIM hypothesis while the circles above and below it indicate that the response was not conforming to RC1 but was conforming to CMK, SIM and SLO. Responses were sorted by conformity to hypotheses.

As this graph shows, roughly half of the non-conforming responses for SLO followed SIM and a large number also followed CIM. The somewhat "complementary" distribution of the Compound hypotheses and RC1 is simply due to the fact that RC1 always predicted unaccentedness and the Compound hypotheses nearly never did. From this graph, it certainly appears that CMK and possibly SIM may have had an artificially inflated conformity percentage due to overlap with SLO.

Lastly, if we consider SLO to be the best predictor, then a high percentage of conformity for its closest competitor, SIM, may be easily explained. Predictions made from these hypotheses overlap for responses where the second source word has the default accent pattern of the language. This may have occurred many times since most of the source words are simplex and are likely to have the default accent. This appears to be true as about 27% (103/387) of the second source words which are not of the unaccented form ([HLL] or [LLLL]) have the default accent. Following this, about 82% (85/103) of the blends made from these source words (none of which are of the unaccented form) retain the default accent. This essentially accounts for more than half (57%, 85/148) of the accented SIM-conforming responses. Additionally, a key prediction of the Simplex analysis is that blends of the form [HLL] or [LLLL] will become unaccented. However, only 46% (48/104) of the blends of this type were unaccented (32 of which had unaccented second source words). This indicates that the high conformity to SIM does not necessarily mean that it is a close competitor to SLO but that this is a result of overlapping predictions made by both of these hypotheses.

In summary, the evidence presented above strongly indicates that SLO is the best predictor of accent from Group A and for the remainder of the thesis I will consider this to be the case. In the next section, I will present an overview of the results from Group B.

3.4.3 Summary of Group B Hypotheses

As discussed previously, the hypotheses from Group B only provided predictions for a subset of the responses. As a result of this, they are less straightforwardly comparable to each other or to the hypotheses from Group A. Nevertheless, (65) provides a summary of the conforming responses for each hypothesis with the percent conforming given out of the subset of relevant responses. To reiterate, these values do not reflect the percent conforming form the total set of responses and they cannot be directly compared to the values from Group A.

(65) Responses for Group B

Hypothesis	**Conforming Responses**	
Strong Head Faith (SHF)	159/266	59.8%
Weak Linear Order (WLO)	168/284	59.2%
Weak Head Faith (WHF)	90/164	54.9%
Reduced Compounds II (RC2)	92/254	36.2%

As this data shows, several of the hypotheses from Group B, including Weak Linear Order and Strong/Weak Head Faith, appear to perform well with over 50% of the relevant responses conforming. However, Reduced Compounds II, like its counterpart in Group A, performs the worst with only 36.2% conforming. In order to limit the number of comparisons in this study, the remaining sections will only consider comparisons between the top three hypotheses from Group B with the best from Group A: Strong Linear Order.

3.4.3 Strong vs. Weak Linear Order

The first opposing hypothesis that I will consider here is the Weak Linear Order (WLO) hypothesis. As stated in section 2, the crucial cases that distinguish these two hypotheses are those in which the second source word is unaccented. Strong Linear Order (SLO) predicts that those blends will be faithful to the unaccented source words and therefore will be unaccented as well. In contrast, WLO predicts that these crucial cases will be subject to an alternative accent pattern. It should be noted that this does not imply that none of the blends will be unaccented as the alternative pattern could also include predictions of unaccentedness (i.e. the Simplex pattern). As it is difficult to predict exactly what this alternative pattern would be, the response conformity to WLO was calculated by excluding blends with unaccented second source words. In summary, if WLO is a better predictor of blend accent, then we would expect to see several indications of this. Firstly, we would see a much lower rate of conformity to SLO for the crucial cases. Secondly, we would expect the crucial cases to all conform to a specific accent pattern different from that of SLO. Lastly, we would expect the conforming crucial cases to follow a specific pattern (e.g. have the form of unaccented simplex words, [LLLL] or [HLL]). However, if SLO is the stronger hypothesis, then there should be a high rate of conformity for crucial cases and this rate should be roughly the same as for the non-crucial cases.

How do the data fare with respect to this comparison? The relevant values are given in (66) below.

(66) Conformity to SLO by accent of SW2

	Unaccented SW2	**Accented SW2**	**Total**
Preserves SW2	93 (64%)	168 (59%)	261
Doesn't Preserve SW2	53 (36%)	116 (41%)	169
Total	146	284	430

As the table above shows, 64% of crucial cases preserve the accent of the second source word. Interestingly, this is actually slightly higher than the percentage of non-crucial cases that conformed to SLO. A logistic regression analysis run in SAS accounting for multiple observations in subjects further provides evidence that the conformity to SLO by crucial and non-crucial cases is not significantly different ($p=0.4677$). This evidence indicates that SLO is indeed a more accurate predictor of blend accent than WLO.

Next, an analysis of what alternative pattern these crucial cases could conform to does not provide support for WLO either. Ignoring Head Faith (to be discussed in the next sections) and Reduced Compounds (not relevant here as they only predict unaccented blends), the remaining hypothesis with the highest conformity is SLO. In comparison, the Simplex pattern only has a 36% conformity and both Compound hypotheses have only 25% conformity. As there does not appear to be a viable alternative accent pattern for the crucial cases, this is further evidence that SLO is a better predictor than WLO.

Lastly, an analysis of which of the crucial cases did follow the unaccentedness of the second source word also supports SLO over WLO. Relevant data for these blends are given in (67) below with items ranked by number of conforming responses to SLO. Only 31 out of 40 of the blends had responses with unaccented second source words and all of these are provided below. Information such as the headedness of the blend (indicated by a 1, 2, or b for left-headed, right-headed and coordinating respectively), the form of the blend (whether it was a four mora word ending in two light syllables), and the source word length difference (in moras) are also provided.

