



Context-dependent phonetic enhancement of a phonation contrast in San Pablo Macuiltianguis Zapotec

MAYA L. BARZILAI

KATHERINE J. RIESTENBERG 

**Author affiliations can be found in the back matter of this article*

RESEARCH

][ubiquity press

ABSTRACT

This paper presents novel data from San Pablo Macuiltianguis Zapotec, a Sierra Juárez variety spoken in Oaxaca, Mexico. This language is ‘laryngeally complex’ (Silverman, 1997), exhibiting both contrastive tone and contrastive phonation. We examine the acoustic properties of the modal and checked vowels in this language, showing that this contrast is variably produced with several different acoustic cues. We analyze the distribution of these cues as an instance of phonetic enhancement and further show that the prosodic position of a given vowel determines which of several enhancement patterns is used. We argue that extant theories of phonetic enhancement, while able to explain the patterns described here, fall short of predicting the distribution of enhancing cues. We therefore propose that future models of phonetic enhancement could be expanded to predict patterns in which more than one acoustic dimension is available for phonetic enhancement.

CORRESPONDING AUTHOR:

Maya L. Barzilai

Georgetown University, US

mlb290@georgetown.edu

KEYWORDS:

Sierra Juárez Zapotec;
phonation; voice quality;
phonetic enhancement

TO CITE THIS ARTICLE:

Barzilai, Maya L. and
Katherine J. Riestenberg. 2021.
Context-dependent phonetic
enhancement of a phonation
contrast in San Pablo
Macuiltianguis Zapotec. *Glossa:
a journal of general linguistics*
6(1): 18. 1–36. DOI: [https://doi.
org/10.5334/gjgl.959](https://doi.org/10.5334/gjgl.959)

1 INTRODUCTION

This paper examines a phonation contrast in San Pablo Macuiltianguis Zapotec, an Otomanguean language traditionally spoken in San Pablo Macuiltianguis, a community in the Sierra Juárez region of Oaxaca, Mexico. By focusing on a Sierra Juárez variety of Zapotec, this work addresses a gap in modern description of Zapotec phonation contrasts, which have up to now largely been described for varieties outside the Sierra Juárez group. Several Zapotec languages have been described as having a three-way vowel phonation contrast, such as a contrast among checked vowels ($V^?$), rearticulated vowels ($V^?V$), and modal vowels (V) (Pérez Báez 2015; Beam de Azcona 2016; Crowhurst, Kelly, & Teodocio 2016). It has also been shown that this phonological vowel phonation can have diverse phonetic realizations that differ based on a vowel's position in a phrase and can vary by individual speaker (e.g., Arellanes Arellanes 2010; Esposito 2010; Crowhurst, Kelly, & Teodocio 2016).

In the current study, we set out to understand the nature of the phonetic properties associated with the phonation contrast in San Pablo Macuiltianguis Zapotec, henceforth abbreviated MacZ.¹ We show there is a range of (sometimes redundant) phonetic cues to the phonation contrast in MacZ, and we propose that the patterning of this diversity can be straightforwardly understood as phonetic enhancement using a contrast and enhancement approach (Hall 2011). Phonation contrasts have received little attention in discussions of phonetic enhancement, but because they often have diverse phonetic realizations (see Section 3.2), they represent an interesting potential application of the theory. Interestingly, we find that changes in F0 contribute to the phonetic enhancements of this contrast, independently from the contrastive tone in the language that is also cued by F0. We also explore how speech context impacts the phonetic realizations of this contrast, examining words in isolation (referred to in this paper as citation form) and in a frame sentence (referred to in this paper as phrase-medial).

In the next section, we present existing descriptions and theories of phonetic enhancement, as well as criticism and an alternative approach to explaining redundant phonetic cues offered by Boersma (1998). Section 3 presents the phonological and phonetic facts about phonation contrasts in Zapotec languages. Our methods are presented in Section 4 and data on the acoustics of the phonation contrast in MacZ in citation form and phrase-medial environments are reported in Section 5. In Section 6, we analyze the data as exhibiting phonetic enhancements that vary according to the speech context in which words are produced. We show that an approach that combines contrast and enhancement theory (Hall 2011) with the ability to stochastically rerank Optimality-theoretic constraints according to different speech contexts (Boersma 1998; 2009; Goldwater & Johnson 2003) best explains the data. We also explore the acoustic evidence related to the phonological status of rearticulated vowels in MacZ, suggesting they may not comprise a third phonation type, as previously described for this and other Zapotec languages, but may instead be better analyzed as a sequence of a checked vowel followed by a modal vowel. In Section 7, we conclude with a summary, implications of the findings, and notes on future related work.

2 PHONETIC ENHANCEMENT

Phonological inventories of the world's languages reveal the preference for contrastive elements to be maximally perceptually distinct (e.g., Liljencrants & Lindblom 1972; Flemming 2002). When elements that participate in a contrast are not distinct enough to ensure accurate perception, additional cues to the contrast are often produced. This phenomenon has been termed “phonetic enhancement” (Stevens, Keyser & Kawasaki 1986; Keyser & Stevens 2006; Stevens & Keyser 1989). Broadly, phonetic enhancement describes the presence of any cue that increases the perceptibility of a contrast between sounds. Examples of phonetic enhancement of phonological contrasts are abundant and well-documented. For instance, the obstruent voicing contrast in American English is enhanced by lengthening of the preceding vowel, such that the vowel in ‘had’ is longer than the vowel in ‘hat’ (Denes 1955). In a perception study, Lisker (1986) further found 16 acoustic cues (such as closure duration and glottalization) that, when manipulated, affected listeners’ judgments of the ‘rapid/rabid’ contrast. While voicing or

¹ The abbreviation MacZ follows the convention in previous work (e.g., Foreman, 2006; Riestenberg, 2017; Tejada, 2012).

lack of voicing of the obstruent is the most reliable cue to the phonological contrast in these examples, one or more additional phonetic cues are used to enhance the contrast and make voiced and voiceless obstruents more perceptually distinct from one another.

