

HEAD FAITHFULNESS IN LEXICAL BLENDS: A POSITIONAL APPROACH TO BLEND FORMATION

Katherine Shaw

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
2013

Approved by:

Elliott Moreton

Jennifer Smith

Katya Pertsova

ABSTRACT

KATHERINE SHAW: Head Faithfulness in Lexical Blends: A Positional
Approach to Blend Formation
(Under the direction of Elliott Moreton)

This thesis applies Positional Faithfulness theory (Beckman 1998) to the problem of lexical blending in English. Lexical blends, like *brunch* or *motel*, contract multiple source words into a single lexical item shaped by competing sets of phonological and psycholinguistic constraints. Existing studies of blend structure (e.g., Bat-El & Cohen 2012, Gries 2004a,b) focus on the contributions of each source word relative to their linear order, positions that have little relevance outside of blend formation.

I present both corpus and experimental data to argue that previously observed right-word faithfulness effects are actually due to head faithfulness (Revithiadou 1999). This has two major implications: it provides evidence for the existence of positional faithfulness and of head faithfulness in particular; second, it demonstrates that blend formation is subject to independently motivated, broadly applicable constraints. In addition, the discovery of left-headed blends in the corpus argues that blending is a distinct process from compounding.

ACKNOWLEDGEMENTS

It is often said that many hands make light work, and many hands deserve recognition for making this process a much lighter undertaking. I am deeply indebted to my advisor, Elliott Moreton, for shaping this thesis since its inception into a far more interesting project than I would have devised alone, for balancing enthusiasm with patience, and for teaching me to think through problems from a wider range of perspectives. Jen Smith and Katya Pertsova have provided helpful feedback as members of my thesis committee, and lab meeting suggestions from Fabian Monroe and Andy White have also improved the project tremendously. Thanks also go to Josiah McCoy for his help with corpus assembly, and to Chris Wiesen for an admirable job implementing and explaining the statistical analyses employed.

I'm grateful to the family, friends, and fellow grad students who volunteered their time as pilot subjects and pronunciation consultants, or who at least forgave me for conscripting them. I'm grateful to everyone who not only participated in these experiments but recruited others to participate as well. I'm grateful for constant encouragement from parents, friends, colleagues, and my extended family at Grace Community Church throughout this process. I'm especially grateful to Amy Reynolds for generously sharing ideas, advice, experience, and cupcakes, all at just the right moments.

An exhaustive list of acknowledgements would easily add a chapter or three to an already unwieldy document, so let me summarize: to the many individuals whose contributions are reflected in this thesis—or in the mere fact of its completion—thank you. All of you are godsend, and I'm grateful for every one of you.

TABLE OF CONTENTS

List of Tables	vi
List of Figures	vii
1 Introduction	1
1.1 Positional Faithfulness	3
1.2 Blend Formation	6
1.3 A Note on Representations	9
2 Stress Placement	12
2.1 Predictions	12
2.1.1 Head faithfulness predictions	13
2.1.2 Linear order predictions	15
2.1.3 Compound stress predictions	17
2.2 Stress Placement in Attested Blends	19
2.2.1 Blends with equal length source words	21
2.2.2 Full corpus analysis	22
2.3 Stress Placement in Novel Blends	24
2.3.1 Methodology	25
2.3.2 Results	29
2.4 Discussion	31
2.4.1 Headedness as a determinant of blend structure	32
2.4.2 Unpacking FAITH(STRESS)	33
2.4.3 Blending as distinct from compounding	35
3 Segmental Content	37

3.1	Predictions	38
3.1.1	Head faithfulness predictions	38
3.1.2	Linear order predictions	40
3.2	Segment retention in attested blends	41
3.2.1	Source words of equal length	41
3.2.2	Full corpus	43
3.3	Segment Deletion in Novel Blends	45
3.3.1	Methods	46
3.3.2	Results	50
3.4	Discussion	52
4	Conclusion	54
4.1	Positional faithfulness implications	54
4.2	Blend formation implications	55
4.3	Concluding Remarks	57
A	Left-Headed Blends	58
	References	60

LIST OF TABLES

1.1	Linear blend structures	1
1.2	Nonlinear blend structures	2
1.3	Privileged positions	3
2.1	Blend categories excluded from corpus analysis	20
2.2	Length-controlled blends by head and stress preservation	21
2.3	Left-stressed blends	22
2.4	All blends by head and stress preservation	23
2.5	Logistic regression coefficients: Stress placement	23
2.6	Stress placement experiment stimuli	26
2.7	Stress placement results by item	30
2.8	Blends with secondary stress	34
3.1	Length-controlled blends by head and segment retention	42
3.2	Length-controlled blends by head and full word retention	42
3.3	All blends by head and segment retention	43
3.4	Multiple regression coefficients: Segment retention	44
3.5	Blends with attested variation	45
3.6	Segment deletion experiment stimuli	48
3.7	Segment deletion results by item	51
A.1	Left-headed blends	59

LIST OF FIGURES

2.1	A typical trial: Stress placement	28
2.2	Stress placement results by subject	31
3.1	A sample rating calculation	47
3.2	A typical trial: Segment deletion	49
3.3	Segment deletion results by subject	52

CHAPTER 1

INTRODUCTION

Despite its frequent use in lighthearted wordplay, lexical blending is a complex morphophonological process that combines elements of subtractive and concatenative morphology, shaped by competing sets of phonological and psycholinguistic constraints. Prototypical blends like *brunch* and *motel* contract phonological material from two source words into a single output through a combination of truncation and overlapping. In some cases, the blend's source words are phonologically similar enough that they overlap without deleting any material. This yields blends like *examnesia* and *kangarooster* that contain both source words in their entirety. More often, one or both source words appears as a splinter, a truncated form that contains enough phonological material to identify the original source word (Lehrer 1996). Tables 1.1 and 1.2 give examples of each of the blend structures found in the literature. Although the nonlinear blends in Table 1.2 are all attested, the linear blends in Table 1.1 are by far the more common and more uniform class; they consequently form the

Type	Overlap	No Overlap
Splinter + Splinter	<i>motel</i> <i>motor</i> + <i>hotel</i>	<i>brunch</i> <i>breakfast</i> + <i>lunch</i>
Word + Splinter	<i>opinionnaire</i> <i>opinion</i> + <i>questionnaire</i>	<i>jazzetry</i> <i>jazz</i> + <i>poetry</i>
Splinter + Word	<i>decathlete</i> (<i>decathlon</i> + <i>athlete</i>)	<i>jamocha</i> (<i>java</i> + <i>mocha</i>)
Word + Word	<i>examnesia</i> <i>exam</i> + <i>amnesia</i>	(compounds)

Table 1.1: Linear blend structures

Type	Example	
Embedded Blend	<i>adver<u>te</u>asement</i>	<i>ad<u>ver</u>tisement + <u>te</u>ase</i>
Orthographic Blend	<i>sin<u>e</u>ma</i>	<i>sin + <u>c</u>inema</i>
Three Source Words	<i><u>b</u>runner</i>	<i>breakfast + <i>lunch</i> + <u>dinner</u></i>
Metathesized Blends	<i>smo<u>k</u>olotive</i>	<i>smo<u>k</u>e + <u>l</u>ocomotive</i>

Table 1.2: Nonlinear blend structures

basis of the present discussion.

The relative contributions of each source word are a key question in blend formation studies. Blends that undergo truncation do not delete material at random, but it is far from obvious what constraints determine the resulting blend’s phonological form. Some factors are phonological: phonological similarity determines how much the two source words can overlap (e.g., Gries 2004a), and source word length determines what proportion of the source word can be deleted without sacrificing the ability to decode the blend (Kau-nisto 2000). When these factors are insufficient—when two source words have no segments in common, for instance, or when they are the same length—authors look at the contributions of the left and right words (e.g., Bat-El & Co-hen 2012, Gries 2004a,b, Kubozono 1990). However, neither “left” nor “right” is a privileged position in any domain beyond blend formation.

In this thesis, I propose that blend formation should be analyzed in terms of existing positional faithfulness constraints. This theoretical goal serves two purposes: enriching our knowledge not only of blending but also of positional faithfulness. First, it contributes to our understanding of blend formation by illuminating not only the mechanics of blending but their relationship to other morphological phenomena and to the linguistics literature more broadly. Second, it demonstrates that positional faithfulness constraints are active in English despite its lack of strong positional effects outside of unstressed vowel reduction, arguing for their inclusion in the constraint set.

The remainder of this chapter provides relevant background on both po-

Privileged Position	Non-Privileged Position	
Initial syllables	Non-initial syllables	Beckman 1998
Stressed syllables	Unstressed syllables	Chomsky & Halle 1968
Onsets	Codas	Lombardi 1999
Roots	Affixes	Alderete 2001
Heads	Non-heads	Revithiadou 1999
Lexical morphemes	Functional morphemes	Casali 1996
Nouns	Verbs	Smith 2011
Proper nouns	Common nouns	Jaber 2011

Table 1.3: Privileged positions

sitional faithfulness (§1.1) and blend formation (§1.2). In the following chapters I argue that blending is subject to head faithfulness constraints, positional faithfulness constraints that privilege morphological heads. Chapter 2 demonstrates the effects of head faithfulness on stress placement in blends; Chapter 3 extends the analysis to blends' segmental content. In Chapter 4 I summarize the head faithfulness effects observed and present suggestions for further research.

1.1 Positional Faithfulness

Positional faithfulness theory states that phonetically and psycholinguistically prominent positions are privileged over less prominent positions and consequently are subject to more stringent faithfulness requirements (e.g., Alderete 2001, Beckman 1998, Lombardi 1999). This insight, formalized in Optimality Theory as part of faithfulness constraint schemata that target particular privileged positions, facilitates a unified analysis of a variety of phenomena in a wide range of languages. Some of these positions are listed with their non-privileged counterparts in Table 1.3.

Phonologically privileged positions are often resistant to contrast-neutralizing phenomena like vowel reduction or assimilation (Beckman 1998); in other cases segments in privileged positions serve as the trigger for processes like cluster assimilation (Lombardi 1999) or vowel harmony (Beckman 1997,

1998). Morphologically privileged positions—heads (Revithiadou 1999, Roon 2006), roots (Alderete 2001, McCarthy & Prince 1995), and privileged lexical categories (Casali 1996, Jaber 2011, Smith 2011)—are more relevant to the present study. In particular, Revithiadou (1999) demonstrates that morphological heads are privileged over non-heads in Greek and Russian stress assignment. In these languages, both roots and suffixes can carry lexical accent, realized as stress. In words with multiple accents, however, only the accent on the head receives stress. This is illustrated in (1), from Revithiadou’s example 10 (1999, p.180):

(1) stafíð-ón raisin-GEN.pl

/stafíð- + -ón/	HEADFAITH(LA)	FAITH(LA)
☞ a. stafíðon		*
b. stafíðón	*!	*

For Revithiadou, roots are heads of words that have inflection only; derivational affixes, when present, are heads by virtue of determining the word’s syntactic and semantic features. When an accented derivational affix is present, it receives stress even when attached to an accented root:

- (2) a. papayál-os parrot-NOM.sg
b. papayal-ák-u parrot-DIM-GEN.sg

Alderete (2001), in contrast, examines lexical accent in terms of root rather than head faithfulness, identifying root faithfulness effects in Cupeño, Russian, and Japanese. Roon (2006) notes that root faithfulness is uninformative when confronted with words like compounds, composed of multiple roots. Nevertheless, he applies both Revithiadou’s and Alderete’s accounts to compound stress in Russian, finding that head faithfulness to the surface form of the words in a compound correctly predicts the compound stress data. This

is the approach taken in this thesis: while applying Revithiadou's head faithfulness proposal to English blend formation, I assume that the input to which blends are faithful is the surface, not underlying, representation of their source words.

Other morphologically privileged positions are not directly addressed here, though they are certainly germane to the exploration of positional faithfulness in blend formation. Given the head faithfulness effects outlined in subsequent chapters, it is not unreasonable to look for blends that display faithfulness to nouns over verbs (Smith 2001, 2011)¹ or to proper nouns over common nouns (Jaber 2011), and indeed the methods used here could easily be adapted for such an investigation.² The majority of the phenomena influenced by these morphological positions are prosodic, with effects appearing in phenomena such as stress and accent placement, tone, and syllable structure (Smith 2011). This need not always be the case, though, as Chapter 3 demonstrates head faithfulness effects in determining a blend's segmental content in addition to the head-faithful stress placement presented in Chapter 2.