(67) Conformity to SLO by crucial cases (blends with unaccented SW2)

	Blend	Head	[LLLL] or [HLL]?	SW1-SW2 (moras)	Proportion of Conforming Responses
1	peraido	2		0	9/9
2	homodati	2	✓	-2	8/8
3	erogu	2		2	7/9
4	kotona	2		0	7/8
5	agufure	2	✓	3	7/7
6	nekama	2		0	7/7
7	saikyoo	b		0	6/8
8	kanzibiki	2		0	5/7
9	hetaria	2	✓	-1	4/9
10	pianika	b	✓	-2	4/8
11	kookoo	2		0	4/4
12	kanageki	b	✓	1	3/4
13	asamuke	2	✓	-1	3/3
14	uppurami	b		1	3/3
15	seizigoku	2		0	2/4
16	matubomi	2	✓	-1	1/9
17	tibitaria	2		-3	1/8
18	gozira	b		0	1/7
19	okinagame	2		1	1/3
20	osyabari	b	✓	0	1/3
21	apasiki	b	✓	1	1/2
22	baho	b		0	1/2
23	biniron	b		0	1/2
24	dasukin	b		2	1/2
25	syameraman	2		-1	1/2
26	abenomikusu	2		-4	1/1
27	ottyen	1		-1	1/1
28	potetoruneedo	1		-2	1/1
29	zyabitto	2	✓	1	1/1
30	monyaki	b	✓	0	0/3
31	yakimorokosi	2		-4	0/1

In the table above, values under *Head* represent the headedness of the blend. '1' represents left-headed, '2' represents right-headed and 'b' represents coordinating. The column *[LLLL] or [HLL]* indicates with whether or not the blend was four moras long ending in two light syllables. The column *SW1-SW2* indicates the difference in length between the two source words by moras.

As mentioned previously, if there was another factor involved such that the rate of conformity in crucial cases was higher due to other reasons, this could indicate that WLO

is actually the better predictor. However, no factor is apparent from the data. It would be reasonable to suspect that blends of the form [LLLL] or [HLL], those that are predicted to be unaccented by the Simplex hypothesis, might be more likely to inadvertently conform to SLO. This does not appear to be the case as slightly less than half of the responses for this type of blend (24/57, 42%) are non-conforming. In fact, for some blends of this form, such as *matubomi*, nearly all responses were non-conforming. Alternatively, we might expect that blends where the second source word was longer than the first source word (indicated by a negative number in the SW1-SW2 column) might also be more likely to follow SLO. Several studies have found this to be the case with English blends (Bat-El & Cohen 2012; Cannon 1986; Gries 2004a,b; Shaw 2013) and it would be reasonable to expect the same pattern in Japanese. However, this too does not appear to have an effect as roughly half of the responses for blends of this type (26/51, 51%) do not, in fact, conform to SLO. Interestingly, there does appear to be a correlation between headedness and conformity to SLO, but that will be discussed further in the next section.

In summary, as an alternative version of SLO, WLO provides no predictive benefit. The findings of this study provide no evidence that unaccentedness of the second source word influences a blend's conformity to SLO. Both a comparison of conformity to SLO between crucial and non-crucial cases as well as a look at the individual conforming crucial cases provide strong evidence that SLO is a better predictor than WLO. If conformity to the accent of the second source word is relevant in blend formation, then the conformity to unaccentedness is equally important. In the next section I will compare the Strong versions of the Linear Order and Head Faith hypotheses.

3.4.4 Strong Linear Order vs. Strong Head Faith

As discussed in section 2, the Strong Linear Order and Strong Head Faith (SHF) hypotheses differed by 1) which source word the blend should follow for headed blends and 2) whether coordinating blends should also follow the accent of the second source word. Their predictions were identical for right-headed blends and the crucial cases for this comparison are the left-headed and coordinating blends. SLO predicts that each blend will follow the second source word regardless of headedness. In contrast, HFS predicts that left-headed blends will follow the accent of the first source word and that coordinating blends will be subject to an alternative accent pattern. Again, as it is difficult to predict what the alternative pattern is, the response conformity to HFS given in section 3.4.3 was calculated by excluding the coordinating blends. In summary, if HFS is a better predictor, then we would expect to see a high rate of conformity for left-headed blends. Also, assuming that the coordinating blends are following an alternative accent pattern other than that of SLO, we would expect to see a much lower rate of conformity to SLO for those blends. However, if SLO is the stronger hypothesis, then the rate of conformity should be just as high for each group of blends regardless of headedness.

Which hypothesis do the critical cases conform to? The table in (68) summarizes the number of responses by blend head and accent matching.

(68)

	Left-Headed	**Right-headed**	**Coordinating**	**Total**
Preserves Left	1	16	46	63
Preserves Right	20	91	46	157
Preserves Both	10	57	37	104
Preserves Neither	6	65	35	106
Total	37	229	164	430

As this data shows, evidence from the left-headed blends appears to support SLO over SHF. For these blends, only 30% (11/37) of the responses followed the accent of the first source word. In comparison, 81% (30/37) of those responses followed the accent of the second source word. To further examine these blends, the responses for all three left-headed blends in the corpus are shown in (69) below. Each response for the accent of the blend is provided with '0' representing unaccented and all other numbers representing the number of the syllable from the right edge that contains the accent (e.g. *po.te.tó.ru.nee.do* is represented by '4').

(69) Conformity to SHF by left-headed blends

	Blend	SW1-SW2 (moras)	Accent of blend	
			Non-conforming	Conforming
1	potetoruneedo	-2	0,1,1,2,2,2,2,2,2,2,2,4	6
2	ottyen	-1	0,2,2,2,2,2,2,2,2	1,2,2
3	mamagon	-2	1,2,3,3	2,3,3,3,3,3,3,3

In the table above, the column *SW1-SW2* indicates the difference in length between the two source words by moras. The numbers provided above under *accent of blend* represent the number of the syllable from the right edge of the word that contains the accent. Each number represents an individual response. The number of responses differed slightly for each blend depending on how many participants were assigned to the group that contained that blend.

As this data shows, two out of three of these blends have a large number of non-conforming responses. One possible explanation for the nearly perfect non-conformity in *potetoruneedo* is that choosing the head-faithful accent (initial accent in this case) would result in an accent placement that is very far from the right edge. As noted by Kubozono (2008), initial accent is extremely rare in blends longer than 4 moras. Yet, one subject still chose this accent placement. As for *ottyen*, an unusual difficulty arises. This is the only blend for which one accent placement in the source word did not allow a particular

matching accent placement in the blend. As shown in (70), the second mora in *otoko* ‘man’, aligns with a coda in the blend (assuming the mora-count procedure discussed earlier) which cannot carry the accent. Only two subjects chose the problematic accent placement for the source word (otóko) and in both cases the subject chose initial accent in the blend (although this accent also matched their choice for source word 2). For this analysis, those responses were coded as non-conforming. However, treating them otherwise only increases conformity for left-headed blends to 35%. Regardless, the number of blends used in this study was too few to make definitive conclusions with respect to left-headed blends.