Early accounts of phonetic enhancement such as Stevens and Keyser (1989) saw phonological features as either universally primary or universally secondary, with the secondary features always serving to perceptually enhance the primary features. For example, [rounding] was seen as a universally secondary feature that is available to enhance the primary feature [coronal]. The contrast between [-coronal] /w/ and [+coronal] /j/ is enhanced respectively by [+round], which lengthens and narrows the lip opening thus lowering the frequency of the second formant (F2), and [-round], which keeps F2 raised. Thus the features [coronal] (primary) and [round] (secondary) work together to create the phonological contrast between /w/ and /j/. Though Stevens and Keyser's (1989) proposal showed how several articulations can work together to contribute to the distinctiveness of a contrast, their claim that all features are universally either primary or secondary leaves little room for explaining language-specific phonetic enhancement patterns like those that we will present in this paper.

More recent approaches, such as Hall's (2011) contrast and enhancement theory, do not assume that any distinctive feature is inherently primary or secondary, but instead use language-specific phonological patterning to offer predictions about what feature underlies a given phonological contrast and what types of cues will contribute to the enhancement of that feature's percept in the context of the language's broader phonological system. Under Hall's (2011) approach, a phonological inventory consists of a hierarchy of ordered phonological features (e.g., Dresher 2009), with the relative scope of these active features determining which phonemes are specified for which features. Phonetic enhancement occurs when a phoneme is produced with any additional phonetic implementation that enhances the perceptibility of one of its feature specifications, making that phoneme more perceptually distinct from other phonemes in the inventory. For instance, given a 3-vowel inventory /i a u/, the features [low] and [back] can be hierarchically ordered in two different ways, as shown in [Figure 1](#), adapted from Hall (2011 p. 13).

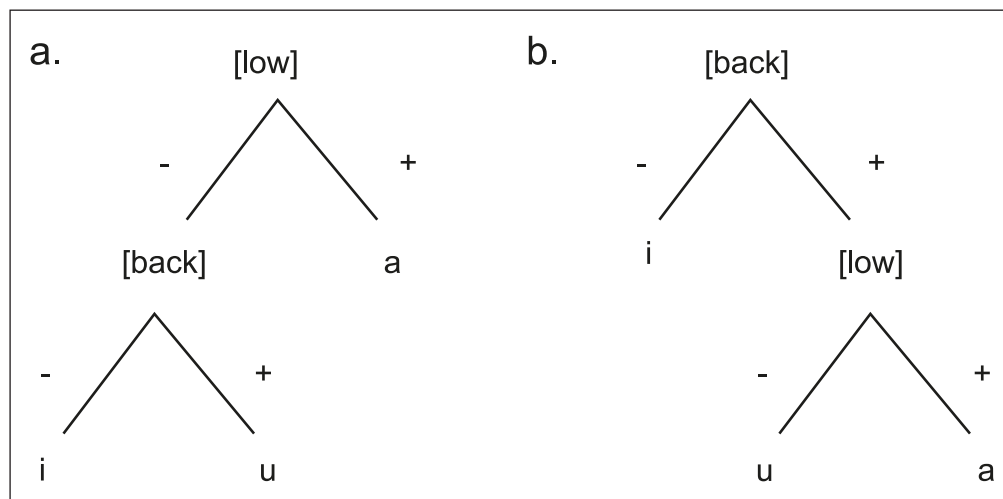


Figure 1 Possible orderings of [low] and [back] in a 3-vowel system.

In the feature hierarchy in 1a., /a/ is specified as [+low] and has no specification for the [back] feature, whereas in 1b., /a/ is specified as both [+back] and [+low]. Given the latter feature hierarchy, /a/ is likely to be produced with not only a backed tongue but also with lip rounding, which serves to phonetically enhance the low F2 percept of the [+back] feature specification. However, given the same vowel inventory with the feature hierarchy in 1a., lip rounding would not be an expected enhancement for the /a/ vowel, as in this case /a/ has no [back] specification, and therefore a low F2 percept is phonologically irrelevant. Contrast and enhancement theory (Hall 2011) thus explicitly links phonetic enhancements to the phonological features active in the contrasts of an inventory and frames the realization of additional phonetic cues to the contrast not as phonological but rather as a phonetic phenomenon. This helps to prevent contrasts that are overly distinct, e.g., imaginary but very perceptually distinct vowel inventories such as /i ẽ ɑ̃ ɔ̃ ũ/ (Lindblom & Maddieson 1988). As Hall (2011) explains, redundant cues are

“much more likely to be present if they reinforce a contrastive feature” (p. 39). While previous approaches to phonetic enhancement only suggested that “enhancement may take place whenever a given distinction can be made more salient than it might otherwise be” (Keyser & Stevens 2006: 42), Hall’s contrast and enhancement approach both constrains and helps to predict which such distinctions are eligible for enhancement.

Other phonetically-driven approaches to phonology, such as Boersma’s (1998) Functional Phonology, have argued against characterizing such processes as phonetic enhancement. Boersma (1998) agrees that contrasts rely on distinctive phonological features whose elements should be maximally perceptually distinct (e.g., Liljencrants & Lindblom 1972; Flemming 2002), but instead of hierarchies of binary features, Boersma adopts “constraint-ranking grammars that contain direct translations of principles of minimization of articulatory effort and perceptual confusion” (p. 148), using an Optimality Theory (OT) framework (Prince & Smolensky 1993). For Boersma, “phonetic explanations can be expressed directly in the production grammar as interactions of gestural and faithfulness constraints” (p. 467), as in this explanation of the lip rounding that tends to accompany contrastively back vowels:

In the auditory spectrum, the front-back distinction is represented by the second formant (F2) [...] Specifying the value “max” for F2 means that F2 should be at a maximum, given a fixed value of F1; this is most faithfully rendered by producing a front vowel with lip spreading. The value “min” specifies a minimum value of F2 given F1; this is most faithfully implemented as a rounded back vowel. No “enhancement” of an allegedly distinctive feature [back] by an allegedly redundant feature [round], as proposed by Stevens, Keyser & Kawasaki (1986) for reasons of lexical minimality, is implied here: the two gestures just implement the same perceptual feature symmetrically (p. 21).

