Gradient clash, faithfulness, and sonority sequencing effects in Russian compound stress[*]

Maria Gouskova and Kevin Roon

Abstract

Russian normally does not have secondary stress, but it is variably realized in compounds. We examined the factors that contribute to secondary stress realization in a rating study, where listeners were asked to rate compounds pronounced without secondary stress and with secondary stress in various locations. We refine some generalizations from impressionistic descriptions: in compounds whose left-hand stems have mobile lexical stress, acceptability of secondary stress decreases with token frequency of the compound, and acceptability of pronunciations without stress increases with frequency. Ratings improve as distance between stresses increases, and this effect is gradient rather than categorical. We also identify new generalizations about secondary stress that relates to the properties of the left-hand stem. First, we identify a faithfulness effect: stress realization is optional on lexically stressed stems, but stress movement is strongly penalized. Second, we identify a sonority sequencing effect: secondary stress is not tolerated well on linker vowels in compounds, but acceptability improves significantly when the linker is the only vowel in a stem with a falling sonority cluster. Thus, the stress system distinguishes clusters with falling sonority from other types.

1 Introduction

Stress plays a central role in the phonology of Russian, conditioning vowel reduction and interacting with several other rules (Halle 1973, Halle & Vergnaud 1987a, Melvold 1989, Crosswhite 1999, Crosswhite et al. 2003, Barnes 2003, Padgett & Tabain 2005, inter alia). Russian stress is also complex: it is fully contrastive and morphologically conditioned. It is surprising, therefore, that so little attention has been paid to secondary stress, which occurs variably in compounds and in certain prefixes. The conditions under which secondary stress surfaces are not very well understood, so the goal of this paper is to elucidate this aspect of Russian phonology. We report on an experimental study of compound stress where we asked Russian speakers to rate pronunciations of compounds without secondary stress and with secondary

[*]For valuable feedback on this and related work, we would like to thank Christina Bethin, Lisa Davidson, Amanda Dye, Gillian Gallagher, Greg Guy, Pavel Iosad, Darya Kavitskaya, Alec Marantz, Bruce Morén-Duolljá, Tuuli Morrill, Kathryn Pruitt, Jason Shaw, the audience at CASTL, Tromsø and the audience of FASL 17 at Yale. Special thanks to Michael Becker, Gillian Gallagher and Jennifer Nycz for discussion of statistical issues, and to our undergraduate research assistant Erika Harris for help with the acoustic analysis of the stimuli. The article was greatly improved as a result of the constructive comments by the anonymous reviewers, Jennifer Cole, and Mirjam Ernestus. We would also like to thank all the Russian native speakers in New York City and at the Eastern Generative Grammar School in Debrecen, Hungary for participating in our study. This study was conducted under NYU UCAIHS #6472 and was supported in part by NSF BCS grant #1224652.

stress in various locations. We test some generalizations known from previous work (Avanesov 1964, Yoo 1992, Roon 2006, Gouskova & Roon 2009, Gouskova 2010), such as the sensitivity of secondary stress to the distance between stresses and to the lexical stress type of the left-hand stem. We also identify some new generalizations. We believe that acceptability judgment studies of this sort tap into the knowledge that people use in assigning stress in their production grammar, but they avoid some of the pitfalls of studying variable phonology in production experiments. We discuss some of the limitations of production experiments, such as the tendency for people to produce too much secondary stress on less familiar words in the lab setting.

The literature on Russian secondary stress identifies a number of factors, some of which favor stress and others which disfavor it. For example, frequently mentioned are the effects of stress clash. Avanesov (1964:52–53) is the first to observe that secondary stress is more likely to appear in Russian the farther it is from primary stress. Yoo (1992) casts this generalization in categorical terms: stress is more likely to appear when stresses are two syllables apart. According to this categorical characterization, clash would arise in (1a) if the left-hand compound stems were stressed, but stresses are sufficiently far apart in (1b).[1] The question of whether the anti-clash constraint is gradient or categorical is of general interest in metrical stress theory: there are several proposals for differential anti-lapse constraints that penalize lapses of longer lengths more severely than lapses of shorter lengths (Steriade 1997, Gordon 2005, McCarthy 2007), but there has been relatively little discussion of differential anti-clash constraints (Liberman & Prince 1977, Nespor & Vogel 1989, Kager 1994, Pater 2000, Alber 2005). Our proposal, based on the Russian pattern, is in section 5.5.

(1) Secondary stress in Russian compounds: effect of clash (Yoo 1992, Gouskova & Roon 2009)

a. No secondary stress: one syllable would separate stresses			
lʲis-ʌ-párk	forest-*linker*-park	‘forest park’	lʲés ‘forest’
vʲir-ʌ-lómstvə	faith-*linker*-breaking	‘treachery’	vʲér-ə ‘faith’
b. Secondary stress: two or more syllables separate stresses			
lʲès-ə-kulʲtúrə	forest-*linker*-culture	‘forest cultivation’	
vʲèr-ə-ispəvʲidánʲijə	faith-*linker*-confession	‘religious denomination’	

Other factors that have been implicated include low frequency, which favors secondary stress: the lower the token frequency of the compound, the more likely is secondary stress to surface (Avanesov 1964, Yoo 1992, Comrie et al. 1996, Gouskova & Roon 2009). We will show that frequency interacts with the lexical stress type of the left-hand stem, which has also been reported to affect secondary stress realization.

Russian is famous for seemingly not restricting word-initial clusters by sonority sequencing (Clements 1990 and others). Gouskova & Roon (2009) find that people produce more secondary stress in words with sonority sequencing-violating clusters such as [lʐ-è-naúkə] ‘pseudoscience’, but their findings are not conclusive because it is not clear whether the effect is due to the low frequency of such words or to their sonority profile. Onset-sensitive stress patterns are relatively unusual (Davis 1988, Smith 2002), so the Russian pattern would be typologically interesting if it holds up.

Disentangling these various factors is difficult because they sometimes favor the same outcome. Secondary stress, moreover, is quite variable: Yoo's (1992) generalizations are based on prescriptive pronunciation dictionaries of Russian (Ageenko & Zarva 1984, Borunova et al. 1988), but even these dictionaries do not agree on whether secondary stress is obligatory in many cases. Our study is the first to consider the many interacting factors systematically and to examine them experimentally.

The rest of the paper is structured as follows. We start with a quick overview of the structure of Russian compounds, and then provide background on the lexical stress system in section 2. This section motivates the hypotheses that we test in our study. Next we discuss the design of the grammaticality judgment study (section 3), including an acoustic analysis of the natural stimuli we used (section 3.4). The results of our study are presented in section 4, which is broken down into four subsections. The first subsection discusses the nature of the anti-clash constraint in Russian, and the next three discuss compounds with different types of left-hand stems: fixed stress stems (section 4.2), mobile stress stems (4.3), and final stress stems (4.4). A general discussion follows in section 5, and section 6 concludes.

2 Preliminaries

2.1 The structure of Russian compounds

Russian has several types of morphological compounds (see Molinsky 1973, Townsend 1975, Yoo 1992, Gouskova & Roon 2009), but we concentrate on subordinating compounds, which are formed by combining two stems connected by a linking vowel /-e-/ or /-o-/ (phonetically realized as one of [ə, i, o, ʌ, e]). The linking vowel is underlined in the examples in (2); throughout, we will show the boundaries to either side of the linker but not internal boundaries within the compound. The compounds are right-headed both morphologically and phonologically: the right-hand stem is always stressed (Roon 2006) and has the stronger prominence if there are two stresses. Even though Russian single-root words have at most one stress, we assume that compounds consist of single prosodic words even when they have two stresses. In parsing compounds as single prosodic words, Russian differs from languages such as Dutch and English (Nespor & Vogel 1986 and others). Gouskova (2010) develops the argument for a single prosodic word analysis in detail, based on segmental rules such as consonant devoicing and pretonic vowel reduction. Devoicing applies at the prosodic word edge, whereas vowel reduction does not. In compounds, devoicing does not apply before the compound linker vowel, suggesting there is no boundary to the left of it. The linker vowel itself reduces to [ə] or [ʌ] depending on whether it is followed by stress, suggesting that it is in the same prosodic word as the following right-hand stem syllable. This pattern of reduction persists whether the left-hand stem has stress on it or not, indicating that the two stems are always grouped into a single phonological word.

(2) Morphological boundaries in subordinating compounds

a.	lʲis-ə̲-s-vál	‘lumber cutting’	cf. lʲés ‘lumber,’ s-vʌlʲ-ítʲ ‘to fell, saw down’
b.	gələv-ʌ̲-lóm-k-ə	‘puzzle’	cf. gəlʌv-á ‘head,’ lʌm-átʲ ‘to break, wrack’
c.	ʌbʌròn-ə̲-s-pʌ-sób-n-əstʲ	‘defense capability’	cf. ʌbʌrón-ə ‘defense,’ s-pʌ-sób-n-əstʲ ‘capability’
d.	səm-ʌ̲-lʲòt-ə̲-strʌj-énʲ-i-jə	‘airplaine building’	cf. sám ‘self,’ lʲit-átʲ ‘to fly,’ strʌj-énʲ-i-jə ‘building’

Primary stress on the right-hand stem is obligatory, and its assignment is quite consistent. Whether and where secondary stress surfaces depends on the stress phonology of the left-hand stem, which we discuss in more detail in the following section. Since stress is manipulated in our experiment, we overview the main features of the system as well.

2.2 Lexical stress in Russian

In the lexical stress system of Russian (Halle 1973, Halle & Vergnaud 1987a, Melvold 1989, Idsardi 1992, Brown et al. 1996, Revithiadou 1999, Chew 2000, Alderete 1999, 2001, Crosswhite et al. 2003), words can follow one of three major patterns: fixed, final, and mobile. The mobile pattern can be further subdivided into subtypes, but the number of words in each category is quite small (Zaliznjak 1977, 1985). We will focus on nouns here, but the same patterns can be found in other syntactic categories. According to Zaliznjak, in the vast majority of nouns—92%—stress falls on one of the syllables of the stem and does not alternate throughout the inflectional paradigm. As the examples in (3) show, the location of stress within the stem itself is contrastive and unpredictable. We will refer to this pattern as *fixed stress,* following Melvold’s (1989) terminology.

(3) Russian stress Pattern A: stress fixed on stem in inflectional paradigms

sg	/komnat-/	/tʲetradʲ-/	pl	
nom	kómnət-ə	tʲitrátʲ-∅	kómnət-i	tʲitrádʲ-i
acc	kómnət-u	tʲitrátʲ-∅	kómnət-i	tʲitrádʲ-i
gen	kómnət-i	tʲitrádʲ-i	kómnət-∅	tʲitrádʲ-ij
dat	kómnətʲ-i	tʲitrádʲ-i	kómnət-əm	tʲitrádʲ-əm
inst	kómnət-əj	tʲitrádʲ-ju	kómnət-əmʲi	tʲitrádʲ-əmʲi
loc	kómnətʲ-i	tʲitrádʲ-i	kómnət-əx	tʲitrádʲ-əx
	‘room’	‘notebook’		

The next largest type contains only 6% of nouns, which have stress on the inflectional suffix if one is present. Words with null suffixes have stem-final stress. We will refer to this type as *final stress*—the only context where stress is not final is in the instrumental plural, with the suffix [-ami], where stress falls on the penult. The forms in (4) demonstrate another well-known feature of Russian phonology, unstressed vowel reduction (Crosswhite 1999, Barnes 2004, Padgett & Tabain 2005, Gouskova 2010). In stressed syllables, there is a five-way contrast, [i u o e a], which is reduced to [i u ʌ] in the immediately pretonic syllable and to [i u ə] elsewhere.

(4) Russian stress type Pattern B: final stress

	/kʌrablʲ-/	/ʧʲert-/	/vʲeʃʃʲ-estv-/
nom sg	kʌráblʲ-∅	ʧʲirt-á	vʲiʃʃʲistv-ó
acc sg	kʌráblʲ-∅	ʧʲirt-ú	vʲiʃʃʲistv-ó
gen sg	kərʌblʲ-á	ʧʲirt-í	vʲiʃʃʲistv-á
dat sg	kərʌblʲ-ú	ʧʲirt-ʲé	vʲiʃʃʲistv-ú
inst sg	kərʌblʲ-óm	ʧʲirt-ój	vʲiʃʃʲistv-óm
loc sg	kərʌblʲ-é	ʧʲirt-ʲé	vʲiʃʃʲistv-ʲé
nom pl	kərʌblʲ-í	ʧʲirt-í	vʲiʃʃʲistv-á
acc pl	kərʌblʲ-í	ʧʲirt-í	vʲiʃʃʲistv-á
gen pl	kərʌblʲ-éj	ʧʲért-∅	vʲiʃʃʲéstf-∅
dat pl	kərʌblʲ-ám	ʧʲirt-ám	vʲiʃʃʲistv-ám
inst pl	kərʌblʲ-ámi	ʧʲirt-ámʲi	vʲiʃʃʲistv-ámʲi
loc pl	kərʌblʲ-áx	ʧʲirt-áx	vʲiʃʃʲistv-áx
	‘ship’	‘feature’	‘substance’

The remaining 2% of nouns have stress alternations: in some case forms, stress falls on the suffix, and in others, stress is initial (Pattern C) or stem-final (Pattern D). We will call these types *mobile stress*. As shown in (5), mobile stress words exhibit the same vowel reduction alternations as final stress stems. In longer words, certain vowels always surface as unstressed: for example, the first vowel in [kəlbʌs-á] ‘sausage’ is always either a [ə] or a [ʌ], and the second vowel of [kóləkəl] ‘bell’ is always [ə].

(5) Russian stress types C and D: mobile stress

	C	C	D	D
nom sg	gəlʌv-á	kóləkəl-∅	dir-á	kəlbʌs-á
acc sg	góləv-u	kóləkəl-∅	dir-ú	kəlbʌs-ú
gen sg	gəlʌv-í	kóləkəl-ə	dir-ú	kəlbʌs-ú
dat sg	gəlʌv-ʲé	kóləkəl-u	dir-ʲé	kəlbʌs-ʲé
inst sg	gəlʌv-ój	kóləkəl-əm	dir-ój	kəlbʌs-ój
loc sg	gəlʌv-ʲé	kóləkəlʲ-i	dirʲé	kəlbʌs-ʲé
nom pl	góləv-i	kələkʌl-á	dír-i	kʌlbás-i
acc pl	góləv-i	kələkʌl-á	dír-i	kʌlbás-i
gen pl	gʌlóf-∅	kələkʌl-óf	dír-∅	kʌlbás-∅
dat pl	gəlʌv-ám	kələkʌl-ám	dír-əm	kʌlbás-əm
inst pl	gəlʌv-ámʲi	kələkʌl-ámi	dír-əmʲi	kʌlbás-əmʲi
loc pl	gəlʌv-ám	kələkʌl-áx	dír-əx	kʌlbás-əx
	'head'	'bell'	'hole'	'sausage'

There is broad agreement that fixed stress stems must be analyzed as underlyingly stressed, but theories disagree on the proper analysis of the alternating stress stems. Some take the final/poststem pattern to be the default (Nikolaeva 1971, Alderete 1999, Crosswhite et al. 2003), because this type is more numerous in the lexicon than patterns C and D, and because final stress stems tend to have slightly higher token frequency (Cubberly 1987). Crosswhite et al. (2003) take these frequency facts as one kind of support for their claim that stem-final stress is the phonological default for Russian. Halle (1973) and Melvold (1989), on the other hand, set up a special final/post-stem stress rule for Pattern B stems, but stress in Pattern C stems is decided in part by lexical stress and in part by phonological default rules. Thus, in [gəlʌv-á] 'head (nom sg)', the suffix is underlyingly stressed, and its stress surfaces. When neither the stem nor the suffix are underlyingly stressed, stress is assigned to the first syllable: /golov-i/ [góləv-i] 'head (nom pl)'. Pattern D stems are sometimes treated as a subtype of final stress/pattern B stems: they follow the final stress pattern in the singular, but stress is retracted one syllable to the left, onto the last syllable of the stem, in the plural.

There is generally no secondary stress in words with one root in Russian, no matter how long they are. Secondary stress only occurs in words with certain loan prefixes, such as [psʲèvdə]- 'pseudo' and [sùpʲir-] 'super', which we do not treat in this paper, and in compounds (Avanesov 1964, Wade 1992). According to the existing literature on secondary stress in compounds, its appearance is controlled by several factors:

- *Lexical stress type:* The presence of secondary stress depends on the lexical stress type of the left-hand stem (Yoo 1992). Yoo reports, based on a survey of dictionaries that transcribe stress (Ageenko & Zarva 1984, Borunova et al. 1988, Zaliznjak 1977), that secondary stress is fairly likely to surface on fixed stress stems, somewhat less likely in mobile stress stems, and even less likely in final stress stems—although he notes that there are exceptions and inconsistencies both between and within his sources.