(70) Mora alignment in *ottyen* from left-edge

SW1: (o)(to)(ko)
Blend: (o.t)(tye.n)

Next, I will consider a comparison between right-headed and coordinating blends. The table in (71) summarizes the data for this comparison.

(71) Responses for Right-headed and Coordinating blends by accent preservation

	Right-headed	**Coordinating**	**Total**
Preserves Right	148 (65%)	83 (51%)	231
Doesn't Preserve Right	81 (35%)	81 (49%)	162
Total	229	164	393

The data here appear to show a different pattern for conformity to head faith. 65% (148/229) of the responses for right-headed blends preserved the right source word compared to only 51% (83/164) of responses for coordinating blends. A logistic regression model was also run in SAS accounting for multiple observations in subjects to determine whether the two groups were significantly different. The values from the statistical model are given in (72) below.

(72) Statistical values from logistic regression for headedness effect

Coefficient	S.E	χ^2	P values
0.36	0.1514	11.73	0.0006

As the table above shows, the two groups are very significantly different (p=0.0006), indicating that right-headed blends are significantly more likely to preserve the accent of the second source word than coordinating blends. This finding is in direct conflict with the previously stated prediction of SLO that blends should be equally likely to preserve the accent of source word 2 regardless of headedness. The pattern found in this data cannot be accounted for by SLO and suggests that blends are actually making reference to headedness rather than simply the linear ordering of source words for determining accent placement.

In addition to finding difference between the patterns found in headed and coordinating blends, if SHF is the more accurate predictor of accent, then we would expect that the coordinating blends would consistently follow an alternative accent pattern. In fact, the data show that the largest percentage of coordinating blends conform to SIM with 57% (94/164) as compared to only 51% (83/164) conforming to SLO. A breakdown of the responses for coordinating blends is shown in (73) below with simplex responses indicated by bold.

(73) Conformity to SLO by coordinating blends

	Blend	[LLLL] or [HLL]?	SW1-SW2 (moras)	Accent of blend	
				Non-Conforming	Conforming
1	gopan		1	**2,2,2,2,2,2,2,2,2,2**	1,1
2	dasukin		2	2,**3,3,3,3,3,3,3,3**	0,**3**
3	apasiki	✓	1	**0,0,0,0**,3,3,3,3	**0**
4	gozira		0	**3,3,3,3,3,3,3**	0,**3**
5	monyaki	✓	0	**0,0,0**,3,3,3,3	2,2
6	osyabari	✓	0	**0**,2,3,3,3,3,3	**0**,2,2,3,3,4
7	pianika	✓	-2	3,3,3,3,3,3,3	**0,0,0,0**,4
8	uppurami		1	0,0,0,0,2,**3,3**	0,0,0,1,2,**3**

9	kanageki	✓	1	**0,0,0,0,0**,1	**0,0,0**
10	baho		0	**2,2,2,2**	0,1,1,**2,2**
11	biniron		0	0,2,**3**	0,1,**3,3,3,3,3,3,3,3,3**
12	saikyoo		0	0,**2,2**	0,0,0,0,0,0
13	ryoteru		0	0,**3**	1,**3,3,3,3,3**
14	faburetto		-3	**2**	3,4,4,4,4,4,4,4
15	mukku		1		**2,2,2,2,2,2,2,2,2**
16	pomato		0		1,**3,3,3,3,3,3,3,3**

In the table above, the column *[LLLL] or [HLL]* indicates with whether or not the blend was four moras long ending in two light syllables. The column *SW1-SW2* indicates the difference in length between the two source words by moras. The numbers provided above under *accent of blend* represent the number of the syllable from the right edge of the word that contains the accent. Each number represents an individual response. Bolded numbers indicate simplex accent. The number of responses differed slightly for each blend depending on how many participants were assigned to the group that contained that blend.

Additionally, 57% (94/164) of the coordinating blends follow the default antepenultimate accent rule without the pattern of unaccentedness in Simplex nouns (which occurs in four mora blends ending in two light syllables). In cases of blends with this form, subjects appear split on whether to treat them as simplex or not. Of the 9 subjects who saw more than one of these (*apasiki* and *monyaki*), 4 of them gave different responses for each. Only 30% of the blend responses cannot be accounted for by either the Simplex or Antepenultimate accent pattern. This result suggests that coordinating blends are distinct from right-headed blends in that their accent placement is primarily simplex rather than being determined by the second source word.

In summary, even though SLO is a strong contender in terms of the total number of responses correctly predicted, it fared significantly worse for the crucial cases in the comparison between SLO and SHF. A brief overview of the conformity to each hypothesis by headedness is given below in (74).

(74) Summary of conformity to SLO and SHF by headedness (with conformity to SHF by coordinating blends given as a range between the conformity to SIM and SIM plus the antepenultimate responses)

	Right-headed	Left-headed	Coordinating	Total	Total (excluding left-headed)
SLO	148 (65%)	30 (80%)	83 (51%)	261 (60%)	231 (58%)
SHF	148 (65%)	11 (30%)	94 (57%)-116 (71%)	253 (58%)-275 (64%)	242 (62%)-264 (67%)
Total	229	37	164	430	393

As (74) shows, the hypotheses correctly predicted the accent of the same percentage of right-headed blends (65%). For left-headed blends, SLO correctly predicted a much larger percentage of the blends. However, since there were only 3 left-headed blends in the corpus and the poor performance of SHF could be contributed to other factors such as distance from the right edge of the word, no conclusions about this can be made. Lastly, for the coordinating blends, SLO only accounted for 51% of responses which was significantly worse than its performance on right-headed blends as shown previously in (72). On the other hand, the alternative pattern of SIM accounts for 57% percent of responses and as many as 71% if the pattern of antepenultimate accent is included. With this alternative pattern accounting for coordinating blends, SHF actually accounts for between 62% and 67% of the total number of responses (excluding left-headed blends) as compared to 58% for SLO. Thus, the evidence presented above has shown that accent placement in blends is actually determined by head faithfulness rather than simply linear order.