In Functional Phonology, faithfulness constraints specify the value of “Max” or “Min” for a given distinctive feature with a corresponding perceptual continuum (F2, in the example above) and thus drive the articulatory implementation of that feature. Of course, this is not completely unlike how Hall (2011) explains similar data. Some enhancements, Hall states, work to push the cue to the contrastive feature further along a given acoustic continuum. For example, “a contrastively [–back] vowel can be enhanced by being realised as front rather than merely central” (Hall, 2011, p. 20), or, as explained above, “redundant rounding can enhance contrastive backness, because both have the effect of lowering F2” (Hall, 2011, p. 20). Also similarly to Hall (2011), Boersma’s (1998) proposal accounts for the lack of the imaginary /i ē ɔ u^ɨ/ vowel inventory described above. Nasality on an /e/ vowel does not increase the percept of any feature active in the contrast between /e/ and the other cardinal vowels, so an /ẽ/ candidate would not emerge as optimal given a constraint ranking in the style of Functional Phonology.

However, Hall (2011) differs from Boersma (1998) in explicitly allowing for phonetic enhancement along multiple acoustic and articulatory dimensions and not just along a single, linear, perceptual continuum, such as in the examples regarding F2 given above. Hall (2011: 20) provides five different types of enhancement, ranging from directly related to the acoustic and articulatory correlates of a feature to more indirectly related. The example of lip rounding as an enhancement of a back vowel, as above, is one of the most direct relationships between the phonological feature and its phonetic enhancement. In the most indirect type of enhancement, “A feature with a particular acoustic/auditory correlate can be enhanced by a separate acoustic/auditory effect that increases the relative salience of that correlate” (Hall, 2011: 20). For example, because low vowels have a relatively high F1, a [+low] vowel that contrasts with a [–low] vowel may be enhanced by lowering F0, which increases the height of F1 relative to F0 thereby making the distinction more salient (Hall, 2011: 20). The properties of voicing contrasts in American English obstruents, discussed above, can also be understood in this way. Some cues to the voicing contrast, such as vowel duration, closure duration, and glottalization (Denes 1955; Lisker 1986; Chong & Garellek 2018), are phonetically distinct from the cues to obstruent voicing but are nonetheless explainable by the articulatory and perceptual facts associated with the voicing contrast. It is not clear how an approach like that of Boersma (1998) applies to these cases, in which additional phonetic cues to a contrast can clearly be argued to enhance the percept of a contrastive feature, but this enhancement does not fall neatly along a given unidimensional perceptual continuum, like F2, that can be set at minimum or maximum.

Hall (2011) and Boersma (1998) thus offer two different approaches to explaining the redundant phonetic cues that emerge to increase the perceptual distinctiveness of contrastive elements in a phonological inventory. Both Functional Phonology (Boersma 1998) and contrast and enhancement theory (Hall 2011) focus on distinctive contrastive features (e.g., Steriade 1987) to offer predictions about which phonological features of a contrast will lead to the emergence of redundant phonetic cues. However, the two approaches differ in their explanations of how the associated phonetic cues are realized. For Hall (2011: 20), phonological features are abstract and redundant acoustic and articulatory cues can enhance the overall contrastive system along multiple dimensions and in direct or indirect ways. For Boersma, phonological features correlate to a particular perceptual continuum and faithfulness constraints drive the “Max” or “Min” implementation of a feature along that continuum.

While these differing approaches have all relied on segmental contrasts as test cases (Stevens, Keyser & Kawasaki 1986; Stevens & Keyser 1989; Boersma 1998; Keyser & Stevens 2006; Hall 2011), we know that redundant phonetic cues also occur with suprasegmental contrasts. Vowel duration has been shown to be an additional cue to tonal contrasts in Mandarin (Dreher & Lee 1968) and Thai (Gandour 1977). There are also many documented instances of voice quality as an additional cue to tone contrast (e.g., Morén & Zsiga 2006; Nguyen & Macken 2008; Yu and Lam 2014; Uchihara 2016; Kuang 2017). For instance, Yu & Lam (2014) show that the low tone in Cantonese is produced with creaky voice in addition to low pitch, and that Cantonese listeners are sensitive to the presence of this phonation difference when identifying this tone. The articulatory configurations required to produce low pitches are similar to those required to produce creaky phonation, and therefore the association between the two is cross-linguistically common, both diachronically and synchronically (Gordon & Ladefoged 2001; Kingston 2011).

Phonation contrasts, in comparison to tone contrasts, have received little attention in discussions of phonetic enhancement, but because they often have diverse phonetic realizations (see Section 3.2), they offer an interesting opportunity to explore the opposing viewpoints presented here. We revisit this in our discussion, where we show how an approach informed by both contrast and enhancement (Hall 2011) and Functional Phonology (Boersma 1998) can account for the diverse phonetic realizations of the phonation contrast in MacZ.

3 MACUILTIANGUIS ZAPOTEC

The Zapotec language family is a group of at least 20 distinct languages distributed across five regions in the Mexican states of Oaxaca and Veracruz (Beam de Azcona 2016). Together with the Chatino languages, Zapotec languages form the Zapotecan branch of the Otomanguean family. Internal classification of Zapotecan languages is complex and ongoing (see Beam de Azcona 2016; Campbell 2017 for overviews). We therefore follow the convention of referring to Zapotec varieties by the names of the communities where they are traditionally spoken. Work that supports the conservation of this language family is urgent, as nearly all varieties are in danger of disappearing within a generation or two. MacZ, the variety that is the focus of this paper, joins with a handful of other Zapotec varieties spoken in the area to form a group classified as Sierra Juárez Zapotec (Northern) (ISO 639-3, *zaa*) (Smith-Stark 2007; Simons and Fennig 2018) or Sierra Zapotec (West) (INALI 2008).

The 2010 Mexican census reported that Macuiltianguis had 897 residents over age three, nearly a quarter of whom were over age 65 (INEGI 2010). This includes residents of Macuiltianguis proper as well as those in San Juan Luvina, a Zapotec community under the municipal jurisdiction of Macuiltianguis and located about 7.5 kilometers to the south. Foreman (2006) describes a sharp population decline in Macuiltianguis since 1960 and in Luvina since 1980 due to migration to locations throughout Mexico and the United States, with several hundred community members now living in or around Oaxaca City, Mexico City, and parts of the United States. The 2010 census report showed a rapid decrease in the number of Indigenous language speakers in the Macuiltianguis municipality over the last two generations, with 96% of people over age 45 but only 36% of people ages 5–14 reporting that they spoke an Indigenous language (INEGI 2010). However, the second author, who resided in Macuiltianguis from 2015–2016, has encountered no children between the ages of 5 and 14 who fluently speak MacZ during several years of linguistic fieldwork and activism with speakers living both within and outside of the community. The young people identified as speakers in the 2010 census may all be residents of Luvina,

or they may be newer migrants to the community who do not speak the traditional Zapotec language but rather a neighboring Chinantec language, as the census data do not specify which Indigenous language a given individual speaks.