- *Clash avoidance:* According to Avanesov (1964), the farther apart the two stresses, the more likely secondary stress is to surface. Yoo (1992) modifies this characterization, arguing that it is sufficient for stresses to be separated by two unstressed syllables (σ). Thus, $\grave{\sigma}\sigma\acute{\sigma}$ and $\grave{\sigma}\acute{\sigma}$ have a stress clash, but $\grave{\sigma}\sigma\sigma\acute{\sigma}$ does not. This type of stress clash prohibition is discussed by Nespor & Vogel (1989), who mention English examples such as *Mìssissippi múd* (cf. *Mississíppi*) (see also Alber 2005, Elenbaas & Kager 1999, Gouskova 2010).

- *Token frequency and register:* it is often noted that low-frequency, bookish words are more likely to have secondary stress. Comrie et al. (1996) observe that professionals who use compound terms of art frequently will use them without secondary stress, whereas laypeople would put secondary stress on such words. Gouskova & Roon (2009) confirm the effect of frequency in a small production study, and they suggest that secondary stress helps to signal morphological complexity, which is more of an issue in lower-frequency words.

- *Vowelless stems*: Gouskova & Roon (2009) find that CC- stems, such as [$\underline{\text{l}^{\text{j}}\text{n}}$-ʌ-vót] 'linen grower', are fairly likely to be pronounced with stress on the linker vowel, [l^{j}n-ò-vót]. They cannot conclude with certainty, however, whether this effect is due to the lack of a vowel in the stem, the marked sonority of the word-initial cluster, or the low token frequency of the stems tested in their study.

As can be seen from this discussion, more than one of these factors can be at play in any given compound, and they can conflict, since some of them favor secondary stress and others disfavor it. Unsurprisingly, it can be hard to disentangle these factors. Take, for example, the role of meaning in stress realization. Avanesov (1964) suggests that the "farther apart in meaning" the two stems in a compound, the more likely they are both to be stressed—but his examples are confounded by grammatical factors that disfavor stress. For example, he cites [blag-o-dúʂnij] 'placid' (lit., "good soul-adj"), but the first stem, [blag-ój] 'good', has mobile stress, and if stress appeared on the stem, there would be a clash: [blàg-ʌ-dúʂnij]. A more systematic investigation of secondary stress in compounds is undertaken in the remainder of the paper.

3 The design of the study

3.1 Introduction

At first blush, the obvious way to collect data on a variable phenomenon might seem to be a large-scale production study. Such a production study of secondary stress, however, would produce an inflated estimate of how common it is in informal speech: if speakers are asked to read infrequent long words, they sometimes read the word off the page "syllable-by-syllable," i.e., without vowel reduction, producing more secondary stress than they probably would outside the lab. Such a study also might not get the full range of possible pronunciations from each speaker, since speakers sometimes develop a strategy that limits variation (Albright & Hayes 2003 use a rating design for their study of English past tense formation precisely for this reason). The acoustic analysis in a production study would run into difficulties, as well: the phonetic correlates of secondary stress have not been studied extensively in a quantitatively rigorous way (Sleptsov 1975, Avanesov 1964, Zaliznjak 1977, Kuznetsova 2006; Gouskova 2010 has a small acoustic study of just three speakers). Experimenters would have to rely on their own perception to decide whether stress is present. While stressed mid and low vowels can be reliably identified as such because of unstressed vowel reduction, the difference between stress and lack thereof is harder to hear on high vowels.

Therefore, instead of a production study, we decided to present listeners with pronunciations of Russian words without secondary stress and with secondary stress in various locations within the word. If everybody hears the same range of pronunciations, we can make explicit comparisons between acceptable and outright impossible pronunciations as well as intermediate ones. This design can disentangle in a more controlled way the factors that control the realization of secondary stress in the grammar.

Rating studies have several advantages over production studies of variable phenomena, but it is not immediately obvious that rating studies tap the same knowledge as production studies. There is some evidence that rating studies might tell us more than production studies about the actual grammar that speakers use to assign ratings to words and to pronounce the words. Kawahara and Wolf's (2010) elicitation study of the accentual properties of the Japanese suffix [-zu] found that some produced only antepenultimate accents, others produced only initial accents, and still others alternated between the two patterns. Kawahara and Wolf construct a grammar that allows variation for some speakers but not others. However, in a series of follow-up judgment studies (both rating and forced choice), Kawahara & Kao (2012) found that Japanese speakers rated both types of patterns as acceptable, though the antepenultimate pattern is rated considerably higher. Thus, the production study indicated that the pattern is variable, but it did not make it clear that one of the variants is preferred. There are other differences between these studies that preclude a direct comparison, but the overall picture is analogous to variable stress in Russian compounds: production results might make it appear that some patterns are categorical, whereas rating studies uncover finer-grained distinctions.

3.2 Hypotheses tested in the study

The following hypotheses are based on previous descriptions of Russian compound stress in the descriptive literature and in the generative analyses of Yoo (1992), Gouskova & Roon (2009), and Gouskova (2010). First, we expect to find an effect of token frequency:

- H1: *Frequency effects.* Ratings should reflect an inverse correlation of token frequency and secondary stress realization: compounds with secondary stress should be rated as more acceptable as frequency decreases, whereas compounds without secondary stress should be rated as more acceptable as frequency increases.

We expect an effect of stress clash. Depending on the definition of clash, however, two competing hypotheses can be formulated. H2A is based on Yoo's characterization, whereas H2B is suggested by the traditional descriptions of Russian stress.

- H2A: *Categorical stress clash.* Pronunciations in which secondary stress is separated from the primary stress by zero or one syllables ($\grave{\sigma}\sigma\acute{\sigma}$, $\grave{\sigma}\acute{\sigma}$) should be rated as less acceptable than pronunciations in which stresses are separated by two or more syllables: [gòləv-ʌ-lómkə] ≻ [gʌlòv-ʌ-lómkə] 'puzzle.'

- H2B: *Gradient stress clash.* Ratings of pronunciations with secondary stress should get better as distance between stresses increases: $\grave{\sigma}\sigma\sigma\sigma\acute{\sigma} \succ \grave{\sigma}\sigma\sigma\acute{\sigma} \succ \grave{\sigma}\sigma\acute{\sigma} \succ \grave{\sigma}\acute{\sigma}$

We expect secondary stress realization and location to depend on the lexical stress status of the left-hand stem. Our hypotheses for compounds with fixed stress stems are as follows:

- H3: *Fixed stem stress preference:* Compounds whose left-hand stems have fixed stress in inflectional paradigms should be rated as more acceptable with secondary stress than without, because these are considered to be lexical stresses: [bʲitòn-ə-mʲiʂálkə] ≻ [bʲitən-ə-mʲiʂálkə] 'concrete mixer'.

- H4: *Stress movement*: Moving stress from its lexical position in fixed stress stems should be penalized: [bʲitòn-ə-mʲiʂálkə] ≻ [bʲètən-ə-mʲiʂálkə].

The motivation for H4 is as follows: moving stress is potentially advantageous as a way of avoiding a stress clash (Gouskova & Roon 2009), but if the main reason for the stress's appearance is to make the left-hand stem easily recognizable, moving stress would defeat that purpose (see the discussion in section 5.3).

As far as stems with stress alternations, Yoo's (1992) generalizations lead us to expect a difference between mobile and final stress stems. In mobile stress stems, secondary stress is supposed to be truly optional: [zʲèmlʲi-vlʌdʲélʲit͡s] ≈ [zʲimlʲi-vlʌdʲélʲit͡s] 'land owner'. Since we do not expect a difference in acceptability, this is a null hypothesis. If ratings correlate with the presence of stress, the null hypothesis will be disconfirmed. On the other hand, we do expect to find a difference in final stress stems. The alternative hypothesis is that both final and mobile stress stems should be rated as less acceptable when pronounced with secondary stress, since their stress patterns are assigned without a reference to lexical stress specification (Gouskova 2010).

- H5: *Final stem stress dispreference*: Final stress stem compounds should be more acceptable when pronounced without secondary stress than with: [rəb-ə-vlʌdʲélʲit͡s] ≻ [ràb-ə-vlʌdʲélʲit͡s] 'slave owner'.

Recall that vowelless (CC-) stems, which necessarily follow the final stress pattern ([lʲón] 'linen (nom sg)', [lʲn-á] (gen sg)),[2] were more likely to be pronounced with stress than without stress in the production study of Gouskova & Roon (2009). Thus, we expect stress to be more acceptable on the linker in such compounds:

- H6A: *Vowelless stem stress preference*: Secondary stress on the linker should be rated as more acceptable in vowelless stems than in longer final stress stems: [zl-ò-rádstvə] 'schadenfreude' ≻ [plʌd-ò-zbór] 'fruit harvest'.

Since there are several possible explanations for why these vowelless stems show this pattern, including sonority and frequency, we formulate the following hypothesis regarding the effects of sonority. The rationale for this hypothesis is that marked clusters in non-prominent positions might be avoided (see Smith 2002 for reasoning along these lines, as well as discussion section 5.4).

- H6B: *Sonority stress preference*: Secondary stress should be rated as more acceptable on vowelless stems with marked (falling sonority) clusters than on rising sonority clusters: [lʐ̨-è-dmʲítrʲij] 'impostor' ≻ [zl-ò-dʲéjstvə] 'evil-doing'.

3.3 Methods

3.3.1 Participants

Twenty-two native speakers of Russian participated in the study. All participants knew at least some English; they were recruited in New York City and at a linguistics summer school in Debrecen, Hungary. The participants ranged in age from 20 to 47, with a mean age of 27.68. There were 10 male and 12 female participants. None reported hearing or speech problems. Each participant received a small payment for his or her time.

3.3.2 Materials

The word list consisted of 60 compound words built from 35 left-hand stems that ranged from one syllable ([lʲn-ʌ-vót] 'linen grower') to four syllables in length (/kartofʲelʲ-e-xranʲilʲiʃʃʲe/ 'potato storage'), including the linker vowel. The right-hand stems ranged from one syllable to four syllables. Most of the left-hand stems contained at least some mid vowels. A native speaker of Moscow Russian, the first author, read each word with several different stress patterns: with no secondary stress on the first stem, and then with secondary stress on each of the available syllables of the first stem. Thus each word appeared in the experiment in N+1 unique forms, where N is the number of syllables in the first stem (see Table 1 for example paradigms). Primary stress was not manipulated but fixed in its normal lexical position. In a study such as this, with naturally produced stimuli, it is important to be explicit about the correlates of secondary stress that we believe were salient to the listeners, so we present an acoustic analysis of the stimuli in section 3.4. A full list of the stimuli is in the Appendix; our stimuli, results, and statistical analyses can be viewed at http://files.nyu.edu/mg152/public/russian/compounds/.

Table 1: Stimulus design

	1-σ stem	2-σ stem	3-σ stem	4-σ stem
	/lʲen-o-vod/	/zʲemlʲ-e-vladʲelʲets/	/golov-o-lomka/	/kartófʲelʲ-e-kopalka/
no $\grave{\sigma}$	lʲnʌvót	zʲimlʲivlʌdélʲits	gələvʌlómkə	kərtəfʲilʲikʌpálkə
1st σ	lʲnòvót	zʲèmlʲivlʌdélʲits	gòləvʌlómkə	kàrtəfʲilʲikʌpálkə
2nd σ		zʲimlʲèvlʌdélʲits	gʌlòvʌlómkə	kʌrtòfʲilʲikʌpálkə
3rd σ			gəlʌvòlómkə	kərtʌfʲèlʲikʌpálkə
4th σ				kərtəfʲilʲèkʌpálkə
	'linen grower'	'land owner'	'puzzle'	'potato digger'

The total number of test stimuli was 186. There were 53 right-hand stems, each of which appeared in 1.13 compounds on average, ranging from 1 to 3. There were 35 left-hand stems, each of which appeared in 1.7 compounds on average, ranging from 1 compound to 7 compounds. Each left-hand stem exhibited one of the three major stem stress patterns (16 fixed stress stems in 23 compounds, 8 vowelless final stems in 20 compounds, 6 longer final stress stems in 8 compounds, and 5 mobile stems in 8 compounds). In Table 1, *l^{j}n-* follows the final stress pattern, *golov-* is mobile, and *kartof^{j}el^j-* is a fixed stress stem. The set of 20 vowelless stem compounds can be further broken down into 7 compounds with CC- stem that starts with a sonorant, and 13 CC- compounds whose first stem starts with an obstruent. CC- stems are significantly underattested even among yer words in Russian (Gouskova & Becker to appear, Becker & Gouskova 2012), and many of them do not form any compounds (e.g., there are no compounds with the stem [rot/rt-] 'mouth' or [lob/lb-] 'forehead'). Thus the number of stems we could test was limited by the lexicon of Russian.

We obtained frequencies for each compound in the list from a Russian-language search engine Yandex (`http://yandex.ru`). The search engine returns the number of hits for the word, summing over its various case-inflected forms, unlike search engines such as Google, which did so for the most common Russian words but not for less frequent ones (at least this was the case when the experiment was run and the frequencies were obtained). The technique of using search engines for frequency estimates is discussed by Blair et al. (2002). There are a few small frequency dictionaries and corpora for Russian, including Zasorina (1977), Sharoff (2005), and the Russian National Corpus, but they do not include a wide range of compound words. For example, the word [golov-o-lomka] 'puzzle' occurs only 407 times in the RNC, and [dn-o-uglubitel^j] 'dredger' does not occur at all. We counterbalanced compounds of different length and stress type by token frequency, so that, for example, vowelless stem compounds included both relatively frequent words such as [sn-ə-v^{j}id^{j}en^{j}ijə] 'dream' and relatively infrequent ones such as [l^{j}n-ə-vót] 'linen grower'.

The stimuli were interspersed with 124 fillers. The fillers were both frequent and infrequent nouns of varying length and morphological complexity. Since the stimulus list contained a number of long compound words, we used polysyllabic fillers. The fillers were recorded with correct stress (e.g., [d^{j}jivʌl^{j}ónək] 'little devil'), with stress on the wrong syllable (e.g., [d^{j}ávəl^{j}inək]), and with secondary stress (e.g., [d^{j}àvəl^{j}ónək]), which is normally not found on single-root words in Russian.

3.3.3 Procedure

The experiment was conducted using PsyScope on a Macintosh computer. Listeners were given written instructions in Russian to listen to words of Russian that would be pronounced with more or less natural pronunciations. They were asked to rate the pronunciations according to how natural they were, on a scale of 1 (most natural) to 7 (least natural) (Weskott & Fanselow 2011). Participants were presented with an example of a perfect 1, a normally stressed pronuncation of the word [sʌbákə] "dog," and a 7, the word [d^{j}ilóvitəst^j] "business," which has stress on the wrong syllable (the correct pronunciation is [d^{j}ilʌvítəst^j]). There was a short training session with more examples. One of the training words was /tvorog/ 'cottage cheese', which most Russians know to have freely varying stress, presented as both [tvórək] and [tvʌrók].

Unlike some other words with variable stress, this one is not prescriptively stigmatized (see Avanesov 1968, Borunova et al. 1988). A Russian-speaking experimenter invited the participants to ask any clarification questions about the instructions before proceeding. Participants who asked about /tvorog/ were told that it was okay to rate both pronunciations as the same or to prefer one to the other.

After training, the participants heard the stimuli via headphones. The stimuli were presented in a random order by Psyscope. Each person heard each stimulus and filler exactly once; everyone heard every pronunciation of every word. Participants were given one break in the middle of the experiment. The entire experiment took about 30 minutes.

3.4 Acoustic analysis of the stimuli

To check that our naturally produced audio stimuli consistently cued primary and secondary stress as intended, we analyzed the vowels in the words for intensity and duration, which are common acoustic correlates of stress (Fry 1955 et seq.). Gouskova (2010) shows in a small acoustic study that in terms of vowel duration, primary stress vowels are longer than secondary stress vowels, which are in turn longer than pretonic vowels (see Bethin 2006) and unstressed non-pretonic vowels. We expect intensity to decline throughout the word (Trouvain et al. 1998), but stressed syllables should have greater intensity than unstressed ones. The vowels should also be reduced in unstressed syllables but not in stressed ones, consistent with the pattern followed by other native speakers.

The digital recordings of the stimuli were analyzed in Praat (Boersma & Weenink 2009) as follows. An unbiased coder labeled the vowels in every syllable of every word according to their supposed stress level (unstressed, primary stressed, secondary stressed, or pretonic) as well as orthographic/underlying vowel quality. The coder was not told of the purpose of the study until after all the labeling was completed and was not a native speaker of Russian. The segmentation criteria for vowels were established on the basis of the waveform and the spectrogram. Vowel boundaries were marked at the onset/offset of clear formant structure and amplitude peaks; in the vicinity of sonorant consonants, boundaries were marked at points of change in amplitude and formant frequencies. The duration and intensity of the stressed and unstressed vowels were then collected by script.

