Tone, stress, quantity, and quality: prosodic patterns and tonal wug-tests in Žiri Slovenian

Gašper Beguš[1] and Peter Jurgec[2]

[1]Department of Linguistics, University of California, Berkeley
[2]Department of Linguistics, University of Toronto

Abstract

The prosodic systems of the world's languages vary widely in their complexity. In this paper, we report on a prosodic pattern in Žiri Slovenian, which displays an intricate set of interactions between tone, stress, quantity, and vowel quality. Some of the prosodic patterns are well-motivated (e.g. the preference of High tone syllables to be stressed) while others lack phonetic or phonological motivation (e.g. the preference of long vowels to fall on the second syllable of a trochaic foot). Žiri also displays a rare case of interaction of prosody with vowel quality. To confirm productivity of the observed patterns, we conducted a wug experiment that tested the dependence of stress on vowel quality, word length, and tone. All in all, this paper brings forth new instances of phonetically unmotivated phonological processes at the suprasegmental level, which appear less discussed than at the segmental level. We also discuss methodological issues arising from artificial experiments on tonal processes.

Keywords: pitch accent, Slovenian, Autosegmental Phonology, floating features, laboratory phonology

1 Introduction

Many prosodic systems incorporate both stress and tone. Ma'ya, for instance, has a tonal contrast on the final syllable as well as lexically contrastive stress (Remijsen 2002). What is more common, however, is that stress and tone interact with one another. In Ayutla Mixtec, the position of stress is predictable from the distribution of tone, generally targeting the highest tone (de Lacy 2002). In Choguita Rarámuri, stress is contrastive, and the stressed syllables contrast three tonal patterns; there are no tonal contrasts in unstressed syllables (Caballero & Carroll 2015). A more complex situation is found in Serbo-Croatian, where stress is assigned to the syllable preceding a High tone, which targets ends of morphological domains, or else the word-initial syllable (Zec 1999).

In this paper, we report novel data from Žiri Slovenian, which also displays interactions between stress and tone. Stress falls on the leftmost High tone (e.g. *pɒˈmɒ́ɾaánʧɒ* 'orange'). In words without a lexical High tone, stress is penultimate, but nevertheless surfaces with a High tone (e.g. *plɛnitsɒ* → *plɛˈnítsɒ*). What sets Žiri apart from other patterns reported in the literature, are the further interactions of tone and stress with vowel length and quality. Long vowels must have a High tone, and stress falls either on the long vowel (e.g. *kɒˈwáaʧ* 'blacksmith') or on the preceding short vowel (e.g. *ˈbɒ́ɫáan* 'sick'). These two stress patterns are lexical and dependent on tone, with one exception: a long [ɛ] is always stressed. Other long vowels impose no restriction on stress.

The complexity of the Žiri prosodic system, which requires reference to three suprasegmental properties, has been noted in the Slavic literature (Stanonik 1977; Greenberg 1987; Beguš 2011). In the context of the South Slavic pitch accent, Žiri is considered unique in the number of prosodic contrasts. In disyllables, for instance, Serbo-Croatian contrasts a falling and rising pitch accent, but stress is always initial. Standard Slovenian distinguishes falling and rising pitch accent, but stress can fall on either syllable, resulting in four different combinations. In Žiri, there are seven patterns in total, five of which have initial stress. The larger number is because Žiri allows two non-adjacent pitch peaks, while the other South Slavic varieties only allow one. This also distinguishes Žiri from other Indo-European pitch accent systems, such as in West and North Germanic, where disyllabic words typically have two distinctive pitch accents.

The prosodic patterns in Žiri are also dependent on vowel quality. Interactions between segmental and suprasegmental properties are cross-linguistically attested, albeit in very different terms than what we find in Žiri. For instance, obstruent voicing can affect tone, which is raised by voiceless obstruents and lowered by voiced obstruents. Vowel sonority can also affect tone (see Becker & Jurgec 2017 for an extensive overview of these patterns). For instance, in Uspanteko tone realization in disyllables depends on the sonority difference between the two syllables (Bennett & Henderson 2013). de Lacy (2006) and Blumenfeld (2006) make an explicit claim that sonority, but not segmental features, can interact with stress assignment and tone, and Rasin (2017), Shih (2016, 2018), Shih & de Lacy (2020) argue that no segmental property can affect stress placement.

Žiri offers evidence that is not related to sonority: it is the markedness of [ɛ] that affects the position of stress and tone. This not only complements our understanding of possible interactions between segmental and suprasegmental properties, but also adds to the literature on the phonetically unmotivated and unnatural processes. These have been discussed substantially in the literature (Blevins, 2004; Kiparsky, 2006, 2008; Coetzee & Pretorius, 2010; Beguš, 2018, 2019, 2020), but most of the discussion focuses on segmental patterns. Substantially fewer work exist on phonetically unmotivated suprasegmental phonology. In this paper, we present a phonetically unmotivated process that targets the distribution of tone and stress.

Our primary description of the Žiri pattern is based on the available literature (particularly Stanonik 1977), complemented by elicitation of the forms from native speakers and subsequent acoustic analysis. We complement the fieldwork data experimentally with a perception tonal wug-test. Nonce word experiments involving tone are particularly rare. The present study contributes to the body of literature on tonal wug-tests in various Chinese languages (Zhang et al. 2006; Zhang & Lai 2010; Zhang & Liu 2016). To our knowledge, the experiment conducted in this study is the first tonal wug-test that looks at pitch accent in particular.

The paper consists of two parts. In Section 2, we describe the prosodic patterns in monosyllables and disyllables, which are based on extensive fieldwork. We back the formal analysis with representative pitch tracks and discuss the alternations. In Section 3, we test the asymmetries and gaps we find in the elicitation data in a perception wug-test. We find that the participants distinguish alternations based on word length, tonal, and segmental patterns.

2 Prosodic patterns in Žiri

The variety of Slovenian described in this paper is that of Žiri [ʒiˈri], a town in Western Slovenia, population 4,920 (Statistical Office of the Republic of Slovenia 2021). The number of speakers of the dialect can be approximated from the population size, but the precise number of native speakers is unknown.

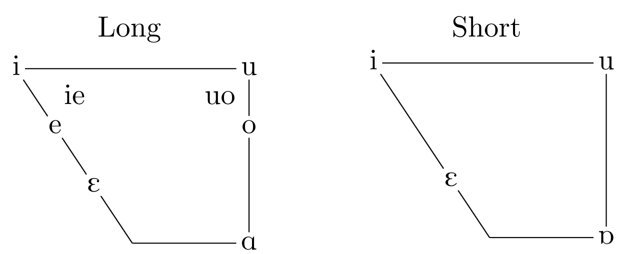
Žiri can be characterized as a pitch-accented variety of Slovenian. This means that tone interacts with other prosodic properties (Zec 1999; Hyman 2006; van der Hulst 2011). In the case of Žiri, tone is dependent on vowel length, vowel quality, and word size. Tone furthermore determines the position of stress. Žiri diplays a unique combination of prosodic factors that are not found to the same extent in other other varieties of Slovenian or the closely related Serbo-Croatian.

This paper is based on material and analysis in Stanonik (1977), Beguš (2011), and original fieldwork, conducted in the Žiri Valley in the summer of 2018. Participants consented to recording by signing informed consent forms approved by the Institutional Review Board (IRB) committee. Recordings were conducted in participants' homes using Sound Devices USB Pre 2 pre-amplifier (with 44.1 kHz sampling) and AKG C544L cardioid condenser head-mounted microphone. Participants were shown pictures of objects chosen for elicitation from an English picture dictionary that contained no Slovenian words. They were asked to name objects and use them in sentences in order to elicit responses with a preposition. The interviewer (one of the authors) instructed the participants in the Žiri dialect in order to avoid influences of Standard Slovenian. A total of 12 participants (mean age = 68.7, median = 84, range from 23 to 92; the age of one participant is undisclosed) were recorded in elicitation sessions that lasted for approximately 2 hours each. To maintain uniformity across tokens, the illustrations of tonal trajectories in this paper are based primarily on the recordings from two speakers.

2.1 Vowel inventory and vowel length

Žiri distinguishes long and short vowels (1). There are eight distinctive long vowels and four short vowels (Stanonik 1977). Notice also the rounding contrast in the low back vowels (long [ɑ] versus short [ɒ]) and the vowel height asymmetry between front and back vowels.

(1) Žiri vowel inventory



Long and short vowels can appear in any position of the word (initial, medial, final), but there can only be one long vowel per word. Long vowels can be stressed or not, and the same is true for short vowels (Stanonik 1977). Vowel length is thus not predictable, so we posit that the distinction between long and short vowels is underlying. This makes Žiri similar to Serbo-Croatian (Inkelas & Zec 1988; Zec 1999; Langston 1997) and some varieties of Slovenian (Lenček 1966; Bidwell 1969; Toporišič 1976/2000; Herrity 2000).

We will further see that long vowels impose additional restrictions on tone and stress: all long vowels carry a tone, and must be either stressed or immediately posttonic. Short vowels must bear a tone only when stressed.

2.2 Tone in monosyllables

Next we look at the tone patterns in monosyllables. Among the words with long vowels, there are two distinct patterns. The first group of monosyllables have long vowels with a falling tone pattern. We transcribe these words with a High tone on the first mora of the long vowel, as in [ée]. The second pattern has a rising contour, which we transcribe with a High tone on the second mora of the long vowel. In the Slavic literature, the falling pattern is termed the circumflex, while the rising is termed the acute (Toporišič 1976/2000; Greenberg 2000; Zec 1999). Finally, short vowels are pronounced with a High tone. The words illustrating the tonal contrasts on all vowels of the inventory are presented in (2).