3.1 PHONETICS OF CONTRASTIVE PHONATION IN ZAPOTEC LANGUAGES

Zapotec languages are “laryngeally complex” (Silverman 1997), meaning they exhibit both contrastive phonation and lexical tone. Phonation contrasts may make use of a range of voice quality configurations, including glottal closures, creaky voice, breathiness, and so on. In the current paper, we use the term “phonation” to refer to the phonological phenomenon that relies on these voicing differences to create contrast, and we use the term “voice quality” to refer to the phonetic properties of voicing. Many Zapotec languages have been described as having a three-way vowel phonation contrast, such as a contrast among checked vowels ($V^?$), rearticulated vowels ($V^?V$), and modal vowels (V) (Pérez Báez 2015; Beam de Azcona 2016; Crowhurst, Kelly, & Teodocio 2016). The glottalization in these checked and rearticulated vowels has most often been analyzed as the realization of a contrastive phonation type and not a consonant segment, as codas are often prohibited or highly restricted in Zapotec languages, and glottal stops never appear as onsets. As demonstrated in [Table 1](#), contrastive phonation has phonologically diverse realizations across Zapotec varieties and can also have quite phonetically diverse realizations within a single variety (Arellanes Arellanes 2010; Esposito 2010; Crowhurst, Kelly, & Teodocio 2016).

ZAPOTEC VARIETY	REPORTED BY	PHONATION CONTRAST	PHONETIC REALIZATIONS DESCRIBED
San Melchor Betaza	Crowhurst et al. (2016)	Modal /a/ Checked /a [?] / Rearticulated /a [?] a/	Phrase-final modal and rearticulated vowels terminate in breathiness. Phrase-final checked vowels can have a full glottal closure, creakiness, or aperiodic voicing. Non-final checked vowels can be creaky or have reduced or absent laryngealization. Phrase-final rearticulated vowels have medial aperiodicity and sometimes creakiness. Non-final rearticulated vowels can be creaky and resemble checked vowels.
San Pablo Güilá	Arellanes Arellanes (2010)	Modal /a/ Strongly laryngealized /a [?] / Weakly laryngealized /a/	Speaker 1: Strongly laryngealized vowels are checked [a [?]] or rearticulated [a [?] a]. Weakly laryngealized vowels are stiff [a] or weakly checked [a [?]]. Speaker 2: Strongly laryngealized vowels are checked [a [?]] or creaky [aqa]. Weakly laryngealized vowels are stiff [a] or creaky checked [aa].
Santa Ana del Valle	Esposito (2010)	Modal /a/ Breathy /a/̤/ Creaky /a/̰/	Breathy and creaky vowels always carry a falling tone and become more modal in focus position.

Table 1 Phonation contrasts and phonetic realizations in three Zapotec varieties.

Taken together, the studies represented in [Table 1](#) suggest that the three-way glottalization contrast on vowels in Zapotec languages has a range of phonetic realizations depending on context and individual speaker and that the factors conditioning the different realizations are also diverse and not always categorical. The MacZ phonation contrast that is the focus of the current paper (and is described in the next section) is most similar to that of San Melchor Betaza Zapotec (Crowhurst, Kelly, & Teodocio 2016). One goal of the current study is thus to investigate whether two varieties with the same phonological characterization of phonation also mirror each other in their phonetic realizations of the contrast. If so, we expect to see that in citation form, word-final modal and rearticulated vowels terminate in breathiness, as found for Betaza Zapotec. We also expect word-final checked vowels in citation form in MacZ to vary between creakiness and a full glottal closure. For phrase-medial checked vowels, we expect to see fewer full glottal closures but more creakiness.

3.2 PHONOLOGICAL CHARACTERISTICS OF MACUILTIANGUIS ZAPOTEC

The focus of this paper is the phonation contrast in MacZ, exemplified in (1), where it can be seen that this contrast is largely independent from the lexical tone contrast that is described in more detail below. MacZ has five vowel qualities /a/, /e/, /i/, /o/, and /u/, and each of these can exhibit a contrast between modal, checked, and rearticulated phonation in open syllables.

(1) Contrastive phonation in MacZ

Modal	Checked	Rearticulated
nǎ COPULA	nǎʔ ‘hand’	naʔá ‘where’
ʕi.lá ‘sister’	ʕi.láʔ ‘cotton’	laʔá ‘gourd’

As noted in the previous section, most descriptions of Zapotec languages consider this type of glottalization to be contrastive phonation, i.e., a feature of vowels rather than an independent consonant segment (see e.g., Chávez Peón 2010: 212–214). We follow such an analysis here, primarily because /ʔ/ surfaces in places few consonants appear, and conversely, does not surface in places we might expect consonants to appear.²

A more precarious question has to do with the status of so-called rearticulated vowels (V^ʔV) and whether they are best analyzed as single vowels with a glottal interruption, as in [Table 1](#) and example (1), or as disyllabic sequences of a checked vowel followed by a modal vowel (V^ʔ.V). The single vowel analysis can be easily motivated for Zapotec languages in which the post-tonic vowel of disyllabic roots has historically been deleted resulting in almost exclusively monosyllabic roots (e.g., Chávez Peón 2010). In such cases, it would not make sense to propose that (C)V^ʔ.V is the only type of disyllabic root allowed. MacZ, however, retains post-tonic vowels and allows plenty of disyllabic roots, so this is not a reason to count out (C)V^ʔ.V as a possible root type. There are also some examples of words of the type (C)V₁^ʔ.V₂, such as *ruʔa* ‘mouth,’ and *niʔa* ‘foot,’ which cannot be analyzed as rearticulated vowels due to the change in vowel quality. If these are sequences of a checked vowel followed by a modal vowel, then V^ʔV sequences in which the vowel quality does not change could be analyzed the same way.