Means and standard deviations for duration and intensity are shown in Figures 1 and 2. The barplot error bars indicate standard error of the mean. In terms of duration, there were reliable distinctions between vowels of different stress levels: primary>secondary>pretonic>unstressed. In terms of intensity, the order was secondary>pretonic>primary>unstressed; recall that secondary stress syllables always precede primary stress syllables, and intensity normally decreases towards the end of the word. This intensity pattern is not atypical. Newlin-Łukowicz (2012) similarly finds that in Polish compounds, the first (secondary) stress is louder than the second (primary) stress, even though the primary stress vowel is longer than the secondary stress vowel. This is not apparently the case in English; Plag et al. (2011) show that primary and secondary stress syllables have a similar intensity even when primary stress occurs later in the word; their study investigated secondary stress in non-compound words in English. Intensity is also a cue for English compound stress, though not in all sentential contexts (Morrill 2012). As a reviewer points out, for English, the second syllable of two prominent syllables in a word can be perceived as more prominent even if it has the same intensity as the first one; languages can obviously differ on the phonetics of stress realization (see Gordon 2011 for a recent overview).

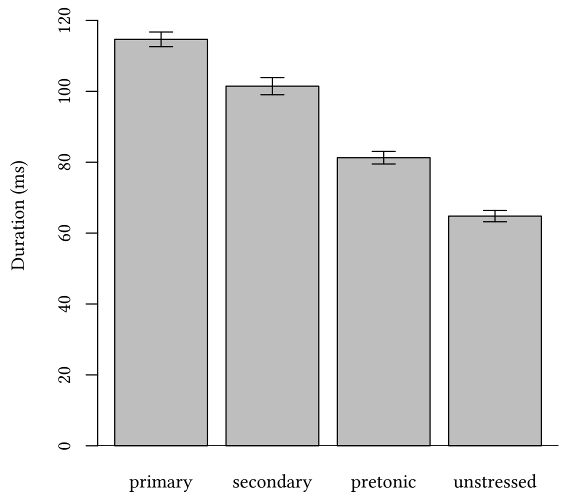

Figure 1: Duration (in milliseconds) as a function of stress level

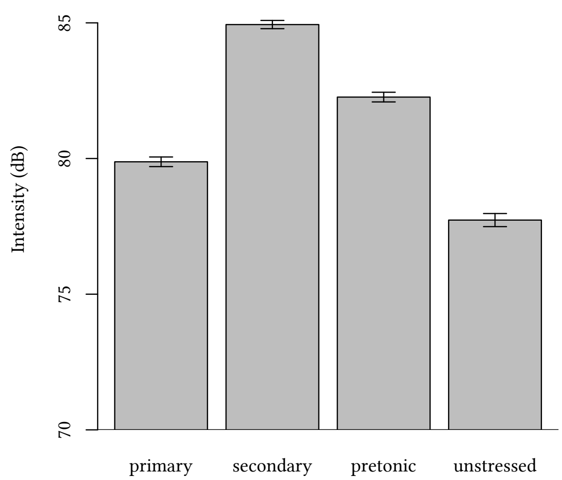

Figure 2: Intensity (dB) as a function of stress level

An analysis of variance with duration as a dependent variable and stress level (primary, secondary, pretonic, and unstressed) as an independent variable showed a main effect of stress level, $F(3, 962) = 137.6, p < .000$. A Tukey HSD post-hoc test confirmed that all four stress levels were reliably distinguished from each other by duration, as shown in Table 2. Therefore, secondary stress should have been reliably distinguishable from other levels by duration.

Because some of the compounds were pronounced with less natural stress patterns than others, we tested whether the natural pronunciations were different from unnatural ones in duration. To test whether duration was different for pronunciations of "correct" vs. "moved" stress, we fit a regression model with the stress level predictor and various other factors that affect vowel duration: frequency of the word, quality of the vowels, number of syllables in the word, position in the word, stress level, interaction of position and stress level, interaction of vowel quality and stress level, the presence of a sonority-reversed cluster in the syllable, the length of the first stem in syllables, and the type of stress manipulation. Frequency is log-transformed raw frequency based on Yandex counts; the vowel quality factor has five levels (a, e, o, u, i) for the orthographic vowels of Russian; position in the word has three levels (initial, medial, final); number of syllables is a numeric factor ranging from 2 to 8, the length of the first stem is a numeric factor ranging from 1 to 4, sonority fall is a logical factor that is true if the left-hand stem is vowelless and has a sonorant-initial cluster (e.g., [lʐ-e-naukə] 'pseudoscience') and false otherwise, and type of stress manipulation is a five-level factor that reflects whether stress was on the lexically stressed syllable, removed from a lexically stressed syllable, or moved to another syllable, and for compounds with mobile and final stress left-hand stems, whether stress was present or absent, since for those, there is no correct lexical location. We compared this model with one that lacked the type of stress manipulation predictor.[3] The predictor of stress correctness does not add much to this model of duration: $F(4, 931) = 1.96,\ p = 0.1$ (the Bonferroni-adjusted *p* value would have to be less than 0.0125 to count as significant).

Table 2: Results of a Tukey HSD post-hoc test for duration (ms) as a function of stress level

	Estimate	Standard Error	*t* value	Pr(>|*t*|)
pretonic - secondary	-16.19	3.44	-5.87	$< .000$
primary - secondary	13.21	3.52	3.76	.001
unstressed - secondary	-36.66	3.08	-11.91	$< .000$
primary - pretonic	33.39	3.05	10.94	$< .000$
unstressed - pretonic	-16.48	2.54	-6.49	$< .000$
unstressed - primary	-49.87	2.68	-18.91	$< .000$

An analysis of variance with intensity as a dependent variable and stress level as an independent variable showed a main effect of stress level ($F(3, 962) = 140.6$, $p < .000$). A Tukey HSD post-hoc test confirmed that all four stress levels were reliably distinguished from each other by intensity (see Table 3). Therefore, secondary stress should have been distinguishable from other levels by intensity. To test whether intensity was different for pronunciations of "correct" vs. "moved" stress, we fit a regression model with a number of factors that could affect vowel intensity: stress level, vowel quality, number of syllables in the word, and position in the word. We compared this model with one that included the additional predictor of stress correctness; the stress correctness predictor did even less in this model than in the model for duration ($F(4, 951) = 0.47$, $p = 0.75$), which means that correct and moved stresses were not detectably different from each other in intensity.

Table 3: Results of a Tukey post-hoc test for intensity (dB) as a function of stress level

	Estimate	Std. Error	t value	P(>|t|)
pretonic - secondary	-2.67	0.44	-6.04	$< .000$
primary - secondary	-5.06	0.45	-11.19	$< .000$
unstressed - secondary	-7.21	0.40	-18.21	$< .000$
primary - pretonic	-2.38	0.40	-6.08	$< .000$
unstressed - pretonic	-4.53	0.33	-13.9	$< .000$
unstressed - primary	-2.15	0.34	-6.34	$< .000$

Thus, all of the stress levels were reliably distinguished in terms of their acoustics in our stimuli, both by duration and by intensity. The correlates of stress in our stimuli are furthermore consistent with those of non-linguist speakers of Russian analyzed by Gouskova (2010).

4 Results of the rating study

All the participants rated the same set of words, but we split the words by the stress type of the left-hand stem for statistical analysis, since a number of factors are not useful for some of the stress types. For example, only fixed stress stems have a meaningful "correct" lexical stress location, and only final stress stems can be vowelless, since final stress is the only logically possible stress type for vowelless stems (e.g., [lʲón]∼[lʲn-á] 'linen (nom sg)/(gen sg)'). While it might be possible to fit a single statistical model to cover all of the stems' behavior, it is likely to be too complex to be presented and interpreted straightforwardly. We therefore analyze each stress type separately in detail in the subsequent sections. First, we present our results regarding the nature of stress clash in Russian.

4.1 Definition of stress clash

We start with the competing hypotheses regarding stress clash, H2A and H2B:

- H2A: *Categorical stress clash.* Pronunciations in which secondary stress is separated by zero or one syllables from the primary stress ($\grave{\sigma}\sigma\acute{\sigma}$, $\grave{\sigma}\acute{\sigma}$) should be rated as less acceptable than pronunciations in which stresses are separated by two or more syllables: [gòləv-ʌ-lómkə] ≻[gʌlòvʌlómkə] 'puzzle.'

- H2B: *Gradient stress clash.* Ratings of pronunciations with secondary stress should get better as distance between stresses increases: $\grave{\sigma}\sigma\sigma\sigma\acute{\sigma} \succ \grave{\sigma}\sigma\sigma\acute{\sigma} \succ \grave{\sigma}\sigma\acute{\sigma} \succ \grave{\sigma}\acute{\sigma}$

To test the effects of stress clash on ratings, we fit a linear hierarchical (mixed effects) model using the *lmer()* function in the *lme4* package (Bates & Maechler 2009) in R (R Development Core Team 2012). Before we discuss the model, a few comments are in order about all of the statistical analyses of people's ratings in section 4. We used Austin Frank's *mer-utils* and *regression-utils* code from `https://github.com/aufrank/R-hacks` to ensure that we had acceptably low collinearity in the models and to "center" variables where necessary. Numerical and binary predictors were centered when the condition number (κ) exceeded 15 and/or when the Variance Inflation Factor (VIF) exceeded 5 (for an explanation of κ and VIF, Baayen 2008:182, for a discussion of centering, see Belsley et al. 2004). A reviewer points out that centering predictors makes the estimates harder to interpret and may interfere with the condition number (Belsley 1984, Echambadi & Hess 2007); however, it yields an analogous model, and it sometimes allows a model to be computable in R ("converge") where it otherwise would not be possible (Gelman & Hill 2007). See also Gelman and Hill's chapter 4.2 on centering for models with interaction terms. Since there is no uncontroversial method for obtaining p values in mixed effects models at the moment, we estimated p values directly from t scores. At present, there are no established practices or clear guidelines as to whether it is appropriate to use ANOVA model comparison for deciding on whether to include random effects in a hierarchical model; cf. Baayen et al. (2008) vs. Barr et al. (2013). We use a design-driven approach recommended by Barr et al.: our models have the most complex random effect structure that is justified by the design, whether the inclusion of each random slope and intercept is justified by model comparison or not. Whenever a random effect is included in the reported model, it includes a random intercept in addition to random slopes. When the maximally specified models did not converge, we simplified the random effect structure by removing the terms with the smallest variance. All of our analyses and data are available for inspection on the first author's website.

We used a subset of our data that included only ratings for pronunciations of compounds with secondary stress; we furthermore excluded ratings for compounds with fixed stress stems in which stress was moved from its lexical location (as we show in section 4.2, pronunciations with moved stress get rated worse for reasons that have nothing to do with stress clash). The model coefficients are summarized in Table 4. The dependent variable is *rating* (an integer value ranging from 1 "best" to 7 "worst"), and the fixed factors are *interstress distance*, which is an integer value ranging from 0 (adjacent stresses, $\ldots\grave{\sigma}\acute{\sigma}\ldots$) to 4 (stresses separated by four syllables, $\ldots\grave{\sigma}\sigma\sigma\sigma\sigma\acute{\sigma}\ldots$), and *stress type*, which is a four-level factor for final, fixed, mobile, and vowelless stem compounds. Of these, final stress words are the baseline, i.e., model coefficients indicate predicted adjustments to the acceptability rating compared to the baseline condition. The model includes a by-participant slope for the interaction of *interstress distance* and *stress type*, allowing for the possibility that people use the rating scale differently and that their ratings are affected by these variables to a different extent. There is a by-word slope for interstress distance (where "word" is a full compound regardless of secondary stress condition). For fixed stress stems, interstress distance does not vary within words in this subset, but it does vary within words for other stem types. We also included a by-stem slope for interstress distance for left-hand stems. On average, each left-hand stem appeared in 1.71 compounds in this subset of the data, just as in the full dataset. The right-hand stem recurs in a few of our compounds, but it is reused so rarely that hierarchical grouping by this stem in addition to left-hand stems and words does not seem to be justified. We confirmed this indirectly by trying to fit models with all of these random effects in addition to a random effect by participant; none of them converged. A fully crossed model did not converge.[4]

As can be seen from the negative estimate for *interstress distance*, ratings improve the farther the stresses are from each other. This is consistent with both the gradient and the categorical view of clash, but we will show shortly that the gradient view accounts for the data better. There was a significant interaction between interstress distance and stem stress type. The nature of the interaction becomes clearer in Figure 3, which plots ratings by interstress distance for the four stem types.

Table 4: Model for stress clash, gradient definition

	Estimate	SE	t	p
Intercept (final stress, $\grave{\sigma}\acute{\sigma}$)	3.46	0.46	7.56	
interstress distance	−1.92	0.31	−6.23	0.000
stress type=fixed	−0.86	0.51	−1.70	0.09
stress type=mobile	−0.39	0.59	−0.66	0.51
stress type=vowelless	−0.56	0.56	−1.00	0.32
interstress distance:fixed	1.31	0.48	2.76	0.006
interstress distance:mobile	0.05	0.44	0.11	0.92
interstress distance:vowelless	1.63	0.48	3.42	0.001

N of observations = 1804 (82 pronunciations of 60 words, rated by 22 people)
Collinearity measures: $\kappa = 9.66$, VIF$= 3.57$, maximal correlation$= -0.67$.

Figure 3 shows the ratings (1="best", 7="worst") given to pronunciations arranged by the number of syllables intervening between stresses, and broken down by the four stem types. This figure is a beanplot, which is a vertical density plot with horizontal bars to indicate means. The area of each bean is based on the number of ratings per pronunciation type; the plot visualizes the same ratings that are modeled in Table 4 (i.e., it includes all the pronunciations with secondary stress except for fixed stress stems in which stress was moved from its lexical location). The plot in the upper right-hand corner shows ratings of longer final stress compounds (e.g., [jestʲestvo-vʲédʲenʲije] 'natural science') as a function of interstress distance. When three syllables intervene between stresses, the compounds got mostly ratings of "1", i.e., most acceptable, whereas pronunciations with adjacent stresses (interstress distance = 0) got worse ratings (the mean is above "5"). The final and mobile stress stems show the same trend: as distance between stresses increases, the ratings get better, in an approximately linear relationship. The trend in fixed stress stems is less linear; the graph suggests a more categorical division between $\grave{\sigma}\sigma\acute{\sigma}$ and longer distances. Notice that the farthest distance in fixed stress stems, 4 syllables, still receives better average ratings than the shorter distances. In the vowelless stem type (e.g., [sn-o-təlkʌvánʲijə] 'dream interpretation'), however, the relationship between stress distance and ratings is qualitatively different. The reasons for this are explored in section 4.4.

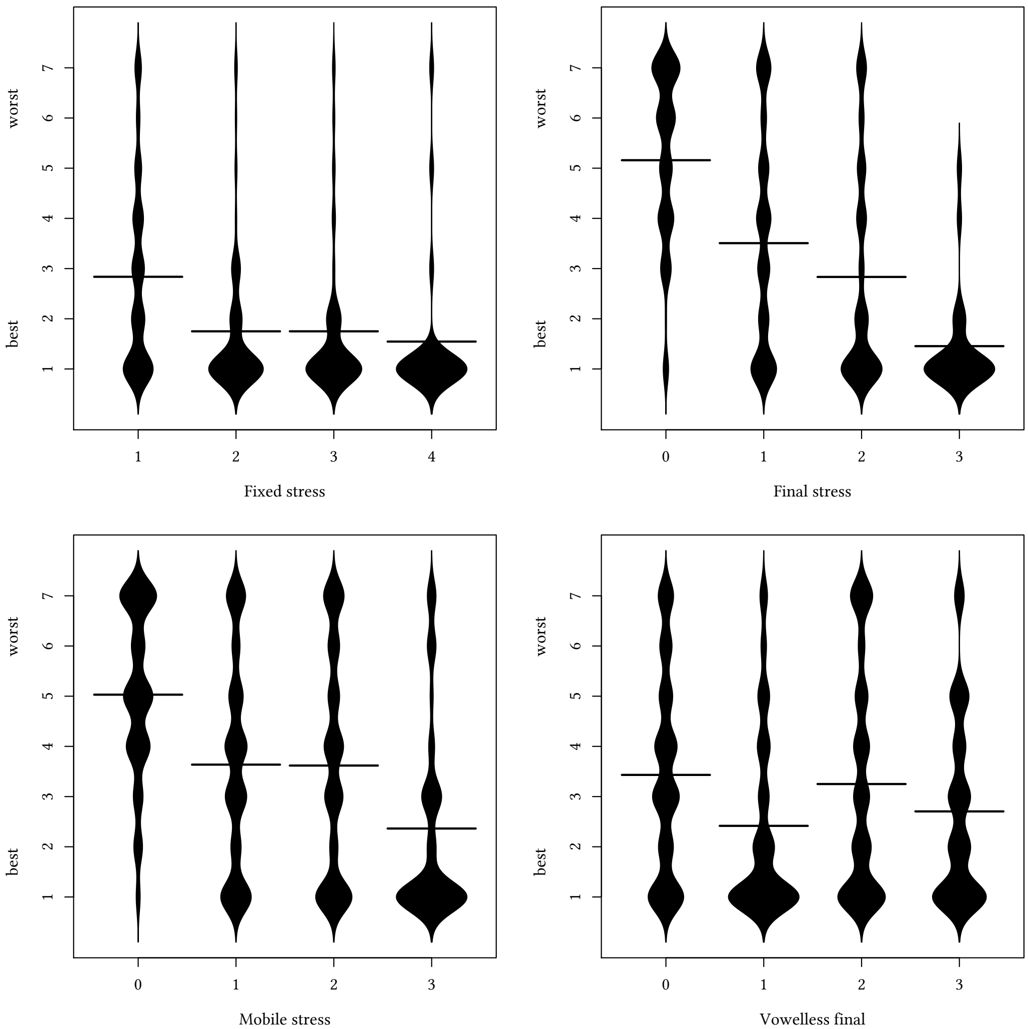

Figure 3: Ratings as a function of number of syllables between stresses, for the four different types of stems. (For fixed stress, only ratings of forms with correct stress are shown; an interstress distance of 0 is not possible for fixed stress stems unless the stress has been moved from its lexical location.)