3.4.5 Strong vs. Weak Head Faith

Lastly, we will examine the differences between Strong and Weak Head Faith. Both of these hypotheses have predictions only for headed blends. The crucial cases that distinguish them are those blends where the second source word is unaccented. SHF predicts that the

blend will follow the accent (or unaccentedness) of the head while WHF predicts that these cases will follow a different accentual pattern. If WHF is the better predictor, then we would expect that the percentage of crucial cases that preserve the head to be much lower than the percentage of non-crucial cases. We would also expect that the crucial cases would conform to a different accentual pattern than Head Faith. The relevant values for this comparison are given in (75) below.

(75) Conformity to SHF by SW accent (SW2 for Right-headed, SW1 for left-headed)

	Unaccented SW	**Accented SW**	**Total**
Preserves Head	69 (66%)	90 (56%)	159
Doesn't Preserve Head	36 (34%)	71 (44%)	107
Total	105	161	266

As this data shows, the percentage of crucial cases that preserve the head is actually larger than that of non-crucial cases. A logistic regression of the data accounting for multiple observations in subjects also confirms that these groups are not statistically different ($p=0.2195$). This data indicates that blends with unaccented source words do not behave differently than those with accented source words and thus that WHF does not provide better predictions than SHF.

Next, if WHF is better than SHF then we might expect that all of the crucial cases would follow an alternative accentual pattern. A look at the data shows that 70% of these cases follow SLO (unsurprisingly slightly higher than SHF due to inclusion of left-headed blends). The next highest competitor is SIM at 36% following the Compound hypotheses at around 20%. This indicates that the crucial cases do not, in fact, follow an alternative accent pattern and provides further evidence that SHF is the better predictor.

Lastly, (76) below provides an in-depth look at the crucial cases. As we can see, 56% (20/36) of the crucial responses non-conforming to SHF come from only three blends. The exact reason for this non-conformity is unclear, especially considering the fact that two of them are of the simplex unaccented form. Nevertheless, the fact that the non-conforming cases are restricted to this few number of blends provides further support for SHF.

(76) Conformity to SHF by crucial cases

	Blend	Hd	[LLLL] or [HLL]?	SW1-SW2 (moras)	Non-Conforming	
					Accent of blend	Proportion
1	matubomi	2	✓	-1	2,3,3,3,3,2,2,2	8/9
2	tibitaria	2		-3	3,3,3,3,3,4,4	7/8
3	hetaria	2	✓	-1	2,3,3,3,3	5/9
4	ottyen	1		-1	2,2,2	3/3
5	erogu	2		2	2,3	2/9
6	kanzibiki	2		0	3,4	2/7
7	seizigoku	2		0	3,4	2/4
8	okinagame	2		1	2,3	2/3
9	mamagon	1		-2	3,3	2/2
10	kotona	2		0	1	1/8
11	syameraman	2		-1	3	1/2
12	yakimorokosi	2		-4	4	1/1
13	peraido	2		0		0/9
14	homodati	2	✓	-2		0/8
15	agufure	2	✓	3		0/7
16	nekama	2		0		0/7
17	kookoo	2		0		0/4
18	asamuke	2	✓	-1		0/3
19	abenomikusu	2		-4		0/1
20	zyabitto	2	✓	1		0/1

In the table above, values under *Head* represent the headedness of the blend. '1' represents left-headed and '2' represents right-headed. The column *[LLLL] or [HLL]* indicates with whether or not the blend was four moras long ending in two light syllables. The column *SW1-SW2* indicates the difference in length between the two source words by moras. The numbers provided above under *accent of blend* represent the number of the syllable from the right edge of the word that contains the accent. Each number represents an individual response. The number of relevant responses differed for each blend depending on how many participants were assigned to the group that contained that blend and also by the number of responses where the relevant source word was judged as accented.

In summary, the evidence discussed above provides strong support for SHF as a better predictor than WHF. Blends with unaccented source words are just as likely (if not more so) to follow the accent of the head. Additionally, the crucial cases were far more likely to follow SHF than other accentual patterns.

3.4.6 Discussion

Out of many different possibilities, the results above have shown that Strong Head Faith is the best predictor of accent placement in Japanese blends. It is able to account for the significant differences between headed and coordinating blends. Additionally, the results show that coordinating blends tend to follow the Simplex accent pattern of the language.

These results are easily accounted for in OT using the constraints discussed in sections 2.1 and 2.5 for Simplex and Head Faithfulness respectively. In particular, the head faithfulness constraints (grouped here as $\text{FAITH}_{\text{HEAD}}$ but including DEP_{HEAD}, MAX_{HEAD}, and $\text{NOFLOP}_{\text{HEAD}}$) must be ranked above the ANTEPENULT constraints. An example of this is shown in the tableau in (77).

(77)

/okinawa⁻ + nagamé*Head*/ 'Okinawa scenergy'	FAITH$_{HEAD}$	INITIALFT	ANTE	WORD PROM	FAITH
☞a. (oki)(naga)mé			*NOLAPSE		
b. (oki)(nága)me	*!NOFLOP$_{HEAD}$				*NOFLOP

In this tableau, the blend is headed and therefore subject to the FAITH$_{HEAD}$ constraints. Candidate b) that violates one of these constraints therefore loses to candidate a), even though it violates ANTEPENULT.

These constraints are also able to account for the results found for coordinating blends which followed the Simplex accent pattern[3]. An example of this is given in the tableau in (78) below.

(78)

/dasutokúroosu + zookín/ (a cleaning company)	FAITH$_{HEAD}$	INITIALFT	ANTE	WORDPROM	FAITH
☞a. (dásu)kin					**
b. (dasu)(kín)			*! NONFIN		*

As this tableau demonstrates, if a blend is coordinating, then the candidate that has simplex accent will be the winner. Candidate a) which is not faithful to the second source word does not incur a violation from FAITH$_{HEAD}$ since the blend is not headed. However, candidate b) which is faithful to the second source word violates ANTEPENULT and thus loses. This demonstrates that an analysis in OT is easily able to account for the patterns found in the results.

[3]It has also been noted that many of these coordinating blends of the form [HLL] or [LLLL] did not follow the unaccented pattern from Simplex but rather had default antepenultimate accent. This variation could possibly be understood as variability in some of the constraints such as INITIALFT or WORDRPOM.

Additionally, an ideal analysis would be able to account for the accentual patterns found in each of the word formation processes in Japanese. Any such account would necessarily include the high-ranking ALIGN and LEXFT constraints which drive accent placement in compounds and reduced compounds, respectively. However, the results from this study indicate that blends are not subject to these constraints. This is somewhat problematic for the current analysis as demonstrated in the tableau below. This tableau contains the high-ranking LEXFT constraint (though a similar discussion could be made with ALIGN). Recall that LEXFT results in a violation when a morpheme does not project its own foot.