Further, V^ʔV is not allowed everywhere that modal (V) and checked (V^ʔ) vowels are allowed. Modal and checked vowels can appear in both stressed and unstressed syllables, while V^ʔV, if analyzed as a single vowel, would be restricted to stressed syllables in root-final position, as in *beláʔa* ‘precipice, cliff.’³ On the other hand, if V^ʔV is analyzed as a sequence of a checked vowel followed by a modal vowel, this would seem to open the door for many more trisyllabic roots in a language which historically has primarily allowed monosyllabic and disyllabic roots. However, such cases can often be shown to have been multimorphemic historically. For example, *beláʔa* ‘precipice, cliff,’ could include a reflection of the proto-Zapotec prefix **pe-*, which is generally analyzed as a type of animacy marker but also often found on plants and other nouns of the natural (and supernatural) world (e.g., Beam de Azcona 2016). Unlike in some other Zapotec languages, tone does not help disambiguate between a monosyllabic and disyllabic analysis of V^ʔV sequences in MacZ because under current analyses, the same sets of tones that are allowed across disyllabic roots are allowed on monosyllabic roots (Foreman 2006; Riestenberg 2017).

If vowels of the type V^ʔV do constitute single vowels with a glottal interruption, then they form part of a three-way contrast among modal (V), checked (V^ʔ) and rearticulated vowels (V^ʔV) in MacZ, as has been proposed for other Zapotec languages (e.g. Pérez Báez 2015 for Isthmus Zapotec; Crowhurst, Kelly, & Teodocio 2016 for Betaza Zapotec). Under such an analysis, the property that allows V^ʔV to create contrast is that the glottalization feature is anchored autosegmentally to the center portion of the vowel rather than the final portion of the vowel as for checked vowels (Arellanes Arellanes 2014; López Nicolas 2014).⁴ In this sense, the contrast has to do with timing rather than with the phonetic cues to glottalization, the latter being the central concern of this study. For this reason, and also due to their ambiguous status, we include words that terminate in V^ʔV in the current study but focus primarily on the contrast

2 The glottal stop does not appear syllable-initially, while all other consonants in MacZ do appear in this context. Only a few consonants can appear in coda position (/l/, /s/, /r/, /n/, /m/), and while this is more segments than other Zapotec languages allow in coda position, words with these codas are rare and can often be traced to historical borrowings. Of these, only /n/ can appear word-finally. On the other hand, /ʔ/ is fairly common syllable-finally and word-finally. Word medially, /ʔ/ can appear before /n/, as in *guʔna* ‘bull,’ before /j/ as in *iʔja* ‘wooded area,’ and before /r/ as in *siʔru* ‘sad.’ However, we have found no examples of consonants appearing before these sounds. Analyses of /ʔ/ as a consonant in Zapotec languages can be found in Foreman (2006) and Avelino Becerra (2004). Other authors have previously proposed a glottal-stop-as-consonant analysis only to later favor a vowel phonation analysis (cf. Pickett, 1953; 1955; Marlett & Pickett 1987).

3 There is one known exception, the verb ‘sleep,’ whose root is *-aʔaθi*.

4 Such approaches assume that rearticulated vowels and other phonetically long vowels, though prosodically bimoraic, have three autosegmental slots or portions eligible to be anchored to suprasegmental phonation features (Arellanes Arellanes 2014; López Nicolas 2014).

between checked and modal vowels. Still, we take advantage of the opportunity to explore the phonetic evidence related to the phonological status of rearticulated vowels and revisit this issue in our discussion.

Along with contrastive phonation, MacZ has a number of lexical tone contrasts, examples of which are given in *Tables 2* and *3*. Monosyllabic word roots with modal phonation exhibit a contrast among low, mid,⁵ falling, rising, and dipping tones (*Table 2*); dashes here represent unattested tone-phonation combinations. Monosyllabic word roots with checked phonation show a slightly restricted set of tone contrasts among low, mid,⁶ rising, and dipping; falling tones are not allowed on checked vowels. Word roots with rearticulated vowels can exhibit mid, high, falling, and rising tones. Across disyllabic roots, there are at least five contrastive tone patterns, regardless of whether the final vowel is modal or checked: low.low, mid.mid, high.high, high.low, and mid.high (*Table 3*). Though not particularly common, a handful of tonal minimal pairs can be found in the language. For instance, the word *íyyà* with low tones means ‘flower’ while the word *íyyá* with high tones means ‘rock.’ The fact that such examples are somewhat rare indicates that tone may have low functional load in the language compared with phonation and other phonological contrasts.

	LOW	MID	HIGH	RISING	FALLING	DIPPING
Modal	bà ‘tomb’	ja ‘tree’	–	jú ‘fig’	sâ ‘day’ jù ‘land’	dà ‘bean’ jò ‘river’
Checked	bà? ‘animal’	ra? ‘above’	–	nǎ? ‘hand’	–	nǎ? ‘female’
Rearticulated	–	kwe?e ‘back’	já?á ‘green’	lle?é ‘conceited’	já?à ‘raw’ llé?è ‘stomach’	–

Table 2 Tone contrasts on monosyllabic (modal, checked, rearticulated) words.⁷

	LOW.low	MID.mid	HIGH.high	HIGH.low	MID.high
Modal	‘jè.là ‘night’	‘je.da ‘ear of corn’	‘ǵú.dí ‘father’	‘jé.dà ‘cigarette’	‘du.sí ‘drunk’
Checked	‘là.sì? ‘allergy’	‘la.sì? ‘seed’	‘ǵí.dá? ‘Chinantec’	gú.sà? ‘mud’	‘lo.sé? ‘tongue’

Table 3 Tone contrasts on disyllabic words with final modal and checked vowels (capitalized tone levels indicate the stressed syllable).

It is important to reiterate that checked vowels do not constitute any particular tone category in MacZ, as modal and checked vowels allow a similar set of contrastive tones. However, because this language has both contrastive phonation and contrastive tone, phonetic enhancement of one of these laryngeal features may have phonological consequences for the other. For instance, as mentioned above, Esposito (2010) found that breathy and creaky vowels in Santa Ana del Valle Zapotec always carry a falling tone. This is independent of a contrast between high and rising tones found on modal vowels. Falling tones, however, do not play a role in the tone contrast in this language. This suggests that enhancing a phonation contrast with a pitch cue could in turn obscure the distinctness of cues to lexical tone. In our discussion, we consider whether the existence of both tone and phonation in the language plays a role in the types of laryngeal cues available for phonetic enhancement of the phonation contrast.