To further explore the subpatterns within the data set, we refit the model with fixed stress stems as the baseline (see Table 5). The effect of interstress distance on the ratings of fixed stress stems goes in the expected direction (the coefficient is negative), although it does not reach significance. The model further shows that fixed stems differ from mobile and final stem compounds in this regard: for final and mobile stress stems, ratings get better with interstress distance, whereas for vowelless stems, ratings get slightly worse than those of fixed stress stems, though not significantly so.

Table 5: Model for stress clash, gradient definition, fixed stress stems as baseline

	Estimate	SE	t	p
Intercept (fixed stress, $\grave{\sigma}\sigma\acute{\sigma}$)	2.60	0.35	7.48	
interstress distance	-0.60	0.37	-1.61	0.11
stress type=final	0.86	0.51	1.70	0.09
stress type=mobile	0.47	0.51	0.93	0.35
stress type=vowelless	0.30	0.48	0.63	0.53
interstress distance:final	-1.32	0.48	-2.76	0.006
interstress distance:mobile	-1.27	0.49	-2.57	0.01
interstress distance:vowelless	0.31	0.52	0.60	0.55

N of observations = 1804 (82 pronunciations of 60 words, rated by 22 people)
Collinearity measures: $\kappa = 9.66$, VIF$= 5.29$, maximal correlation $= -0.76$

Fixed stress stem compounds and vowelless stem compounds pattern together because their ratings depend on additional phonological factors that introduce variability not accounted for by interstress distance alone. Mobile and final stress stems are the ones that show the effect of interstress distance most clearly, but they are also relatively small subsets of our data (there there were 572 ratings of the 8 final stress stem compounds, and 660 ratings of the 9 mobile stress stem compounds). The reason the sets are small is that mobile and final stress stems are relatively uncommon in the lexicon (recall section 3.3.2), and not all roots form compounds that allowed us to balance frequency (see the Appendix). Since our findings about the gradience of stress clash are in line with the traditional descriptions of Russian compound stress, we believe they would hold up with more data.

The overall pattern of ratings is inconsistent with the categorical definition of clash used in Yoo's analysis: there is no obvious cutoff point between $\{\grave{\sigma}\acute{\sigma},\ \grave{\sigma}\sigma\acute{\sigma}\}$ on the one hand and $\{\grave{\sigma}\sigma\sigma\acute{\sigma}\}$ and longer distances on the other. To test this, we fitted a model similar to the one in Table 4, but instead of using the ordinal measure of interstress distance as a predictor, we used a simple binary predictor that was true if stress distance was less than 2 and false otherwise. The interactions and random effect structure of the model in Table 6 are analogous to the model in Table 4.[5] The binary clash model unsurprisingly shows an effect of clash presence, as is expected: it simply cuts the more fine-grained ordinal predictor of interstress distance into two arbitrary groups. The simplification of the clash predictor comes at the cost of explanatory power, which we can assess by comparing the models' Akaike Information Criterion values (Gelman & Hill 2007). A smaller AIC indicates better fit; for the interstress distance model, AIC=6870, whereas AIC=7088 for the model with binary clash. There is another sign that the model underdetermines the data: the interaction term for clash and stem type is not significant, and thus the model misses a difference between mobile and final stems on the one hand and fixed stems on the other.[6]

Table 6: Model for ratings of secondary stress using a binary definition of clash

	Estimate	SE	t	p
Intercept (final stress, bin clash=F)	3.64	0.36	10.24	
binary clash = TRUE	1.41	0.35	4.06	0.000
stress type=fixed	−1.23	0.36	−3.38	0.0007
stress type=mobile	0.03	0.45	−0.07	0.95
stress type=vowelless	−0.59	0.41	−1.46	0.14
binary clash:fixed	−0.52	0.47	−1.11	0.27
binary clash:mobile	−0.21	0.48	−0.43	0.67
binary clash:vowelless	−0.67	0.51	−1.33	0.18

N of observations = 1804 (82 pronunciations of 60 words, rated by 22 people)
Collinearity measures: $\kappa = 7.58$, VIF$= 3.51$, maximal correlation $= -0.68$

There is no reason to separate 0 and 1 on the one hand from greater stress distances, and the statistical analyses support a more complex view of clash in Russian. There is, however, a visible separation between adjacent stresses and other distances in Figure 3, for several compound subtypes. Adjacent stresses are only possible in Russian compounds when secondary stress falls on the linker vowel (as in [lʲn-ò-vót] 'linen grower' or [gəlʌv-ò-tʲáp] 'bungler'), so stress on the linker seems to be particularly dispreferred (although not in vowelless stems, as we will show in section 4.4).

Finally, before accepting the gradient definition of clash, we need to rule out an alternative explanation: that the ratings improve the closer secondary stress is to the left edge of the phonological word. Since secondary stress is more likely to be closer to the beginning of the word as the distance between stresses increases, an initial default for stress might explain some of the pattern (Halle 1973, Melvold 1989). We can distinguish between these explanations by looking at ratings as a function of secondary stress position and by looking at ratings as a function of stress distance in just those words that have initial stress, which already satisfy the initial default. We fit a simple hierarchical model to just a subset of our items that had stress on the first syllable (excluding, again, words with fixed stress stems where stress has been moved). The model includes random by-participant slopes for interstress distance and stress type but not their interaction (the model with the interaction slope did not converge); it also includes a random intercept for words and a random by-left-stem slope for interstress distance. The interaction of interstress distance and stress type did not improve the model and was excluded from the final model. Stress type also did not affect ratings by itself, so it is included only in the by-participant slope. Stress distance does not vary by word in this subset, but it does vary by stem; hence we include the by-stem random slope for this predictor.[7] As shown in Table 7, interstress distance is a predictor of ratings for these words as well. The relationship between ratings and interstress distance in compounds with initial secondary stress is shown graphically in the beanplot in Figure 4. We also tried an analogous model with a binary clash predictor; the predictor was significant ($t = 2.72$, $p = 0.03$) but once again the model did not compare favorably based on AIC (3988 for gradient clash, 3997 for binary clash).[8] Thus, when all stresses are already initial, the binary definition of clash does not explain as much variance in the ratings as gradient clash.

Table 7: Model for gradient stress clash words with initial stress

	Estimate	SE	t	p
Intercept ($\grave{\sigma}\acute{\sigma}$)	2.66	0.25	10.70	
interstress distance	-0.37	0.08	-4.49	0.000
N of observations = 1034 (47 pronunciations of 47 words rated by 22 people)				

Collinearity measures: $\kappa = 1.00$, VIF$= 1$, maximal correlation $= 0$

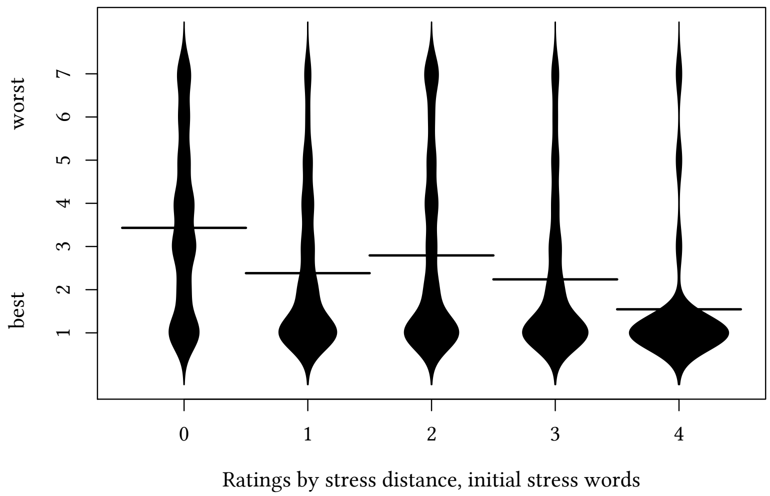

Figure 4: Ratings as a function of distance between stresses for words with initial secondary stress

Thus, we conclude that the correct characterization of stress clash effects in Russian must make reference to stress distance, as in H2B; the categorical cutoff of Yoo's clash is not justified by the data. We discuss the implications of this finding for the definition of the phonological constraint against stress clash in section 5.5.

4.2 Fixed stress stems

The hypotheses relevant to fixed stress stems are H1, H3, and H4:

- H1: *Frequency effects.* Ratings should reflect an inverse correlation of token frequency and secondary stress realization: compounds with secondary stress should be rated as more acceptable as frequency decreases, whereas compounds without secondary stress should be rated as more acceptable as frequency increases.

- H3: *Fixed stem stress preference:* Compounds whose left-hand stems have fixed stress in inflectional paradigms should be rated as more acceptable with secondary stress than without: [bʲitòn-ə-mʲiʂálkə] ≻[bʲitən-ə-mʲiʂálkə] 'concrete mixer'.

- H4: *Stress movement*: Moving stress from its lexical position in fixed stress stems should be penalized: [bʲitòn-ə-mʲiʂálkə]≻[bʲètən-ə-mʲiʂálkə].

We tested these hypotheses in a linear hierarchical model (the patterns are also illustrated graphically below). We used a step-down procedure: we started with a fully crossed model that included all of the predictors of interest, then removed one predictor at a time and compared the resulting model with the full model. The predictors were log frequency, stress manipulation, and the interaction of frequency and stress manipulation. The stress manipulation levels are *deleted* for pronunciations without secondary stress (e.g., [bʲitən-ə-mʲiʂálkə] 'concrete mixer'), *correct* for pronunciations with secondary stress in the lexical location (e.g., [bʲitòn-ə-mʲiʂálkə], cf. [bʲitón] 'concrete'), and *moved,* for pronunciations with secondary stress somewhere other than the lexically stressed syllable (e.g., [bʲètən-ə-mʲiʂálkə]). Neither frequency nor its interaction with stress manipulation improved the model (ANOVA model comparison for the interaction: $\chi^2(2) = 1.46,\ p = 0.48$; for frequency:$\chi^2(1) = 1.01,\ p = 0.31$). The best model for ratings includes only one fixed factor: the three-level factor encoding types of stress manipulation (see Table 8). The model includes random by-word and by-left-stem slopes for stress manipulation and a by-participant slope for the interaction of stress manipulation and frequency. A fully crossed model did not converge.[9] As can be seen from the model in Table 8, moving stress results in significantly worse ratings, confirming H4. On the other hand, pronunciations with secondary stress in the lexical location do not differ from pronunciations without secondary stress—we were unable to confirm H3. Figure 5 plots ratings as a function of stress manipulation.

Table 8: Model for compounds with fixed stress left-hand stems

	Estimate	SE	t	p
Intercept (*deleted stress*)	1.95	0.22	8.81	
stress manipul.=moved	3.14	0.25	12.41	< 0.0000
stress manipul.=correct	0.27	0.30	0.91	0.36

N of observations = 1980 (90 pronunciations of 23 words rated by 22 people)
The collinearity measures for the model are $\kappa = 5.19$, VIF$= 1.01$, maximal correlation$= 0.11$

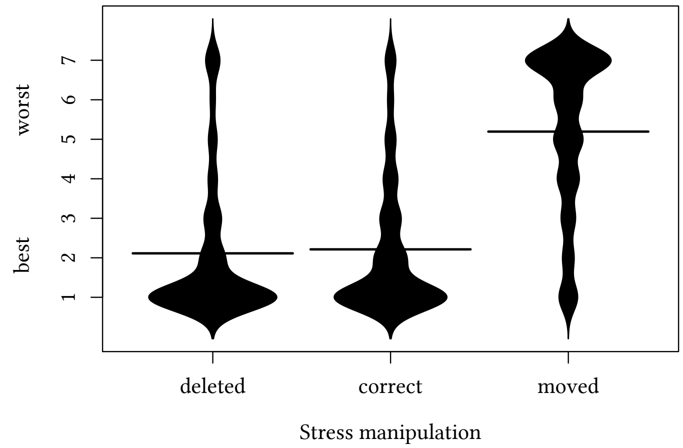

Figure 5: Effect of stress manipulation on the ratings of fixed stress stem compounds

We were unable to confirm the effect of frequency on ratings of pronunciations with secondary stress (H1). This is not surprising when we examine Figure 6, which plots ratings for fixed stress stems as a function of log frequency. Ratings are grouped by stress manipulation; the lines are simple regression lines produced by the *lm()* function in R. The correlation between ratings and frequency is flat in pronunciations with secondary stress on the lexically stressed syllable (leftmost panel). For both pronunciations without secondary stress and pronunciations with moved stress, ratings slightly improve with frequency, though overall, the ratings are much worse when stress has been moved.

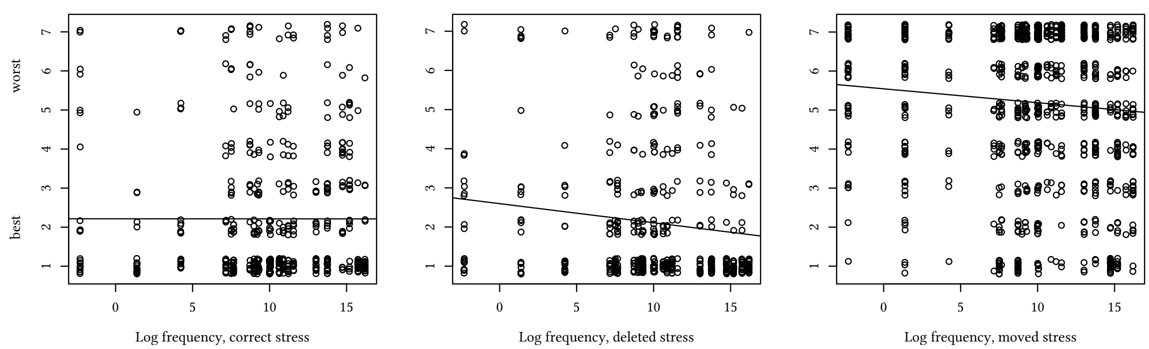

Figure 6: Fixed stress stems: ratings by frequency and stress manipulation

To summarize, we expected that moving stress would have a negative impact on ratings, but we also expected that words with no secondary stress would be rated worse than words with secondary stress on the correct syllable. Instead, we found that secondary stress is optional for fixed stress stem compounds—only moving stress has a negative effect on ratings.

4.3 Mobile stress stems

In compounds whose left-hand stems have mobile stress (such as [gəlʌv-á], [góləv-i], [gʌlóf] 'head (nom sg)/(nom pl)/(gen pl)'), there is no one "correct" stress location, and so the only hypothesis to test for this type of stems is H1 (for compounds with secondary stress, higher frequency means better acceptability, whereas for compounds without secondary stress, higher frequency means worse ratings). We do not look at the effects of clash in this section, since they were already established in section 4.1.