(79)

/$d_a a_a s_a u_a t_a o_a k_a ú_a r_a o_a o_a s_a u_a$ + $z_b o_b o_b k_b í_b n_b$/ (a cleaning company)	LEXFT	$\text{FAITH}_{\text{HEAD}}$	ANTE	FAITH	WORDPROM
💣a. $(d_a a_a s_a u_a)(k_b i_b n_b)$ ¯		*(vacuous)*		**	*
b. $(d_a á_a s_a u_a)(k_b i_b n_b)$		*(vacuous)*	*!NONFIN	**	
c. $(d_a á_a s_a u_a) k_b i_b n_b$	*!	*(vacuous)*		**	

In the data we saw that a coordinating blend like the one above should have default antepenultimate accent. However, as this tableau demonstrates, if the blend is subject to the LEXFT constraint then it is incorrectly predicted to be unaccented, as in candidate a). The alternative candidates which have the correct accent placement will fail as they either violate the high ranking constraint LEXFT, as in the case of c), or their exhaustive footing results in a violation of ANTEPENULT, as in the case of b).

There are two possible methods for remedying the situation. The first is to assume that the LEXFT constraint is indexed to reduced compounds only (in the sense of Pater 2006). As shown in the tableau in (80), this would mean that the constraint was always vacuously

satisfied for blends. Candidate c), which has the correct accent placement is now correctly predicted to be the winner.

(80) Account of blends using indexed LEXFT constraint

/$d_a a_a s_a u_a t_a o_a k_a ú_a r_a o_a o_a s_a u_a$ + $z_b o_b o_b k_b í_b n_b$/ (a cleaning company)	LEXFT(RC)	$\text{FAITH}_{\text{HEAD}}$	ANTE	FAITH	WORDPROM
a. $(d_a a_a s_a u_a)(k_b i_b n_b)$ ¯	*(vacuous)*	*(vacuous)*		**	*!
b. $(d_a á_a s_a u_a)(k_b i_b n_b)$	*(vacuous)*	*(vacuous)*	*!NONFIN	**	
☞c. $(d_a á_a s_a u_a)k_b i_b n_b$	*(vacuous)*	*(vacuous)*		**	

However, this solution is somewhat unsatisfactory as Ito & Mester (2012) state that the idea for a constraint such as LEXFT was first proposed by Poser (1984) in order to account for Sino-Japanese compounds. Indeed, the Ito & Mester (2012) account for compounds assumes that each member minimally projects a foot. While it is possible that this constraint could be indexed to kinds of *compounds* only, this seems highly unlikely given the degree of similarity they have with blends. All three involve combinations of two source words and both blends and reduced compounds involve truncated segments from these source words.

The second possibility for accounting for blends not being subject to ALIGN and LEXFT is that blends do not have the same morphological structure as reduced compounds and compounds. An example of this alternative solution is shown in the tableau in (81).

(81) Account of blends using different internal morphological structure

/$d_a a_a s_a u_a t_a o_a k_a ú_a r_a o_a o_a s_a u_a$ + $z_b o_b o_b k_b í_b n_b$/ (a cleaning company)	LEXFT	$\text{FAITH}_{\text{HEAD}}$	ANTE	FAITH	WORDPROM
a. $(d_a a_a s_a u_a)(k_b i_b n_b)$ ¯		*(vacuous)*		**	*!
b. $(d_a á_a s_a u_a)(k_b i_b n_b)$		*(vacuous)*	*!NONFIN	**	
c. $(d_a á_a s_a u_a)k_b i_b n_b$	*!	*(vacuous)*		**	
☞d. (dásu)kin		*(vacuous)*		**	

In this tableau, candidate d) does not follow the same internal structure as the other candidates. In particular, it does not possess two different morphemes that must project feet in order to satisfy LEXFT. Since it is unclear what the exact morpheme specification of the segments would be, it has been displayed here with no morpheme specification. As this example shows, without a violation of LEXFT, candidate d) with antepenultimate stress is correctly predicted to be the winner.

Of course, this account also raises some difficult questions. What does it mean for a blend to be headed if its morphological structure is simplex? How can a blend be faithful to a "head" when the blend itself has no head? As these questions are outside of the scope of this thesis they will not be addressed here. Future research into this topic is likely required to investigate these issues.

4. CONCLUSIONS

The main goal of this study was to investigate the determination of accent placement in Japanese blends and this thesis has shown that head faithfulness, more than any other factor, is responsible for accent placement. The data obtained demonstrates that there is a disparity in the accentual pattern of headed and coordinating blends and that headed blends are very likely to preserve the accent of the head. Coordinating blends, on the other hand, are likely to follow the simplex accent pattern. These findings have several implications for blends cross-linguistically as well as for their morphological structure.

Firstly, this study provides further support for the claim made in Shaw (2013) that blend formation is subject to head faithfulness. Indeed, while Shaw found that head faithfulness

was a secondary effect in English blends, these findings indicate that no other factor has a stronger effect on accent placement in Japanese blends. This goes against previous claims that it is faithfulness to the right that is relevant for blends in English (Bat-El & Cohen 2012, Gries 2004a,b) as well as in Japanese (Kubozono 1990). Faithfulness to the head also provides a more satisfying explanation for blends since positions like "right" and "left" are not privileged in other domains. The morphologically privileged position of the head, however, is independently motivated and is connected to a whole family of constraints involved in Positional Faithfulness theory (Beckman 1998). This also suggests that factors found to be relevant in blend formation are cross-linguistically relevant. Further studies may find head faithfulness effects in blends from many different languages.

Since head faithfulness has been shown to influence blend prosody, a new analysis of the segmental contributions of source words in Japanese may also indicate that it is subject to head faithfulness. While Kubozono (1989) his findings were based off of a relatively small corpus and he did not address the possibility that headed blends could behave differently than coordinating blends. While many other studies of positional faithfulness involve prosodic faithfulness to morphological categories (Smith 2011, Alderete 2001, Revithiadou 1999), Shaw's (2013) study on English blends found that segment structure was also influenced by head faith. Also, a study on other types of positional faithfulness including noun faithfulness and proper noun faithfulness has found a slightly larger effect for segments over prosody (Moreton et al. in preparation).