Although recognized positional effects in English are limited to unstressed vowel reduction, the universality of the constraint set predicts that positional faithfulness constraints are present in English as well. This thesis presents evidence that head faithfulness is an active pressure in lexical blend formation, demonstrating that English does display faithfulness to morphological positions and offering a new phenomenon through which to explore other aspects of PF theory.

¹ Many cases of noun privilege are augmentation processes—that is, positional markedness, not positional faithfulness. However, if nouns are indeed privileged, we would expect them to be valid targets of positional faithfulness constraints as well.

² Although there are also documented cases of lexical morphemes privileged over functional morphemes (Casali 1996), the fact that blends comprise several lexical morphemes means blending is no more suited to probe the relationship between lexical and functional morphemes than it is to address the question of root faithfulness.

1.2 Blend Formation

Recent work on blends falls into two major lines of inquiry: the mechanics of blend formation and the relationship between blending and other morphological processes like compounding and abbreviation. The present study addresses both of these. First, it demonstrates that head faithfulness accounts for both the corpus data and the experimental results more accurately than existing accounts that rely on the source words' linear order. Second, it identifies two clear but previously unnoted distinctions between blending and compounding.

A blend's phonological form is determined in part by purely phonological considerations; both stress placement and segmental content are influenced, for example, by the relative length of the two source words. Blends tend to take their stress from the longer source word, measured in number of syllables, (e.g., Bat-El & Cohen 2012, Cannon 1986, Gries 2004a,b) but retain a larger percentage of the shorter word's segments so that listeners can identify the words used to form the blend (Gries 2004a,b, Kaunisto 2000). When the two source words are the same length, these studies look at what material typically comes from the left word and what typically comes from the right. In these cases, blends frequently preserve the stress of the right-hand word (Bat-El 2006, Bat-El & Cohen 2012, Gries 2004a,b), even when its stressed segments have been deleted (Arndt-Lappe & Plag 2012); the right word also seems to contribute a larger percent of its segments (Gries 2004a,b). Because most blends are right-headed (Kubozono 1990), these effects are plausible under a head faithfulness account as well. However, there are a substantial number of non-headed blends, and even a few that are left-headed, so the two analyses make testably different predictions in many cases; chapters 2 and 3 present evidence that head faithfulness, not linear order, makes predictions consistent with the facts of blending.

The relationship between blending and other morphological processes is also not obvious. Most authors agree that blending is a distinct process from derivation with bound morphemes (e.g., Algeo 1977, Frath 2005, Lehrer 2007), so forms like *meritocracy*, for example, are more aptly analyzed as *merit* + *-ocracy* than as a blend of *merit* and *aristocracy*. However, it is often difficult to distinguish between the two. The truncated portions of words (“splinters”) that make up blends may be full morphemes (Algeo 1977) or not (Lehrer 2007), and those that are not morphemic may become so if they are used frequently (Frath 2005, Hamans 2010, Lehrer 2007). Examples using these morphologized splinters are listed in (3) below; while the earliest of these forms may be blends, those coined later are likely derived rather than blended.

(3) Splinters morphologized through frequent blending³

a. (Water)- <u>gate</u>	b. (alco)- <u>(a)holic</u>	c. <u>Mc</u> -(Donald’s)
<i>Monicagate</i>	<i>workaholic</i>	<i>McMuffin</i>
<i>Fajitagate</i>	<i>shopaholic</i>	<i>McJobs</i>
<i>Benghazigate</i>	<i>controlaholic</i>	<i>McGarbage</i>

Many authors claim that blends are morphologically compounds that have been phonologically reduced. Arcodia & Montermini (2012), for example, find that blending in Russian and Chinese strongly resembles compounding, based on the strong phonological regularities exhibited by the two processes. However, all of the blends they consider contain two left splinters (i.e., the beginning of both source words) where typical blends contain a left splinter and a right splinter (i.e., the beginning of the first word and the end of the second). Arcodia & Montermini’s examples are analogous to the English examples in (4):

³ All except *Benghazigate* from Lehrer 2007.

- (4) a. *agitprop*: *agitation* + *propaganda*
- b. *sitcom*: *situational* + *comedy*

In English, these reduced compounds behave differently from typical blends: they do not overlap their source words, and they follow the compound stress rule (Chomsky & Halle 1968, Liberman & Prince 1977) rather than blend stress assignment (see section 2.4.3 for more discussion). Consequently, many authors either background them (e.g., Lehrer 2007) or exclude them entirely (e.g., Bat-El 2006, Bat-El & Cohen 2012). Plag (2003) and Mattiello (2013) adopt a semantic criterion that effectively considers all headed blends to be reduced compounds by stipulating that a blend has an “autonomous sense which is entirely retained in the final form” where reduced compounds display “a composite meaning, often of the type determinant-determinatum” (Mattiello 2013). Nearly all explicit definitions of blending still refer to it as a type of compounding (e.g., Bat-El 2006, Lehrer 2007).

Very few studies have investigated blends’ morphological structure, and those that have consider a limited subset of blend types. Kelly (1998) acknowledges that blends, like compounds, can be headed or non-headed; his analysis extends patterns of conjunct ordering to blend source word order, and so he considers only coordinating blends. Kubozono (1990), on the other hand, looks specifically at headed blends to find that they, like English compounds, are always right-headed. As mentioned above, however, this generalization does not hold in a larger dataset: while the majority of headed blends are indeed right-headed, section 2.4 presents a set of left-headed blends, with a reading not available to the equivalent compound. This, along with the stress facts presented in Chapter 2, argues that blending is not simply a form of compounding with truncation, despite the assumptions in much of the literature.

1.3 A Note on Representations

There are two key aspects of blend representation that require clarification: headedness and stress assignment. Throughout this thesis, I determine headedness morphologically where possible; when a blend's source words belong to different parts of speech, the source word that determines the blend's lexical category is considered the head. Thus a blend like *gawkward* ('awkward or stupid', from *gawk* + *awkward*) is right-headed because its morphological behavior is dictated by its right source word, while a blend like *homonymble* ('a clever pun that plays on distinct meanings of the same word', from *homonym* + *nimble*) shares the lexical category of its left source word and is consequently left-headed. Most blends, however, are blends of two nouns and cannot be classified in this manner. In these cases, head is a semantic matter: a *cafetorium* is both a *cafeteria* and an *auditorium* and is therefore non-headed, but a *goditorium* (slang term for a church) is right-headed because it is only an *auditorium* and not typically a *god*.

In analyzing stress assignment, previous studies have employed a narrow standard to determine whether a blend has the same pattern as its source words: a blend's stress matches a source word's when the two words have the same number of syllables, with stress on the same syllable in both (Bat-El & Cohen 2012, Gries 2004b). To maximize the similarity observed between blends and their source words, I follow these analyses in treating stress as a suprasegmental feature, but I assume that is matched by alignment rather than by identity. Many blends consist of a monosyllable superimposed on the stressed syllable of the longer source word, like *bellcony* (*bell* + *balcony*, *jollybean* (*jolly* + *jellybean*), or *yellocution* (*yell* + *elocution*); treating stress as a segmental feature would categorize *yellocution* as preserving both stresses while *bellcony* and *jollybean* preserve only the first. This arbitrarily distinguishes blends that are very much alike in composition and clearly under-

states the degree of similarity between the blends and their source words, so I consider stress abstracted away from the segments that host it.

The rationale for implementing stress matching as alignment is similar. Requiring identical stress patterns means that blends like the three above would match only the second word; the first has only one syllable, while the blend is trisyllabic, so the two are not identical. Although in these cases it suffices to say that both words have initial stress, other blends demonstrate that the stress of second source word should be compared in terms of distance from the end. A clear example is *abhorrible*, which preserves both words in their entirety, without moving, deleting, or reducing the stress in either word. Comparing the entire stress pattern of the blend to that of the source words misses this fact and categorizes *abhorrible* as deleting both source words' stress. Comparing both source word patterns from the left edge of the blend's similarly ignores the commonality between the second source word and the blend. The most informative comparison notes that both *abhor* and *abhorrible* have stress on the second syllable, while *abhorrible* and *horrible* both have antepenultimate stress. Accordingly, in the analyses that follow, I identify which source word stresses appear in the blend by aligning the left source word from the left edge and the right source word from the right edge.⁴

The remainder of the thesis presents evidence that blending is not compounding and that it is subject to independently motivated positional faithfulness constraints, in the form of head faithfulness. Chapter 2 uses stress placement to argue both in favor of positional faithfulness and against the conflation of blending and compounding by comparing the predictions of Revithiadou's (1999) head faith analysis of lexical accent, Bat-El & Cohen's (2012) model of blend stress, and the Compound Stress Rule (e.g., Chomsky & Halle 1968,

⁴ Not coincidentally, this also reflects the segmental properties of these blends, which preserve a left-aligned substring of the left source word and a right-aligned substring of the right source word.

Lieberman & Prince 1977). In Chapter 3 I extend the head faithfulness account to blends' segmental content, demonstrating that morphologically privileged positions appear in segmental constraints as well as prosodic. Chapter 4 summarizes the results and implications of the preceding chapters and suggests avenues for future research.

CHAPTER 2

STRESS PLACEMENT

Because existing head faithfulness accounts address the placement of lexical accent (Revithiadou 1999, Roon 2006), I begin by examining the role of head faithfulness in determining blend stress. Recall that existing analyses of stress placement in blends consider both the source words' length and their linear order to determine which source word's stress will appear in the blend (Bat-El & Cohen 2012, Gries 2004a). In this chapter, I demonstrate through both corpus and experimental data that a blend's phonological form is also affected by its morphological headedness. Section 2.1 outlines the predictions of three different stress assignment theories: positional faithfulness, linear order, and compound stress. Section 2.2 describes the corpus analysis, and section 2.3 presents the experimental results. Section 2.4 summarizes the findings from the preceding sections and discusses their theoretical implications.

2.1 Predictions

This section focuses on the three competing hypotheses concerning stress assignment in blends. The head faithfulness hypothesis (§2.1.1), modeled after Revithiadou (1999), probes the relationship between the morphological and phonological structure of a blend. The linear order hypothesis (§2.1.2), based on the existing body of blend studies, describes the tendencies found in blend corpora without grounding them in broader phonological theory. The compound stress hypothesis (§2.1.3) directly challenges both head faithfulness and linear order while exploring the relationship between blending and com-

pounding. I predict that head faithfulness affects blend stress both in the corpus and in experimental data, while compound stress affects only the group of *sitcom*-type “blends” composed of two left splinters. If head faithfulness affects stress assignment, then the previously observed linear order effects are likely derived from a combination of length and head faithfulness.

2.1.1 Head faithfulness predictions

Revithiadou’s (1999) head faithfulness account of lexical accent applies straightforwardly to blends: when one of the two source words functions as the head of a blend, the blend should preserve that word’s source pattern. This is enforced by the constraints in (5):¹

(5) Lexical Accent Constraints: Head Faithfulness

- a. $\text{MAX}(\text{STRESS})$: assign a violation for every stress in the input that does not have an output correspondent
- b. $\text{MAX}(\text{STRESS})_{\text{Head}}$: assign a violation for every stress in an input head that does not have an output correspondent
- c. $\text{IDENT}(\text{STRESS})$: assign a violation for every primary stress in the input whose output correspondent is a secondary stress
- d. $\text{IDENT}(\text{STRESS})_{\text{Head}}$: assign a violation for every primary stress in an input head whose output correspondent is a secondary stress

The $\text{MAX}(\text{STRESS})$ constraints penalize blends that delete either source word’s stress outright; the $\text{IDENT}(\text{STRESS})$ constraints penalize those that demote a source word’s primary stress to a secondary stress. For space, throughout this

¹ Because stress placement in blends follows the stress placement of at least one source word, $\text{DEP}(\text{STRESS})$ is not relevant.

section I use the combined constraints FAITH(STRESS) and FAITH(STRESS)_{Head} to penalize any change in stress between input and output:

(6) Combined Lexical Accent Constraints: Head Faithfulness

- a. FAITH(STRESS): assign a violation for every primary stress in the input that does not have an output correspondent with primary stress
- b. FAITH(STRESS)_{Head}: assign a violation for every primary stress in an input head that does not have an output correspondent with primary stress

Note that a violation of FAITH(STRESS)_{Head} entails a violation of FAITH(STRESS) and that a blend can only violate FAITH(STRESS) once per source word. Consequently, in blends with two source words, FAITH(STRESS)_{Head} will either assign no violations, fail to apply, or break a tie between candidates that each violate FAITH(STRESS) once. Blends with two source words simply do not provide a means to rank the two.² Blends with more than two source words are, however, exceedingly rare and are not included in any of the present analyses; therefore the inclusion of head faithfulness constraints is sufficient, regardless of the ranking between general and head faithfulness.