Our model for the ratings of mobile stress stem compounds is shown in Table 9. The model has *rating* as a dependent variable and two fixed effects: *secondary stress* (true if there is a secondary stress, false otherwise) and *log frequency*, as well as the interaction of *secondary stress* with log-transformed token frequency; there is a by-subject random slope for the interaction term and frequency and secondary stress, and a by-word slope for secondary stress, as well as a random intercept for the left-hand stem. A fully crossed model did not converge. The frequency predictor was centered to reduce collinearity in the model (without centering, $\kappa = 32.1$, VIF $= 9.79$, maximum correlation$= -0.95$).[10]

Table 9: Model for mobile stress compounds

	Estimate	SE	t	p
Intercept (*secondary stress=FALSE*)	1.33	0.13	10.03	
secondary stress=TRUE	2.50	0.29	8.69	0.000
log frequency	-0.09	0.04	-2.36	0.018
sec. stress=TRUE : log frequency	0.24	0.07	3.20	0.014

660 observations (30 pronunciations of 9 words rated by 22 participants)
The collinearity measures are $\kappa = 4.06$, VIF$= 1.12$, maximum correlation$= 0.26$

The interaction between log frequency and secondary stress for mobile stems is plotted in Figure 7. The regression lines show that ratings get worse as frequency increases for words with secondary stress, but they improve as frequency increases without secondary stress—exactly what we would expect under H1.

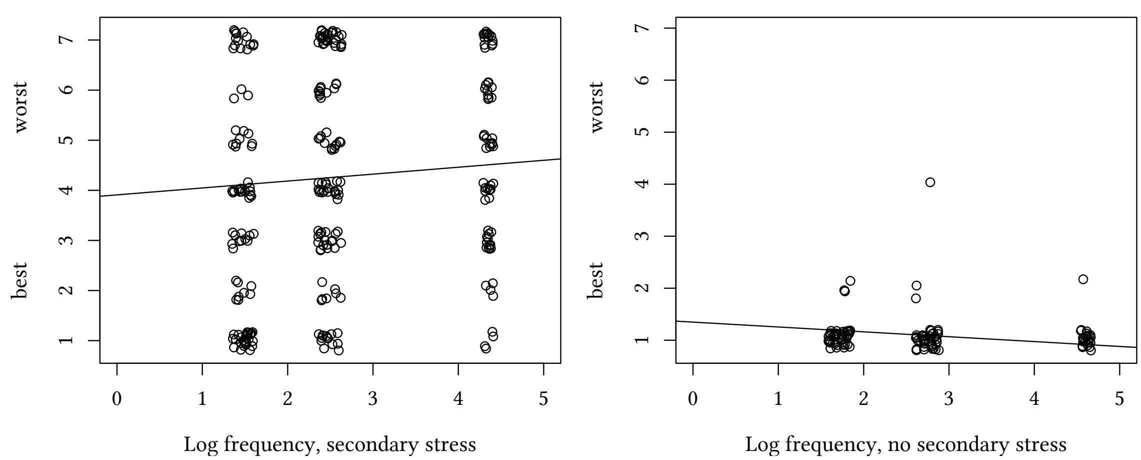

Figure 7: Mobile stress stems: ratings by log frequency (centered) for pronunciations with and without secondary stress

People rated compounds with mobile stress stems as most acceptable when they were pronounced without secondary stress (as in [gələvə-kruʐénʲijə] 'vertigo'): the mean rating for these was 1.35. Pronunciations with secondary stress received an average rating of 3.91 (a difference of 2.56 points). This is shown in the left-hand beanplot in Figure 8. This difference was not expected under Yoo's characterization, since mobile stems are supposed to have freely variable stress. Recall from section 4.1, moreover, that ratings for mobile stress stem compounds get better with the distance between stresses. The right-hand beanplot in Figure 8 shows the ratings as a function of interstress distance and includes the pronunciations without secondary stress. As can be seen from this graph, the worst ratings were given to pronunciations with adjacent stresses (as in [gəlʌv-ò-lómkə] 'puzzle').

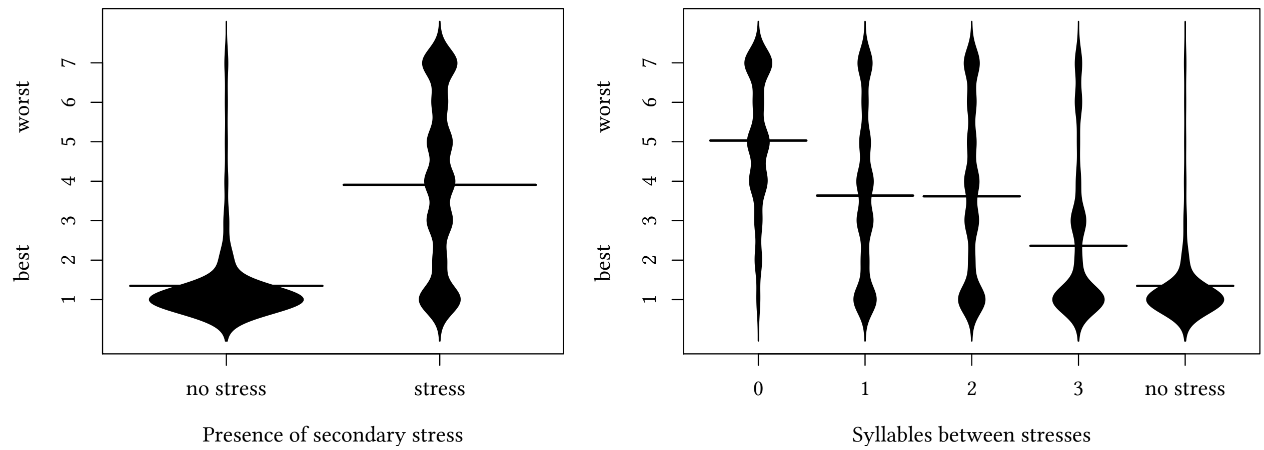

Figure 8: Effects of stress presence and interstress distance on ratings of mobile stress stem compounds

To summarize, we were able to confirm the hypothesis (H1) about the interaction between secondary stress and frequency in compounds with mobile stress stems. Also, for these compounds, secondary stress degrades acceptability (somewhat unexpectedly, given previous descriptions), and, as we showed in section 4.1, the location of stress matters: the closer the stresses are to each other, the less acceptable the pronunciation.

4.4 Final stress stems

Final stress stems in our experiment come in two varieties: vowelless stems such as [sn-ə-vʲidʲénʲijə] 'dream' and [lʲd-ə-əbrəzʌvánʲijə] and longer stems such as [pləd-ʌ-zbór] 'harvest' and [jistʲistv-ʌ-vʲédʲinʲjə] 'natural science'. The hypotheses relevant to all final stress stems are H1 (interaction of secondary stress and frequency) and H5; the hypotheses relevant to vowelless stems were H6A and H6B.

- H5: *Final stem stress dispreference*: Final stress stem compounds should be more acceptable when pronounced without secondary stress than with: [rəb-ə-vlʌdʲélʲit͡s] ≻[ràb-ə-vlʌdʲélʲit͡s] 'slave owner'.

- H6A: *Vowelless stem stress preference*: Secondary stress on the linker should be rated as more acceptable in vowelless stems than in longer final stress stems: [lʐ-è-prʌrók] 'linen thresher'≻[plʌd-ò-zbór] 'fruit harvest'.

- H6B: *Sonority stress preference*: Secondary stress should be rated as more acceptable on vowelless stems with marked (falling sonority) clusters than on rising sonority clusters: [lʐ-è-dmʲítrʲij] 'impostor'≻[zl-ò-dʲéjstvə] 'evil-doing'.

The results of our linear mixed effects model for the ratings of final stress stem compounds are shown in Table 10. The dependent variable is *rating*, as before. The predictors that ended up in the final model were the binary predictor coding the presence of *secondary stress*, *log frequency*, and the interaction of *secondary stress* and *sonority stem type*. *Sonority stem type* was a three-way predictor distinguishing RC- stems with sonority fall (where R stands for "sonorant") from CC- stems without sonority fall and longer stems. We included random by-participant slopes for stress and sonority stem type, a random by-word intercept, and a random by-left-stem slope for stress and frequency. Models with more complex random effect structure did not converge.[11] As can be seen from the coefficient estimates, the presence of secondary stress correlates with significantly worse ratings, as expected under H5. Higher frequency words are rated slightly better, as can be seen from the negative coefficient. We were, however, unable to confirm H1 for final stress stems—the interaction of secondary stress and log frequency did not improve the model ($\chi^2(1) = 0.07$, $p = 0.79$). Vowelless stems are rated worse than longer stems when they consist of a cluster with falling sonority (in our data, [lʐ-], [lʲn-] and [lʲd-]; there were 7 compounds with these stems), but, as expected under H6B, they are more acceptable with secondary stress than stems that consist of CC- clusters. Just being vowelless does not degrade a final stress stem's acceptability, though, with or without stress—the presence of secondary stress in CC- stems is associated with slightly better ratings than stressless CC- stems (the coefficient for *secondary stress=TRUE:CC-* is negative), but this predictor does not reach significance. Thus, we were unable to confirm H6A.

Table 10: Model for compounds with final stress left-hand stems

	Estimate	SE	t	p
Intercept (*stressless long stem*)	1.53	0.18	8.30	
log frequency	−0.10	0.02	−4.90	0.000
secondary stress=TRUE	1.91	0.29	6.55	0.000
sonority stem type = RC-	0.59	0.22	2.63	0.009
sonority stem type = CC-	0.11	0.21	0.55	0.58
secondary stress=TRUE : RC-	−1.08	0.38	−2.79	0.005
secondary stress=TRUE : CC-	−0.20	0.37	−0.54	0.59

1452 observations (66 pronunciations of 28 compounds, rated by 22 people)
Collinearity measures: $\kappa = 12.36$, VIF$= 1.77$, maximum correlation $= -0.55$

To illustrate these effects graphically, we first plot the relationship between ratings and log frequency for pronunciations with and without secondary stress in Figure 9. The regression lines have negative slopes in both scatterplots: more frequent words are rated as more acceptable, whether they have secondary stress or not. Pronunciations with secondary stress receive overall worse ratings than pronunciations without secondary stress, since the very presence of secondary stress introduces additional factors that impact acceptability (such as clash violations). This can be seen from the higher intercept of the regression line in the left-hand scatterplot in Figure 9. The overall relationship between secondary stress and frequency in final stress stems is not consistent with H1, however, since H1 predicts a positive slope in the left-hand graph (i.e., ratings should get worse for pronunciations with secondary stress as frequency increases).

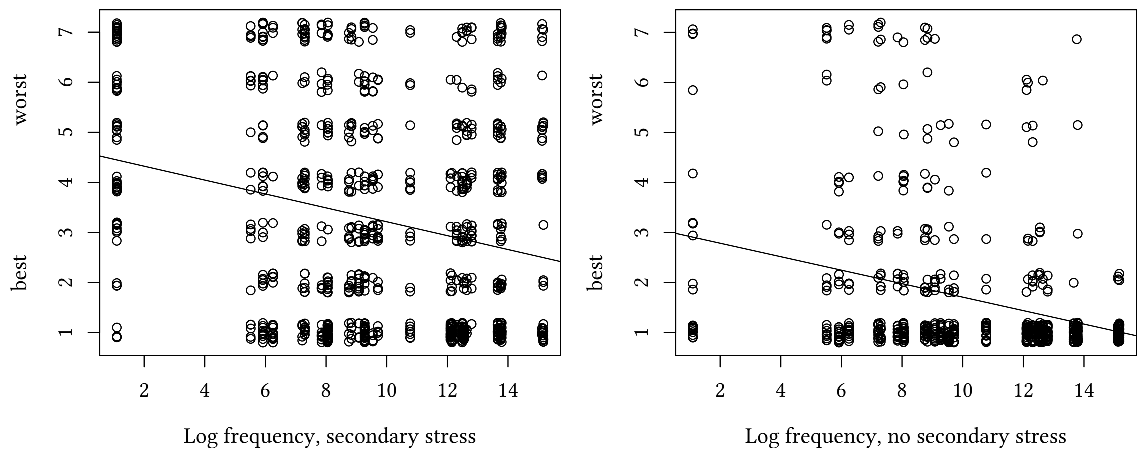

Figure 9: Final stress stems: ratings by frequency for pronunciations with and without secondary stress

The effect of secondary stress in the three subtypes of final stress stems and the interaction between stem type and stress is plotted in Figure 10. For both longer stems such as [jistʲistv-ʌ-vʲédʲinʲijə] 'nature science' and obstruent-initial (CC-) vowelless stems such as [sn-ə-təlkʌvánʲijə] 'dream interpretation', the pattern is similar: pronunciations without secondary stress are rated better (1.60 and 1.56 respectively) than pronunciations with secondary stress (3.54 and 3.34 respectively). For sonorant-initial (RC-) vowelless stems, pronunciations with secondary stress (such as such as [lʲd-ò-tʲéxnʲikə] 'ice technology') are rated as more acceptable (average rating of 2.57) than stressed pronunciations of longer stems (such as [jèstʲistv-ʌvʲédʲinʲijə]) and obstruent-initial stems (such as [sn-ò-təlkʌvánʲijə]). Pronunciations of RC- stems without secondary stress, on the other hand, are rated slightly worse (average rating of 2.01) than stressless pronunciations of longer and obstruent-initial stems. It is clear from this plot that the difference between having stress and not having it is far less pronounced in RC- stems than in other final stress stems.

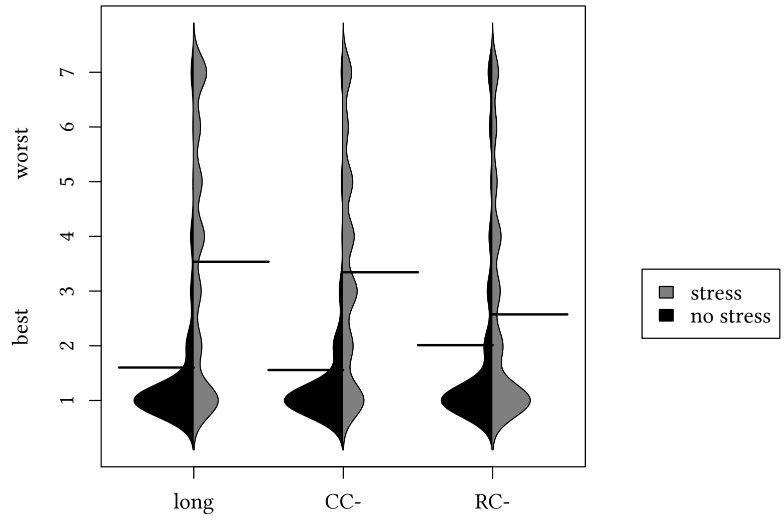

Figure 10: Effect of secondary stress on final stress stems (long, obstruent-initial vowelless, and sonorant-initial vowelless)

To summarize, we could not confirm that frequency interacts with secondary stress for final stress stems, although frequent compounds were rated better than less frequent ones. We found that secondary stress in general depressed ratings for final stress stems, consistent with previous descriptions. We were also able to explain why secondary stress is more likely to surface on vowelless stems. These are non-uniform: obstruent-initial stems pattern with longer final stress stems, whereas sonorant-initial stems with marked sonority clusters (e.g., [lʐ-ə-nʌúkə] 'pseudoscience') are rated as more acceptable than other types of stems when they are pronounced with secondary stress on the linker vowel. This is interesting because compounds with falling sonority RC- stems are independently rated lower than other types of final stress stem compounds. Finally, recall that interstress distance is an additional factor in final stress stems, and they again pattern in a non-uniform way.

5 Discussion

5.1 Task effects

We found that in fixed stress stems, secondary stress was optional—there was no penalty for its absence, even though it has been claimed that secondary stress should be the default for these stems (Yoo 1992, Gouskova & Roon 2009). For mobile and longer final stress stems, pronunciations of secondary stress got worse ratings than pronunciations without secondary stress. Why did our findings differ from those of previous studies? One possible reason is the "syllable-by-syllable" pronunciation characteristic of formal settings and speech in the lab, which encourages more stress. Recall that Yoo's sources are prescriptive pronouncing dictionaries (Ageenko & Zarva 1984, Borunova et al. 1988). It is possible that the prescriptive norm favors more stress, whereas in normal colloquial usage, secondary stress is not required. Lagerberg

(2007) has noted that prescriptive dictionaries tend to underreport variability in Russian stress even for primary stress, where prescriptivists tend to pick one pronunciation as the norm. In fixed stress stems, the location of secondary stress is fairly obvious—it should fall on the lexically stressed syllable—and so that is where it is usually transcribed in the dictionaries.

Our methodology allows us to verify the optionality of secondary stress in Russian because the listeners are given the pronunciations that they might not necessarily produce in a lab setting. We showed that listeners do not accept just any pronunciation they are given—they do reject certain forms as unacceptable, suggesting that a grammar is at work in guiding the decisions. Listeners' ratings of the filler stimuli show that this is the case (and also that they understood the task). Considering only non-compound filler words, Figure 11 shows that listeners rated words with stress on the correct syllable as highly acceptable (mean rating score= 1.28) and rated words with stress on the incorrect syllable highly unacceptable (mean rating score $= 6.00$). A simple mixed-effects model with subject and word as random effects and whether the primary stress had been moved as a fixed effect showed that moving the primary stress of the word had a significant negative impact on ratings ($t = -64.62$).