Next, this study found that blends are just as faithful (if not more so) to the unaccentedness of their source words as they are to accentedness. In terms of OT this means that a DEP constraint is just as highly ranked as the MAX and NOLAPSE constraints.

Interestingly, this is contrary to the pattern found in Kubozono (2008) whereby compounds may be faithful to the non-final accents of their second source words but not to their unaccentedness. On the other hand, this may be more consistent with the Ito & Mester (2012) account where compounds may be unaccented if their second words are unaccented and longer than four moras. The fact that these three faithfulness constraints are treated as being equally relevant according to the findings of this study provides some support for the Ito & Mester (2012) account of compounds.

These findings indicate that, despite their fundamental similarities, blending is a distinct word formation process from both compounding and reduced compounding. Under the assumption that each of these processes can be explained with the same grammar and one set of constraints, this has several implications for the morphological structure of blends. Firstly, this study indicates that the internal structure of blends is different from that of reduced compounds. If the Ito & Mester (2012) account is to be believed then each morpheme minimally projects its own head thus resulting in unaccented words. They have used this claim to account for both unaccentedness in reduced compounds as well as the vast number of unaccented native words which tend to be smaller and morphologically complex. However, the fact that blends do not follow this pattern indicates that they do not, in fact, possess separate morphemes. Similarly, both analyses of compound accent involve an ALIGN constraint which is satisfied when accent is placed near a word-internal morpheme boundary. Once again, the fact that blends do not follow the accentual pattern of compounds indicates that they do not have an internal morpheme boundary. This idea is further supported by the fact that coordinating blends generally followed the simplex

accent pattern of the language and indicates that blends are unique among word formation processes of Japanese.

Finally to summarize, this thesis has shown that accent in Japanese blends is determined by head faithfulness. Additionally, it has provided support for the claim that head faithfulness exists and that it is an important factor in blend formation. Further, this has shown that factors affecting blends are cross-linguistically relevant. It has also demonstrated that blends are an entirely different word formation than both compounds and reduced compounds in Japanese. Lastly, it has shown that unlike the two other word formations, blends do not have an internal morpheme boundary. Further investigations on blends in Japanese may reveal that they demonstrate segmental faithfulness to heads as well as prosodic faithfulness and investigations into blends of different languages may show that they share this pattern of head faithfulness.

APPENDIX A

CORPUS SURVEY ITEMS

Each item is presented with the first line in each group being the Japanese text as written in the survey, the second line being a transcription of the Japanese words with accent as provided from the native speaker for source word 2 and the third line being a translation into English. Items are sorted first by headedness, second by accent placement and third alphabetically.

Blend	Hd	Source Words		Definition
ポテトルネード	1	ポテト	トルネード	トルネードみたいなポテトのスナック
*poteto*runeedo		*poteto*	torunéedo	
		'potato'	'tornado'	'a potato snack that looks like a tornado'
オッチェン	1	男	メッチェン	女性のような男性
*o*ttyen		*otoko*	méttyen	
		'man'	'girl'	'a girl-like man'
ママゴン	1	ママ	ドラゴン	とても厳しい母親
*mama*gon		*mama*	dóragon	
		'mother'	'dragon'	'a very strict mother'
アグフレ	2	アグリカルチャー	インフレ	農産物の価格上昇
*agu*fure		*agurikarutyaa*	infure¯	
		'agriculture'	'inflation'	'the increase in prices in agricultural products'
エログ	2	エロチック	ブログ	色っぽいブログ
*e*rogu		*erotikku*	burogu¯	
		'erotic'	'blog'	'a sexy blog'
ヘタリア	2	ヘタレ	イタリア	第二次大戦時の能がないイタリア軍
*heta*ria		*hetare*	itaria¯	
		'incompetence'	'Italy'	'Italy's incompetent army during WWII'
ホモ達	2	ホモ	友達	ホモセクシュアル同士の友達
*homo*dati		*homo*	tomodati¯	
		'homosexual'	'friend'	'a homosexual friend'
高校	2	高等	学校	高等学校の略称
*koo*koo		*kootoo*	gakkoo¯	
		'high grade'	'school'	'an abbreviation of high school'
ことな	2	子供	大人	子供のような心を持った大人
*k*otona		*kodomo*	otona¯	
		'child'	'adult'	'an adult with the heart of a child'
ネカマ	2	ネット	オカマ	ネット上で男性が女性を装うこと及び装っている人
*ne*kama		*netto*	okama¯	
		'internet'	'effeminate man'	'a man on the internet who is pretending to be a woman'

ペライド	2	ペラペラ	プライド	外国語が自然に話せることの自慢
peraido		*perapera*	puraido¯	
		‘fluent’	‘pride’	‘the pride of being able to speak a language fluently’
チビタリア	2	ちび	イタリア	「Axis Powers ヘタリア」に登場するイタリアという人物のチビバージョン
tibitaria		*tibi*	itaria¯	
		‘little person’	‘Italy’	‘a Chibi version of a man named Italy that appeared on the show "Axis Powers Hetalia"’
あさむけ	2	朝	寒気	寒い朝の寒気
**asamuke*		*asa*	samuké	
		‘morning’	‘chills’	‘chills on a cold morning’
かんじびき	2	感じ	字引	漢字の辞書
**kanzibiki*		*kanzi*	zibikí	
		‘kanji’	‘dictionary’	‘a Kanji dictionary’
まつぼみ	2	松	蕾	松の蕾
**matubomi*		*matu*	tubomí	
		‘pine’	‘bud’	‘a pine bud’
おきながめ	2	沖縄	眺め	沖縄の美しい景色
**okinagame*		*okinawa*	nagamé	
		‘Okinawa’	‘scenery’	‘the beautiful scenery of Okinawa’
せいじぞく	2	政治	地獄	政治に関わる地獄
**seizigoku*		*seizi*	zigokú	
		‘politics’	‘hell’	‘the hell involved in politics’
写メラマン	2	写メール	カメラマン	スマートフォンに付属されているカメラ機能で写真を撮る人
syameraman		*syameeru*	kameráman	
		‘picture text’	‘camera man’	‘someone that takes pictures using the camera feature on a smartphone’
チャリダー	2	チャリ	ライダー	自転車に乗る人
tyaridaa		*tyari*	ráidaa	
		‘bicycle’	‘rider’	‘a person who rides a bike’
アベノミクス	2	安倍	エコノミクス	安倍晋三首相の経済政策
abenomikusu		*abe*	ekonomíkusu	
		‘Abe’	‘economics’	‘Prime Minister Abe Shinzo's economic policy’
ジャビット	2	ジャイアンツ	ラビット	野球チームのジャイアンツのマスコット
zyabitto		*zyaiantu*	rábitto	
		‘giant’	‘rabbit’	‘the mascot of the Giants baseball team’
コミケット	2	コミック	マーケット	漫画を売買する特別なマーケット
komiketto		*komikku*	máaketto	
		‘comic’	‘market’	‘a special market where you can buy and sell manga’
ウァシュレット	2	ウァシュ	トイレット	温水洗浄便座
wasyuretto		*wasyu*	tóiretto	
		‘wash’	‘toilet’	‘a toilet that cleans with warm water’