(2) Distribution of tones in monosyllables (Stanonik 1977)

Long Falling		Long Rising		Short High	
ɣrάɑt	‘castle’	mɑάʃk	‘male’	krɒ́x	‘bread’
dάɑn	‘day’	pɑάlk	‘spider’	kɒ́ʃ	‘basket’
ʒíir	‘beechnut’	ʒiír	‘Žiri’	zít	‘wall’
líet	‘ice’	spiét	‘bound’		
ɣréex	‘sin’	deéts	‘male’		
ʧέεl	‘forehead’	εέnː	‘one’	zέt	‘son-in-law’
drúujx	‘others’	uúsk	‘wax’	ɬúʧ	‘light’
púot	‘path’	ɣuópts	‘muzzle’		
nóoʧ	‘night’	boóxk	‘God’		

The three-way contrast becomes clear when we examine the pitch tracks. Figure 1 presents the pitch tracks and the corresponding spectrograms, waveforms and segmentation. The Long Falling pitch has a clear peak towards the beginning of the vowel, while the Long Rising pitch has a peak towards the end of the vowel. Finally, the Short vowel has a relatively flat high pitch. The vowel is also significantly shorter.

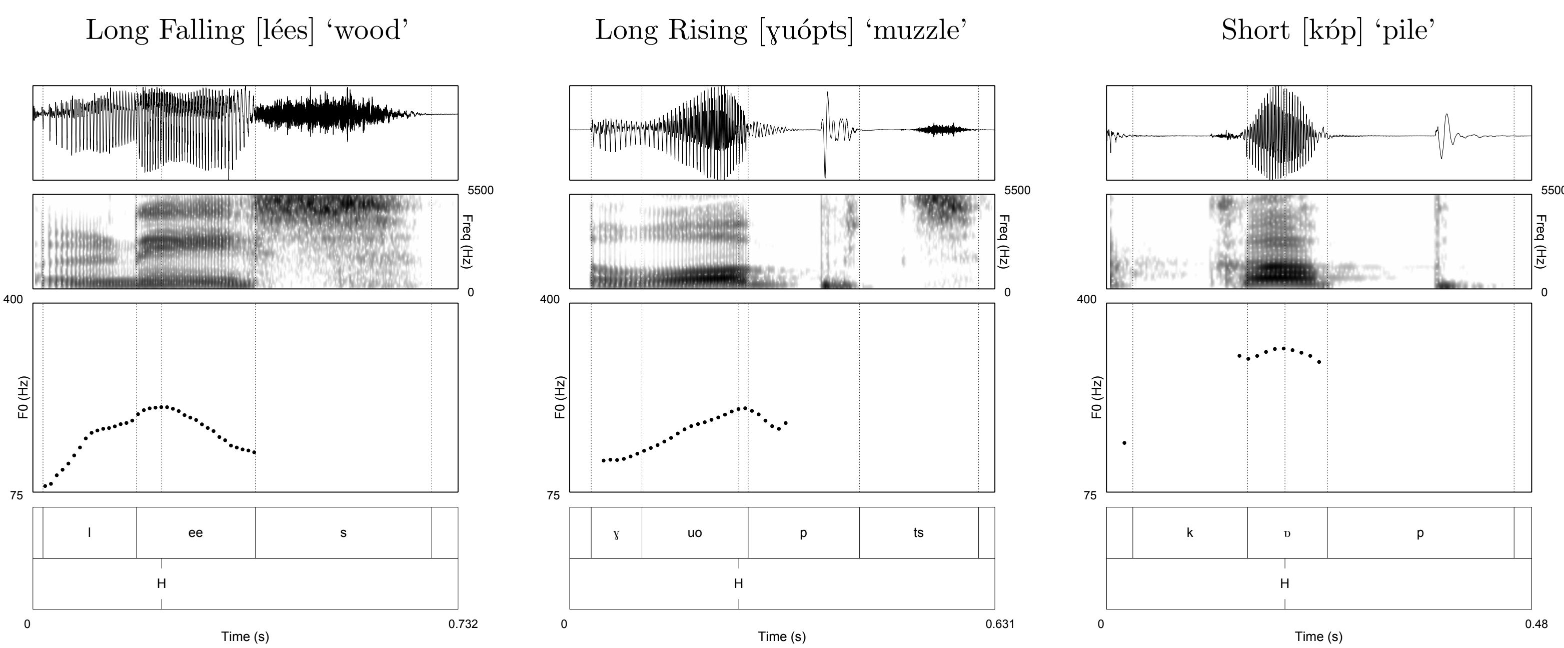


Figure 1: Phonetic realization of the monosyllabic prosodic contrasts.

To analyze these contrasts, we make use of autosegmental representations (Goldsmith 1976 et seq.). We propose that only High tone is phonologically active (3), as in Serbo-Croatian. For the falling contours, the High tone is associated with the first mora, whereas for the rising contours, the High tone is linked to the second mora. Short vowels are monomoraic and have a single High tone.

(3) Tonal contrasts in monosyllables

Long Falling	Long Rising	Short
H	H	H
lees	ɣuopts	kɒp

Monosyllables without a High tone are unattested. As we will see below, we will connect this gap with the interactions between stress and tone: stress falls on the leftmost syllable with a High tone, and in words without High tones, High tone is inserted on the penultimate syllable, which is then stressed.

2.3 Tone in disyllables

In disyllables and longer words, stress also plays a role. Stress is cued primarily as having the highest pitch and perhaps increased intensity, and can be readily identified by native speakers. As we well see, the position of stress is entirely predictable from vowel length and tone.

As with most other dialects of Slovenian, most Žiri words are rather short. In this paper, we look at monosyllables and disyllables. The patterns described can be extended to longer words, but distributions in longer words are more limited (see also Stanonik 1977).

The tonal contrast seen in monosyllables is also found in disyllables with initial long vowels. These disyllables have both attested patterns, as expected (4). Notice that in these words all end in a consonant, but the same pattern is observed in words ending in a vowel (e.g. *bréezɒ* 'birch tree')

(4) Distribution of tone in disyllables with initial long vowels (Stanonik 1977)

Falling		Rising	
ˈskáɑkɒt	'to jump'	ˈmɛɛ́dʋɛt	'bear'
ˈstáɑrɒst	'age'	ˈeézɛɾ	'lake'
ˈʒáɑłɒst	'sadness'	ˈjɛɛ́ʧmɛn	'barley'
ˈɣáɑjtrɒʃ	'rose'	ˈsɾɛɛ́bɒɾ	'silver'

A more complex situation is found in disyllables with the final long vowel. Here we find four different attested patterns (5). The first two groups of words have final stress (a), and show the usual distribution of tone seen so far. The second group of words, in contrast, have initial stress (b). This situation is somewhat unusual as it involves a short stressed followed by a long posttonic vowel. The stressed vowel bears a High tone.

(5) Distribution of tone in disyllables with final long vowels (Stanonik 1977)

a. Final stress

Falling		Rising	
ʒɛˈlíes	'iron'	zɒˈklɛɛ́nt	'to lock'
kɒˈléen	'knee'	kɒˈzuóʦ	'haystack'
pɒˈléen	'log'	uˈɣɑɑ́nt	'to guess'
kɒˈwáɑʧ	'blacksmith'	ɒˈtɾɑɑ́ʦ	'children'

b. Initial stress

Falling		Rising	
ˈkɒ́kúoʃ	'chicken'	ˈkɒ́ɾwɑɑ́w	'bloody'
ˈʋɛ́ʧíeɾ	'evening'	ˈpɒ́dɫoóʃk	'bolt washer'
ˈbɒ́ɫɑ́ɑn	'sick'	ˈlɛ́sién	'wooden'
ˈɣɒ́ɫúop	'dove'	ˈɒ́tɾuóp	'bran'

These data suggest that words can have multiple underlying tones in Žiri. We represent the contrast between the two types of words with long vowels by positing a different number of underlying High tones (6). In the words with final stress, there is only one underlying tone on the vowel. In the words with initial stress, however, there are two underlying tones, one on each syllable.

(6) Tone contrasts in disyllables with final long vowels

```
Falling       Rising

    H              H
    |              |
kɒleen        kɒzuots

 H   H        H    H
 |   |        |    |
kɒkuoʃ        ɒtruop
```

As we will see below, the leftmost High tone becomes stressed, resulting in initial stress in words with two High tones.

Finally, in disyllables with only short vowels, stress is always initial, and realized with a High tone.

(7) Distribution of tone in disyllables with short vowels (Stanonik 1977)

ˈɒ́ku	'eye'
ˈkúʃɛɾ	'lizard'
ˈstɒ́bɒɾ	'pillar'
ˈlípɒ	'linden tree'

In summary, stress in disyllables can be initial or final, which depends on the distribution of tone. Stress is initial if High tone is present underlyingly on the first vowel. Else, stress is final if that vowel is long or initial in words without long vowels. To support these transcriptions, we include the annotated pitch tracks for disyllables in Appendix A.

In longer words, stress follows a similar set of facts. In words with a long vowel, stress falls either on the long vowel (e.g. *kɫɒˈbaásɒ* 'sausage', *ɣɫɒˈbaákɒ* 'deep-FEM') or on the preceding short vowel (e.g. *pɒˈmɒ́ɾaánʧɒ* 'orange', *tɛlɛˈvíziíjɒ* 'television'). In words with only short vowels, stress is penultimate in all words we could find (*plɛˈnítsɒ* 'diaper').

As we have seen in Section 2.2, High tone needs to be specified on a specific mora to distinguish rising and falling tones. In the data discussed in this section, we see that the first syllable also differs as to whether it has a High tone or not. Žiri thus turns out to be a language in which the position of stress is generally predictable from tone, which we demonstrate next. de Lacy (2002) uses a set of constraints on the combinations of prosodic heads and tone. Here we adapt the approach to privative features and use the constraint in (8).

(8) Hd/H
Stressed syllables must have a High tone.

The constraint on the combination of stress and tone Hd/H is never violated in Žiri: all stressed syllables have a High tone. Feet in Žiri is trochaic (Trochee $\gg$ Iamb). Words with long vowels are always footed, whether they are in the first or the second syllable. To capture this generalization, we make use of the alignment constraint Align-R($\sigma_{\mu\mu}$, Foot), which is similar to the constraint used for prosodification in Karok (Sandy 2017). This constraint makes sure that heavy syllables, which contain long vowels, are footed if preceded by a High tone. In Žiri coda consonants are not moraic.