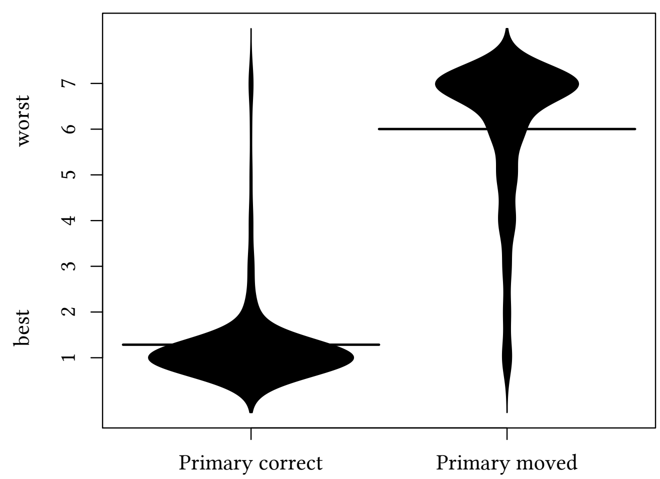

Figure 11: Ratings for fillers with stress in the correct vs. moved location

Pronunciations of compounds with secondary stress were not rejected outright by our experiment participants, however, which suggests that secondary stress is definitely within the range of acceptable—if more marked—pronunciations for compounds. We conclude that rating studies are a fruitful method of studying variable phonological phenomena.

5.2 Effects of frequency

Some of our findings diverge from previous descriptions of Russian compound stress. First of all, we found limited evidence of the oft-mentioned effect of frequency on Russian secondary stress: secondary stress was only more likely to be rated higher on less frequent compounds if the left-hand stem had mobile stress.

In the other stress types, we either found no effect of frequency (for fixed stems) or we found that frequency affected the ratings independent of stress. This was the case for longer final stress compounds: frequent compounds were rated as more acceptable than infrequent compounds, whether they were pronounced with or without secondary stress. In this case, people might have been rating these words as less acceptable simply because they were less familiar with them, not because of how the words were pronounced.

We cannot rule out that frequency interacts with the presence of stress on the basis of its not being a consistently significant predictor of ratings, especially since the effects of frequency are often small (Baayen 2010); we may simply not have enough power to detect it. It could be the case that our measure of token frequency is imperfect, and that the expected effect would be found in subsequent studies that measure frequency differently or include more items and participants. If our findings are confirmed in other studies, however, then our result seems to put into question our earlier hypothesis (Gouskova and Roon 2009) that secondary stress helps listeners parse less frequent compounds morphologically; instead, secondary stress might be a feature of the production grammar only. Speakers simply produce more secondary stress on less frequent compounds for reasons we discuss in section 5.1.

A reviewer suggests that ratings may be influenced by frequency in another way: people use frequency in assigning worse ratings to *pronunciations* that are rare. Thus, people rate [gòləvʌlómkə] 'puzzle' as less acceptable than [gələvʌlómkə] simply because they hear the pronunciation with secondary stress less often. This hypothesis is difficult to test directly at present because Russian corpora that include stress information are still relatively small and do not include enough compounds (see section 3.3.2). The search engine Yandex does not include information about secondary (or indeed any) stress, either. Still, we doubt that this interpretation of the effect of frequency is right for our results. Recall that for both fixed and final stress compounds, ratings get better with frequency, though the effect is only statistically significant in final stems. Fixed stress compounds with moved stress are rated as more acceptable when frequency increases, but there is no apparent correlation between ratings and frequency for pronunciations with correct stress. According to traditional descriptions, pronunciations with correct stress are the norm for fixed stress compounds; this is what people pronounce and hear most often. But in our study, ratings seem to be unaffected by frequency precisely in this case. If people were guided by how often they heard various pronunciations, they would presumably rate more frequent compounds with correct secondary stress as more acceptable simply because they hear such pronunciations more often. This is speculative because we do not have a good way to measure the frequency of pronunciations; still, in order to look for an explanation along the lines the reviewer suggested, one would have to show an effect of frequency in fixed stress stems.

5.3 Differences between lexical stress types: the role of faithfulness

We identified a different set of generalizations for the three stress stem types from what Yoo (1992) reported. Yoo (1992) makes a distinction between fixed stress stems, which prefer to have secondary stress, mobile stress stems, which have it optionally, and final stress stems, which prefer to have no secondary stress. We found instead that people rate fixed stress stems as more or less equally acceptable when they are pronounced with or without secondary stress, provided that stress is in the correct position. This is a case of free variation: secondary stress is optional in fixed stress stems.[12]

For both mobile and final stress stems, the presence of secondary stress meant worse ratings. Gouskova (2010) suggests a phonological explanation for this difference between fixed stress stems on the one hand and mobile/final stress stems on the other. Fixed stress stems are underlyingly specified for stress location, whereas mobile and final stress are not stressed underlyingly. The difference in surface stress pattern in inflectional paradigms (reviewed in section 2) arises because stress location in fixed stress stems is underlyingly specified (and is therefore governed by faithfulness constraints, which prohibit insertion, movement, and deletion of stress), whereas the location of stress in mobile and final stress stems is determined by the constraints that restrict surface phonological patterns (in Gouskova's analysis, this is lexically specific phonological markedness constraints, which prohibit certain phonological structures in a subset of morphemes; see Ito & Mester 1995, Pater 2000, 2006, Inkelas et al. 1997, Becker 2009, inter alia). These constraints impose an initial stress in mobile stems and final stress in final stems, but the requirements may be overridden in compound grammar: stress is optionally assigned to the left-hand stem to highlight its morphological status, but the location is determined by factors such as clash avoidance. Indeed, the two stem types show a similar effect of interstress distance in compounds, which sets them apart from fixed and vowelless RC- stems (recall Table 4). Second, both mobile and longer final stress stems suffer worse ratings when secondary stress is present. We cannot conclude with certainty that this effect arises because the insertion of stress is problematic or because this stress introduces a stress clash. Either explanation is consistent with our results.

5.4 Vowelless stems and sonority in clusters

We found that vowelless stem compounds such as [lʲn-ò-vót] 'linen grower' are fairly acceptable with stress on the linker vowel, in a departure from the general pattern found in longer final stress stems. Gouskova & Roon (2009) identified that stress is more likely to surface on the linker vowel in such stems than in other compounds, but it wasn't clear in that study whether this was because the compounds were of low frequency or because of their phonological properties. We were able to show in the present study that this effect is due to the marked sonority profile of the word-initial clusters in these compounds rather than low token frequency.

The vowelless stems are split into two subpatterns in our study. Obstruent-initial CC- stems (e.g., [sn-ə-təlkʌvánʲijə] 'dream interpretation') behave like longer final stress stems in that people prefer them without secondary stress. Sonorant-initial RC- stems,[13] on the other hand, are in their own category: people do not mind them without secondary stress, but they do rate them higher when they have secondary stress on the linker vowel, which otherwise is the worst possible location for secondary stress. Why would this be? Gouskova & Roon (2009) speculate that clusters with marked falling sonority are preferentially placed in prominent positions. It is cross-linguistically common to restrict marked structure to prominent positions such as the stressed syllable or for stress to be attracted to prominent marked structure (Zoll 1998, Beckman 1998, Crosswhite 1999, Smith 2002). For example, Beckman (1998) cites languages that have click consonants in root-initial syllables but not elsewhere; many languages have a full range of consonantal place contrasts in syllable onsets but not codas (Ito 1986 et seq.); in English, aspirated consonants and

[h] occur only in word-initial and stressed syllables (Davis 1999). So the Russian pattern is not surprising (although the clusters themselves are cross-linguistically unusual enough that evidence of their interaction with stress is that much more rare).

We can reject the explanation that the RC- stems are special because they belong to the "yer" class in Russian, which shows lexically restricted vowel deletion before vowel-initial morphemes (e.g., /lʲo̲d-a/ [lʲd-a] 'ice gen sg'; cf. [lʲot] 'ice nom sg').[14] If yer deletion had special consequences for the stress grammar, then RC- stems would pattern with CC- stems such as [son]/[sn-] 'sleep', but the latter behave much like the longer final stress stems. It appears that sonority is responsible.

This finding has larger implications for the role of sonority in phonology. Russian is often cited as an example of a language where sonority sequencing is not respected in word-initial clusters (Clements 1990, Blevins 1995), the implication being that sonority might not matter in the system. Our results indicate that even in Russian, clusters that violate sonority sequencing are phonologically distinguished from those that do not. There is evidence from syllable structure rules that consonant clusters are not all treated equally (Bethin 1992, Szpyra 1992, Yearley 1995, Gouskova 2012, Gouskova & Becker to appear, and others), but our study is the first, we believe, to find evidence for the more marked clusters being special in the stress system. It is not the first study to find onset-sensitive stress in Russian, however: Ryan's (2013) study of the Russian lexicon found that words that begin with CC clusters are more likely to have initial stress than words that begin with one consonant, at least in the fixed stress subset (we considered the distribution of stress on RC- vs. CC-initial words in the electronic version of Zaliznjak (1977), but did not find anything conclusive). Our findings indicate that it is not just the number of consonants but their arrangement that matters. Ryan suggests that the longer duration of clusters could explain why they attract stress; this would be the an excellent question for future research.

5.5 Implications for the phonological treatment of stress clash

We found some evidence to decide between the conflicting characterizations of stress clash in Russian compounds. Traditional descriptions of Russian describe stress clash as a gradient effect: the farther apart the stresses, the more likely secondary stress is to surface (Avanesov 1964). Yoo (1992) casts the effect of clash in categorical terms: stresses are not allowed to be adjacent or be separated by one syllable, but two or more syllables between stresses are acceptable. This definition of clash is attractive because this type of clash has been claimed to be active outside of Russian: Nespor & Vogel (1989) cite examples such as the English *M̲ì̲ssissippi M̲ú̲d* (cf. *Missis̲s̲í̲ppi* in isolation), and Kager (1994) defines a version of clash as penalizing adjacent feet, not just two stressed syllables (i.e., $(\sigma\grave{\sigma}_{Ft})(\sigma\acute{\sigma}_{Ft})$ violates clash as much as $(\grave{\sigma}_{Ft})(\acute{\sigma}_{Ft})$ does). Kager's concern is ternary rhythm in languages such as Finnish, which is sometimes analyzed as a resolution of two conflicting pressures: avoidance of stress lapses/unfooted syllable strings and avoidance of foot clashes. The constraints against stress lapses have been argued to come in several varieties, some of which penalize short lapses and others—longer ones (Steriade 1997, Elenbaas & Kager 1999, McCarthy 2003, Gordon 2005), but there has not been as much discussion of gradience in stress clash effects. Russian is an interesting case because there is no general avoidance of stress lapses in its stress phonology, so the effects of stress clash are independently visible.

The anti-clash preference in Russian is easy to describe, but formalizing it as a phonological constraint is not straightforward. Let us assume that acceptability ratings reflect a form's performance on the phonological constraint hierarchy (Hayes & Wilson 2008, Daland et al. 2011, Becker et al. 2011, 2012, Gouskova & Becker to appear, Coetzee & Pater 2008, 2011). The ratings of Russian compounds with mobile or final left-hand stems depend primarily on interstress distance. This fact presents a challenge for any formally precise theory of clash avoidance. For example, in constraint-based phonological frameworks (e.g., Optimality Theory, Prince & Smolensky 1993/2004), this could be expressed as the extent of satisfaction of *Clash. A typical definition of *Clash is "assign a violation mark for every pair of adjacent stresses" (Prince 1983, Selkirk 1984, Nespor & Vogel 1989). Two constraints against clashes are illustrated in (6), where violations of constraints are shown as numerical penalties, and full satisfaction is indicated with a zero.[15] The first two words have no stress clashes under any definition of standard constraints against clashes: neither feet nor stressed syllables are adjacent in (6a) or (6b). The word in (6c) has adjacent feet, and so it receives one violation of the version of the constraint that penalizes adjacent feet. The form in (6d) has adjacent feet and adjacent syllables, violating any definition of *Clash. Crucially, the standard definitions of anti-clash constraints do not distinguish (a) and (b), even though stresses are farther apart in (6b) than in (6a).

(6) Violations of *Clash constraints, as standardly defined

	*Clash(stressed syllables)	*Clash(Feet)
a. $(\sigma\grave{\sigma})\sigma\sigma(\sigma\acute{\sigma})$	0	0
b. $(\sigma\grave{\sigma})\sigma(\sigma\acute{\sigma})$	0	0
c. $\sigma(\sigma\grave{\sigma})(\sigma\acute{\sigma})$	0	-1
d. $\sigma(\sigma\grave{\sigma})(\acute{\sigma})$	-1	-1

To capture the gradient clash effect in Russian, *Clash would need to assign *fewer violations* the more syllables separate the stresses. It is difficult to define such a constraint negatively because it is unclear what candidate receives zero violations: the distance between stresses can be unboundedly large. The constraint can, however, be defined as assigning *rewards* in proportion to interstress distance, as in (7). This constraint would assign zero marks to the worst kind of clash, where stresses are adjacent (see 8); stresses separated by one syllable receive one reward point, stresses separated by two syllables receive two points, and so on. Thus, the distinction between (8a) and (8b) is captured by this definition of *Clash. Russian compounds almost never include more than two roots, so two is the maximum number of stresses. We assume that positively defined *Clash only looks at adjacent stresses; thus, the second and fourth syllables in $\grave{\sigma}\sigma\grave{\sigma}\sigma\acute{\sigma}$ only get one reward each, for the stressed syllables that are nearest to them in the metrical grid.

(7) *Clash+ (positive gradient definition): 'Assign a reward for every unstressed syllable that separates two stresses.'

(8) Violations of *Clash+, defined as a gradient positive constraint

		*Clash+
a.	$(\sigma\grave{\sigma})\sigma\sigma(\sigma\acute{\sigma})$	$+3$
b.	$(\sigma\grave{\sigma})\sigma(\sigma\acute{\sigma})$	$+2$
c.	$\sigma(\sigma\grave{\sigma})(\sigma\acute{\sigma})$	$+1$
d.	$\sigma(\sigma\grave{\sigma})(\acute{\sigma})$	0

Constraints against stress clashes have figured in metrical stress theory since the early 1980's, and they have been defined in several ways in parameter-based theories (Prince 1983, Nespor & Vogel 1989, Yoo 1992) and in later constraint-based frameworks (Kager 1994, Elenbaas & Kager 1999).[16] Work in metrical stress theory has aimed to generate all and only the attested stress patterns since at least Hyman (1977) and Hayes (1980), but it is now becoming increasingly clear that there are counterexamples to typological trends that were once considered to be universal (e.g., Altschuler 2006). At the same time, Hayes & Wilson (2008) demonstrate that metrical stress constraints can be learned from distributional evidence in the lexicon, as long as the learner is provided with the means of encoding these constraints at the appropriate level. Our findings suggest that clash avoidance is more complex than previously thought, and that the options for defining constraints against stress clashes should include gradience.

6 Conclusion

We presented a rating study of secondary stress in Russian compounds. The study confirmed and refined some previously known generalizations about secondary stress: its location and realization depends on the lexical stress pattern of the left-hand stems, and secondary stress is dispreferred when it is close to the main stress. We also identified some new generalizations. We found that clash avoidance is gradient in Russian: the farther apart the stresses, the better. We suggest that these results are compatible with defining *Clash as assigning a reward for every syllable that separates stresses; the facts cannot be captured with the standard view that *Clash only penalizes adjacent stresses. We found that secondary stress in general degrades ratings, but for lexically stressed stems, secondary stress is optional: people rate them as equally acceptable without stress and with it, though moving stress from its lexical location is penalized. We found that secondary stress is tolerated well on linker vowels connecting vowelless (yer) stems such as [lʲd-] in [lʲd-ò-əbrʌzovániʲə] 'ice formation', but that this effect was limited to stems with marked falling sonority clusters. This is a subtle effect of syllable markedness in a somewhat unexpected domain, secondary stress assignment.