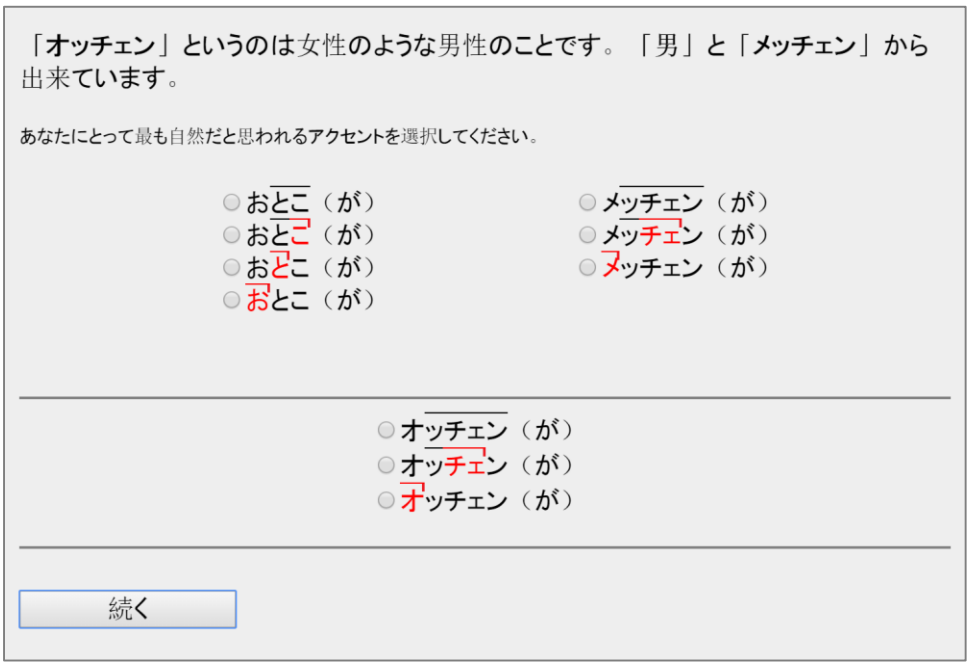
焼きもろこし	2	焼き	トウモロコシ	焼いたトウモロコシ
*yaki*morokosi		*yaki*	toomórokosi	
		‘fried’	‘corn’	‘baked corn’
ゴジラ	b	ゴリラ	くじら	ゴリラのような強さとクジラのような形をした映画の怪獣の名称
*go*zira		*gorira*	kuzira¯	
		‘gorilla’	‘whale’	‘the name of a movie monster with the strength of a gorilla and the shape (size) of a whale’
おしゃばり	b	お喋り	出しゃばり	おしゃべり且つ出しゃばりな人のこと
*osyab*ari		*osyaberi*	desyabari¯	
		‘chattering’	‘being intrusive’	‘a person who is both chatty and intrusive’
ピアニカ	b	ピアノ	ハーモニカ	ピアノとハーモニカのような楽器
*pia*nika		*piano*	haamonika¯	
		‘piano’	‘harmonica’	‘an instrument that is a mix of a piano and a harmonica’
埼京	b	埼玉	東京	東京都から埼玉県までを結ぶ運転系統
*sai*kyoo		*saitama*	tookyoo¯	
		‘Saitama’	‘Tokyo’	‘a route connecting Tokyo and Saitama’
アパシキ	b	アパート	屋敷	屋敷にある集合住宅
**apasiki*		*apaato*	yasikí	
		‘apartment’	‘mansion’	‘apartments that are in a 'yashiki' mansion’
バホ	b	ばか	アホ	愚か者
*b*aho		*baka*	ahó	
		‘idiot’	‘idiot’	‘a fool’
ダスキン	b	ダストクロス	雑巾	掃除会社の名前
*dasu*kin		*dasutokurosu*	zookín	
		‘dustcloth’	‘dustcloth’	‘the name of a cleaning company’
ゴパン	b	ご飯	パン	米パンを作るホームベーカリー機器
*go*pan		*gohan*	pán	
		‘rice’	‘bread’	‘a home bakery device that makes rice bread’
かなげき	b	悲しみ	嘆き	悲しい気持ち
**kanageki*		*kanasimi*	nagekí	
		‘sadness’	‘grief’	‘a sad feeling’
もんやき	b	文句	ぼやき	害を受けたことに対する不平
**monyaki*		*monku*	boyakí	
		‘complaint’	‘complaint’	‘a complaint against harm’
うっぷらみ	b	うっぷん	恨み	心にたまった怒りや不満
**uppurami*		*uppun*	uramí	
		‘grudge’	‘resentment’	‘anger and frustration accumulated in the heart’
ビニロン	b	ビニール	ナイロン	ビニールとナイロンのような人工的な素材
*bini*ron		*biniiru*	náiron	
		‘vinyl’	‘nylon’	‘an artificial material that is like vinyl and nylon’

ムック *m*ukku	b	マガジン *magazin* ‘magazine’	ブック búkku ‘book’	マガジンとブックのミックス ‘a combination of a magazine and a book’
ポマト *p*omato	b	ポテト *poteto* ‘potato’	トマト tómato ‘tomato’	ポテトとトマトのハイブリッド ‘a hybrid of a potato and a tomato’
旅テル *ryo*teru	b	旅館 *ryokan* ‘Japanese hotel’	ホテル hóteru ‘hotel’	旅館とホテルのような宿 ‘a mix of a Japanese hotel and a western hotel’
ファブレット *f*aburetto	b	フォン *fon* ‘phone’	タブレット táburetto ‘tablet’	電話とタブレットの機能がある機器 ‘a device that has features of both a phone and a tablet’

*Novel blends created for the corpus survey

APPENDIX B

CORPUS SURVEY SCREEN SHOT



REFERENCES

Akinaga, Kazue (1985). Kyootuugo no akusento [Accent of Standard Japanese]. In *NHK pronunciation and accent dictionary*, ed. NHK, 70-116. Tokyo: Nippon Hoso Shuppan Kyokai.