In blends that preserve both stresses, of course, neither constraint assigns violations to the winning candidate. In headed blends, when one source word's stress is deleted, FAITH(STRESS)_{Head} eliminates the candidate that preserves non-head rather than head stress:

² Given an appropriate blend with three source words, it would be possible to devise scenarios that do provide a ranking argument; for example:

$\sigma\acute{\sigma}\sigma + \sigma\acute{\sigma} + \acute{\sigma}\sigma\sigma_{\text{HEAD}}$	FAITH(STRESS) _{Head}	FAITH(STRESS)
a. $\sigma\acute{\sigma}\sigma$	*	*
b. $\acute{\sigma}\sigma\sigma$		**

(7)

$\acute{\sigma}\sigma + \sigma\acute{\sigma}_{\text{HEAD}}$	FAITH(STRESS) _{Head}	FAITH(STRESS)
☞ a. $\sigma\acute{\sigma}$		*
b. $\acute{\sigma}\sigma$	*!	*

In non-headed blends, FAITH(STRESS)_{Head} assigns no violations even when the blend deletes stress, leaving FAITH(STRESS) to distinguish between them. However, when the two source words have conflicting stress patterns, FAITH(STRESS) assigns one violation to each candidate:

(8)

$\acute{\sigma}\sigma + \sigma\acute{\sigma}$	FAITH(STRESS) _{Head}	FAITH(STRESS)
☞ a. $\sigma\acute{\sigma}$		*
☞ b. $\acute{\sigma}\sigma$		*

In non-headed blends, then, head faithfulness predicts stress assignment according to default stress rules. Alternately, the failure of head faithfulness to distinguish between the two may indicate that non-headed blends should display more variation than headed blends do.

2.1.2 Linear order predictions

In a linear order account, like that of Bat-El & Cohen (2012), the second source word should contribute its stress to the blend whenever length differences between source words do not dictate left stress. The appropriate combined constraints are listed in (9); as before, each of these includes the relevant MAX(STRESS) and IDENT(STRESS) constraints.

- (9) Lexical Accent Constraints: Linear Order (Bat-El & Cohen 2012)
- a. FAITH(STRESS)-L: assign a violation for every primary stress in the left source word that does not have a primary-stressed correspondent in the blend

- b. FAITH(STRESS)-R: assign a violation for every primary stress in the right source word that does not have a primary-stressed correspondent in the blend
- c. Ranking: FAITH(STRESS)-R \gg FAITH(STRESS)-L

Because left and right are the only positions referenced by these constraints, this predicts the same blend regardless of morphological context. For right-headed blends, linear order makes the same predictions as the head faithfulness analysis: the blend will exhibit the right-hand source word's stress pattern, at the expense of the left word stress if necessary:

(10)

$\acute{\sigma}\sigma + \sigma\acute{\sigma}_{\text{HEAD}}$	FAITH(STRESS)-R	FAITH(STRESS)-L
☞ a. $\sigma\acute{\sigma}$		*
b. $\acute{\sigma}\sigma$	*!	

In the linear order analysis, however, this has nothing to do with the right word's status as head. Non-headed blends are also predicted to show right-stress:

(11)

$\acute{\sigma}\sigma + \sigma\acute{\sigma}$	FAITH(STRESS)-R	FAITH(STRESS)-L
☞ a. $\sigma\acute{\sigma}$		*
b. $\acute{\sigma}\sigma$	*!	

This does not allow for variation, instead predicting a single output for every blend. It predicts that blends should follow the left word's stress only when the left word is longer, and that all blends should pattern together regardless of their morphological structure.

2.1.3 Compound stress predictions

Stress in compounds usually falls on the left element (e.g., Chomsky & Halle 1968, Cinque 1993, Liberman & Prince 1977). This is especially true of headed compounds; non-headed compounds, like *singer-sóngwriter*, have a greater tendency to be right-stressed than headed compounds like *bláckbird* do (Plag 2006). Accordingly, the compound stress rule could apply in a weak version or a strong version:

(12) Compound Stress Rule

- a. Strong CSR: assign a violation for every compound without primary stress on the left word
- b. Weak CSR: assign a violation for every headed compound without primary stress on the left word

Liberman & Prince (1977) and those following them integrate the Compound Stress Rule and the Nuclear Stress Rule governing stress placement in phrases by noting that the right-hand element of a phrase or compound is stressed *iff* it branches. This characterization is substantially harder to apply to blends: if the constraints assigned violations to blends that placed stress on the second source word, it is unclear whether this would penalize blends that preserve the stress of both source words, or whose stressed syllable contains segments from both source words. Here I use the versions in (12); if a more theoretically sound formulation penalizes full prosodic faithfulness to both source words, that fact provides more evidence that the CSR is either ranked below the relevant faithfulness constraints or not relevant to blending.

In most cases, these formulations of the CSR serve as inverses of the linear order and head faithfulness analyses, respectively.³ In a strong CSR analysis,

³ In cases where preserving the sequence of stressed and unstressed syllables found in a right head would place stress on the segments of the first source word, all four accounts

blends should be left-stressed regardless of their morphological structure. As long as the stressed syllable of the left word is not deleted, it assigns exactly the same violations as FAITH(STRESS)-L, which reverses the predictions of the linear order analysis:

(13)

$\acute{\sigma}\sigma + \sigma\acute{\sigma}$	S-CSR	FAITH(STRESS)
a. $\sigma\acute{\sigma}$	*	*
☞ b. $\acute{\sigma}\sigma$		*

A weak CSR analysis, on the other hand, predicts stress on the first source word when the blend is headed ((14a)), with no prediction when it is not ((14b)). This is a theoretically strange constraint, acting as a positional markedness constraint that reduces rather than augments the target position. This inverts the head faithfulness predictions by demanding that heads not be stressed:

(14) a. Headed blends

$\acute{\sigma}\sigma + \sigma\acute{\sigma}_{\text{HEAD}}$	W-CSR	FAITH(STRESS)
a. $\sigma\acute{\sigma}$	*!	*
☞ b. $\acute{\sigma}\sigma$		*

b. Non-headed blends

$\acute{\sigma}\sigma + \sigma\acute{\sigma}$	W-CSR	FAITH(STRESS)
☞ a. $\sigma\acute{\sigma}$		*
☞ b. $\acute{\sigma}\sigma$		*

Both the weak and strong forms of the CSR predict a large proportion of blends with left stress, among right-headed blends under the weak version and in all blends according to the strong version. This depends, however,

would predict the same form. This is far from a typical case, however, as it does not occur at all within my corpus.

not only on the CSR itself but on its applicability to blends. A lack of CSR effects criticizes not the CSR itself but its relevance in blending; failure to conform to these predictions likely indicates that blends are *not* morphologically compounds, and therefore the CSR does not apply at all.

2.2 Stress Placement in Attested Blends

In this section I test each of these hypotheses against a corpus of attested blends. Head faithfulness predicts that right-headed blends will match the stress of their right source word, while non-headed blends may display variation or default stress assignment; linear order predicts right stress across the corpus, and CSR predicts left stress. Previous studies have presented corpus data matching the linear order hypothesis (Gries 2004a,b), but not in a corpus of this size. The present corpus is one of the largest blend corpora analyzed to date, containing 1,387 blends from Thurner’s (1993) blend dictionary. Thurner’s collection contains closer to 1,600 terms; I excluded those that were onomatopoeic or made up of combining forms. There are 292 blends that appear in the corpus but were not used in analyses. Following most existing work, I consider only linear blends with two source words—those that begin with segments from the first source word then switch to segments from the second, with some medial segments optionally belonging to both source words. These are prototypical blends, like *brunch* and *motel*, and comprise 95.9% of the corpus. I also excluded brand names, which tend to rely on orthography more than other blends do (Cannon 1986) and may follow their own phonological principles; for example, in many brand names, only degemination distinguished the “blend” from a corresponding compound, as in *Crunchips* or *Infantoy*. Table 2.1 lists the blend categories that were consistently excluded from analysis, leaving 1,095 blends for use in this study.

Category	Blend Count	Example
Brands	221	
Embedded blends	20	<i>ubookquitous</i> : a widely publicized book
Orthographic puns	16	<i>fizzician</i> : a soda jerk
Two word-initial splinters	15	<i>cockapoo</i> : cocker spaniel/poodle cross
Degeminated blends	27	<i>atomechanics</i> : the mechanics of atoms
Three source words	5	<i>brunner</i> : all three meals eaten at once
Metathesized blends	1	<i>smokolotive</i> : a locomotive train
Total Excluded:	292	

Table 2.1: Blend categories excluded from corpus analysis
Some brand names also belong to other excluded categories, so the total excluded is less than the sum of each category.

Thurner’s (1993) dictionary does not indicate pronunciations; where possible, pronunciations were taken from an existing dictionary, typically the CMU Pronouncing Dictionary (CMUdict 1998), the online edition of the OEDOnline (2012), or the online edition of Merriam-Webster.com (2011). Many of the blends were not listed in dictionaries, however, so I and another linguistically trained native English speaker, who was not aware of the experimental hypothesis, provided phonetic transcriptions of our pronunciations. These transcriptions disagreed in just 38 cases; for these, 2-4 linguistically naive English speakers were asked to read the blend, and the consensus pronunciation was added to the corpus. Of particular interest, 17 of these disagreements were about the stress placement in right-headed blends whose left source word is longer, like *geri**atr**ickster* or *ubi**qu**inone*; in 12 of those cases, most speakers pronounced the blend with head stress.

Section 2.2.1 looks at the subcorpus of blends whose source words are the same length, in number of syllables; section 2.2.2 turns to an analysis of the full corpus.

	Right-headed	Non-headed	Left-headed	Total
Preserves both	59	129	7	191
Preserves right	59	22	0	81
Preserves left	4	2	0	6
Total	122	149	7	278

Table 2.2: Length-controlled blends by head and stress preservation

2.2.1 Blends with equal length source words

Because longer source words are known to contribute their stress to blends (Bat-El & Cohen 2012, Gries 2004a,b), I look first at only the subset of blends whose source words have the same syllabic length. This subcorpus contains just 278 blends, 191 of which preserve both source words’ primary stress. While this is not a sufficiently large sample for statistically significant conclusions, examining the numerical results provides a useful example of the patterns seen in the full corpus, without the confounding influence of length. To that end, table 2.2 summarizes the number of blends in each stress and headedness category.

This subcorpus, perhaps surprisingly, includes six left-headed blends. These are blends like *smealth*, a person’s *smell* used as a determinant of their *health*, and *tambooo*, a *tambour* made of *bamboo*. Five of the six preserve both source words’ stress. The sixth, *támbooo*, preserves the stress of its head, although in dialects that forestress nouns like *bamboo*, both words’ stresses are preserved here as well. At any rate, there are few enough examples of this category that they are not significant in any analysis, though I discuss them in §2.4 and provide a full list in Appendix A.

It is interesting to observe what does *not* appear in the data. First, there are no blends with a unique stress pattern. That is, there are 95 blends whose stress pattern is not identical to either source word’s; when matching by alignment, however, all 95 preserve one or both words’ primary stress. There are

Blend	Source Words		Head
<u>cí</u> trange <u>quat</u>	<i>cí</i> trange	k <u>úm</u> quat	both
<u>pó</u> locrosse	<i>pó</i> lo	<u>lac</u> rósse	both
<u>fan</u> dá <u>ng</u> led	<i>fan</i> dángo	n <u>ew</u> -fangled	right
<u>hú</u> ski <u>ing</u>	<i>hú</i> sky	<u>ski</u> ing	right
<u>pá</u> triot <u>ute</u>	<i>pá</i> triot	<u>pr</u> ostíute	right
<u>pó</u> sitron	<i>pó</i> sitive	<u>el</u> éctron	right

Table 2.3: Left-stressed blends

also very few blends that preserve only the stress of the left source word: six in this subcorpus, listed in Table 2.3. The fact that so few the blends that display a stress conflict preserve the left word’s stress argues strongly that the compound stress rule is irrelevant in determining blend stress placement.