Appendix

Transcriptions indicate orthographic vowel quality, without vowel reduction, since each vowel in the left-hand stems appeared with unreduced quality when secondary stress was placed on it. Secondary stress is shown for fixed stress stems only.

syllable count in stem 1	Word	Gloss	Frequency in Yandex	Left-hand stem type	Zaliznjak's type
0	dno-uglubʲítʲelʲ	dredger	517	final/vless	1b
2	jestʲestvo-vʲédʲenʲije	natural science	10631	final	1b
1	kʲino-atelʲjé	movie studio	3158	final	0
1	kʲino-lʲénta	reel of film	855678	final	0
1	kʲino-zvʲezdá	film star	947406	final	0
2	korablʲe-nósʲets	ship carrier	3	final	2b
0	lʲdo-obrazovánʲije	ice formation	6333	final/vless	1b
0	lʲdo-tʲéxnʲika	ice technology	248	final/vless	1b
0	lʲno-vód	linen grower	6886	final/vless	1b
0	lʐe-dmítrij	impostor	180828	final/vless	8b'
0	lʐe-naúka	pseudoscience	280585	final/vless	8b'
0	lʐe-prʲisʲága	false oath, lie	3119	final/vless	8b'
0	lʐe-prorók	false prophet	189105	final/vless	8b'
1	plodo-zbór	fruit harvest	371	final	1b
1	rabo-vladʲélʲets	slave driver	266248	final	1b
1	skoto-promíʂlennostʲ	cattle industry	1474	final	1b
0	sno-tolkovánʲije	dream interpretation	1360	final/vless	1b
0	sno-vidʲénʲije	dream vision	3712791	final/vless	1b
0	sto-lʲétnik	aloe/agave	47809	final/vless	b (no qual. for numerals)
0	tʃʃʲe-dúʂije	feebleness	2572	final/vless	1b (tʃʃʲeta)
0	tʃʃʲe-slávʲije	vanity	977225	final/vless	1b (tʃʃʲeta)
0	zlo-dʲéjstvo	villainy	362944	final/vless	1b
0	zlo-jazítʃʲije	slander	8788	final/vless	1b
0	zlo-nrávʲije	depravity	16491	final/vless	1b
0	zlo-pixátʲelʲ	spiteful critic	221232	final/vless	1b
0	zlo-rádstvo	Schadenfreude	308888	final/vless	1b
0	zlo-ʃʃʲástʲije	misery	13927	final/vless	1b
0	zlo-upotrʲeblʲénʲije	abuse	3863020	final/vless	1b
2	bʲetòno-mʲeʂálka	concrete mixer	941468	fixed	1a
1	bòmbo-derʐátelʲ	bomb rack	8069	fixed	1a
3	kartòfʲelʲe-kopálka	potato digger	23378	fixed	2a
3	kartòfʲelʲe-tʃʲístka	potato peeler	104452	fixed	2a
3	kartòfʲelʲe-xranʲílʲiʃʃʲe	potato storage	22123	fixed	2a
3	kukurùzo-vód	corn grower	6174	fixed	1a
2	maʂìno-strojénʲije	machine building	10834069	fixed	1a
2	mʲetàllo-lóm	scrap metal	2511009	fixed	1a
2	mʲetàllo-rʲézka	metal cutter	1839	fixed	1a
3	oboròno-promíʂlʲennostʲ	defense industry	4	fixed	1a
3	oboròno-sposóbnostʲ	defense capability	452314	fixed	1a

syllable count in stem 1	Word	Gloss	Frequency in Yandex	Left-hand stem type	Zaliznjak's type
1	pòʧʲvo-utomlʲénʲije	soil fatigue	1288	fixed	1a
1	pòʧʲvo-vʲéd	soil scientist	73838	fixed	1a
2	sàxaro-varénʲije	sugar refining	9942	fixed	1a
2	sobàko-vódstvo	dog breeding	950341	fixed	3a
1	tànko-dróm	tank training area	53525	fixed	3a
1	tànko-nósʲets	tank carrier	69	fixed	3a
1	tànko-strojénʲije	tank building	41226	fixed	3a
2	tovàro-poluʧʲátelʲ	goods receiver	2130	fixed	1a
1	vʲèro-ispovʲedánije	denomination	6867839	fixed	1a
1	xlòpko-vód	cotton grower	10933	fixed	3a
2	zakòno-dátʲelʲ	lawmaker	3979956	fixed	1a
2	zòloto-lovʲítelʲ	gold catcher	0	fixed	1a
2	golovo-kruʐénʲije	vertigo	1547889	mobile	1f'
2	golovo-lómka	puzzle	10729164	mobile	1f'
2	golovo-tʲáp	bungler	14854	mobile	1f'
1	maslo-bójʃʃʲik	butter churner	4596	mobile	1c
1	volko-dáv	wolfhound	1814754	mobile	3e
1	zʲemlʲe-dʲélʲets	farmer, tiller	560677	mobile	2d'
1	zʲemlʲe-vladʲélʲets	land owner	646438	mobile	2d'
1	zʲerno-obrabótka	grain processing	1715	mobile	1d
1	zʲerno-vʲédʲenʲije	grain science	1515	mobile	1d

Notes

[1] All examples are transcribed in the IPA and reflect surface forms in Moscow Russian as pronounced by the first author, unless otherwise indicated. We use [ʂ] for "ш", [ʐ] for "ж", [ʧʲ] for "ч", and [ʃʃʲ] for "щ"; see Hamann (2004), Zygis & Padgett (2010). We indicate primary stress with an acute accent on the stressed vowel, and secondary stress with a grave accent. The distinction between orthographic "и" and "ы" is indicated as palatalization or lack thereof on the preceding consonant "и" is Cʲi, and "ы" is Ci (Padgett 2003). The following abbreviations are used in glosses: "nom" is "nominative", "acc" is "accusative", "gen" is "genitive", "inst" is "instrumental", "loc" is "locative" (a.k.a. prepositional case), "sg" is "singular", "pl" is "plural".

[2] These morphemes contain the famous so-called yer (jer) vowels, which have been the subject of much attention in the literature on Slavic (Lightner 1965 et seq). These are mid vowels that delete in a select set of morphemes in the context of most affixes, including compound linker vowels. It should be noted that vowelless stems are usually classified as part of the final stress type, but they are not entirely homogeneous: for example, /lʲod/ 'ice' is only stressed on the stem vowel when it is the only vowel (cf. [lʲót]/[lʲd-á] 'ice (nom/gen sg)), but feminines have stem stress when they occur with the instrumental suffix [-ju] ([lóʂ], [lʐ-í] but [lóʐ-ju] 'lie (nom/gen/inst sg)). Gouskova 2010 provides some arguments that the linker is phonologically and morphologically grouped with the left-hand stem rather than with the right-hand stem.

[3] The full model can be seen in the R (R Development Core Team 2012) file on the first author's website.

[4] We report the results of model comparisons for each random effect, which we conducted by taking the model we report and comparing it with the model minus the relevant random effect. We then took the resulting model and compared it with a model

that had only a random intercept or, in case of interaction random slopes, models with individual terms. For full details, see the R file on the first author's website. For the model in Table 4, the comparisons were: the reported model minus the by-participant slope, $\chi^2(15) = 486$, $p < 0.0001$; the latter model plus a by-participant intercept, $\chi^2(1) = 445$, $p < 0.0001$. The reported model minus the by-word slope, $\chi^2(3) = 80$, $p < 0.001$; the latter model plus by-word intercept, $\chi^2(1) = 78$, $p < 0.0001$. The reported model minus the by-left-stem slope, $\chi^2(3) = 12,\ p < 0.007$; the latter model plus by-stem intercept, $\chi^2(1) = 1.47,\ p = 0.23$.

[5]Model comparisons for random effects: the reported model minus by-participant interaction slope, $\chi^2(15) = 404,\ p < 0.0001$; the latter model plus by-participant intercept, $\chi^2(1) = 378,\ p < 0.0001$. The reported model minus by-word slope, $\chi^2(3) = 14,\ p = 0.003$, the latter model plus by-word intercept, $\chi^2(1) = 3.93,\ p = 0.047$. The reported model minus by-left-stem slope, $\chi^2(3) = 25,\ p < 0.0001$; the latter model plus by-left-stem-intercept, $\chi^2(1) = 25,\ p < 0.0001$.

[6]We would like to thank an anonymous reviewer for pointing this out.

[7]The model comparisons for random effect inclusion are: the reported model minus the by-participant slope for stress distance and stress type, $\chi^2(15) = 207,\ p < 0.0001$; the latter model plus a by-participant slope for stress distance, $\chi^2(3) = 192,\ p < 0.0001$, and plus a by-participant slope for stress type, $\chi^2(10) = 200$, $p < 0.0001$. The reported model minus by-word intercept, $\chi^2(1) = 0,\ p = 1$. The reported model minus by-left-stem slope, $\chi^2(3) = 35,\ p < 0.0001$; the latter model plus by-left-stem intercept, $\chi^2(1) = 34,\ p < 0.0001$.

[8]Model comparisons for random effect structure: the reported model minus by-participant slope for binary clash plus stress type, $\chi^2(15) = 205$, $p < 0.0001$; the latter model plus a by-participant slope for binary clash, $\chi^2(3) = 189$, $p < 0.0001$, and plus a by-participant slope for stress type, $\chi^2(10) = 200,\ p < 0.0001$, and plus a by-participant intercept, $\chi^2(1) = 181,\ p < 0.0001$. The reported model minus the by-word intercept, $\chi^2(1) = 0.97,\ p = 0.32$. The reported model minus the by-left-stem slope for binary clash, $\chi^2(3) = 28,\ p < 0.0001$; the latter model plus a by-left-stem intercept, $\chi^2(1) = 28,\ p < 0.0001$.

[9]Model comparisons for random slopes/intercepts: the reported model minus by-participant slope, $\chi^2(10) = 521$, $p < 0.001$; the latter model plus by-participant intercept, $\chi^2(1) = 433,\ p < 0.0001$. The reported model minus by-word slope, $\chi^2(6) = 13,\ p < 0.04$; the latter model plus by-word intercept, $\chi^2(1) = 9.17,\ p = 0.002$. The reported model minus by-left-stem slope, $\chi^2(6) = 24,\ p = 0.0005$; the latter model plus by-left-stem intercept, $\chi^2(1) = 0.73,\ p = 0.39$. We report the maximal model justified by the design whether or not the inclusion of random slopes and intercepts is supported by the model comparison (see Barr et al. 2013 for proof and discussion).

[10]Model comparisons for random effects: the reported model minus by-participant interaction slope, $\chi^2(10) = 148,\ p < 0.0001$, by-participant slope for the interaction vs. just the slope for secondary stress plus frequency, $\chi^2(6) = 139,\ p < 0.0001$, the latter model vs. secondary stress, $\chi^2(3) = 38,\ p < 0.0001$, the stress plus frequency model vs. frequency only, $\chi^2(3) = 2.84,\ p =$0.41, by-participant intercept, $\chi^2(1) = 99,\ p < 0.0001$. The reported model minus the by-word slope, $\chi^2(3) = 24,\ p < 0.0001$, the latter model plus by-word intercept, $\chi^2(1) = 11,\ p = 0.0007$. The reported model minus by-stem intercept, $\chi^2(1) = 0.09,\ p = 0.77$. The by-stem intercept is is justified by the design: each stem appeared on average in two compounds in this subset of the data.

[11]Model comparisons for random effects: reported model minus the by-participant slope, $\chi^2(10) = 190$, $p < 0.0001$; the latter model plus a by-participant slope for stress, $\chi^2(3) = 185$, $p < 0.0001$; plus a by-participant slope for sonority stem type, $\chi^2(6) = 145$, $p < 0.0001$. The reported model minus by-left-stem slope for secondary stress and word frequency, $\chi^2(6) = 23$, $p = 0.0008$; the latter model plus a by-left-stem slope for frequency, $\chi^2(3) = 5.68,\ p = 0.13$, and plus a by-left-stem slope for stress, $\chi^2(3) = 26,\ p < 0.0001$; the model without a random effect for left stems vs. a by-left-stem intercept, $\chi^2(1) = 5.25,\ p = 0.02$. The model without the word intercept did not converge.

[12]The difference in acceptability between deletion and movement of secondary stress in fixed stress stems observed in the present study is also predicted by Gouskova and Roon's (2009) analysis of Russian compound stress. That analysis, formulated in terms of the stress constraints in Optimality Theory (Alderete 1999), proposes that the constraint prohibiting stress movement (No-Flop(Stress)) is ranked above anti-clash constraints, which are in turn ranked above the constraint prohibiting stress deletion

(Max(Stress)). This analysis allows underlying stresses to be deleted to avoid clash violations, but not moved.

[13]The binary distinction between obstruent- and sonorant-initial clusters is a fairly crude one, but we do not have enough data to establish a more subtle gradient effect of the sonority scale on secondary stress, and it might be hard to test this in Russian because frequency affects ratings and these compounds are not all equally common in the language.

[14]Work that specifically discusses the sonority issue and the stress properties of yer morphemes includes, inter alia, Halle (1973), Halle & Vergnaud (1987b), Melvold (1989), Yearley (1995), Gouskova (2012), Gouskova & Becker (to appear).

[15]We assume that normal markedness constraint definitions in Standard OT say "Assign a violation mark for every instance of X", where X is a local structural configuration such as an unstressed syllable at the edge of a word or a pair of adjacent unstressed syllables (Eisner 1997, McCarthy 2003, Pater 2006, see especially McCarthy 2008a 4.4 and 4.5). Markedness constraints in OT and related constraint-based frameworks penalize surface structures that have certain properties; canonically, a marked structure is in some sense difficult articulatorily or perceptually.

[16]In parallel OT, positively stated constraints run into the so-called infinite goodness problem (Prince 2007, fn. 9, Kimper to appear) because word length is unbounded, and because clash violations can in principle be resolved through vowel epenthesis. We do not discuss this in detail, but there is a solution to this problem, discussed in Kimper (to appear). The solution is to abandon parallel OT in favor of a serial, dervational version called Harmonic Serialism (Prince & Smolensky 1993/2004 ch. 2, McCarthy 2000, McCarthy 2008b, inter alia), and to assume that vowel epenthesis and metrification require separate steps, as has been argued for independent reasons (Elfner to appear).

Another property of positive constraints such as our gradient *Clash+ is that they prefer candidates that have secondary stress to candidates that do not, and this is actually the opposite of what people did in our experiment. This does not mean that the constraint is wrong: the dispreference against secondary stress is only pronounced in mobile and final stress stems, which have inserted as opposed to lexical stress. See section 5.3.

References

Ageenko, F. L., & M. V. Zarva. 1984. *Slovarj udarenij dlja rabotnikov radio i televidenja. [A dictionary of stress for TV and radio workers]*, 5th ed. Moscow: Russkij Jazyk.

Alber, Birgit. 2005. Clash, Lapse and Directionality. *Natural Language and Linguistic Theory* 23 (3): 485–542.

Albright, Adam, & Bruce Hayes. 2003. Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition* 90 (2): 119–161.

Alderete, John. 1999. Morphologically-governed accent in Optimality Theory. Ph.D. dissertation, University of Massachusetts, Amherst, Amherst, MA.

Alderete, John. 2001. Dominance effects as transderivational anti-faithfulness. *Phonology* 18: 201–253.

Altschuler, Daniel. 2006. Osage fills the gap: the quantity insensitive iamb and the typology of feet. Ms., Available as ROA-870 on the Rutgers Optimality Archive, `http://roa.rutgers.edu`.

Avanesov, R. I. 1964. *Udarenije v sovremennom russkom jazyke. [Modern Russian stress]*. New York: Macmillan.

Avanesov, R. I. 1968. *Russkoe literaturnoe proiznoshenie [Literary pronunciation of Russian]*, 6th (1984) ed. Moscow: Izdatelstvo Prosvescenie.

Baayen, R. Harald. 2008. *Analyzing linguistic data: A practical introduction to statistics*. Cambridge University Press.

Baayen, R. Harald. 2010. Demythologizing the word frequency effect: A discriminative learning perspective. *The Mental Lexicon* 5 (3): 436–461.

Baayen, R. Harald, Doug J. Davidson, & Douglas M. Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language* 59 (4): 390–412.

Barnes, Jonathan. 2003. Positional Neutralization: A Phonologization Approach to Typological Patterns. Ph.D. dissertation, University of California, Berkeley.

Barnes, Jonathan. 2004. Vowel reduction in Russian: The categorical and the gradient. Boston, MA. Handout from LSA 78. `http://www.bu.edu/linguistics/UG/barnes/LSA%202004%20-%20Russian%20VR.pdf`.

Barr, Dale, Roger Levy, Christoph Scheepers, & Hal Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68 (3): 255–278.

Bates, Douglas, & Martin Maechler. 2009. *lme4: Linear mixed-effects models using S4 classes*. R package version 0.999375-32. `http://CRAN.R-project.org/package=lme4`.

Becker, Michael. 2009. Phonological trends in the lexicon: The role of constraints. Ph.D. dissertation, University of Massachusetts, Amherst, MA.

Becker, Michael, & Maria Gouskova. 2012. Source-oriented generalizations as grammar inference in Russian vowel deletion. Ms. Indiana University and NYU. Available on LingBuzz at `http://ling.auf.net/lingbuzz/001622`.

Becker, Michael, Nihan Ketrez, & Andrew Nevins. 2011. The surfeit of the stimulus: Analytic biases filter lexical statistics in Turkish devoicing neutralization. *Language* 87 (1): 84–125.

Becker, Michael, Andrew Nevins, & Jonathan Levine. 2012. Asymmetries in generalizing alternations to and from initial syllables. *Language* 88 (2): 231–268.