(9) Align-R($\sigma_{\mu\mu}$, Foot), after McCarthy & Prince (1993)
For every Heavy syllable, there must be a foot, such that the Right edge of the foot and then Right edge of the foot coincide.

This constraint creates an unusual situation in which heavy syllables are preferred to be at the right edge of trochees. Yet this constraint is instrumental at positioning stress on the syllable preceding the long vowel. This situation is cross-linguistically marked, as it violates Stress-to-Weight (and Weight-to-Stress, omitted), which requires that stressed vowels be long (and unstressed syllables be short).[1]

The ranking is shown for a tetrasyllabic word in (10). The winning candidate (a) foots the two syllables with High tones. While an iamb would satisfy Stress-to-Weight (b), it is ruled out by the dominant Trochee constraint. Disyllabic feet are preferred over monosyllabic (c), and stressed syllables must have a High tone (d). Finally, moving stress to the right edge (e) violates Align-R($\sigma_{\mu\mu}$,Foot). Note that the ranking between Parse-σ and Align-R($\sigma_{\mu\mu}$,Foot) is relevant in words with initial long vowels; the ranking assumed here prefers disyllabic trochees in those words.

[1]The alternative to Align-R($\sigma_{\mu\mu}$, Foot) would be to propose a constraint that would require High tones to be footed. This constraint would be somewhat unusual from a cross-linguistic perspective. Crucially, this alternative would incorrectly predict the existence of words in which long vowels fall outside the foot, which is never attested in Žiri.

(10) High tones are footed

H H / pɒmɒɾɑɑnʧɒ/	Hd/H	Trochee	Parse-σ	Align-R($\sigma_{\mu\mu}$,Foot)	AllFt-R	STW
H H a. ☞ pɒ(ˈmɒɾɑɑn)ʧɒ			**		*	*
H H b. pɒ(mɒˈɾɑɑn)ʧɒ		*!	**		*	
H H c. pɒmɒ(ˈɾɑɑn)ʧɒ			***!		*	
H d. pɒ(ˈmɒɾɑɑn)ʧɒ	*!		**		*	*
H H e. pɒmɒ(ˈɾɑɑnʧɒ)			**	*!		

This ranking also successfully accounts for disyllables or longer words with Falling tones, words with a single underlying High tone, and finally words with short vowels, where stress is always penultimate, regardless of what the input tones are (11). In this system, then, High tones can be deleted if they occur outside long vowels or pretonic short vowels. They, however, can only be inserted to satisfy the headedness requirement in words with short vowels, as seen below, but not in words with long vowels, since the presence of the pretonic High tone is contrastive (Dep(H) $\gg$ Parse-σ).

(11) Words without long vowels receive penultimate stress

/plɛnitsɒ/	Hd/H	Dep(H)	Parse-σ	AllFt-R
H a. ☞ plɛ(ˈn i tsɒ)		*	*	
H b. (ˈplɛni)tsɒ		*	*	*!
c. plɛ(ˈnitsɒ)	*!		*	
H d. plɛni(ˈtsɒ)		*	**!	

These data show that tone, stress, and quantity interact in complex ways. All stressed vowels must have a High tone, which means that underlying High tone determines the position of stress. In words without an underlying High tone, is has to be inserted. The foot is generally a disyllabic trochee, with long vowels, perhaps counterintuitively, preferring the right edge of the foot. Finally, stress is penultimate in words with only short vowels.

2.4 Stress retraction

As we have seen so far, stress in Žiri is inherently linked to tone and vowel length. In words with a long vowel, stress can fall either on the long vowel or on the preceding short vowel. We now move to alternations. The alternations involve words with initial stress, regardless of word length. We have seen the realization of tone in these words in isolation and carrier sentences. Slavic languages, however, are well-known for their ability to retract stress to the preceding clitic, which is what we examine next.

Cross-linguistically, clitics can have a variety of shapes (see Anderson 2005, 2011 for a cross-linguistic overview and Zwicky & Pullum 1983 for English). In Žiri, however, clitics are monosyllabic, stressless, toneless, and contain only short vowels (as in Standard Slovenian, see Toporišič 1976/2000; Jurgec 2007). Put differently, proclitics form a Prosodic Word with the following content word (Nespor & Vogel 1986; Peperkamp 1997; Dixon & Aikhenvald 2002).

In Žiri, stress can retract from the word with initial stress to the preceding clitic. If the word is monosyllabic, the mechanism behind stress retraction is identical to the distribution of tone in disyllables. Recall that among disyllables with a final long vowel some have initial while others have final stress. We analyzed the distinction between the two types in terms of underlying High tone on the initial, short vowel. Monosyllables preceded by clitics show similar behaviour. With some monosyllables, clitics receive stress (and High tone), whereas with others, stress falls on the monosyllable (12). For example, the Long Falling monosyllable [ˈnɒ́ núos] ’on (the) nose’ shows stress retraction to the clitic, but [nɒ ˈpúot] ’on (the) path’ does not retract.

All clitics behave uniformly in that their prosody is fully predictable from the following word. For instance, the monosyllable [púot] ‘path’ can be preceded by any clitic and it shows the same pattern regardless: [nɒ ˈpúot] ‘on’, [zɒ ˈpúot] ‘for’, [ʒɛ ˈpúot] ‘already’, etc. The same is true for all monosyllables and all clitics. Retraction from short monosyllables is reported and tested for the first time in this paper.

(12) Prosodic patterns with proclitics

a. Retraction

Long Falling	nóoʧ	‘night’	ˈʒɛ́ nóoʧ	‘already night’
	wɾɑ́ɑt	‘neck’	ˈzɒ́ wɾɑ́ɑt	‘around (the) neck’
	núos	‘nose’	ˈnɒ́ núos	‘on (the) nose’
Long Rising	ˈɣuópts	‘muzzle’	ˈnɒ́ ɣuópts	‘on (the) muzzle’
	ˈoóʧk	‘little eye’	ˈú oóʧk	‘into (the) little eye’
Short High	kɾɒ́x	‘bread’	ˈpɒ́ kɾɒx	‘to (the) bread’
	kɒ́p	‘pile’	ˈnɒ́ kɒp	‘on/onto (the) pile’

b. No retraction

Long Falling	ɣɾéex	‘sin’	ʒɛ ˈɣɾéex	‘already sin’
	dʋéex	‘two’	dɒ ˈdʋéex	‘until two’
	púot	‘path’	nɒ ˈpúot	‘on (the) path’
Long Rising	wɾɑɑ́t	‘door-GEN.PL’	dɒ ˈwɾɑɑ́t	‘to (the) door’
Short High	xɒ́ɾpt	‘back’	zɒ ˈxɒ́ɾpt	‘behind (one’s) back’

To support these transcriptions, consider the pitch tracks in Figure 2. The pitch tracks of monosyllables, repeated from Figure 1 are shown on the left, while the prepositions and monosyllables are shown on the right. Notice also the increased pitch and amplitude of the clitic (when compared to the monosyllable), suggesting greater intensity, and hence stress. In all three cases, the pitch is highest on the clitic, and then gradually falls towards the end of the monosyllable.

Figure 3 illustrates the contrast between a monosyllable with stress retraction and one without

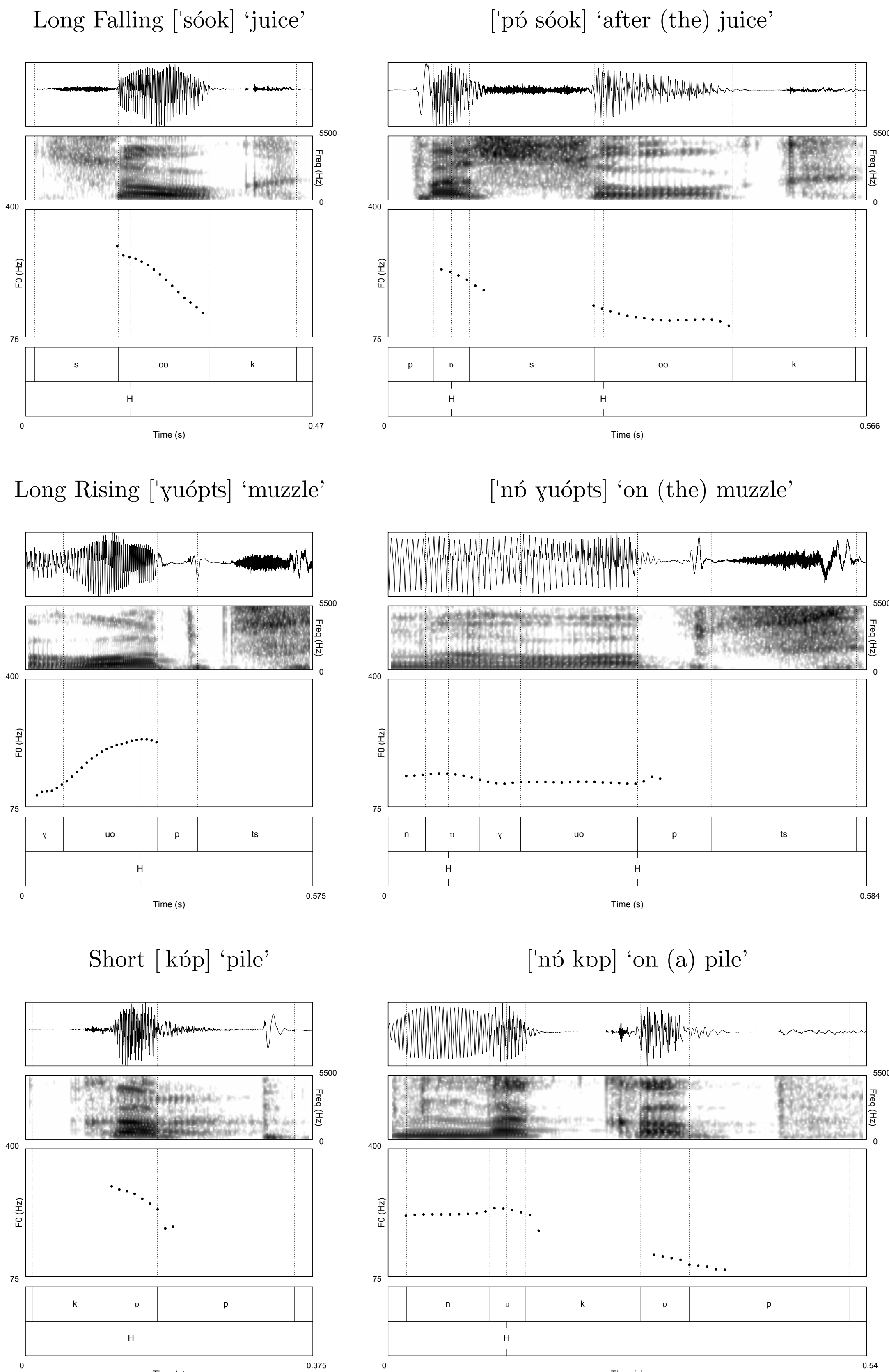


Figure 2: Stress retraction and the corresponding pitch contours.

it. The key distinction is in the relative pitch height: when stress retracts, so does the pitch: we analyze this as a High tone on the mora of the clitic as well as the first mora of the long vowel. In the word without retraction, the tonal peak is at the beginning of the long vowel instead.