Alderete, John (2001). *Morphologically governed accent in optimality theory*. Outstanding dissertations in Linguistics. New York: Routledge.

Arndt-Lappe, Sabine, and Ingo Plag (2012). "The role of prosodic structure in the formation of English blends." *English Language and Linguistics* 17.

Bat-El, Outi & Evan-Gary Cohen (2012). Stress in English blends: A constraint-based analysis. In Vincent Renner, Francios Maniez & Pierre Arnaud (eds.), *Cross-disciplinary perspectives on lexical blending*, 193-212. Berlin: De Gruter Mouton.

Becker, Michael & Jonathan Levine (2014). Experigen – an online experiment platform. Available at http://becker.phonologist.org/experigen.

Beckman, Jill (1998). *Positional faithfulness*. University of Massachusetts Amherst dissertation.

Cannon, Garland (1986). Blends in English word formation. *Linguistics* 24. 725-753.

Gries, Stefan Th. (2004a). Isn't that *fantabulous*? How similarity motivates intentional morphological blends in English. In Michel Achard & Suzanne Kemmer (eds.), *Language, culture, and mind*, 415–428. CSLI Publications.

Gries, Stefan Th. (2004b). Shouldn't it be *breakfunch*? A quantitative analysis of blend structure in English. *Linguistics* 42(3). 639–667.

Gries, Stefan T. (2012). Quantitative corpus data on blend formation: psycho- and cognitive-linguistic perspectives. In: Vincent Renner, François Maniez and Pierre J. L. Arnaud (eds.), Cross-disciplinary perspectives on lexical blending, pp. 145--167. Berlin: Mouton de Gruyter.

Ito, Junko & Armin Mester (1992). Weak layering and word binarity. Ms. University of California, Santa Cruz.

Ito, Junko & Armin Mester (2007). Prosodic Adjunction in Japanese Compounds. In *Formal Approaches to Japanese Linguistics 4*. Cambridge, Massachusetts: MIT Department of Linguistics and Philosophy. 97-111

Ito, Junko & Armin Mester (2012). Unaccentedness in Japanese. Ms. University of California, Santa Cruz.

Ito, Mitsu (2011). Konseigo nitsuite [On Blends]. Kyushu International University studies of liberal arts 17(3), 1-15.

Kawahara, Shigeto (2013). The phonology of Japanese accent. In The Handbook of JapaneseLanguage and Linguistics: Phonetics and Phonology, Haruo Kubozono, ed., Mouton.

Kubozono, Haruo (1989). The Mora and Syllable Structure in Japanese: Evidence from Speech Errors. *Language and Speech* 249-278.

Kubozono, Haruo (1990). Phonological constraints on blending in English as a case for phonology-morphology interface. *Yearbook of Morphology* 3. 1-20.

Kubozono, Haruo (1999). Mora and Syllable. *The Handbook of Japanese Linguistics*. Tsujimura Natsuko, eds., Blackwell Publishers, 31-61.

Kubozono, Haruo (2002a). Prosodic structure of loanwords in Japanese: syllable structure, accent and morphology. The Journal of the Phonetics Society of Japan 6 (1), 79-97.

Kubozono, Haruo (2006). Where does loanword prosody come from? A case study of Japanese loanword accent. *Lingua* 116(7).

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

McCarthy, John, and Alan Prince (1993). Generalized Alignment. In *Yearbook of Morphology*, 79-153.

McCarthy, John, and Alan Prince (1994). The emergence of the unmarked: Optimality in prosodic morphology. NELS 24, vol.2, 333-379.

McCawley, James D. (1968). *The Phonological Component of a Grammar of Japanese*. The Hague: Mouton.

Moreton, Elliott, Jennifer Smith, Katya Pertsova, Rachel Broad, and Brandon Prickett (In preparation). Emergent Faithfulness in novel English blends.

Namiki, Takayasu (2001). *Further Evidence in Support of the Righthand Head Rule in Japanese*. Issues in Japanese Phonology and Morphology

Plag, Ingo (2006). *Word-formation in English*. Cambridge University Press.

Poser, William (1984). *The phonetics and phonology of tone and intonation in Japanese* (Doctoral dissertation, Massachusetts Institute of Technology).

Poser, William (1990). Evidence for foot structure in Japanese. *Language* 66:78-105.

Prince, Alan, and Paul Smolensky (1993/2004). *Optimality Theory: Constraint interaction in generative grammar.* Technical report, Rutgers University, New Brunswick, NJ, and University of Colorado, Boulder. [ROA-537] Revised version, Oxford: Blackwell (2004).

Roon, Kevin (2006). Stress in Russian compound nouns: head dominance or anti-faithfulness? In J. E. Lavine, S. Franks, M. Tasseva-Kurktchieva & H. Filip (eds.), *Proceedings of FASL (Formal Approaches to Slavic Linguistics)*, vol. 14, 319-300. Michigan Slavic Publications.

Revithiadou, Anthi (1999). *Headmost accent wins: Head dominance and ideal prosodic form in lexical accent systems:* University of Leiden dissertation.

Shaw, Katherine (2013). Head faithfulness in lexical blends: a positional approach to blend formation. MA thesis, Department of Linguistics, University of North Carolina, Chapel Hill.

Shaw, Katherine E., Andrew M. White, Elliott Moreton, and Fabian Monrose (2014). Emergent faithfulness to morphological and semantic heads in lexical blends. In John Kingston, Claire Moore-Cantwell, Joe Pater, and Robert Staubs (eds.), Proceedings of the 2013 Meeting on Phonology. Washington, DC: LSA.

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.

Smith, Jennifer L. (2011). Category-specific effects. In Marc van Oostendorp, Colin Ewen, Beth Hume, & Keren Rice (eds.), *The Blackwell Companion to Phonoloogy*, 2439-2463. Wiley-Blackwell.

Williams, Edwin (1981). "On the notions 'lexically related' and 'head of a word.'" *Linguistic Inquiry* 12,245-274.