It does not, however, support the linear order hypothesis over the head faithfulness hypothesis. While most of these blends that preserve only one stress do preserve that of the right word, there is a striking difference across headedness categories in the percentage of blends that preserve both source word stresses: 84% of non-headed blends preserve both stresses (126 of 150), compared to just 48.3% of right-headed blends (59 of 122). According to a two-sample test of proportions conducted in R, this difference is highly significant ($\chi^2=39.3$, $p<0.0001$), which suggests that headedness affects blend formation in even more ways than hypothesized.

2.2.2 Full corpus analysis

The same difference can be found in the full corpus: 261 of 392 non-headed blends (66.6%) have the primary stress of both source words, which is true of only 338 of 707 right-headed blends (47.8%). The difference is again highly significant ($\chi^2 = 35.84$, $p<0.0001$). Unlike the subcorpus of the previous section, the full corpus also shows a significant difference in blends with conflicting input stress. 78.9% of right-headed blends that preserve only one source word

	Right-headed	Non-headed	Left-headed	Total
Preserves both	338	261	15	614
Preserves right	291	87	3	381
Preserves left	78	44	3	125
Total	707	392	21	1120

Table 2.4: All blends by head and stress preservation

Indicator	β	z	p	
Length difference	0.90	9.024	<0.001	***
Right-headed	0.71	2.828	0.005	**

Table 2.5: Logistic regression coefficients: Stress placement
Significance Levels: 0.05: * 0.01: ** 0.001: ***

stress preserve the right word's, compared with 66.4% of non-headed blends ($\chi^2 = 8.12$, $p=0.004$). Table 2.4 provides a breakdown of the full corpus by stress and head.

When analyzing blends whose source words differ in length, it is important to consider length effects as well. Each blend was coded with the length difference between its left and right source word, measured as the number of syllables in the left source word subtracted from the number of syllables in the right. This value ranged from -3 in *acceleread* (from *accelerate* + *read*) to 5 in *Eurosclerosis* (from *Europe* + *arteriosclerosis*). A two-indicator logistic regression on the blends that preserve only once source word's stress, performed with R's GLM procedure, found significant effects of both length and headedness in determining whether a blend preserved the stress of the right-hand source word. Table 2.5 summarizes the logistic regression results.

These results indicate that headedness does affect blend stress: not only do we find the predicted disparity between the number of right stressed blends with right versus no heads, we also see a notable tendency for non-headed rather than right-headed blends to preserve both stresses. However, there are many other phonological factors that may affect the corpus data, particularly the segmental similarity between the two source words; in the next section, I

present an experiment designed to probe the relationship between headedness and stress placement more directly.

2.3 Stress Placement in Novel Blends

The experiment tested whether participants inferred morphological structure from a novel blend's phonological structure through a definition matching task. Participants were presented with a pair of novel blends, identical except in which source word's stress the blend preserves, and asked to map them to a pair of definitions. Each definition pair contained one subordinating (right-headed) definition and one coordinating (non-headed) definition, and participants were forced to assign one blend to each definition. Because the experiment included only right- and non-headed blends, participants using head faithfulness match the subordinating definition with the blend that preserves right word stress. If linear order were the only active principle in blend stress placement, subjects should prefer the right-preserving blend for both definitions, leading to chance performance, with about half of their judgments following the head faithfulness predictions and half in opposition; instead, the results fit the predicted head faithfulness pattern.

Section 2.3.1 outlines the experimental methods; section 2.3.2 presents the results.

2.3.1 Methodology

Stimulus Creation

The stimulus blends were all created from disyllabic noun pairs, where the left word has initial stress and the right has final stress. The two blends in each pair were segmentally identical (aside from the effects of vowel reduction), with one blend preserving the trochaic stress of the first word and one following the iambic stress of the right. To ensure that the blends had a single optimal segmental form, I selected pairs with a shared consonant between the two syllables' nuclei to reduce the number of plausible segmental blends.

I selected eight of these pairs as test items based on the pronunciations in the CMU Pronouncing Dictionary (CMUdict, syllabified by Bartlett et al. 2009). To limit the search space, I considered only the words found in both CMUdict and CELEX (Baayen et al. 1995), which removed most of the proper nouns found in CMUdict. This produced a list of 36,216 words, from which I identified 2,600,363 word pairs that matched the phonological criteria outlined above. For simplicity, I removed all pairs that consisted of another pair on the list plus affixes: because *cookie bouquet* was included, for example, *cookies bouquet*, *cookie bouquets*, and *cookies bouquets* were all excluded. This reduced the list to 704,652 pairs.

From these, I selected 12 noun-noun pairs with plausible coordinating and subordinating definitions. After pilot testing, I discarded two pairs that were subject to southern stress retraction⁴ and two that pilot subjects found difficult to interpret correctly in both coordinating and subordinating contexts. I recorded myself⁵ reading each blend in the remaining eight test pairs in a

⁴ Words like *police* that are pronounced with final stress in standard American English but with initial stress in some southern dialects

⁵ A 22-year-old female, monolingual American English speaker from the southeastern United States, without a strong regional accent

Source Words		Blends	Definitions
<i>zebra</i>	<u>giraffe</u>	<i>zébraffe</i>	a cross between a giraffe and a zebra
		<i>zebráffe</i>	a giraffe with zebra stripes
<i>robin</i>	<u>baboon</u>	<i>róboon</i>	a cross between a baboon and a robin
		<i>robóon</i>	a baboon with a robin-red chest
<i>turkey</i>	<u>raccoon</u>	<i>túrcoon</i>	a cross between a turkey and a raccoon
		<i>turcóon</i>	a raccoon that steals turkey eggs
<i>flounder</i>	<u>sardine</u>	<i>flóundine</i>	a cross between a sardine and a flounder
		<i>floundíne</i>	a type of sardine eaten by flounder
<i>bachelor</i>	<u>valet</u>	<i>báchelet</i>	a valet who is a bachelor
		<i>bachelét</i>	a valet who works for a bachelor
<i>bistro</i>	<u>garage</u>	<i>bístrage</i>	a building containing a garage and a bistro
		<i>bistráge</i>	the delivery garage of a bistro
<i>pygmy</i>	<u>premier</u>	<i>pýgmier</i>	a leader who is a pygmy
		<i>pygmíer</i>	a leader of the pygmies
<i>raisin</i>	<u>dessert</u>	<i>ráissert</i>	a type of raisin eaten for dessert
		<i>raissért</i>	a raisin-filled dessert

Table 2.6: Stress placement experiment stimuli

soundproof booth using a Windows 7 PC, a headset mic, and the Audacity audio editing program with a sampling rate of 44100Hz. To maximize the naturalness of each token while maintaining similarity between the recordings for each pair, I recorded 8-12 audio tokens of each blend and eliminated those that linguistically naïve native speakers of English judged to be unnatural. I measured the duration and the pitch difference between syllables in each remaining token and selected the tokens for each pair whose duration and pitch difference were the most similar. These I concatenated in both orders, with 0.5s of silence on either side of each token. Table 2.6 lists the eight pairs used in the experiment.

Survey Administration

The experiment was administered as an anonymous online survey using LimeSurvey v1.92+. It consisted of four sections: demographics, instructions and example, the test items, and a post-questionnaire. The demographics in-

cluded basic questions about age, sex, education level, native language and dialect, and familiarity with other languages, plus a question about participation in the segment retention experiment described in section 3.3. The example asked subjects to select the correct definitions for *ínsult* and *insúlt*. This was intended to familiarize them with the survey interface, with the use of orthographic stress marking, and with stress as a contrastive feature; because it used real words rather than novel, it included only the orthographic stress marking and no accompanying audio.

Figure 2.1 shows a typical trial. The question text presented the pair of words to be blended, written as a compound. The two blends were listed below, once with the initial-stressed blend first and once with the final-stressed blend first, with accompanying audio; the definitions appeared below the blends. Subjects used radio buttons to assign each blend to a definition and to assign the question a difficulty rating from “Very Easy” (1) to “Very Hard” (5).

There were four permutations of each question, depending on which blend and which definition were presented first. Each subject was randomly assigned to one of four groups to determine which version of each question they answered; all subjects received two questions of each order type. After all eight test items, subjects answered a post-questionnaire about the strategy they employed; whether they made their judgments based on the audio, the orthography, or both; and which pairs they found hardest.

Participants

Forty-three participants completed the survey, 13 of whom also participated in the segment retention experiment. All were recruited via social media and participated as uncompensated volunteers. Ages ranged from 19-82 ($M =$

The definitions below describe two types of **zebra giraffe**. Listen to the audio, then choose the option that correctly matches each blend to its definition.
Choose one of the following answers

☐ **zébraffe**
zebráffe

0:03

☐ **zebráffe**
zébraffe

0:03

_____ : a giraffe with zebra stripes
_____ : a cross between a giraffe and a zebra

How hard was it to decide?
Choose one of the following answers

☐ Very easy
 ☐ Easy
 ☐ Average
 ☐ Hard
 ☐ Very Hard

The definitions below describe two types of **zebra giraffe**. Listen to the audio, then choose the option that correctly matches each blend to its definition.
Choose one of the following answers

☒ **zébraffe**
zebráffe

0:03

☐ **zebráffe**
zébraffe

0:03

zébraffe : a giraffe with zebra stripes
zebráffe : a cross between a giraffe and a zebra

How hard was it to decide?
Choose one of the following answers

☐ Very easy
 ☐ Easy
 ☒ Average
 ☐ Hard
 ☐ Very Hard

Figure 2.1: A typical trial: Stress placement before (top) and after (bottom) answer selection

36.3, S.D. = 14.83), with 30 female and 13 male respondents. One participant was a native English/Spanish bilingual; all others were monolingual American English speakers. Seventeen reported themselves as southern or southeastern dialect speakers, while three more said they speak a standard dialect with “some southern influence.” Thirteen participants answered the example incorrectly; however, I retained their data, as there was no observable difference between their performance and that of the participants who correctly answered the example question ($p=0.22$ on a two-sample means comparison). Further, the 13 who answered incorrectly all reported relying primarily on the audio for their test item judgments, while the example was the one question without an accompanying recording.

2.3.2 Results

The experiment results clearly support the head faithfulness hypothesis: in 65.8% of the 344 individual trials, participants paired the right-stressed blend with the right-headed definition as predicted. This is significantly above chance ($t[294] = 4.77, p<0.001$) according to a generalized linear mixed model with subject as a random effect.⁶ Six of the eight word pairs also had head faith rates significantly above chance, as detailed in Table 2.7. Neither *robin* *baboon* nor *bachelor* *valet* reached significance, although both had a mean predicted response rate of more than 50%. This corresponds neatly with participants’ responses to the post-questionnaire: eight subjects commented that the two *robin* *baboon* blend tokens sounded similar, and fifteen remarked that they found *bachelor* *valet* especially difficult. Despite these comments, *robin* *baboon* had the lowest average difficulty rating, and overall, an item’s average difficulty rating did not correlate with the rate of head faithful definition mapping.

⁶ Analysis performed in using the GLIMMIX procedure in SAS.

Pair	% Head Faithful	t	p
<i>bistro garage</i>	74.7	3.01	0.002
<i>turkey raccoon</i>	70.0	2.48	0.008
<i>raisin dessert</i>	67.7	2.20	0.02
<i>zebra giraffe</i>	65.4	1.92	0.03
<i>flounder sardine</i>	65.4	1.92	0.03
<i>pygmy premier</i>	63.0	1.63	0.05
<i>robin baboon</i>	60.6	1.34	0.09
<i>bachelor valet</i>	58.3	1.04	0.15
Overall	65.8	4.77	<0.001

Table 2.7: Stress placement results by item

Based on the comments in the post-questionnaire, I conclude that these two pairs received the fewest head faithful responses because they were the hardest for participants to distinguish, interfering with participants' ability to form consistent definition mappings.

Twenty-nine subjects answered more than half of the questions as predicted by head faithfulness. Eight answered exactly half as predicted, and six answered fewer than half head-faithfully. Figure 2.2 gives the number of subjects who answered each number of questions as predicted. There was no significant variation between subjects, nor was there any correlation with any of the demographic information collected.