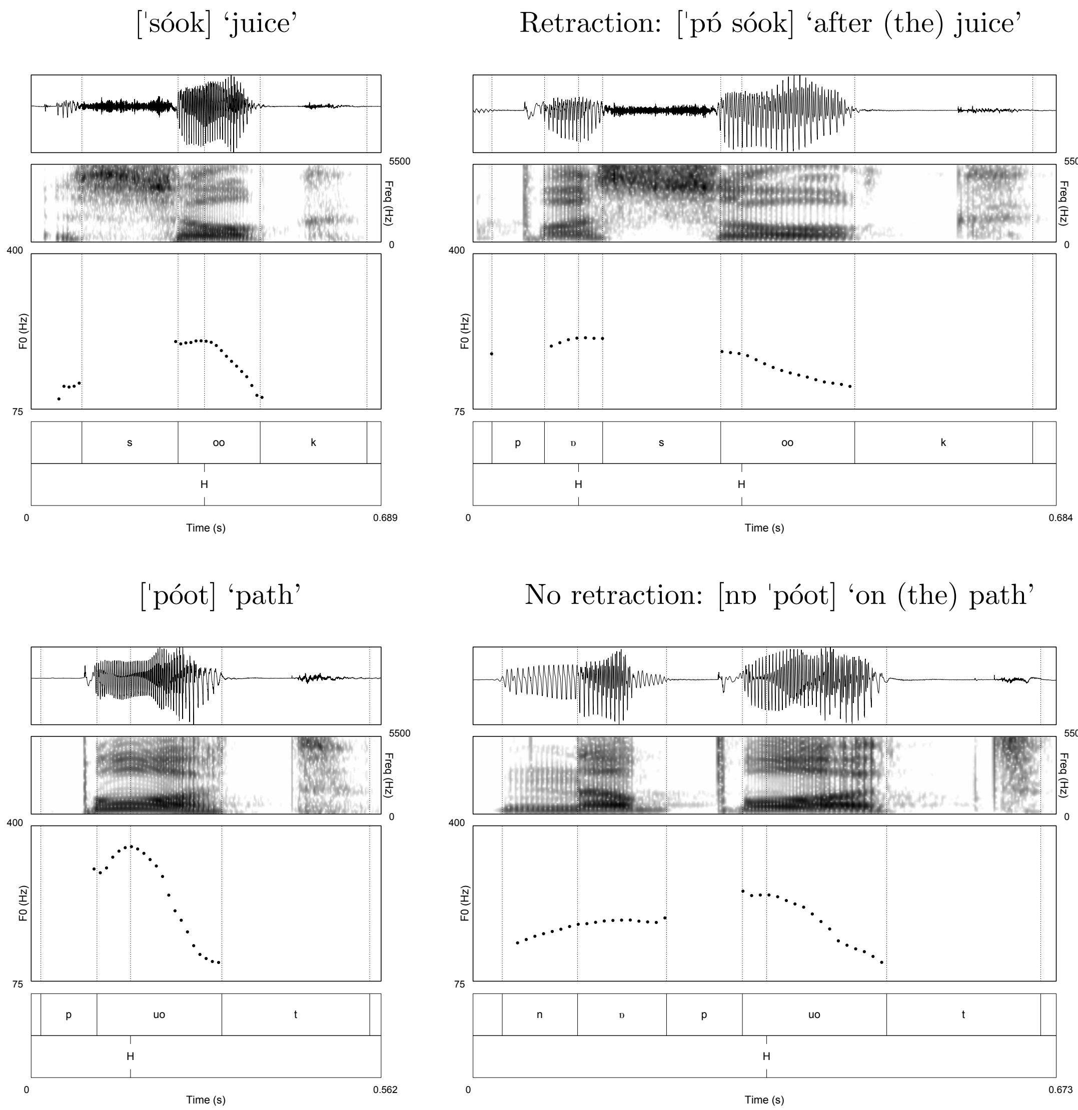


Figure 3: Stress retraction with Long Falling tone.

The distinction between retracting and non-retracting monosyllables is not predictable, and has to be specified lexically. We thus propose the distinction comes from the representation of these words. Recall that the disyllables could differ whether they have stress on the long vowel or the preceding vowel. The latter group had two underlying High tones. We suggest that the same mechanism is at play with clitics. The stems that show retraction have an additional High tone, which is not associated with a particular position in the monosyllable, but is instead floating (Odden 1988; Akinlabi 1996). The floating tone docks onto the clitic when possible, but it not realized otherwise. This allows us to distinguish the retracting and non-retracting monosyllables. The mappings are shown in (13).

(13) Retracting monosyllables have a floating High tone

	Input	Output
	HH	H H
Retraction	nɒ sook	(ˈnɒ sook)
	H	H
No retraction	nɒ puot	nɒ (ˈpuot)

The ranking required for monosyllables is directly parallel to the disyllables: the only difference is that the High tone is floating in monosyllables but linked in disyllables.

Stress can also retract from disyllables and longer words with an initial long vowel, albeit there are some differences (14). Notice there is a slight difference in the retracting words with initial Long Falling pitch. The disyllables attain a rising contour after retraction (e.g. *ˈnɒ ɣlaάwɒ* 'on (top of the) head-ACC' ∼ *ˈɣlάawɒ* 'head-ACC'), while the monosyllables do not (e.g. *ˈnɒ́ núos*) 'on (the) nose'. This can be seen as an affect of avoiding two adjacent High tones in non-final syllables. We will elaborate on this restriction on the distribution of tones in Section 3.3. Stress never retracts from disyllables with only short vowels, which is expected given the preference for penultimate stress in these words.

(14) Disyllables with proclitics

a. Retraction

Long Falling	ˈɣlάɑwɒ	'head-ACC'	ˈnɒ́ ɣlɑάwɒ	'on (top of the) head-ACC'
Long Rising	xɾuúʃkɒ	'pear'	ˈpɒ́ xɾuúʃkɒ	'after (the) pear'

b. No retraction

Long Falling	ˈbɾéezɒ	'birch tree'	pɒd ˈbɾéezɒ	'under (the) birch tree'
Long Rising	ˈkɑάsɒ	'scythe'	pɒ ˈkɑάsɒ	' after (the) scythe'
Short High	xíʃɛ	'house'	u ˈxíʃɛ	'in (the) house'

All in all, the patterns in monosyllables (and disyllables with an initial long vowel) with clitics are similar to disyllables with a final long vowel. In these cases we see five prosodic combinations: retracted and non-retracted, with either a falling or rising pitch on the long vowel. Words with only short vowels will have penultimate stress, except in monosyllables with clitics, which may or may not retract.

Stanonik (1977) does not elaborate on the distribution of retraction. In particular, monosyllables appear equally likely than disyllables to show retraction, and Falling pitch retracts at rates similar to Rising pitch. There is one exception: when it comes to vowel quality, Stanonik (1977) notes that retraction never occurs with [ɛɛ]. These distinctions will play a central role in our experiment.

3 Wug-test experiment

Section 2 established the main generalizations about the distribution of vowel length, tone, and stress in Žiri. We have seen that Žiri has a complex prosodic system that shows alternations, in particular retraction. In this section, we tackle the question how retraction is affected by word length, the tonal pattern on the long vowel, and vowel quality. These questions are difficult to answer given that our data so far is elicitation-based (although see Jurgec 2019 for a counterexample on another dialect of Slovenian). Moreover, the clear restrictions that have been observed, such as the non-retraction from [ɛɛ], may be accidental due to the gaps in the lexicon or the limited amount of data analyzed. We follow a growing body of literature that uses nonce-word experiments to corroborate speakers' grammatical knowledge (Berko 1958). For instance, Ernestus & Baayen (2003) show that Dutch speakers' productions of nonce words reflect distributional characteristics of the Dutch lexicon, with velars in root-final position eliciting relatively more voicing than labials and coronals. For Slovenian, Jurgec & Schertz (2020) show that velar palatalization at the root-suffix boundary is disfavoured when the steam contains another postalveolar due to a consonant co-occurrence restriction confirmed on corpus data (Jurgec 2016).

Even though the tonal generalizations are quite robust in the lexicon, research has shown that not all trends in the lexicon are productively extended to nonce words. In Becker et al. (2011), for instance, the lexicon displays different rates of laryngeal alternation in stops depending on vowel quality, but this was not mirrored in the nonce word task, although many other generalizations were. As for tone, Zhang et al. (2006) show that while several tone-sandhi alternations are completely regular in Taiwanese, they are not extended to nonce words. In Zhang & Lai (2010), participants extended a phonetically motivated Mandarin tone-sandhi to novel words; another type of sandhi, which lacks clear phonetic motivation, however, was applied at significantly lower rates. The productivity is a key issue for Žiri, where [ɛɛ] inhibits stress retraction. While there are some examples of interactions of prosody with vowel quality (see Becker & Jurgec 2017, 2020 for a review) or sonority (de Lacy 2006, 2007), the Žiri type interactions do not have clear phonetic grounding. In particular, while vowels' intrinsic pitch is dependent on vowel height, which could lead to tonogenesis, this analysis is not possible for Žiri. In order for this analysis to work, mid vowels would have to pattern with high or low vowels with respect to the distribution of tone. This is not what we find in Žiri: the mid vowel [ɛɛ] patterns differently than high and low vowels, as well as the two other mid vowels [ee] and [oo]. The same reasoning applies to sonority, and to the relationship between sonority and stress.