⁵ It is worth noting that mid-toned monosyllabic word roots with modal phonation appear to be quite rare. Analyses of monomorphemic nouns and adjectives have revealed only one example: *ja* ‘tree.’ The lexical tone patterns of verb roots have not yet been analyzed and may yield additional examples. However, a number of monosyllabic enclitics with modal phonation and underlying mid tones have been identified, such as the proximal demonstrative =*ni* and the emphatic adverb =*ba* (Foreman, 2006). Also pending is systematic analysis of whether mid-surfacing tones are underlyingly Mid or underlyingly toneless. This has proven difficult, as there is a lack of simple noun morphology that would allow for such testing, and the highly complex verbal morphology makes it difficult to pinpoint the source of tonal changes. For now, we follow Foreman (2006) in not marking mid tones.

⁶ The F0 of the monosyllabic word roots with a mid tone tends to be higher for items with checked phonation than items with modal phonation. However, there does not seem to be evidence of a contrast between mid and high tones among the monosyllabic word roots examined to date. Another possible analysis for these word roots is that they have an underlying falling tone but the low portion of the falling tone is blocked by the checked phonation thus realizing a high tone. These different analyses need to be further explored in future research.

⁷ Contour tones in MacZ have been described as surfacing on phonetically long vowels (Riestenberg, 2017). Because vowel length on its own is not contrastive in this language, we assume that this generalization is due to phonetic or allophonic vowel lengthening and leave the details about vowel duration and its interaction with phonological tone to future work.

The language data presented here are from elicitation sessions conducted with native speakers of Macuiltianguis Zapotec in Oaxaca, Mexico by the second author. Speaker 1 is female, born in 1958, and resides in San Pablo Macuiltianguis. Speaker 2 is male, born in 1942, and resides in Estado de Mexico. Both speakers are also native speakers of Spanish. All recordings were made using an Audio-Technica ATR3350 Omnidirectional Condenser Lavalier Microphone connected to a Tascam DR-40 recorder.

Words terminating in modal (V), checked (V²) and rearticulated vowels (V²V) were elicited in two different contexts: in isolation (which we refer to as citation form) and in a carrier phrase (which we refer to as phrase-medial). The carrier phrase was *Rniya' X, Y, Z nna* 'I say X, Y, and Z,' where Z was the target word and X and Y were other words from the word list (counterbalanced and randomized). Both portions of the rearticulated vowels were analyzed, as V1 and V2 respectively, to observe whether the phonetics of these these vowel portions support either a monosyllabic or disyllabic analysis of the V²V sequences. Investigating the production of the phonation contrast in these two contexts allows us to assess whether the phonetic cues to phonation vary based on the individual speaker and a vowel's position in a phrase, while also allowing some control over other potentially intervening factors (such as following consonant, focus position, and speech rate).

A word list consisting of 144 monomorphemic words was created for elicitation.⁸ The word list contained 82 words terminating in a modal vowel, 44 words terminating in a checked vowel, and 18 words terminating in a rearticulated vowel. The lack of existing documentation and analysis of this language made it difficult to find more monomorphemic words terminating in checked or rearticulated vowels, but this distribution seems to reflect the distribution in the language overall. For the same reason, it was not possible to fully counterbalance the list for tone and phonation type, and not all phonation types allow all tones at the same rates. As noted above, rearticulated vowels never carry dipping or low tones, and falling tones never appear on checked vowels. Some other combinations are possible but rare. [Table 4](#) shows the distribution of phonation type and tone in the word list.

		TONE						TOTAL
		LOW	MID	HIGH	DIP	RISE	FALL	
Phonological Phonation Type	Checked	12	11	9	2	10	N/A	44
	Modal	23	15	16	8	7	13	82
	Rearticulated	N/A	5	2	N/A	3	8	18
Total		35	31	27	10	20	21	144

Table 4 Distribution of word-final phonation type and tone in the word list.⁹

Both sets of data we analyze below, citation form data and phrase-medial data, were examined for their voice quality and F0 properties. We look at voice quality measurements as this is the expected cue to phonation contrasts, and previous work, discussed above, has reported variable patterns in the voice quality phonetics across and within Zapotec varieties. We choose to analyze F0 phonetics due to an observation from previously collected data that some vowels were produced with fluctuations in F0 in the final portion of their vowels, patterning in ways that appeared orthogonal to the F0 cues to phonological tone. An F0 cue to phonation would not be surprising in light of the articulatory considerations; both pitch and voice quality are the result of laryngeal articulations. On the other hand, given that this language uses F0 as a cue to tone, a contrast phonologically independent from the phonation contrast, the presence of F0 cues to phonation would provide interesting evidence that one laryngeal cue, F0, corresponds to two laryngeal contrasts, both tone and phonation.

⁸ The word list included nouns (n = 120), adjectives (n = 16), adverbs (n = 2), interrogatives (n = 2), and number words (n = 4) (see Appendix A). The imbalance among lexical categories here is reflective of the fact that words other than nouns are rarely monomorphemic in MacZ, and monomorphemic words were prioritized in the formation of the word list as an additional control. Verbs in MacZ are never monomorphemic and were thus excluded from the current study.

⁹ Rearticulated vowels are listed here with one tone, following a monosyllabic analysis of V²V sequences as described for other Zapotec languages (Chávez Peón 2010). For mid and high tones, both vowel portions are produced with the same level tone. For contour tones, there is one tone target per vowel portion, creating an overall rise or fall across the sequence. As revealed in the rest of this section, we see no evidence that phonological tone interacts with the phenomena discussed here, and therefore we leave the question of tones on rearticulated vowels to future work.

To measure voice quality properties, the intervals during which each vowel exhibited modal voicing, creaky voicing, and breathy voicing were delineated in a TextGrid in Praat (Boersma & Weenink 2018), and the durations of each interval were measured. Boundaries were marked for each voice quality type using visual examination of the spectrogram and waveform, as well as auditory perception. If, for any one of these voice quality types, no portion of the vowel was produced with it, no corresponding interval was marked. Intervals were also created for the duration of any closure and burst associated with checked vowels. Representative examples of interval segmentation for words of all three phonological phonation types are in [Figure 2](#). The durations of these intervals were extracted and then converted into the percentage of the vowel during which each type of voicing was exhibited.