Because there were no filler items, it is possible that this was a conscious strategy rather than a genuine headedness effect. However, only nine participants reported using a strategy related to headedness. Nearly half – 20 subjects out of 43 – reported that they based their judgments on intuition or aesthetics; eight either did not have or did not report a strategy. A further six reported a strategy, but without enough detail to classify it: four simply responded “Yes,” while two responded more thoroughly but too vaguely to be informative. Further, not all respondents who claimed to match definitions to the blend that stressed the “dominant” or “most important” word did so consistently, making it even less likely that these results are merely due to an

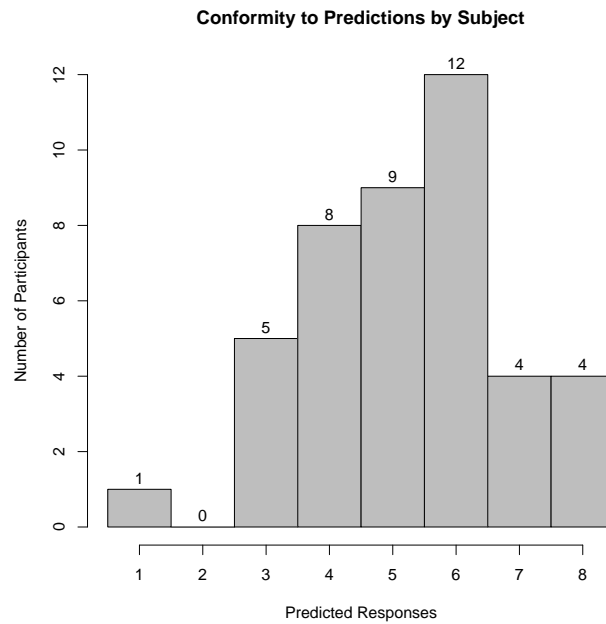


Figure 2.2: Stress placement results by subject

experiment strategy.

2.4 Discussion

The stress placement facts outlined in previous sections respond to both of the major questions in blend studies. First, they contribute to our understanding of the mechanisms that drive blend formation by demonstrating that a blend’s phonological form is determined not only by the source words themselves but by the morphological relationship between them; second, they provide evidence that blending is not merely compounding with truncation but a distinct morphological process by showing that stress placement in blends is not subject to the compound stress rule. I address the role of headedness in section 2.4.1. Section 2.4.2 expands the combined constraint $\text{FAITH}(\text{STRESS})$ to facilitate the discussion of the relationship between blends and compounds in section 2.4.3.

2.4.1 Headedness as a determinant of blend structure

Both experimental and corpus results clearly show that head faithfulness affects stress placement in blends. Although phonological factors such as length and the availability of overlap have a greater influence on stress placement, a multiple regression analysis confirms that whether a blend is right-headed serves as a significant predictor of whether it will preserve its second source word's primary stress. This effect is strong enough that it appears even in the proportions of right- and non-headed blends that preserve right stress only, with significantly higher right stress retention in right-headed blends. This matches the head faithfulness predictions but is not explained by either the linear order or compound stress hypotheses.

Another, more unexpected, piece of evidence for head faith in the corpus is in the drastically higher proportion of non-headed blends than right-headed blends that preserve both source word stresses: 80% of non-headed blends in the subcorpus from section 2.2.1, compared to just 46% of right-headed blends from the same subcorpus. The difference is smaller (66.6% v. 47.8%), but still significant, with the length factors introduced in the full corpus. Preserving both stresses avoids violations of both the positional and general faithfulness constraints, and so both types of blends should preserve both source words' stress when possible. The significant asymmetry between right- and non-headed blends suggests that a blend's morphological structure affects its phonological form even when headedness does not directly determine the optimal candidate. One way to model this asymmetry would be to allow non-headed blends to re-order their source words to maximize overlap. This could be implemented by not specifying an order for non-headed blend- or by including a low-ranked constraint against source word metathesis. Both the existence and the rarity of left-headed blends support this analysis. However, if headed and non-headed blends both undergo this metathesis but at

significantly different rates, it clearly requires a careful implementation.

The experiment results indicate that speakers also use head faithfulness to interpret novel blends. Participants matched the blend that preserved more of the right source word to the right-headed definition in nearly two-thirds of the trials, significantly more than the chance performance predicted if linear order is the key determinant of stress placement. All of the words tested were disyllabic, and each pair had exactly one shared segment, located in the middle of both words, to use as a switch-point; the effect therefore cannot be due to length differences or to the degree or location of segmental overlap, purely phonological factors that influence stress placement in blends. Further, the fact that only one in five participants reported using headedness in their judgments, while nearly half relied on intuition or “guessing,” argues that the observed preference for head-faithful definition mappings is indeed a linguistic effect and not merely a strategy developed for the experiment.

2.4.2 Unpacking Faith(stress)

Before discussing compound stress as it (fails to) apply to blends, it is helpful to look briefly at the role of the individual constraints that I have so far combined into FAITH(STRESS). In 331 of the 497 blends that violate FAITH(STRESS), the violation is a result of reducing one source word’s stress, not removing it completely. In these blends, primary stress comes from one source word; the other source word’s primary stress surfaces as a secondary stress in the blend. Table 2.8 lists a few examples.

Only 166 blends fully delete either source word’s primary stress, indicating that blends which preserve their source word’s stress in any form are preferable to those that delete the stress of one source word. Modelling this requires

Blend	Source Words
<i>àlibi</i> ógraphy	<i>álibi</i> + <i>bi</i> ography
<i>Ìbsénity</i>	<i>Íbsen</i> + <i>obsénity</i>
<i>pópul</i> uxe	<i>pópular</i> + <i>de</i> luxe
<i>sàccharin</i> óceros	<i>sáccharine</i> + <i>rhin</i> óceros

Table 2.8: Blends with secondary stress

using the component constraints of FAITH(STRESS):

(15) Relevant Stress Faithfulness Constraints (repeated from 5)

- a. MAX(STRESS): assign a violation for every stress in the input that does not have an output correspondent
- b. IDENT(STRESS): assign a violation for every primary stress in the input whose output correspondent is a secondary stress⁷

Ranking MAX(STRESS) \gg IDENT(STRESS) produces the partial reduction pattern seen in the blends in Table 2.8:

(16) Partial stress reduction in different-length blends

<i>pópular</i> + <i>de</i> luxe	MAX(STR)	IDENT(STR)
☞ a. <i>pópul</i> uxe		*
b. <i>pópul</i> uxe	*!	

(17) Partial stress reduction in same-length blends

<i>álibi</i> + <i>bi</i> ography _{HEAD}	FAITH(STR) _{head}	MAX(STR)	IDENT(STR)
☞ a. <i>àlibi</i> ógraphy			*
b. <i>álibi</i> ògraphy	*!		*
c. <i>álibi</i> ography	*(!)	*(!)	
d. <i>alibi</i> ógraphy		*!	

⁷ Note that this differs from the IDENT-STRESS constraint used in Pater 2000, which penalizes correspondents when one is stressed and the other is unstressed but does not assign violations for a primary stress that surfaces as secondary stress.

IDENT(STRESS) also prevents the reversal of primary and secondary stresses once they have been assigned, a fact relevant to the comparison between blends and compounds in the next section.

2.4.3 Blending as distinct from compounding

Both the experimental and corpus results further support a distinction between blending and compounding. Despite the predictions of the Compound Stress Rule, there are only a handful of blends that preserve just the left stress in the corpus, and none of the test pairs used in the stress placement experiment show a pattern consistent with either the weak or strong formulations of the CSR. The strong CSR, mandating that all compounds have left stress, incorrectly predicted chance assignment in the experiment by virtue of predicting the same blend for both definitions. The weak CSR demands left stress only from headed blends; this is precisely the opposite of the experimental results, which instead preferentially assign right stress to headed blends, following the head faithfulness predictions outlined in section 2.1.1.

The same problem arises if we try to derive the blends used in the experiment from the corresponding compounds instead of forming them directly from the source words. The tableaux in (18) illustrate the process: in (18a), the compound stress rule applies to *zebra* and *giraffe* to form the compound *zébra giràffe*; when head faithfulness applies in (18b), it guarantees faithfulness to the surface form of the compound used as the input to blending (following Roon 2006). In fact, the double violation of IDENT(STRESS) incurred by the right-stressed candidate in (18b) ensures that the compound analysis predicts the candidate with compound stress even if head faithfulness is irrelevant. This left stressing never occurs in the experimental data and occurs only occasionally in the corpus.

(18) Deriving a blend from an underlying compound

a. Compound stress

<i>zébra</i> + <i>giraffe</i> _{HEAD}	CSR	MAX(STRESS)	IDENT(STRESS)
☞ a. <i>zébra giraffe</i>			*
b. <i>zèbra giraffe</i>	*!		*

b. Subsequent blending

<i>zébra giraffe</i> _{HEAD}	HEADFAITH	MAX	IDENT
a. <i>zèbraffe</i>	*!		**
✗ b. <i>zébraffe</i>			*

The existence of left-headed blends highlights a further difference between blends and compounds: while compounds follow the Righthand Head Rule (Williams 1981) with very few exceptions, the ordering facts mentioned in §2.4.1 suggest that word order in blends is secondary to phonological considerations. In light of these differences, it seems appropriate to abandon the notion that blends are a form of truncated compound and to examine the relationship between blending and compounding more closely. This distinction necessitates a more careful delineation of what should and should not be considered a blend. For example, some authors include words like *agitprop* or *napalm* that contain two left splinters, rather than the typical left and right, in lists of blends, when they share a much greater resemblance to compounds. These reduced compounds rarely if ever display the segmental overlap so common in blends, and they invariably follow the CSR. This provides justification for excluding them from analyses of more typical blends, though a systematic investigation of their similarities to as well as their differences from more typical blends would offer a more detailed view on the morphological status of blending.

CHAPTER 3

SEGMENTAL CONTENT

In the previous chapter, I demonstrated that head faithfulness is one of the factors that affects stress placement in English blends. Although most positional faithfulness effects involving morphological positions involve prosodic phenomena (§1.1), there is no theoretical reason to stipulate that they cannot influence segmental processes. Indeed, this type of positional faithfulness seems ideally suited to determine which segments are retained or deleted in a blend; while most of the possible blends for a given word pair are prohibited by phonological well-formedness constraints, many pairs of source words produce more than one acceptable blend by transitioning from the first word to the second at different points. For example, *blue* and *green* can combine to form either *bleen* or *breen* (Arndt-Lappe & Plag 2012); *Hufflepuff* and *Gryffindor* can form *Huffledor* or *Huffindor*; *alpaca* and *apocalypse* produce both *alpaca*lypse and *alpocalypse*. I argue that this variation is another manifestation of head faithfulness, and that the phonological form of these blends reflects the morphological structure speakers assign them.

The experiment in §3.3 addresses this question directly. Before I discuss the experiment and its results, however, I outline the alternative predictions (§3.1) and apply them to the corpus (§3.2). Section 3.3 presents the results of a second definition matching experiment, this one focusing on the relationship between headedness and segment retention. In section 3.4 I summarize both the corpus and experimental results and their implications.

3.1 Predictions

The segmental content of a blend, much like its stress placement, is determined to large degree by the relative length and potential overlap of its source words. Shorter words generally contribute proportionally more segments to a blend than long words do (Gries 2004a,b, Kaunisto 2000), while overlapping in such a way as to maximize the similarity between the blend and both source words (Gries 2004a,b) minimizes the number of segments that are deleted. However, just as these criteria are not sufficient to determine the stress placement of all blends, they are not always sufficient to determine segmental content. Not all blends have source words of the same length, and not all word pairs have a single blend that would maximize overlap. Head faithfulness and linear order offer competing explanations for which segments are deleted when phonological constraints alone do not determine the optimal blend.

3.1.1 Head faithfulness predictions

As in stress placement, head faithfulness predicts that headed blends will preserve their heads more faithfully than they preserve their other source word. This means we expect higher deletion rates in non-heads, as well as more heads than non-heads that are preserved in their entirety. Formally, MAX and MAX_{head} , defined in (19), can be used to model these predictions.

- (19) Segment retention constraints: Head faithfulness
- a. MAX : assign a violation for every segment in the input that does not have an output correspondent
 - b. MAX_{head} : assign a violation for every segment in the input of the head word that does not have an output correspondent

Unlike in stress placement, it is possible to violate general faithfulness more than once per source word. This makes it possible to create blends that could provide a ranking argument by identifying a source word pair that deletes either a single segment from the head or a pair of segments from the non-head. None of the blends in the corpus display an appropriate overlap to make such a ranking argument, however; nor do those used as test pairs in the segment deletion experiment. Consequently it is again enough for present purposes to include MAX_{head} without ranking it relative to the general MAX , though future work may wish to explore this ranking further.