To assess the productivity of the interaction between vowel quality and prosody, we conduct a perceptual wug-test. Participants were asked to judge nonce-word paradigms. The results show that participants preferred less retraction with [ɛɛ] when compared to the other vowels. Moreover, the results further show that stress retraction is preferred in monosyllables more than compared to disyllables and in words with Rising tone more than with words with Falling tone. These experimental findings complement the elicitation-based study.

3.1 Methods

Materials. Our stimuli consisted of phonotactically valid words that were generated to vary on several parameters (summarized in Table 1). The monosyllables were of the shape CVC whereas the disyllables were ˈCVCɒ. The words could have either Long Falling, Long Rising, or Short pitch accent. However, not all combinations of tone and word length were included. For instance, there were no Long Rising monosyllables, as those are rare in the lexicon. The same is true for disyllables with only short vowels. In short, the gaps in the stimuli reflect the gaps in the lexicon, yet we will still be able to answer the key question about the interaction of stress, tone, quantity, and vowel quality. The final variable was vowel quality: among the Long Rising disyllables, some had the vowel [ɛɛ]. All other long stressed vowels were [ee, ɑɑ, oo], whereas the short vowels were [i, ɛ, ɒ]. About two thirds of the stimuli were minimal n-tuples, as the disyllabic pair in Table 1, while the remaining stimuli were not paired. We provide more detail when discussing individual comparisons in the section below. The stimuli were checked with the native speakers to make sure they are not real words.

Variables			Examples		
Words	Pitch accent	Vowel quality	Without clitic	Non-retracted	Retracted
Disyllables	Long Falling	other	ˈɾéelɒ	nɒ ˈɾéelɒ	ˈnɒ́ ɾeélɒ
	Long Rising	[ɛ] other	ˈɾɛɛ́lɒ ˈbɑɑ́mɒ	nɒ ˈɾɛɛ́lɒ nɒ ˈbɑɑ́mɒ	ˈnɒ́ ɾɛɛ́lɒ ˈnɒ́ bɑɑ́mɒ
Monosyllables	Long Falling	other	móos	nɒ ˈmóos	ˈnɒ́ móos
	Short	other	ˈxɒ́t	nɒ ˈxɒ́t	ˈnɒ́ xɒt

Table 1: Summary of parameters for nonce word generation.

In total, there were 82 words (see Appendix B for a full list). For each stem, we recorded a triplet: a bare form and two forms with a preposition, one with retracted stress and the other without. Stimuli were recorded in anechoic room with a Sound Devices USB Pre 2 pre-amplifier with 44.1 kHz sampling using AKG C544L cardioid condenser head-mounted microphone. We RMS-equalized stimuli in Praat (Boersma & Weenink, 1992) using a Praat script by Beckers (2002)) and converted .wav to .mp3 format with Audacity's LAME mp3 encoder for faster processing.

Eliciting stimuli with tonal trajectories as complex as those in the Žiri dialect poses unique challenges. To achieve the desired tonal trajectories of nonce words, the researcher produced tonal trajectories and then asked a female speaker (age 26) to repeat them. This methodology was necessary due to the speaker's difficulty in producing tonal trajectories in nonce words, and the task was facilitated by the researchers's metalinguistic awareness and native competence in the relevant dialect. Each item was recorded several times in two separate sessions. The final triplets (stimulus and two options) were chosen after a careful acoustic analysis in Praat (Boersma & Weenink 1992). Altogether 246 items were chosen (82 × 3 for each triplet).

Procedure. The experiment was conducted online using Experigen (Becker & Lavine 2012). IRB approval was obtained, and speakers were asked to complete an online consent form. Participants were recruited for the experiment with the help of a local radio station, personal connections, and a paid social media post that targeted inhabitants of the Žiri Valley. At the beginning of the experiment, the participants were asked to volunteer their demographic information, including age and place of residence. Participants were then told that they will hear three words. The first one will be a name of an imaginary object. The visual stimuli of unusual or non-objects were taken from Horst & Hout (2016) and Gonzalez-Gomez et al. (2013). The participants would then hear two versions of the same object in the clitic context. Each trial started with the picture on the left in Figure 4, under which the frame sentence "Sonya sees (button)" was written. When the participants clicked on the button, they heard the stimulus in the accusative and without the preposition. Then an additional picture appeared with an arrow above the object and the sentence "If you put it on something, would you say (button) or (button)?" When selected, the buttons played the stimulus with the preposition [nɒ] 'on'. One of the stimuli had retracted stress and the other had the stress on the stem. Once the stimuli were played, the sentence "What would you say? (You can play the words again.)" appeared as well as two buttons: "like above" or "like below". The participants were free to listen to the materials as many times as they wanted. Once they made a judgment, the next stimulus was played. All text was written using informal conventions (e.g. *vid'* instead of the standard *vidi* 'sees', *rekl* instead of the standard *rekli* 'said') to prime the participants for non-standard stimuli; see Jurgec (2019) for using similar methodology in a production experiment on another dialect of Slovenian. Stimuli were never presented orthographically.

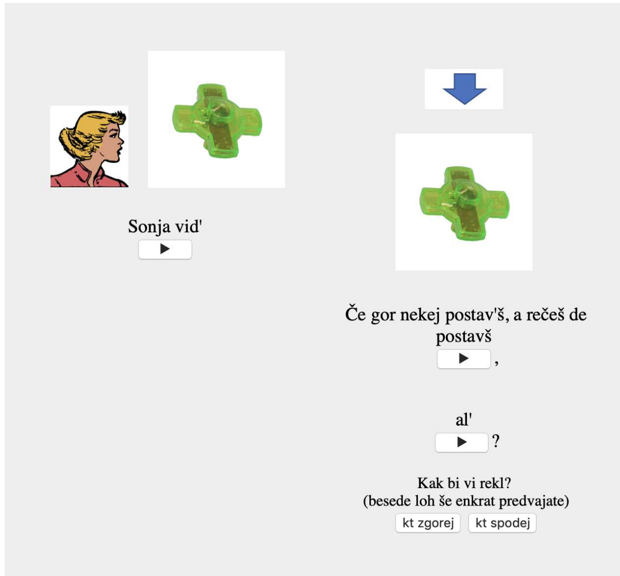


Figure 4: Web interface for the nonce word study using Experigen (Becker & Lavine 2012) and visual stimuli (non-objects) from Gonzalez-Gomez et al. (2013) and Horst & Hout (2016).

Participants. We include in the analysis only those participants who indicated on the demographic survey that they were native speakers of the Žiri dialect without an linguistic training. Among those participants, we include only participants who provided more than 10 judgments. Altogether 56 participants are included in the final analysis. Their self-reported mean age was 32 (the range was between 19 and 57 with a median of 29 years). 24 participants provided judgments to all 82 stimuli. The mean number of judgments for the remaining participants is 24 (ranging from 11 to 74, with a median of 18).

3.2 Results

The aim of the experiment was to explore whether word length, tone, and vowel quality have an effect on retraction. We find that the participants preferred retraction with monosyllables and with Rising pitch, but dispreferred retraction from [εε]. We review these results in turn.

Word length. To test if rates of retraction differ depending on word length, we analyze retraction rates from 11 pairs (22 items) of monosyllabic and disyllabic words with the Long Falling pattern. These minimal pairs differ only in the number of syllables, but are otherwise segmentally and suprasegmentally identical. For example, for the monosyllabic form ['xáɑt], participants chose between the non-retracted [nɒ 'xáɑt] and retracted ['nɒ́ xáɑt]. For the non-clitic form ['xáɑtɒ], the participants chose between the non-retracted [nɒ 'xáɑtɒ] and retracted ['nɒ́ xɑátɒ]. We also include 16 disyllabic words with the Long Falling tone without the monosyllabic pair. We include these additional items because the random slopes in our statistical analysis reveal that segmental material does not influence experimental outputs; we elaborate on these below.

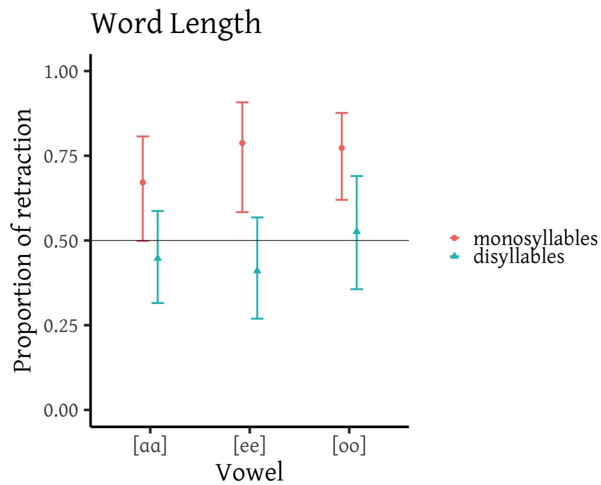


Figure 5: Estimates of a mixed-effects logistic regression model showing preference for retraction in disyllables with long vowels other than [ɛɛ] in the Falling tone condition. Error bars indicate 95% confidence intervals.

We fit the data into a mixed-effects logistic regression model with Response as the dependent variable (retraction coded as success; N = 1259) and two predictors: Word Length and Vowel Quality ([ɑɑ, ee, oo]; sum-coded with [oo] as the reference level). The full model includes an interaction between the two predictors as well as random intercepts for participant and frame with by-participant and by-frame random slopes for Word Length.