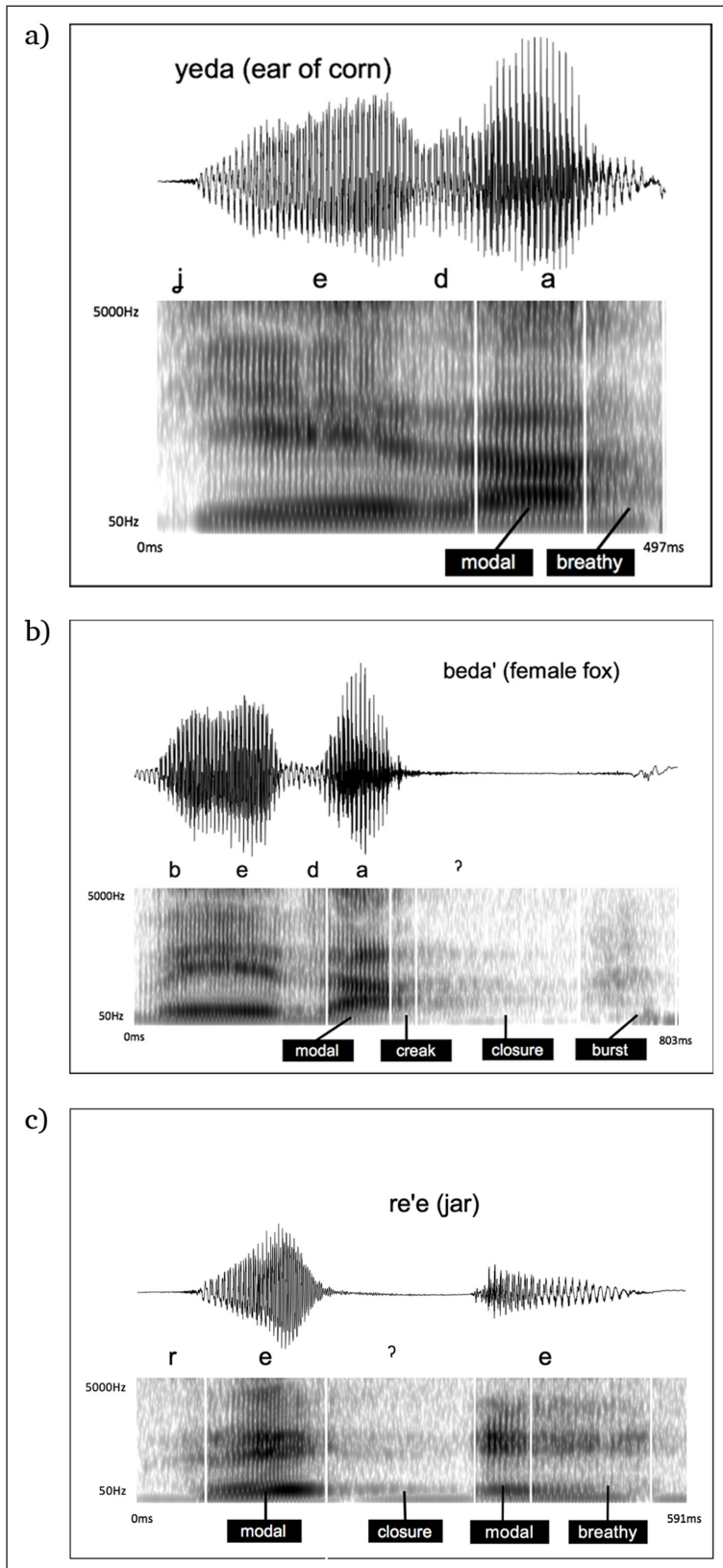


Figure 2 Segmentation of voice quality in phonologically modal (a), checked (b), and rearticulated (c) vowels in citation form.

Changes in pitch trajectory across the duration were not as straightforward to measure as voice quality changes. This is due to the rapid nature of the F0 fluctuations observed, their relatively short duration, and the fact that they occur in addition to the F0 trajectory associated with phonological tone. For these reasons, vowels were coded independently by both authors as exhibiting either a “change” or “no change” in pitch. Changes in pitch manifested as a fluctuation in F0 over the final portion of the vowel or as a dropoff in pitch in the final portion of the vowel that was not a result of the lexical tone of the vowel. Examples of tokens that were coded as exhibiting a fluctuation in F0 are provided in [Figure 3](#). Vowels that were coded as having no F0 change were those for which the pitch track did not deviate from the expected pitch trajectory for the tone of the given word. Vowels for which the pitch track was unreliable were coded as “other”. There were coding discrepancies with 10% of the F0 data, which were resolved through discussion.