In the sort of blends seen in the corpus, then, where this ranking is not crucial, head faithfulness predicts that blends with a morphological head will delete segments from their non-head source word rather than from the head:

(20)

<i>/bu.tik + tæk.si</i> _{HEAD} /	MAX_{head}	MAX
☞ a. but <u>æ</u> ksi		*
b. but <u>i</u> ksi	*!	*

In headed blends composed of a word and a splinter, we expect the head to be the whole word, with the non-head surfacing as a splinter. Headed blends containing two splinters should preserve proportionally more of the head than of the non-head.

Head faithfulness makes no predictions for non-headed blends, leaving them to chance or to speakers' preference:

(21)

<i>/bu.tik + tæk.si</i> /	MAX_{head}	MAX
☞ a. but <u>æ</u> ksi		*
☞ b. but <u>i</u> ksi		*

This leaves many non-headed blends with two optimal outputs predicted, which suggests that the blends with the most variation should be non-headed.

3.1.2 Linear order predictions

Segment deletion based on the linear order of the source words predicts that all blends preferentially preserve the same source word regardless of their morphological structure. As before, this utilizes a pair of positional faithfulness constraints, one for each source word:

(22) Segment Retention Constraints: Linear Order

- a. MAX-L: assign a violation for every segment in the left source word that does not have a correspondent in the blend
- b. MAX-R: assign a violation for every segment in the right source word that does not have a correspondent in the blend

According to Gries (2004a,b), blends usually retain more segments from the second source word. This implies the ranking MAX-R \gg MAX-L, shown in (23a) and (23b).

(23) a. Headed blend

<i>/bu.tik + tæk.si_{HEAD}/</i>	MAX-R	MAX-L
☞ a. but <u>æ</u> ksi		*
b. but <u>i</u> ksi	*!	

b. Non-headed blend

<i>/bu.tik + tæk.si/</i>	MAX-R	MAX-L
☞ a. but <u>æ</u> ksi		*
b. but <u>i</u> ksi	*!	

A preponderance of left-preserving blends, from the reverse ranking MAX-L \gg MAX-R, would provide equal support for the linear order hypothesis, although they would be more surprising in light of Gries' results.

3.2 Segment retention in attested blends

This section examines the relationship between segment deletion and headedness using the corpus described in section 2.2. The corpus contains 188 blends in which both source words surface intact and 37 more that delete the same percentage of both source words; this leaves 897 blends that are relevant to the present investigation. Section 3.2.1 describes the subcorpus of blends whose source words are the same length, while section 3.2.2 presents an analysis of the full corpus.

3.2.1 Source words of equal length

When a blend's source words are the same length, other factors—including morphological structure—play a larger role in determining the blend's phonological form. Blends with same-length source words, measured in number of syllables, are therefore a promising place to look for positional faithfulness effects. Unfortunately, this is a very small subset of the full corpus. Only 82 blends have source words of the same length and delete segments from both of them. This is a sufficiently small sample size that none of the differences identified in this section are significant, but as in the stress analysis, it provides a more intuitive survey of some of the patterns seen in the corpus.

Table 3.1 breaks down the number of blends in each segment deletion and headedness category. Only nine of the 82 blends under examination preferentially preserve a larger percent of their left source words, so a left-favoring linear order hypothesis can be discarded. There are two left-headed blends in this subcorpus; both preserve equal amounts of both source words and are thus uninformative. Because the right-favoring linear order hypothesis and the head faithfulness hypothesis make identical predictions regarding right-

	Right-headed	Non-headed	Left-headed	Total
Preserves both equally	7	22	2	31
Preserves more of right	16	26	0	42
Preserves more of left	3	6	0	9
Total	26	54	2	82

Table 3.1: Length-controlled blends by head and segment retention

	Right-headed	Non-headed	Left-headed	Total
Preserves more of left	8	7	1	16
Preserves more of right	19	9	1	29
Total	27	16	2	45

Table 3.2: Length-controlled blends by head and full word retention

headed blends, the crucial comparison is in the non-headed blends' behavior. Of the 54 non-headed blends, 22 preserve equal amounts of both source words (40.7%), while 26 preserve more of the right source word (48.1%) and six preserve more of the left word (11.1%). By contrast, 16 of the 26 right-headed blends preserve more of the right source word (61.5%), with seven preserving their source words equally (26.9%) and three preserving more of the left word (15.8%). A larger percentage of right-headed blends than of non-headed blends preferentially preserves the right source word—61.5 v. 48.1—but with such a small sample size the difference is not significant ($p=0.26$ in a two-sample test for equality of proportions).

Looking at blends that preserve one source word fully and reduce the other (Table 3.2), we see a stronger pattern: 7 of 16 non-headed blends preserve the right word (43.75%), compared to 19 of 27 (70.4%) of right-headed blends. Again, the effect is not significant ($p=0.08$), and an analysis of all 82 blends¹ reveals no main effects of head ($\chi^2=2.01$, $p=0.157$) or degree of reduction ($\chi^2=0.78$, $p=0.378$), or any interaction between the two.

Still, the numerical patterns are promising for such a small data set: as in

¹ Performed using the GENMOD procedure in SAS

	Right-headed	Non-headed	Left-headed	Total
Preserves both	163	55	6	224
Preserves right	231	183	10	424
Preserves left	313	152	7	472
Total	707	390	23	1120

Table 3.3: All blends by head and segment retention

the analogous stress corpus, right-headed blends tend to exhibit head faithfulness, while non-headed blends are more likely to preserve both source words equally. We also see head faithfulness protecting the word from deletion, just as predicted.

3.2.2 Full corpus

The headedness effects in the full corpus are dwarfed by the effects of source word length and segmental overlap. Table 3.3 gives the number of blends in the corpus with each head and deletion pattern, but in this case the counts alone do little to demonstrate any effect of positional faithfulness, either head faith or linear order. In all headedness categories, there is a nearly even split between blends that preserve more of the left word and blends that preserve more of the right, so a more qualitative exploration of the number of blends in each group serves little purpose.

Instead I performed a multiple linear regression analysis using R’s `lm` procedure to determine the contributions of headedness and of source word length and overlap to the blend’s segmental form. The length difference between the two source words was measured by subtracting the number of segments in the right source word from the number of segments in the left source word; this value is negative when the left word is longer, positive when the right word is longer, and zero when the two are the same length. Overlap location measured the degree of source word preservation in a hypothetical blend

Indicator	β	t	p	
Length difference	-0.043	-17.871	<0.001	***
Overlap location	0.011	4.397	<0.001	***
Right-headed	0.024	1.748	0.08	.
Length * Overlap	-0.001	-2.969	0.003	**

Table 3.4: Multiple regression coefficients: Segment retention
Significance Levels: 0.1 . 0.05: * 0.01: ** 0.001: ***

determined solely by the segments in the two source words. This measure assumes that a phonologically-determined blend would minimize faithfulness violations by retaining as much material as possible from both source words, and that the best way to accomplish this is to overlap the two words as much as possible; after identifying the longest common substring for each pair, it subtracts the number of segments in the left word that follow the longest common substring from the number of segments in the right word that precede it. Blends like *autel* (*auto* + *motel*), where the blend candidate that maximizes overlap between the two source words deletes more of the right word than the left, receive a negative score; a pair whose maximally overlapped blend deletes more of the left word than the right receives a positive score. Blends with an overlap score of zero belong to one of three categories: those whose source words overlap with no need for deletion (e.g., *abhorrible*), those whose source words delete the same number of segments when overlapped (e.g., *argle*, from *argue* + *haggle*), and those whose source words have nothing in common (e.g., *spork*, from *spoon* + *fork*). ²

The regression analysis demonstrates significant main effects of the length difference ($\beta = -0.043$, $p < 0.0001$) and of the overlap measure described above ($\beta = 0.011$, $p < 0.0001$) on which source word is more fully preserved, as well as a significant interaction between the two ($\beta = -0.002$, $p = 0.003$). It also reveals

² In many cases, like those above, the hypothetical blend on which this calculation is based is the same as the attested blend. This is not always the case, however. For example, the source words of the attested blend *dormantory* (*dormant* + *dormitory*) maximally overlap in the hypothetical blend *dormitory*. This is blocked as an actual blend because it is homophonous with one of its source words, but it is nonetheless the form used in the overlap calculation.

a marginal effect of right-headedness ($\beta = 0.024$, $p=0.08$), with right-headed blends preserving more of their second source word. Left-headedness was not significant in any of the models; the model with the highest adjusted r^2 did not even include it. As in the stress model, despite its significance this model is a fairly poor fit for the corpus data ($R^2=0.28$, $F[4,1115]=110.6$, $p<0.0001$), suggesting that further exploration into the determinants of blend shape is necessary.

Another notable result is that there are indeed blends that display variation, and, in line with the head faithfulness predictions, all of the blends with variant forms are non-headed. Table 3.5 lists a few examples.

The full corpus, then, bears out the head faithfulness predictions in two ways: first, right-headed blends preserve more of their right source word, though the effect is marginal after accounting for the influence of length and segmental overlap; second, non-headed blends show variation not found in any of the headed blends in the corpus. Because the main effect of headedness in the corpus is marginal, however, I turn to experimental data for additional evidence that head faithfulness shapes speakers' decisions about which segments to delete when forming blends.

3.3 Segment Deletion in Novel Blends

If blend formation is sensitive to head-faithfulness, speakers should infer a right-headed structure from a blend that preferentially preserves the right-

Source Pair		Blend 1	Blend 2
<i>recollect</i>	<u>remember</u>	<i>recol<u>lem</u>ber</i>	<i>recom<u>em</u>ber</i>
<i>fraternity</i>	<u>sorority</u>	<i>frator<u>ity</u></i>	<i>fraror<u>ity</u></i>
<i>tiger</i>	<u>lion</u>	<i>tig<u>l</u>on</i>	<i>tig<u>o</u>n</i>

Table 3.5: Blends with attested variation

hand source word. I used a second definition-matching experiment to test this, presenting speakers with *alpaca* apocalypse-type ambiblendable word pairs that differed only in which source word contributed the stressed vowel. Subjects were tasked with matching the blends and definitions—one blend per definition and one definition per blend. Following the head faithfulness hypothesis, I predict that subjects would pair the blend that preserves more of the right-hand source word with the subordinating definition, in which the right-hand word is the morphological head of the blend.

3.3.1 Methods

Stimulus Creation

I selected eight test pairs using the same procedures outlined for the stress experiment in section 2.3.1, with a few small differences. Most importantly, for this experiment the relevant pairs all had shared consonants on either side of a unique stressed vowel. This ensured that the only difference between the two blends was a preferential preservation of one source word; the *alpacalpyse*-type blends and the *alpocalypse*-type blends are phonologically identical in all respects except the vowel under observation. After identifying all relevant word pairs from the 36,216-word intersection of CMUdict (1998) and CELEX (Baayen et al. 1995), for the sake of source word recoverability I removed pairs where the left word had initial stress or the right word had final stress. This excluded pairs like *copy* coupon, which produce one blend that is unique but hard to interpret (copon) and one that is identical to one of the two source words (coupon). This left 1,324,471 pairs. Filtering out pairs derived from other pairs on the list left 469,190 word pairs.

I manually selected 32 noun-noun pairs with plausible coordinating and

<i>boutique taxi</i>		Length	word 1 position	word 2 position
Coordinating:		6	2 from right	2 from left
a taxi with on-board boutique shopping				
Subordinating:		6	1 from right	2 from left
a taxi to the local boutiques				
Rating: 1	Differences:	0	1	0

<i>lampoon punishment</i>		Length	word 1 position	word 2 position
Coordinating:		6	1 from right	1 from left
punishing someone by printing a lampoon				
Subordinating:		6	1 from right	1 from left
punishing someone for printing a lampoon				
Rating: 0	Differences:	0	0	0

Figure 3.1: A sample rating calculation

subordinating definitions. I rated the structural similarity of the definitions in each pair to ensure that any observed effects were due to the difference in morphological headedness rather than superficial differences such as the order in which the source words appeared in each definition. These numerical ratings summed the difference between the two definitions in length (in number of words), left-hand source word position (in number of words from nearest edge), and right-hand source word position (in number of words from nearest edge). Figure 3.1 gives a sample rating calculation.

I selected eight of the 32 as test pairs: the four with the best similarity ratings, and four pairs of plant or animal words. Subjects in pilot testing had trouble correctly identifying the source words of two pairs—*champagne pineapple* and *stiletto platform*—which were replaced by the two pairs with the next best similarity ratings. The plant and animal pairs had worse overall similarity ratings, but all four had identical definition structures, and the pilot participants reported it was easier to understand the difference in definitions in the plant and animal pairs than in the others. Table 3.6 lists the word pairs, blends, and definitions used as test items.