To control for effects of segmental material and neighborhood density, we include random intercepts and slopes not by item, but by *frame*. For example, [ˈxɑ́ɑt] and [ˈxɑ́ɑtɒ] are coded as one frame, because they involve the same segmental material, except that [ˈxɑ́ɑtɒ] includes an additional syllable—but the number of syllables is a fixed effect and as such already controlled for in the model. Items that are not part of the mono- and disyllabic pairs are coded as unique frames. We model random intercepts for frames rather than items, because this allows random slopes for the predictor of interest (see also Beguš 2017 for a similar approach). Had we instead chosen items and random slopes for the predictors of interest, the models do not converge as each item has only one level of the predictor of interest.

The interaction between Word Length and Vowel Quality is not significant. In particular, the Akaike Information Criterion (AIC), a measure of the quality of statistical models, is 1385.6 in the model without the interaction and 1387.8 with it. Word Length is a significant predictor (AIC is 1385.6 in the model with the predictor and 1399.5 in the model without the predictor). The estimate of the difference between disyllables and monosyllables (β) is 1.24, while the z-score is 5.04, which is highly statistically significant ($p < 0.0001$). Figure 5 illustrates the effects with 95% confidence intervals.

By-participant and by-frame random slopes reveal relatively little variability across participants and frames. Random slope for the difference between disyllables and monosyllables is positive for all frames and for all participants. This also suggests that segmental material and neighborhood density do not crucially affect the results of the experiment.

Tone. To test if retraction preference differs between Long Falling and Long Rising tones, we analyze the data from 46 disyllables. The first vowel in the disyllables was either [ee] or [ɑɑ]. The

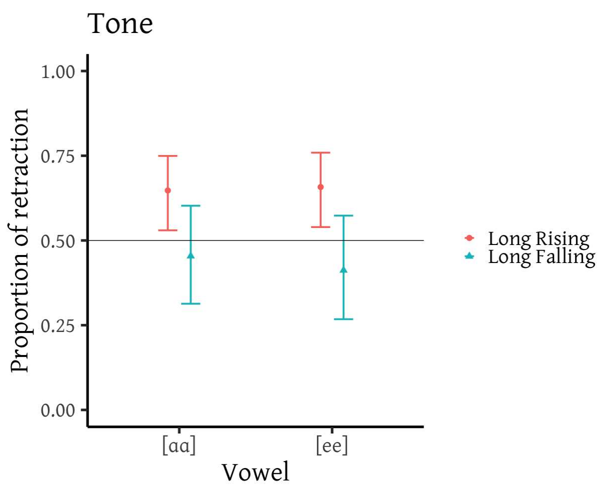


Figure 6: Estimates of a mixed-effects logistic regression model showing preference for retraction for Long Rising (compared to Long Falling) with long vowels other than [ɛɛ]. Error bars indicate 95% confidence intervals.

data included 17 minimal pairs (34 items in total) in which the segmental material was constant, the only difference between them was the tonal trajectory of the stimulus (Long Falling versus Long Rising). For example, for the non-clitic form [ˈbɑɑ́mɒ], speakers chose between the non-retracted [nɒ ˈbɑɑ́mɒ] and retracted [ˈnɒ́ bɑɑ́mɒ]. For the non-clitic form [ˈbɑ́ɑmɒ], speakers choose between the non-retracted [nɒ ˈbɑ́ɑmɒ] and retracted [ˈnɒ́ bɑɑ́mɒ]. In addition to 34 such minimal pairs, the data also includes 12 items in which the Long Falling and Long Rising patterns were not segmentally identical (5 are Long Falling, while 7 are Long Rising). Including these items is justified because the random slopes in our model reveal that segmental material does not influence experimental outputs (see below for details).

We fit the data into a mixed-effects logistic regression model with Response as the dependent variable (retraction coded as success; N = 1516) and two predictors: Tone (Long Falling and Long Rising; treatment-coded with Long Falling as the reference level) and Vowel Identity ([aa] vs. [ee], sum-coded with [ee] as reference). The full model includes an interaction between the two predictors as well as random intercepts for participant and frame with a by-participant and by-frame random slopes for tone. The interaction is not significant (AIC is 1701.5 in the model with the interaction and 1699.9 without it). The Tone is a significant predictor (AIC is 1699.9 in the model with the predictor Tone and 1710.9 in the model without it). The estimate is 0.90, with the z-score of 3.914, which is statistically significant $p < 0.0001$. Due to convergence, the model without predictor Tone had to be fit with the Nelder-Mead optimizer (all other models are estimated with the BOBYQA optimizer). Figure 6 illustrates the effects with 95% confidence intervals.

Estimates of the random slopes for the difference between High Rising and High Falling is positive for all frames. The same estimates are negative for only three out of 56 subjects. This again suggests that segmental material or neighborhood density do not crucially affect the results and that subjects are highly uniform in their responses.

Vowel quality. To test the effect of vowel quality on retraction, we analyze disyllables with the Long Rising pattern and vowels [ee, ɑɑ, ɛɛ] to maximally control for other influences. For example, for the non-clitic form [ˈbɛɛ́mɒ], participants chose between the non-retracted [nɒ ˈbɛɛ́mɒ] and retracted [ˈnɒ́ bɛɛ́mɒ]. For [ˈbeémɒ], the participants chose between the non-retracted [nɒ ˈbeémɒ] and retracted [ˈnɒ́ beémɒ].

The items include 7 minimal triplets (a total 21 items) that differ only in the first vowel [ˈbɛɛ́mɒ, ˈbeémɒ, ˈbɑɑ́mɒ]. The two items with mid vowels are coded as one frame; the item with [ɑɑ] is coded as a separate frame. Additionally, we analyze five item-pairs (10 in total) that differ only in [ee] and [ɛɛ] as one frame each. Finally, we include five items without pairs that include Long Rising tone and vowel [ɑɑ] (separate frames) for a total of 36 items (with corresponding retracted and non-retracted options).

To test if rates of retraction differ between the three vowels, we fit the data to a mixed-effects logistic regression model with Response as the dependent variable (retraction coded as success, N = 1201) and the predictor Vowel Quality ([ee, ɑɑ, ɛɛ]; treatment-coded with [ɛɛ] as the reference level). The full model includes random intercepts for participant and frame with by-participant and by-frame random slopes for vowel quality. As Table 2 shows, retraction is significantly less frequent than chance for vowel [ɛɛ]. Retraction is also significantly more frequent for vowel [ee] when compared to [ɛɛ] and more frequent for vowel [aa] when compared to [ɛɛ]. Figure 7 illustrates the effects with 95% confidence intervals. Mild underdispersion was detected in all three models: 0.82, 0.83, and 0.77 (using the *overdisp_fun()* from Bolker 2019), but because underdispersion is mild and it can cause conservative estimates, we leave the models uncorrected for underdispersion.

	β	SE	z	p
(Intercept)	−1.006	0.265	−3.802	0.0001
Vowel [ee] (versus [ɛɛ])	1.740	0.305	5.698	0.0000
Vowel [ɑɑ] (versus [ɛɛ])	1.739	0.285	6.108	0.0000

Table 2: Estimates of a mixed-effects logistic regression model showing preference for retraction from [ee] and [ɑɑ] compared to [ɛɛ] (Intercept) in disyllabic words with the Long Rising pattern. SE stands for Standard Error.

Again, random intercepts and slopes suggest that participants are highly uniform in their responses. All estimates for by-frame random intercepts and slopes conform to the estimates of the main effects with no deviations. The difference between [ɛɛ] and [ee] is not positive for only three participants (out of 56), and the difference between [ɛɛ] and [ɑɑ] for only one participant.

To summarize, the nonce-word experiment results suggest that retraction is preferred in monosyllables when compared to disyllables and in words with Falling pitch when compared to Rising pitch. Finally, [ɛɛ] inhibits retraction when compared to other vowels.

3.3 Discussion

The experiment allowed us to tackle the questions that go beyond the analysis of the data reported in Section 2. In particular, we confirmed that retraction is preferred in monosyllables when compared to disyllables, and in words with Rising tone when compared to Falling tone. Participants also strongly dispreferred retraction from [ɛɛ]. In this section, we offer an explanation of these patterns.

Word length. A greater preference for retraction was displayed in monosyllables than in disyllables. This result is not surprising, since monosyllables often behave differently than disyllables. For instance, while Serbo-Croatian contrasts rising and falling tones in disyllables and longer words, monosyllables allow only falling tone (Inkelas & Zec 1988; Zec 1999; Langston 1997). When it comes to segmental patterns, monosyllables are entirely protected by initial syllable faithfulness,

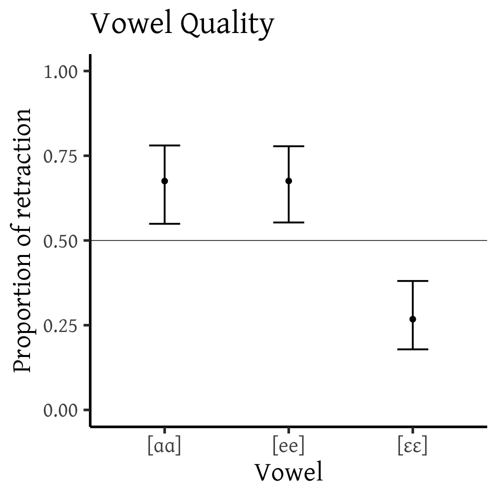


Figure 7: Estimates of a mixed-effects logistic regression model showing preference for retraction from [ee] and [ɑɑ] compared to [ɛɛ] in disyllabic words with the Long Rising pattern. Error bars indicate 95% confidence intervals.

but disyllables or longer words are not (Beckman 1997). In Tamil, coda nasals of the stem assimilate to the onset of the suffix, except in monosyllabic stems.