Beckman, Jill. 1998. *Positional faithfulness*. New York: Routledge.

Belsley, David. 1984. Demeaning conditioning diagnostics through centering. *The American Statistician* 38 (2): 73–77.

Belsley, David, Edwin Kuh, & Roy E. Welsch. 2004. *Regression diagnostics: Identifying influential data and sources of collinearity*. New York: Wiley.

Bethin, Christina. 1992. *Polish syllables: The role of prosody in phonology and morphology.* Columbus: Slavica Publishers.

Bethin, Christina. 2006. Stress and tone in East Slavic dialects. *Phonology* 23 (3): 125–156.

Blair, Irene V., Geoffrey R. Urland, & Jennifer E. Ma. 2002. Using Internet search engines to estimate word frequency. *Behavior Research Methods, Instruments, & Computers* 34 (2): 286–290.

Blevins, Juliette. 1995. The syllable in phonological theory. In *The handbook of phonological theory*, ed. John A. Goldsmith, 206–244. Cambridge, MA, and Oxford, UK: Blackwell.

Boersma, Paul, & David Weenink. 2009. Praat: Doing phonetics by computer, version 5.1.

Borunova, S. N., V. L. Vorontsova, & N. A. Yesjkova. 1988. *Orfoepicheskij slovar' russkogo jazyka: proiznoshenije, udarenije, grammaticheskije normy. [A pronunciation dictionary of Russian: pronunciation, stress, grammatical norms]*, 4th ed. Moscow: Russkij Jazyk.

Brown, Dunstan, Greville Corbett, Norman Fraser, Andrew Hippisley, & Alan Timberlake. 1996. Russian noun stress and network morphology. *Linguistics* 34 (1): 53–107.

Chew, Peter A. 2000. A computational phonology of Russian. Ph.D. dissertation, University of Oxford.

Clements, George N. 1990. The role of the sonority cycle in core syllabification. In *Papers in Laboratory Phonology 1: Between the grammar and physics of speech*, eds. John Kingston & Mary Beckman, 283–333. New York: Cambridge University Press.

Coetzee, Andries, & Joe Pater. 2011. The place of variation in phonological theory, 2nd edition ed. In *The Handbook of Phonological Theory*, eds. John Goldsmith, Jason Riggle, & Alan Yu, 401–434. Oxford: Blackwell. Available as ROA-946 on the Rutgers Optimality Archive, http://roa.rutgers.edu.

Coetzee, Andries W., & Joe Pater. 2008. Weighted constraints and gradient restrictions on place co-occurrence in Muna and Arabic. *Natural Language and Linguistic Theory* 26 (2): 289–337.

Comrie, Bernard, Gerald Stone, & Maria Polinsky. 1996. *The Russian language in the twentieth century*, 2nd edition ed. Oxford: Clarendon Press.

Crosswhite, Katherine. 1999. Vowel reduction in Optimality Theory. Ph.D. dissertation, UCLA, Los Angeles, CA.

Crosswhite, Katherine, John Alderete, Tim Beasley, & Vita Markman. 2003. Morphological effects on default stress placement in novel Russian words. In *Proceedings of WCCFL 22*, eds. Gina Garding & Mimu Tsujimura, 151–164. Somerville, MA: Cascadilla Press.

Cubberly, Paul. 1987. Stress patterns in high-frequency Russian nouns and verbs. *Russian Language Journal* 41: 31–44.

Daland, Robert, Bruce Hayes, James White, Marc Garellek, Andrea Davis, & Ingrid Norrmann. 2011. Explaining sonority projection effects. *Phonology* 28 (2): 197–234.

Davis, Stuart. 1988. Syllable onsets as a factor in stress rules. *Phonology* 5: 1–19.

Davis, Stuart. 1999. The parallel distribution of aspirated stops and /h/ in American English. In *Indiana university working papers in linguistics*, eds. Karen Baertsch & Daniel Dinnsen, Vol. 1, 1–10. Bloomington, IN: IULC Publications.

Echambadi, Raj, & James Hess. 2007. Mean-centering does not alleviate collinearity problems in moderated multiple regression models. *Marketing Science* 26 (3): 438–445.

Eisner, Jason. 1997. Efficient generation in Primitive Optimality Theory, Philadelphia. Ms. Hopkins University. Available as ROA-206 on the Rutgers Optimality Archive..

Elenbaas, Nine, & René Kager. 1999. Ternary rhythm and the Lapse constraint. *Phonology* 16: 273–330.

Elfner, Emily. to appear. Stress-epenthesis interactions in Harmonic Serialism. In *Harmonic Grammar and Harmonic Serialism*, eds. John J. McCarthy & Joe Pater. London: Equinox Publishing.

Fry, Dennis. 1955. Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America* 35: 765–69.

Gelman, Andrew, & Jennifer Hill. 2007. *Data analysis using regression and multilevel/hierarchical models*. New York, NY: Cambridge University Press.

Gordon, Matthew. 2005. A perceptually-driven account of onset-sensitive stress. *Natural Language and Linguistic Theory* 23 (3): 595–653.

Gordon, Matthew. 2011. Stress: Phonotactic and phonetic evidence. In *Blackwell companion to phonology*, eds. Marc van Oostendorp, Colin Ewen, Elizabeth Hume, & Keren Rice, Vol. II, 924–948. Chichester, UK & Malden, MA: Wiley-Blackwell.

Gouskova, Maria. 2010. The phonology of boundaries and secondary stress in Russian compounds. *The Linguistic Review* 17 (4): 387–448.

Gouskova, Maria. 2012. Unexceptional segments. *Natural Language and Linguistic Theory* 30 (1): 79–133.

Gouskova, Maria, & Michael Becker. to appear. Nonce words show that Russian yer alternations are governed by the grammar. *Natural Language and Linguistic Theory*. Available on Lingbuzz at `http://ling.auf.net/lingBuzz/001456`.

Gouskova, Maria, & Kevin Roon. 2009. Interface constraints and frequency in Russian compound stress. In *Proceedings of FASL 17*, eds. Jodi Reich, Maria Babyonyshev, & Darya Kavitskaya, 49–63. Ann Arbor, MI: Michigan Slavic Publications.

Halle, Morris. 1973. The accentuation of Russian words. *Language* 49 (2): 312–348.

Halle, Morris, & Jean-Roger Vergnaud. 1987a. *An essay on stress*. Cambridge, MA: MIT Press.

Halle, Morris, & Jean-Roger Vergnaud. 1987b. Stress and the cycle. *Linguistic Inquiry* 18 (1): 45–84.

Hamann, Silke. 2004. Retroflex fricatives in Slavic languages. *Journal of the International Phonetic Association* 34 (01): 53–67.

Hayes, Bruce. 1980. A Metrical Theory of Stress Rules. Ph.D. dissertation, MIT, Cambridge, MA. Published by Garland Press, New York, 1985.

Hayes, Bruce, & Colin Wilson. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39 (3): 379–440.

Hyman, Larry. 1977. On the nature of linguistic stress. In *Studies in stress and accent*, ed. L. Hyman. Los Angeles: University of Southern California.

Idsardi, William. 1992. The computation of prosody. Ph.D. dissertation, MIT, Cambridge, MA.

Inkelas, Sharon, C. Orhan Orgun, & Cheryl Zoll. 1997. The implications of lexical exceptions for the nature of grammar. In *Derivations and constraints in phonology*, ed. Iggy Roca, 393–418. Oxford and New York: Oxford University Press.

Ito, Junko. 1986. Syllable Theory in Prosodic Phonology. Ph.D. dissertation, University of Massachusetts, Amherst. Published 1988. Outstanding Dissertations in Linguistics series. New York: Garland.

Ito, Junko, & Armin Mester. 1995. The core-periphery structure of the lexicon and constraints on reranking. In *Papers in Optimality Theory II (University of Massachusetts occasional papers in linguistics)*, eds. Jill Beckman, Laura Walsh Dickey, & Suzanne Urbanczyk, 181–210. Amherst, MA: GLSA Publications.

Kager, René. 1994. Ternary rhythm in alignment theory, University of Utrecht. Ms. Available on the Rutgers Optimality Archive as ROA-35.

Kawahara, Shigeto, & Sophia Kao. 2012. The productivity of a root-initial accenting suffix, [-zu]: judgment studies. *Natural Language and Linguistic Theory* 30 (3): 837–857.

Kawahara, Shigeto, & Matthew Wolf. 2010. On the existence of root-initial-accenting suffixes: An elicitation study of Japanese [-zu]. *Linguistics* 48 (4): 837–864.

Kimper, Wendell. to appear. Positive constraints and finite goodness in Harmonic Serialism. In *Harmonic Grammar and Harmonic Serialism*, eds. John J. McCarthy & Joe Pater. London: Equinox Publishing.

Kuznetsova, E. V. 2006. Acoustic parameters of secondary stress in Russian. In *Proceedings of the XVIII meeting of the Russian Acoustical Society*, 547–550. http://www.akin.ru/Docs/Rao/Ses18/R18.PDF. (accessed on Jan. 4, 2007).: GEOS.

Lagerberg, Robert. 2007. Variation and frequency in Russian stress. *Australian Slavonic and East European Studies* 21 (1–2): 165–176.

Liberman, Mark, & Alan Prince. 1977. On stress and linguistic rhythm. *Linguistic Inquiry* 8 (2): 249–336.

Lightner, Theodore. 1965. Segmental phonology of Modern Standard Russian. Ph.D. dissertation, MIT, Cambridge, MA.

McCarthy, John J. 2000. Harmonic serialism and parallelism. In *Proceedings of the North East Linguistics Society 30*, ed. Masako Hirotani, 501–524. Amherst, MA: GLSA Publications.

McCarthy, John J. 2003. OT constraints are categorical. *Phonology* 20 (1): 75–138.

McCarthy, John J. 2007. *Hidden generalizations: Phonological opacity in Optimality Theory*. London: Equinox.

McCarthy, John J. 2008a. *Doing Optimality Theory*. Cambridge: Blackwell.

McCarthy, John J. 2008b. The serial interaction of stress and syncope. *Natural Language and Linguistic Theory* 26 (3): 499–546.

Melvold, Janis. 1989. Structure and stress in the phonology of Russian. Ph.D. dissertation, MIT, Cambridge, MA.

Molinsky, Steven J. 1973. *Patterns of ellipsis in Russian compound noun formation*. The Hague: Mouton.

Morrill, Tuuli. 2012. Acoustic correlates of stress in English adjective-noun compounds. *Language and Speech* 55 (2): 167–201.

Nespor, Marina, & Irene Vogel. 1986. *Prosodic phonology*. Dordrecht: Foris.

Nespor, Marina, & Irene Vogel. 1989. On clashes and lapses. *Phonology* 6 (1): 69–116.

Newlin-Łukowicz, Luiza. 2012. Polish stress: looking for phonetic evidence of a bidirectional system. *Phonology* 29 (2): 271–329.

Nikolaeva, T. M. 1971. Mesto udareniia i foneticheskii sostav slova (rasstanovka udareniia v neizvestnykh slovakh inostrannogo proiskhozhdeniia). [Place of stress and the phonetic composition of the word (stress placement in unknown words of foreign origin)]. In *Fonetika. Fonologiia. Grammatika. K semidesiatiletiiu A. A. Reformatskogo. [Phonetics. Phonology. Grammar. For the 70th Birthday of A.A. Reformatskii]*, ed. F. P. et al. Filin, 59–69. Moscow: Nauka.

Padgett, Jaye. 2003. Contrast and postvelar fronting in Russian. *Natural Language and Linguistic Theory* 21 (1): 39–87.

Padgett, Jaye, & Marija Tabain. 2005. Adaptive dispersion theory and phonological vowel reduction in Russian. *Phonetica* 62 (1): 14–54.

Pater, Joe. 2000. Nonuniformity in English secondary stress: The role of ranked and lexically specific constraints. *Phonology* 17 (2): 237–274.

Pater, Joe. 2006. The locus of exceptionality: Morpheme-specific phonology as constraint indexation. In *Papers in Optimality Theory III*, eds. Leah Bateman, Michael O'Keefe, Ehren Reilly, & Adam Werle, 259–296. Amherst, MA: GLSA.

Plag, Ingo, Gero Kunter, & Mareile Schramm. 2011. Acoustic correlates of primary and secondary stress in North American English. *Journal of Phonetics* 39: 362–374.

Prince, Alan. 1983. Relating to the grid. *Linguistic Inquiry* 14 (1): 19–100.

Prince, Alan. 2007. The pursuit of theory. In *Cambridge handbook of phonology*, ed. Paul de Lacy, 33–60. Cambridge: Cambridge University Press.

Prince, Alan, & Paul Smolensky. 1993/2004. *Optimality Theory: Constraint interaction in generative grammar*. Malden, MA, and Oxford, UK: Blackwell. Available as ROA-537 on the Rutgers Optimality Archive, http://roa.rutgers.edu.

R Development Core Team. 2012. R: A language and environment for statistical computing. Vienna, Austria. `http://www.R-project.org`.

Revithiadou, Anthi. 1999. *Headmost accent wins: Head dominance and ideal prosodic form in lexical accent systems*. The Hague: Holland Academic Graphics.

Roon, Kevin. 2006. Stress in Russian compound nouns: head dominance or anti-faithfulness? In *Proceedings of FASL 14*, eds. James E. Lavine, Steven Franks, Mila Tasseva-Kurktchieva, & Hana Filip, 319–330. Michigan Slavic Publications.

Ryan, Kevin. 2013. Onsets contribute to syllable weight: Statistical evidence from stress and meter. Ms., Harvard University. Accessed on 10 January, 2013 at `http://www.people.fas.harvard.edu/~kevinryan/ryan_onsetweight_june2012draft.pdf`..

Selkirk, Elisabeth. 1984. *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, MA: MIT Press.

Sharoff, Serge. 2005. Methods and tools for development of the Russian Reference Corpus. In *Corpus linguistics around the world*, eds. Andrew Wilson, Dawn Archer, & Paul Rayson. *Language and computers*, 167–180. Amsterdam and New York: Rodopi. `http://www.comp.leeds.ac.uk/ssharoff/frqlist/frqlist-en.html`.

Sleptsov, P. A. 1975. *Russkie leksicheskie zaimstvovaniia v iakutskom iazyke: poslerevoliutsionnyi period. [Russian loanwords in Yakut: The postrevolutionary period]*. Moscow: Nauka.

Smith, Jennifer. 2002. Phonological augmentation in prominent positions. Ph.D. dissertation, University of Massachusetts, Amherst, Amherst, MA.

Steriade, Donca. 1997. Lexical conservatism and its analysis. Ms., UCLA.

Szpyra, Jolanta. 1992. Ghost segments in nonlinear phonology: Polish yers. *Language* 68 (2): 277–312.

Townsend, Charles. 1975. *Russian word-formation*, Corrected reprint (1980) ed. Columbus: Slavica Publishers.

Trouvain, Jürgen, William J. Barry, Claus Nielsen, & Ove Andersen. 1998. Implications of energy declination for speech synthesis. In *Proceedings of the 3rd ESCA/COCOSDA workshop on speech synthesis*, 47–52. Australia: Jenolan Caves.

Wade, Terence. 1992. *A comprehensive Russian grammar*. Cambridge, MA: Blackwell.

Weskott, Thomas, & Gisbert Fanselow. 2011. On the informativity of different measures of grammaticality. *Language* 87 (2): 249–273.

Yearley, Jennifer. 1995. Jer vowels in Russian. In *Papers in Optimality Theory II (University of Massachusetts occasional papers in linguistics)*, eds. Jill Beckman, Laura Walsh Dickey, & Suzanne Urbanczyk, 533–571. Amherst, MA: GLSA Publications.

Yoo, Seung-Nam. 1992. Subsidiary stress in Russian compound words. Ph.D. dissertation, University of Illinois at Urbana-Champaign.

Zaliznjak, Andrej Anatoljevich. 1977. *Grammatičeskij slovar' russkogo jazyka. [A grammatical dictionary of the Russian language]*. Moscow: Russkij Jazyk.

Zaliznjak, Andrej Anatoljevich. 1985. *Ot praslavjanskoj akcentuacii k russkoj. [From Proto-Slavic to Russian accentuation.]*. Moscow: Nauka.

Zasorina, L. N. 1977. *Chastotnyj slovar' russkogo jazyka. [A frequency dictionary of Russian]*. Moscow: Russkij Jazyk.

Zoll, Cheryl. 1998. Positional asymmetries and licensing. Ms., MIT. Available as ROA-282 on the Rutgers Optimality Archive, `http://roa.rutgers.edu`.

Zygis, Marzena, & Jaye Padgett. 2010. A perceptual study of Polish sibilants, and its implications for historical sound change. *Journal of Phonetics* 38 (2): 207–226.