Source Words		Blends	Definitions
<i>baboon</i>	<u>bandit</u>	<i>bab<u>oo</u>ndit</i>	a baboon who steals like a bandit
		<i>ba<u>b</u>andit</i>	a baboon-stealing bandit
<i>buccaneer</i>	<u>narrator</u>	<i>buccane<u>e</u>rrator</i>	someone who tells pirate stories
		<i>buccan<u>a</u>narrator</i>	a pirate who tells stories
<i>lampoon</i>	<u>punishment</u>	<i>lamp<u>oo</u>nishment</i>	punishing someone by printing a lampoon
		<i>lamp<u>p</u>unishment</i>	punishing someone for printing a lampoon
<i>boutique</i>	<u>taxi</u>	<i>bout<u>i</u>xi</i>	a taxi with on-board boutique shopping
		<i>bout<u>a</u>xi</i>	a taxi to the local boutiques
<i>impala</i>	<u>polecat</u>	<i>imp<u>a</u>lcat</i>	a hybrid of a polecat and an impala
		<i>imp<u>p</u>olcat</i>	a polecat that hunts impalas
<i>armadillo</i>	<u>dolphin</u>	<i>armad<u>i</u>lphin</i>	a hybrid of a dolphin and an armadillo
		<i>armad<u>a</u>dolphin</i>	a dolphin with an armadillo's leathery skin
<i>rhododendron</i>	<u>dandelion</u>	<i>rhodod<u>e</u>ndelion</i>	a cross between a dandelion and a rhododendron
		<i>rhodod<u>a</u>ndelion</i>	a dandelion that grows in rhododendron-like clusters
<i>flamingo</i>	<u>mongoose</u>	<i>flam<u>i</u>ngoose</i>	a hybrid of a mongoose and a flamingo
		<i>flam<u>m</u>ongoose</i>	a mongoose that preys on flamingos

Table 3.6: Segment deletion experiment stimuli
(coordinating definitions above, subordinating definitions below)

Survey Administration

The survey administration was identical to the stress experiment with the exception of audio; unlike in the stress experiment, pilot participants had no trouble interpreting the blends using only orthography. The survey contained the same four sections (demographics, example, test items, and post-questionnaire), collected the same demographic information, and presented the same number of test items. Figure 3.2 depicts a typical trial.

Define That Blend

0% 100%

The definitions in the grey box below describe two types of **buccaneer narrator**. One of them is called a **buccaneerrator**, and the other one is called a **buccanarrator**. Use the answers below to match each blend with its definition.

Choose one of the following answers

☐ buccanarrator
buccaneerrator

☐ buccaneerrator
buccanarrator

_____ : a pirate who tells stories

_____ : someone who tells pirate stories

How hard was it to decide?

Choose one of the following answers

☐ Very easy

☐ Easy

☐ Medium

☐ Hard

☐ Very Hard

Exit and clear survey

Resume later

Next >>

Define That Blend

0% 100%

The definitions in the grey box below describe two types of **buccaneer narrator**. One of them is called a **buccaneerrator**, and the other one is called a **buccanarrator**. Use the answers below to match each blend with its definition.

Choose one of the following answers

☐ buccanarrator
buccaneerrator

☒ buccaneerrator
buccanarrator

buccaneerrator : a pirate who tells stories

buccanarrator : someone who tells pirate stories

How hard was it to decide?

Choose one of the following answers

☐ Very easy

☐ Easy

☒ Medium

☐ Hard

☐ Very Hard

Exit and clear survey

Resume later

Next >>

Figure 3.2: A typical trial: Segment deletion before (top) and after (bottom) answer selection

Participants

Seventy-two native English speakers participated in this experiment. All were recruited via social media and participated as uncompensated volunteers. Fifty-three answered all eight test items; twelve others were only shown seven questions, due to a software bug, and seven exited the survey without completing it. All subsequent discussion considers only the 53 participants who answered all questions. One was a native English/Spanish bilingual; the others were monolingual English speakers. Thirty-four of the respondents analyzed were female. Participants' ages ranged from 18-68 ($M = 32.9$, $S.D. = 14.14$). None had more than introductory linguistic training.

3.3.2 Results

The results support the hypothesis that morphological heads resist deletion: among the responses from completed surveys, subjects answered 66% of trials (280 of 424) as predicted by the head faith hypothesis, assigning the blend that preserved more segments from the right word to the right-headed definition. Table 3.7 details the results of a generalized linear mixed model analysis with subject as a random effect, using the GLIMMIX procedure from the SAS statistical package. The rate of head-faithful responses was significant overall ($t[364]=6.52$, $p<0.001$) as well as for six of the eight individual items. A seventh, *rhododendron dandelion*, is nearly significant; the effect may be weaker here because the two blends, *rhododendelion* and *rhododandelion*, are phonetically less distinct than the blends in other pairs. The final pair, *impala polecat*, was treated as predicted in only 43% of responses, neither significantly above ($p=0.83$) nor below ($p=0.16$) chance. It may be that speakers parse *impalcat* and *impolcat* as blends of *impala* and *cat*, especially due to the blends' spelling. In that case, both blends preserved the right-hand word in its entirety, so its

Pair	% Predicted	z	p
<i>boutique taxi</i>	81.1	4.53	<0.001
<i>flamingo mongoose</i>	73.6	3.43	<0.001
<i>lampoon punishment</i>	71.7	3.16	<0.001
<i>baboon bandit</i>	67.9	2.61	0.005
<i>armadillo dolphin</i>	67.9	2.61	0.005
<i>buccaneer narrator</i>	62.3	1.79	0.037
<i>rhododendron dandelion</i>	60.4	1.51	0.065
<i>impala polecat</i>	43.4	-0.96	0.831
Overall	66.0	6.52	<0.001

Table 3.7: Segment deletion results by item

status as head would be irrelevant.³

Forty of the 53 subjects answered more than half of the test items according to the head-faith predictions. Seven answered at chance (exactly half as predicted), and six answered below chance. Figure 3.3 gives a more detailed breakdown. As in the stress experiment, the lack of filler items raises the possibility that subjects' responses reflect a conscious strategy rather than a genuine effect. Here, 19 people reported using a strategy based on headedness—usually expressed as “whichever word was dominant” or “whichever word was more important”—for all questions, while three more reported using headedness only in the plant and animal blends. Fifteen participants relied on aesthetics or intuition in their judgments. Twelve either did not have or did not report a strategy; the remaining four reported strategies that were not easily classified.⁴ Thus it seems speakers are able to consciously manipulate

³ Neither potentially misleading spelling nor the availability of an alternate decomposition alone is sufficient to account for the *impala polecat* responses: the two pairs that most conform to the head-faith hypothesis each display one of these characteristics. *Boutixi* and *boutaxi* also employ consonant spellings that may mask the source word, though notably unlike in *polecat* this does not interrupt the contiguity of either source word. In *flamingoose* and *flamongoose*, it is possible to assume *goose* as the right-hand source word, but this is perhaps less likely for the monomorphemic *mongoose* than for the compound *polecat*.

⁴ Two of these were too vague to classify (“I can’t really put it into words”; “Looked at how the words were blended”), one responded affirmatively without providing detail (“Yes—developed one after the first few questions”), and one considered morphological structure but did not specify how that affected her decision (“Try both options—think of meanings of pre- and suffixes”).

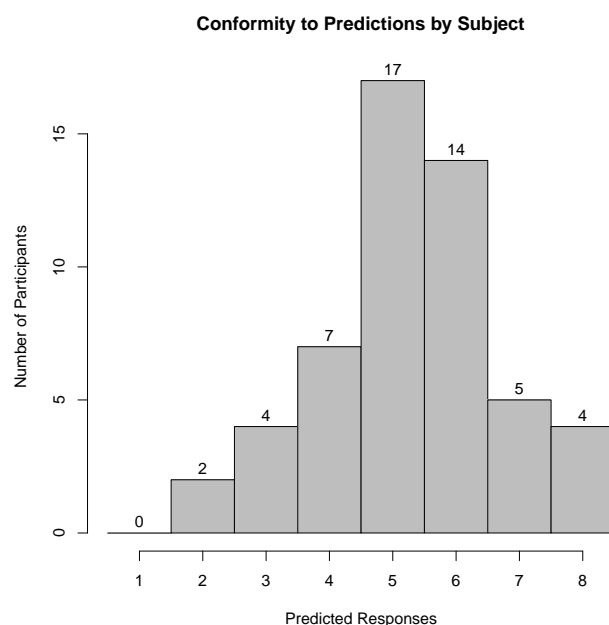


Figure 3.3: Segment deletion results by subject

the segmental content of blends more easily than their stress placement, but there is a definite preference for preserving segments of the head even among speakers who are not using it as a test strategy.

None of the demographic factors collected—age, gender, or education level—significantly affected participants’ conformity to the head-faith hypothesis.

3.4 Discussion

The experiment results show that head faithfulness influences speakers’ interpretation of novel blends, while the corpus analysis shows that established blends also reflect their morphological structure in their segmental content. The corpus analysis supports all three aspects of the head faith predictions outlined in section 3.1.1. First, among blends in the same-length subcorpus that delete segments from both source words, a higher percentage of right-headed than non-headed blends preserve more of the right word (61.5% v.

48.1%). Second, more right-headed than non-headed blends contain all of the second source word but only part of the first (70.4% v. 43.8%). These trends are reflected in the multiple regression results indicating correlation between right-headedness and preservation of the second source word, following the prediction that the head of a blend resists deletion as a result of occupying a privileged position. The third characteristic of the corpus, variation among non-headed blends, reflects the fact that head faithfulness constraints do not distinguish between candidates in a non-headed blend. While there are only a few blends in the corpus that have variants listed, they are all non-headed; all the right-headed blends Thurner (1993) lists have a single segmental form.

The experiment results closely resemble the results from the stress experiment, with most subjects, six of eight items, and the overall results significantly supporting the head faithfulness predictions. More subjects reported using a head-based strategy, which suggests that speakers are more aware of the relationship between segment retention and headedness.⁵

The presence of head faithfulness effects perhaps more interesting than the stress results in light of the fact that faithfulness to morphological positions more often triggers prosodic than segmental phenomena. This not only demonstrates that positional faithfulness effects are active in the creation of English blends, and thus in English, it also shows that morphological positions need not be limited to affecting prosody.

⁵ This observation matches personal reports of blend usage I have gathered during this project, including one individual who distinguishes more fork-like *sporks* from more spoon-like *spoorks* and a family who has *brinner* when they eat pancakes for dinner and *brenner* when they eat just one large meal in a day. I have found no such distinctions that hinge on stress placement.

CHAPTER 4

CONCLUSION

The primary aim of this thesis has been to demonstrate the mutual applicability of positional faithfulness theory and blend formation in English. To that end, I have shown that blends exhibit head faithfulness effects both in their stress placement and in the determination of which segments to delete or retain. Although they are secondary to phonological influences and the demand for source word recoverability, these effects are pervasive: they appear in corpus data from a dictionary of attested blends and in experimental data on interpretation of novel blends. These results have implications both for positional faithfulness theory and for blend formation.

4.1 Positional faithfulness implications

The findings presented here have several implications for positional faithfulness theory. First, they demonstrate that PF constraints are active in English. Unstressed vowel reduction (e.g., Chomsky & Halle 1968) provides ample evidence that stressed syllables are a privileged position in English, but other positions have little effect in the regular grammar. By showing that blend formation involves at least one more type of positional faithfulness, this thesis offers more evidence that PF, and head faithfulness more specifically, are viable members of the universal constraint set.

Second, it shows that morphological positions can influence segmental, not just prosodic, phenomena. Most instances of faithfulness to morphological po-

sitions involve tone (Smith 2011), stress and accent (Alderete 2001, Gouskova & Roon To Appear, Revithiadou 1999, Roon 2006, Smith 2011, Ussishkin 2005, among others), or syllable structure (Smith 2011). It is therefore unsurprising to find PF effects in blend stress placement; it is more interesting to observe the effects of head faithfulness on segment deletion. This suggests that other privileged positions usually implicated in prosodic phenomena may also interact with segment structure.