In Žiri, however, it seems that stress retraction in monosyllables has a simpler explanation, one that has to do with footing. Recall that Žiri has a trochaic foot. In monosyllables, retracting stress to the preceding clitic creates a disyllabic trochaic foot—e.g. *(ˈnɒ́ báam)*—while retaining stress on the stem results in a monosyllabic foot—*nɒ (ˈbáam)*. Put differently, it is more important to parse the entire prosodic word (clitic and the following noun) than not stressing the clitic, which creates a marked configuration as seen in Section 2.3. In disyllables no such preference exists because disyllables will have a non-parsed syllable regardless of retraction: *(ˈnɒ́ baá)mɒ* versus *nɒ (ˈbáamɒ)*. We leave aside whether these generalizations are made by the speakers at the level of underlying representations or on the surface. In the former case, the speakers would prefer the monosyllables to have a floating High tone and disyllables not to have it. In the latter case, the ranking of Parse-σ would have to be above Dep(H); otherwise the ranking would be the same as in (11).

An alternative way of interpreting these results involves Output-to-Output Correspondence (Benua 1997). Recall that in disyllables the Falling tone there is an alternation in the long vowel: the non-clitic form *(ˈbáamɒ)* retains its falling contour when preceded by an unstressed clitic, but changes the contour to rising when preceded by a stressed clitic: *(ˈnɒ́ baá)mɒ*. No such alternation is found in monosyllables. Hence the participants might prefer no tonal changes, and thus disfavour retraction in disyllables which necessarily leads to a tonal alternation on the long vowel. This reveals that an output-to-output faithfulness constraint on tonal identity inhibits stress retraction in disyllables, but have no effect on monosyllables. Under this view, retraction is a function of word length only indirectly: participants prefer to retain the contour on the long vowel, which applies equally to all words. Disyllables with stress retraction neutralize tone on the posttonic vowel while monosyllables do not.

Tone. We also saw that rising tone prefers more retraction than falling tone. In non-retracted configurations both types of tones are distinctive. In retracted configurations, however, the contrast is neutralized: only rising tone is possible on the surface. For instance, *(ˈnɒ́ baá)mɒ* is the derived from *ˈbaámɒ* or *ˈbáamɒ*. The tonal change in the words with falling tone can be explained as an Obligatory Contour Principle (OCP) effect: two adjacent tones are a marked situation, which is resolved by shifting the High tone from the first mora to the second one when preceded by a High tone on the clitic. Cross-linguistically, the prohibition on adjacent identical tones may be resolved by deleting one of the tones or shifting one of the tones so that they are no longer adjacent (Myers 1997). The latter situation is what happens in Žiri.

It is worth noting that the Falling tone becomes Rising only in very limited circumstances: when it is preceded by a stressed vowel with a High vowel and when the another syllable follows. The former is explained by the OCP: tone only shifts from the first mora to the second to avoid two High tones on adjacent moras. The latter is explained by the fact that High tones are generally avoided in word-final positions of derived words. This restriction is found elsewhere in Žiri, for instance in words without long vowels: *ˈkɒ́p* 'pile' alternates with *ˈnɒ́ kɒp* 'on the pile' and not **ˈnɒ́ kɒ́p*. We leave a full account of why retraction leads to a tonal change in disyllables but not monosyllables with a long vowel to future work.

The patterns just discussed also show another point about Žiri tone: when two High tones surface, this situation does not involve spreading. Instead, the two instances of High tone are independent autosegments. We know this because when an extra syllable is available at the end of the word, the second High tone will move one mora to the right to satisfy OCP. If spreading were involved in these examples, we would expect no alternation or additional spreading. This analysis also allows a better insight into the retraction with a falling tone, which are attested in real words: the phonetic data suggests that the second High tone is realized lower than the first, as from Figure 2. We attribute this to stress as well as downstep: a phonetic process in which the second High tone is realized lower than the first (Connell 2011). Downstep does not usually apply to tone spreading. All in all, the experimental results corroborate the representational analysis of the falling tone as two separate High tones.

Returning to the asymmetry between the Rising and Falling tones, Output-to-Output Correspondence becomes relevant yet again. Participants preferred more retraction with Rising tones, which retain the same contour regardless of whether they are retracted or not. With the Falling tone, retraction is dispreferred. This is because a retracted form would lead to tonal change, violating the pressure for the uniformity of tone across all instances of the same word. This complements the results seen in monosyllables.

Vowel quality. We finally move to vowel quality which also interacts with retraction. This interaction is tied to a specific vowel quality: retraction from [ɛɛ] is dispreferred compared to the other vowels. This type interaction between segmental and prosodic properties is very unusual. Most other types involve sonority, but in Žiri this is clearly not the case. High sonority vowels may prefer stress (Kenstowicz 1997), but [ɛɛ] is intermediate in sonority between [ee] and [ɑɑ] which we had in our nonce words. de Lacy (2006) and Blumenfeld (2006) make an explicit claim that sonority, but not segmental features, can interact with stress assignment and tone, and Rasin (2017); Shih (2016, 2018); Shih & de Lacy (2020) argue that no segmental property can affect stress placement. Our results offer clear evidence that vowel quality interacts with prosody, and that this is unrelated to sonority. Retraction from [ɛɛ] is dispreferred in production as well. The speaker who recorded the stimuli had difficulties producing the retracted pattern with this vowel.

A similar pattern is found in Standard Slovenian. There are both tonal and stress restrictions. Lax mid vowels cannot have a falling tone in Standard Slovenian (Becker & Jurgec 2017): in native words, this means that the tone will change to accommodate the vowel quality, whereas in loanwords, which generally only have falling tone, the vowel tenses. There is a second interaction between vowel quality and stress: lax vowels are marked and avoid stress. Becker & Jurgec (2020) analyze this situation in terms of positional faithfulness, which is satisfied by shifting stress to the following vowel in disyllables. When stress is stem-final, the vowel quality changes instead, and monosyllables are faithful.

Žiri offers another piece of evidence that interactions between vowel quality and prosody are attested. When compared to Standard Slovenian, however, there are a few key differences. First, it is only [εε] that is affected, since there is no corresponding back mid lax vowel. Second, [εε] in Žiri cannot be posttonic, regardless of its tone. When stressed, [εε] can have a Rising or Falling tone. When stress retracts from the Falling tone, the resulting tone is Rising in non-final syllables, and Falling in the word-final syllable. Because retraction is strongly disfavored, both tonal patterns are ruled out on unstressed [εε] and both are possible on stressed [εε], so the restriction in Žiri cannot be related to tone, but has to do with footing. One way to capture that is to say that [εε] has to be aligned with the left edge of the foot. This type of constraint complements the two other patterns in Standard Slovenian: a markedness constraint on a combination of tone and vowel quality (Becker & Jurgec 2017) or a positional faithfulness constraint on a specific prosodic position and vowel quality (Becker & Jurgec 2020). In Žiri the positional faithfulness is not available, because shifting stress can occur to another instance of [ε].

The interactions of tone and vowel quality (height or sonority) has a clear grounding in intrinsic fundamental frequency of vowels: the higher the vowel, the higher its fundamental frequency, which is attested in a wide variety of languages (see Becker & Jurgec 2017 for a review). However, the Žiri pattern singles out the mid lax vowel [εε] while other vowels behave identically. There is no obvious grounding for the pattern. Phonetically unmotivated processes are fairly well understood on the segmental level (Blevins 2004; Kiparsky 2006, 2008; Coetzee & Pretorius 2010; Beguš 2018, 2019, 2020). For instance, in Ojibwe [n] alternates with [ʃ]: *ki-naːn̠-aː* 'you fetch him' versus *ki-naːʃ̠-im-i* 'you fetch us' (Buckley 2000). There is no apparent phonetic motivation for this process, but is not obviously antagonistic to some universal phonetic tendency (Beguš 2019). Žiri provides an example of a phonetically unmotivated process that involves prosody and a segmental feature. The results of the wug-test clearly document such an interaction in Žiri, complementing the data in other varieties of Slovenian.

4 Conclusions

This paper provided a description of complex prosodic interactions in Žiri Slovenian. We have seen that Žiri displays an interaction of tone, stress and quantity, which further interact with vowel quality. Stress in Žiri docks to the leftmost syllable with a High tone. Stress falls on the long vowel or the syllable immediately preceding it. In words with only short vowels, stress is penultimate.

The complexity of the system defines a new frontier of what a pitch accent language might look like. Although pitch accent as a term is controversial (see Hyman 2006, 2009; Hyman & Leben 2020), it has been used to characterize prosodic patterns in languages ranging from Swedish, Latvian, Japanese to Basque, Luganda, and Yaqui. Compared to these languages, Žiri displays a greater number of prosodic contrasts. Disyllables, for instance, fully contrast seven different prosodic shapes based on the combination of stress, tone, and quantity. All prosodic contrasts fall within a single disyllabic foot, even in trisyllabic or longer words, contrasting Žiri from tonal languages like Mandarin, Vietnamese or Mazatec.

Žiri displays active alternations in which stress retracts to the preceding clitic. These alternations display several preferences, the productivity of which we corroborate in a perception experiment. We find that the participants prefer more retraction in monosyllables when compared to disyllables, and more from rising tone than from falling tone. These preferences are a consequence of the tonal alternations seen in retracted forms. Formally, we attribute these asymmetries to a combination of constraints on the distribution of tone (such as OCP) as well as output-to-output faithfulness.

Complicating things further, Žiri displays interactions between prosody and vowel quality. The particular restriction in Žiri involves the mid lax vowel [ɛɛ], which displays lower rates of retraction when compared to tense mid vowels and low vowels. This interaction is thus not related to vowel height or sonority. It is also not related to tone, because the retracted forms can bear either tone, albeit this is dependent on the position of the vowel within a word. This interaction is phonetically unmotivated, and adds to the growing body of work on possible interactions between segmental and suprasegmental properties.

Our final contribution is methodological. Tonal wug-tests are rare and our work shows how they can be used to corroborate the data based on elicitation. In the paper, we outline a variety of challenges in recording the stimuli and interpreting the results.

References

Akinlabi, Akinbiyi (1996). Featural affixation. *Journal of Linguistics* **32**. 239–289.

Anderson, Stephen R. (2005). *Aspects of the theory of clitics*. Oxford: Oxford University Press.

Anderson, Stephen R. (2011). Clitics. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren D. Rice (eds.) *The Blackwell Companion to Phonology*, Malden, MA: Blackwell. 2002–2018.

Becker, Michael & Peter Jurgec (2017). Interactions of tone and ATR in Slovenian. In Wolfgang Kehrein, Björn Köhnlein, Paul Boersma & Marc van Oostendorp (eds.) *Segmental Structure and Tone*, Berlin: De Gruyter. 11–26.

Becker, Michael & Peter Jurgec (2020). Positional faithfulness drives laxness alternations in Slovenian. *Phonology* **37**. 335–366.

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

Becker, Michael & Jonathan Lavine (2012). Experigen – an online experiment platform. Available at https://github.com/tlozoot/experigen.

Beckers, Gabriel J.L. (2002). `rms equalize`. Praat script.

Beckman, Jill N. (1997). Positional faithfulness, positional neutralization and Shona vowel harmony. *Phonology* **14**. 1–46.

Beguš, Gašper (2011). Relativna kronologija naglasnih pojavov govora Žirovske kotline poljanskega narečja. *Slovenski jezik / Slovene Linguistic Studies* **8**. 19–33.

Beguš, Gašper (2017). Effects of ejective stops on preceding vowel duration. *The Journal of the Acoustical Society of America* **142**. 2168–2184.

Beguš, Gašper (2018). *Unnatural phonology: A synchrony-diachrony interface approach*. Ph.D. dissertation, Harvard University.

Beguš, Gašper (2019). Post-nasal devoicing and the blurring process. *Journal of Linguistics* **55**. 689–753.

Beguš, Gašper (2020). Estimating historical probabilities of natural and unnatural processes. *Phonology* **37**. 515–549.

Bennett, Ryan & Robert Henderson (2013). Accent in Uspanteko. *Natural Language & Linguistic Theory* **31**. 589–645.

Benua, Laura (1997). *Transderivational Identity: phonological relations between words*. Ph.D. dissertation, University of Massachusetts, Amherst.

Berko, Jean (1958). The child's learning of English morphology. *Word* **14**. 150–177.

Bidwell, Charles A. (1969). *Outline of Slovenian morphology*. Pittsburgh, PA: University Center for Interantional Studies, University of Pittsburgh.

Blevins, Juliette (2004). *Evolutionary Phonology: The Emergence of Sound Patterns*. Cambridge: Cambridge University Press.

Blumenfeld, Lev A. (2006). *Constraints on phonological interactions*. Ph.D. dissertation, Stanford University, Palo Alto, CA. Available on Rutgers Optimality Archive, ROA 877, http://roa.rutgers.edu.

Boersma, Paul & David Weenink (1992). Praat. http://www.praat.org.

Bolker, Ben (2019). Glmm faq. Retrieved on June 22, 2022.

Buckley, Eugene (2000). On the naturalness of unnatural rules. *Proceedings from the Second Workshop on American Indigenous Languages. UCSB Working Papers in Linguistics* **9**. 1–14.

Caballero, Gabriela & Lucien Carroll (2015). Tone and stress in Choguita rarámuri (Tarahumara) word prosody. *International Journal of American Linguistics* **81**. 457–493.

Coetzee, Andries W. & Rigardt Pretorius (2010). Phonetically grounded phonology and sound change: The case of tswana labial plosives. *Journal of Phonetics* **38**. 404–421. ISSN 0095-4470.

Connell, Bruce (2011). Downstep. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.) *The Blackwell Companion to Phonology*, Malden, MA: Blackwell. 824–847.

de Lacy, Paul (2002). The interaction of tone and stress in Optimality Theory. *Phonology* **19**. 1–32.

de Lacy, Paul (2006). *Markedness: reduction and preservation in phonology*. Cambridge: Cambridge University Press.

de Lacy, Paul (2007). The interaction of tone, sonority, and prosodic structure. In Paul de Lacy (ed.) *The Cambridge handbook of phonology*, Cambridge: Cambridge University Press. 281–307.

Dixon, R. M. W. & Alexandra Y. Aikhenvald (2002). *Word: A cross-linguistic typology*. Cambridge: Cambridge University Press.

Ernestus, Mirjam & Harald Baayen (2003). Predicting the unpredictable: Interpreting neutralized segments in Dutch. *Language* **79**. 5–38.

Goldsmith, John A. (1976). *Autosegmental Phonology*. Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge.

Gonzalez-Gomez, Nayeli, Silvana Poltrock & Thierry Nazzi (2013). A "bat" is easier to learn than a "tab": Effects of relative phonotactic frequency on infant word learning. *PLOS ONE* **8**. 1–11.

Greenberg, Marc L. (1987). Prozodične možnosti v slovenskem knjižnem jeziku in slovenskih narečjih. *Slavistična Revija* **35**. 171–186.

Greenberg, Marc L. (2000). *A Historical Phonology of Slovene*. Heidelberg: C. Winter Universitätsverlag.

Herrity, Peter (2000). *Slovene: A Comprehensive Grammar*. London, New York: Routledge.

Horst, Jessica S. & Michael C. Hout (2016). The novel object and unusual name (noun) database: A collection of novel images for use in experimental research. *Behavior Research Methods* **48**. 1393–1409.

Hyman, Larry M. (2006). Word-prosodic typology. *Phonology* **23**. 225–257.

Hyman, Larry M. (2009). How (not) to do phonology typology: the case pof pitch-accent. *Language Sciences* **31**. 213–238.

Hyman, Larry M. & William R. Leben (2020). Tone systems. In Carlos Gussenhoven & Aoju Chen (eds.) *The Oxford Handbook of Language Prosody*, Oxford: Oxford University Press.

Inkelas, Sharon & Draga Zec (1988). Serbo-Croatian pitch accent: the interaction of tone, stress, and intonation. *Language* **64**. 227–248.

Jurgec, Peter (2007). *Novejše besedje s stališča fonologije: primer slovenščine [Neologisms in Phonology: The Case of Slovenian]*. Ph.D. dissertation, University of Ljubljana.

Jurgec, Peter (2016). Velar palatalization in Slovenian: Local and long-distance interactions in a derived environment effect. *Glossa* **1**. 24.

Jurgec, Peter (2019). Opacity in Šmartno Slovenian. *Phonology* **36**. 265–301.

Jurgec, Peter & Jessamyn Schertz (2020). Postalveolar co-occurrence restrictions in Slovenian. *Natural Language & Linguistic Theory* **38**. 499–537.

Kenstowicz, Michael (1997). Quality-sensitive stress. *Rivista di Linguistica* **9**. 157–187.

Kiparsky, Paul (2006). Amphichronic program vs. Evolutionary Phonology. *Theoretical Linguistics* **32**. 217–236.

Kiparsky, Paul (2008). Universals constrain change, change results in typological generalizations. In Jeff Good (ed.) *Linguistic universals and language change*, Oxford: Oxford University Press. 23–53.

Langston, Keith (1997). Pitch accent in Croatian and Serbian: Towards an autosegmental analysis. *Journal of Slavic Linguistics* **5**. 80–116.

Lenček, Rado L. (1966). *The verb pattern of contemporary standard Slovene: with an Attempt at a generative description of the Slovene verb by Horace G. Lunt*. Bibliotheca Slavica, Wiesbaden: Otto Harrassowitz.

McCarthy, John J. & Alan Prince (1993). Generalized Alignment. In Geert Booij & Jaap van Marle (eds.) *Yearbook in Morphology 1993*, Dordrecht: Kluwer. 79–153.

Myers, Scott (1997). OCP effects in Optimality Theory. *Natural Language and Linguistic Theory* **15**. 847–892.

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

Odden, David (1988). Floating tones and contour tones in Kenyang. *Studies in African Linguistics* **19**. 1–34.

Peperkamp, Sharon (1997). *Prosodic words*. The Hague: Holland Academic Graphics.

Rasin, Ezer (2017). The stress-encapsulation universal and phonological modularity. Unpublished manuscript. Massachusetts Institute of Technology.

Remijsen, Bert (2002). Lexically contrastive stress accent and lexical tone in Ma'ya. In Carlos Gussenhoven & Natasha Warner (eds.) *Laboratory phonology 7*, Berlin and New York: Mouton de Gruyter. 585–614.

Sandy, Clare Scoville (2017). *Prosodic Prominence in Karuk*. Ph.D. dissertation, Univeristy of California, Berkeley, Berkeley.

Shih, Shu-hao (2016). Sonority-driven stress does not exist. In Gunnar Ólafur Hansson, Ashley Farris-Trimble, Kevin McMullin & Douglas Pulleyblank (eds.) *Supplemental Proceedings of the 2015 Annual Meeting on Phonology*, Washington, DC: Linguistic Society of America.

Shih, Shu-hao (2018). On the existence of sonority-driven stress in Gujarati. *Phonology* **35**. 327–364.

Shih, Shu-hao & Paul de Lacy (2020). Evidence for sonority-driven stress. *Catalan Journal of Linguistics* **18**. 9–40.

Stanonik, Marija (1977). Govor Žirovske kotline in njenega obdobja. *Slavistična Revija* **25**. 293–309.

Statistical Office of the Republic of Slovenia (2021). SI-STAT data portal demographics.

Toporišič, Jože (1976/2000). *Slovenska slovnica*. Maribor: Obzorja.

van der Hulst, Harry (2011). Pitch accent systems. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren D. Rice (eds.) *The Blackwell Companion to Phonology*, Malden, MA: Blackwell. 1003–1026.

Zec, Draga (1999). Footed tones and tonal feet: rhythmic consistency in a pitch-accent language. *Phonology* **16**. 225–264.

Zhang, Jie & Yuwen Lai (2010). Testing the role of phonetic knowledge in mandarin tone sandhi. *Phonology* **27**. 153–201.

Zhang, Jie, Yuwen Lai & Craig Turnbull-Sailor (2006). Wug-testing the 'tone circle' in Taiwanese. *WCCFL* **25**. 453–461.

Zhang, Jie & Jiang Liu (2016). The productivity of variable disyllabic tone sandhi in tianjin chinese. *Journal of East Asian Linguistics* **25**. 1–35.

Zwicky, Arnold M. & Geoffrey K. Pullum (1983). Cliticization *vs.* inflection: English *n't*. *Language* **59**. 502–513.