This study also illustrates the value of blend formation as a laboratory for exploring the impact of other privileged positions. Blending manipulates the segmental and prosodic structure of and the morphological relationship between source words, making it an ideal vehicle to investigate the role of both morphological and phonological privileged positions. Blending is not amenable to probing all privileged positions, of course; lexical versus functional categories and roots versus affixes, in particular, would be virtually impossible to investigate given that the conflict in blending is between faithfulness to two words. However, the study design used in this thesis could easily be modified to look at lexical category effects (Smith 2011) or proper noun privilege (Jaber 2011). It may also be possible to use blend formation to investigate relative privilege between positions. For example, it may be possible to find ambiblendable word pairs that produce a linear blend that preserves one source word's initial syllable or an embedded blend that preserves the same word's stressed syllable. Few other processes involve enough conflicting faithfulness constraints to support such an inquiry.

4.2 Blend formation implications

This study contributes to the blend formation literature in three key ways. First, it relates the analyses developed specifically to account for patterns in

blending to the PF theoretical framework, which is applicable to a much wider range of linguistic phenomena. This highlights the similarities between blend formation and other phonological processes along with the regularities that underlie the linguistic creativity seen in blends.

Second, it demonstrates some of the differences between blending and compounding, which are commonly assumed to be closely related morphological processes. I have identified two ways in which blending and compounding are significantly different: their patterns of stress assignment and their adherence to the Righthand Head Rule. While this does not negate the similarities between the two—most blends can be expressed as compounds with little change in meaning, and the precise distinction between blends and reduced compounds is still unclear—it does suggest the need for more investigation into the relationship between blend formation and the morphological processes it most resembles. This includes concatenative processes like compounding and derivation as well as acronym formation and other subtractive processes.

Finally, the data presented here does contribute to our knowledge of blend formation. By establishing a connection between a blend's morphological structure and its phonological form, it situates the blending process along the phonology-morphology interface. This dimension offers many interesting ideas for future research, especially the related investigations into left-headed blends and the morphophonological factors that motivate source word ordering. Kelly's (1998) discussion of source word order only considered non-headed blends, and it only looked at factors that also affect conjunct ordering. The possibility of overlap in blend formation, and indeed the pressure to maximize it, add additional phonological factors that should be included in a full analysis, and the head-based asymmetries in degree of overlap suggest morphological factors at play as well. To my knowledge, no other studies have

mentioned the existence of left-headed blends, much less conducted a systematic analysis to determine what they have in common or how to coin more; discussions both of head faithfulness and of ordering effects would benefit tremendously from the inclusion of a significant number of left-headed blends.

4.3 Concluding Remarks

This thesis brings the broad theoretical applicability of Positional Faithfulness theory to bear for the first time on the abundance of competing faithfulness constraints at work in blend formation. The present findings are certainly of interest in both areas, with consequences for both positional faithfulness and blending. While we have seen the effects of one type of PF constraint at all levels of the blending process, however, the real contribution of this work is not simply the demonstration of head faithfulness effects in English morphophonology. Rather, the most promising implication of the new data presented here is the potential it reveals for future investigations into the breadth of positional faithfulness, into the complex process of blend formation, and into the nature of the morphology-phonology interactions that shape it.

APPENDIX A

LEFT-HEADED BLENDS

Blend	Source Words		Definition
1 <i>abnormous</i>	<i>abnormal</i>	<u>enormous</u>	abnormal or irregular
2 <i>acceleread</i>	<i>accelerate</i>	<u>read</u>	the act of increasing one's reading speed
3 <i>administr<u>al</u>ia</i>	<i>administration</i>	<u>Australia</u>	the upper echelon of corporate management, a realm that is often as remote and distant as a foreign land
4 <i>banjorine</i>	<i>banjo</i>	<u>tambourine</u>	a short-necked banjo whose belly resembles a tambourine
5 <i>bovie</i>	<i>book</i>	<u>movie</u>	a book based on the screenplay of a popular movie
6 <i>chaitroplane</i>	<i>chair</i>	<u>aeroplane</u>	an amusement park ride consisting of seats hanging from a wheel so that they spin out when the wheel is revolved
7 <i>Druriolanus</i>	<i>Drury</i>	<u>Coriolanus</u>	a theatrical nickname for the Drury Lane Theatre in London
8 <i>entrem<u>anure</u></i>	<i>entrepreneur</i>	<u>manure</u>	the operators of an experimental power plant that burns dried cattle chips for fuel
9 <i>expunctu<u>ation</u></i>	<i>expunction</i>	<u>punctuation</u>	the deletion of material from written copy during the process of editing
10 <i>Fil<u>am</u>erican</i>	<i>Filipino</i>	<u>American</u>	a Filipino who exhibits sympathetic or loyal feelings toward Americans
11 <i>gimble</i>	<i>gambol</i>	<u>nimble</u>	to frolic nimbly

Blend	Source Words		Definition
12 <u>homonymble</u>	<i>homonym</i>	<u>nimble</u>	a clever pun that plays on distinct meanings of the same word
13 <u>newzak</u>	<i>news</i>	<u>Muzak</u>	news coverage, the original impact of which has been dulled by repetition
14 <u>pinkermint</u>	<i>pink</i>	<u>peppermint</u>	having the pink color of peppermint candy
15 <u>poetlariat</u>	<i>poet laureate</i>	<u>lariat</u>	Will Rogers, a humorist who delivered social commentary while performing rope tricks
16 <u>rendezwoo</u>	<i>rendezvous</i>	<u>woo</u>	a meeting arranged between two lovers
17 <u>satelloon</u>	<i>satellite</i>	<u>balloon</u>	a satellite launched from a high-altitude balloon
18 <u>smealth</u>	<i>smell</i>	<u>health</u>	a person's scent as a determinant of their health
19 <u>taboob</u>	<i>taboo</i>	<u>boob</u>	exaggerated standards of moral extremism
20 <u>tambo</u>	<i>tambour</i>	<u>bamboo</u>	a small drum made of bamboo
21 <u>tilge</u>	<i>tea</i>	<u>bilge</u>	unpleasantly tepid tea
22 <u>unperor</u>	<i>umpire</i>	<u>emperor</u>	an imperious umpire

Table A.1: Left-headed blends

REFERENCES

- Alderete, John. 2001. *Morphologically governed accent in optimality theory* Outstanding dissertations in Linguistics. New York: Routledge.
- Algeo, John. 1977. Blends, a structural and systemic view. *American Speech* 52(1). 47–64.
- Arcodia, Giorgio Francesco & Fabio Montermini. 2012. Are reduced compounds compounds? Morphological and prosodic properties of reduced compounds in Russian and Mandarin Chinese. In Vincent Renner, François Maniez & Pierre Arnaud (eds.), *Cross-disciplinary perspectives on lexical blending*, 193–212. Berlin: De Gruyter Mouton.
- Arndt-Lappe, Sabine & Ingo Plag. 2012. Phonological variability in English blends. Handout from at the 20th Manchester Phonology Meeting, Manchester.
- Baayen, R. Harald, Richard Piepenbrock & Leon Gulikers. 1995. The CELEX lexical database. Linguistic Data Consortium, University of Philadelphia, Philadelphia, PA.
- Bartlett, Susan, Grzegorz Kondrak & Colin Cherry. 2009. On the syllabification of phonemes. In *Proceedings of human language technologies: The 2009 annual conference of the north american chapter of the association for computational linguistics*, 308–316. Boulder, Colorado: Association for Computational Linguistics. <http://www.aclweb.org/anthology/N/N09/N09-1035>.
- Bat-El, Outi. 2006. Blend. In Keith Brown (ed.), *Encyclopedia of language and linguistics*, 66–70. Oxford: Elsevier 2nd edn.
- Bat-El, Outi & Evan-Gary Cohen. 2012. Stress in English blends: A constraint-based analysis. In Vincent Renner, François Maniez & Pierre Arnaud (eds.), *Cross-disciplinary perspectives on lexical blending*, 193–212. Berlin: De Gruyter Mouton.
- Beckman, Jill. 1997. Positional faithfulness, positional neutralization, and Shona vowel harmony. *Phonology* 14(1). 1–46.
- Beckman, Jill. 1998. *Positional faithfulness*: University of Massachusetts Amherst dissertation.
- Cannon, Garland. 1986. Blends in English word formation. *Linguistics* 24. 725–753.
- Casali, Roderic. 1996. *Resolving hiatus*: University of California Los Angeles dissertation.
- Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.
- Cinque, Guglielmo. 1993. A null theory of phrase and compound stress. *Linguistic Inquiry* 24(2). 239–297.
- CMUdict. 1998. The CMU pronouncing dictionary v0.6d. <http://www.speech.cs.cmu.edu/cgi-bin/cmudict/>.

- Frath, Pierre. 2005. Why is there no ham in a hamburger? A study of lexical blends and reanalyzed morphemes. *Recherches Anglaises et Nord Americaines* 38. 99–112.
- Gouskova, Maria & Kevin Roon. To Appear. Gradient clash, faithfulness, and sonority sequencing effects in Russian compound stress. *Laboratory Phonology*.
- 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.
- Hamans, Camiel. 2010. The productivity of blending: Linguistic or cognitive? Or how to deal with *administrivia* and *ostalgia*. In Roman Kalisz, Danuta Stanulewicz, Tadeusz Z. Wolański & Joanna Redzimska (eds.), *Lingua terra cognita*, vol. 2, 455–478. Gdańsk: Wydawnictwo Uniwersytetu Gdańskiego.
- Jaber, Aziz. 2011. High vowel syncope in Jordanian Arabic: A positional faithfulness treatment. PhD qualifying paper, Department of Linguistics, University of North Carolina, Chapel Hill.
- Kaunisto, Mark. 2000. Relations and proportions in the formation of blend words. In Paper presented at QUALICO, Prague.
- Kelly, Michael H. 1998. To “brunch” or to “brench”: some aspects of blend structure. *Linguistics* 36(3). 579–590.
- Kubozono, Haruo. 1990. Phonological constraints on blending in English as a case for phonology-morphology interface. *Yearbook of Morphology* 3. 1–20.
- Lehrer, Adrienne. 1996. Identifying and interpreting blends. *Cognitive Linguistics* 7(4). 259–290.
- Lehrer, Adrienne. 2007. Blendalicious. In Judith Munat (ed.), *Lexical creativity, texts and contexts*, 115–133. John Benjamins.
- Lieberman, Mark & Alan Prince. 1977. On stress and linguistic rhythm. *Linguistic Inquiry* 8(2). 249–336.
- Lombardi, Linda. 1999. Positional faithfulness and voicing assimilation in Optimality Theory. *Natural Language and Linguistic Theory* 17(2). 267–302.
- Mattiello, Elisa. 2013. *Extragrammatical morphology in English: Abbreviations, blends, reduplicatives, and related phenomena* chap. Blends, 111–140. De Gruyter Mouton.
- McCarthy, John & Alan Prince. 1995. Faithfulness and reduplicative identity. In Jill Beckman, Laura Walsh Dickey & Suzanne Urbanczyk (eds.), *University of Massachusetts occasional papers in linguistics 18: Papers in Optimality Theory*, 249–384. Amherst: GLSA.
- Merriam-Webster.com. 2011. <http://www.merriam-webster.com>.
- OEDOnline. 2012. <http://www.oed.com>.

- Pater, Joe. 2000. Non-uniformity in English secondary stress: The role of ranked and lexically specific constraints. *Phonology* 7(2). 237–274.
- Plag, Ingo. 2003. *Word-formation in English*. Cambridge University Press.
- Plag, Ingo. 2006. The variability of compound stress in English: structural, semantic, and analogical factors. *English Language and Linguistics* 10(1). 143–172.
- Revithiadou, Anthi. 1999. *Headmost accent wins: Head dominance and ideal prosodic form in lexical accent systems*: University of Leiden dissertation.
- 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, 319–330. Michigan Slavic Publications.
- Smith, Jennifer L. 2001. Lexical category and phonological contrast. In Robert Kirchner, Joe Pater & Wolf Wikely (eds.), *PETL 6: Proceedings of the Workshop on the Lexicon in Phonetics and Phonology*, 61–72. Edmonton, Alberta: University of Alberta.
- Smith, Jennifer L. 2011. Category-specific effects. In Marc van Oostendorp, Colin Ewen, Beth Hume, & Keren Rice (eds.), *The Blackwell Companion to Phonology*, 2439–2463. Wiley-Blackwell.
- Turner, Dick. 1993. *Portmanteau dictionary: Blend words in the English language, including trademarks and brand names*. Jefferson, NC: McFarland & Co.
- Ussishkin, Adam. 2005. A fixed prosodic theory of nonconcatenative morphology. *Natural Language and Linguistic Theory* 23. 169–218.
- Williams, Edwin. 1981. On the notions “lexically related” and “head of a word”. *Linguistic Inquiry* 12(2). 245–274.