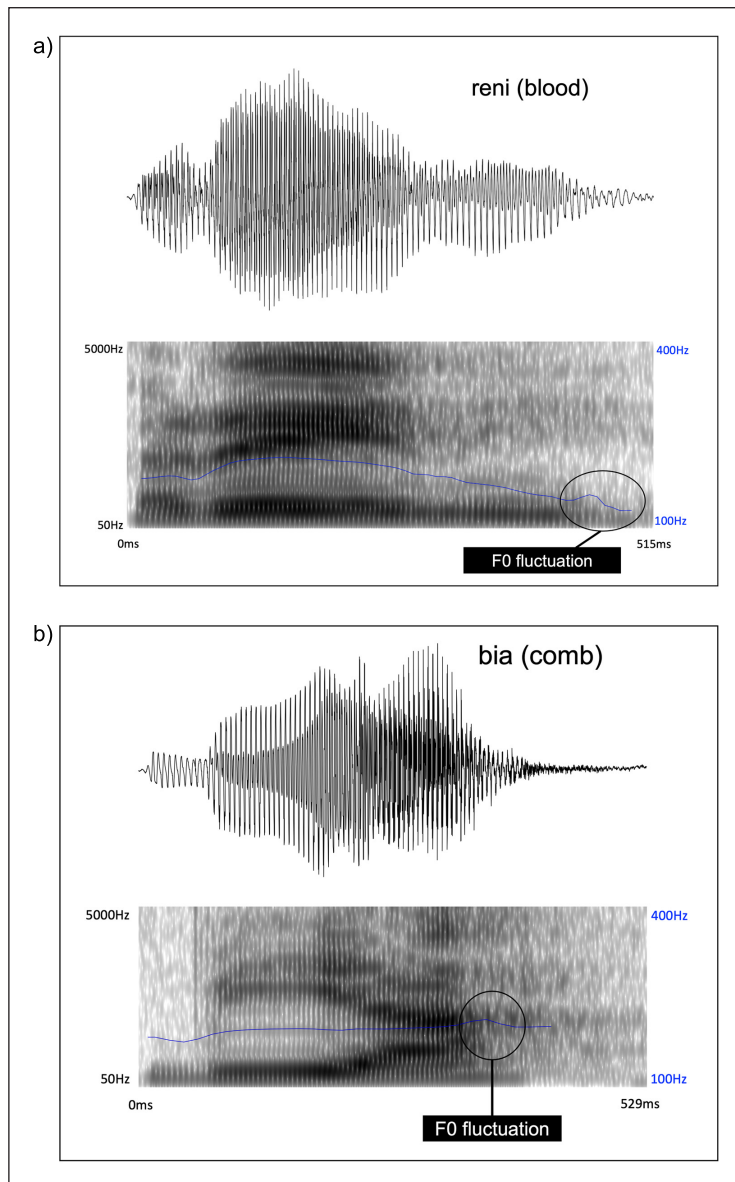


Figure 3 Two instances of F0 fluctuation.

5 RESULTS

This section presents data on the voice quality and F0 acoustics of checked, modal, and rearticulated vowels in MacZ in the two different speech contexts (citation form and phrase-medial). We first present the acoustic results of phonologically modal and checked vowels in both contexts, showing that these vowels differ both in the voice qualities with which they are produced and in the rates at which they exhibit fluctuations in F0. We then address the V1 and V2 portions of rearticulated vowels, showing that V1 vowels are largely acoustically similar to checked vowels and V2 vowels are acoustically similar to modal vowels with respect to both the voice quality and F0 measures.

5.1 MODAL AND CHECKED VOWELS

Tables 5 and 6 show the number of tokens of modal and checked vowels in each speech context and for each speaker that exhibited each voice quality type.

CONTEXT	VOWEL TYPE	VOICE QUALITY MEASURES				
		MODALITY	CREAKINESS	BREATHINESS	CLOSURE	RELEASE
citation	modal	84/84 (100%)	0/84 (0%)	56/84 (67%)	0/84 (0%)	0/84 (0%)
	checked	44/47 (94%)	15/47 (32%)	0/47 (0%)	33/47 (70%)	29/47 (62%)
medial	modal	90/91 (99%)	0/91 (0%)	2/91 (2%)	0/91 (0%)	0/91 (0%)
	checked	51/51 (100%)	24/51 (47%)	2/51 (4%)	36/51 (71%)	0/51 (0%)

Table 5 Number (percentage) of tokens with each quality measure, Speaker 1.

CONTEXT	VOWEL TYPE	VOICE QUALITY MEASURES				
		MODALITY	CREAKINESS	BREATHINESS	CLOSURE	RELEASE
citation	modal	84/87 (97%)	0/87 (0%)	58/87 (67%)	0/87 (0%)	0/87 (0%)
	checked	46/46 (100%)	17/46 (37%)	0/46 (0%)	43/46 (93%)	45/46 (98%)
medial	modal	84/94 (89%)	0/94 (0%)	15/94 (16%)	1/94 (1%)	0/94 (0%)
	checked	51/51 (100%)	11/51 (22%)	0/51 (0%)	49/51 (96%)	0/51 (0%)

Table 6 Number (percentage) of tokens with each voice quality measure, Speaker 2.

The counts of tokens exhibiting the different voice quality types are largely similar across the speakers, except that Speaker 2 produced glottal closures for checked vowels in both speech contexts at a higher rate than Speaker 1, and with a higher rate of audible glottal releases in the citation form context. Speaker 2 also produced more modal vowels in phrase-medial context with breathy voicing than Speaker 1, and Speaker 1 produced more checked vowels in phrase-medial context with creaky voicing than Speaker 2. While the overall pattern for the two speakers is similar, we include *Speaker* as a random effect in our analyses below to account for this slight inter-speaker variation.

Note that nearly all vowel tokens, regardless of whether they are phonologically checked or modal, display some duration of modal voicing, typically the beginning and middle of the vowel. The different types of non-modal voicing typically occur in the final portion of the vowel. Therefore, the count data in Tables 5 and 6 do not on their own fully represent the voice quality differences across different phonation types. In order to provide more informative data, we also analyzed the differences in proportional durations of these voice quality types across vowels in the two speech contexts. Mean proportional durations were calculated for each type of voice quality interval (measured as described in Section 4 above) for modal and checked phonation vowels in both citation form and in phrase-medial contexts. These mean proportional durations are presented in Table 7 and visualized as boxplots of proportional duration in Figure 4 with the data from the two speakers collapsed.

Phonologically modal vowels in both citation form and medial contexts are on average produced with modal voicing for at least 75% of the duration of the vowel; in citation form contexts, the modal portion is slightly shorter and the vowel tends to be followed by a relatively long period of breathiness. Checked vowels, on the other hand, tend to have a much shorter mean duration of modal voicing. In citation form, on average 41% of the total duration of a checked vowel consists of a glottal closure; in medial position this closure lasts for an average of 36% of the vowel. Phonologically checked vowels also tend to have periods of creaky voicing in both speech contexts.

CONTEXT	VOWEL TYPE	VOICE QUALITY MEASURES				
		MODALITY	CREAKINESS	BREATHINESS	CLOSURE	RELEASE
citation	modal	0.75	0	0.25	0	0
	checked	0.39	0.05	0	0.41	0.12
medial	modal	0.92	0	0.07	0.004	0
	checked	0.56	0.08	0.006	0.36	0

Table 7 Mean proportional durations of voice quality measures for vowels of each phonation type and produced in each context.¹⁰

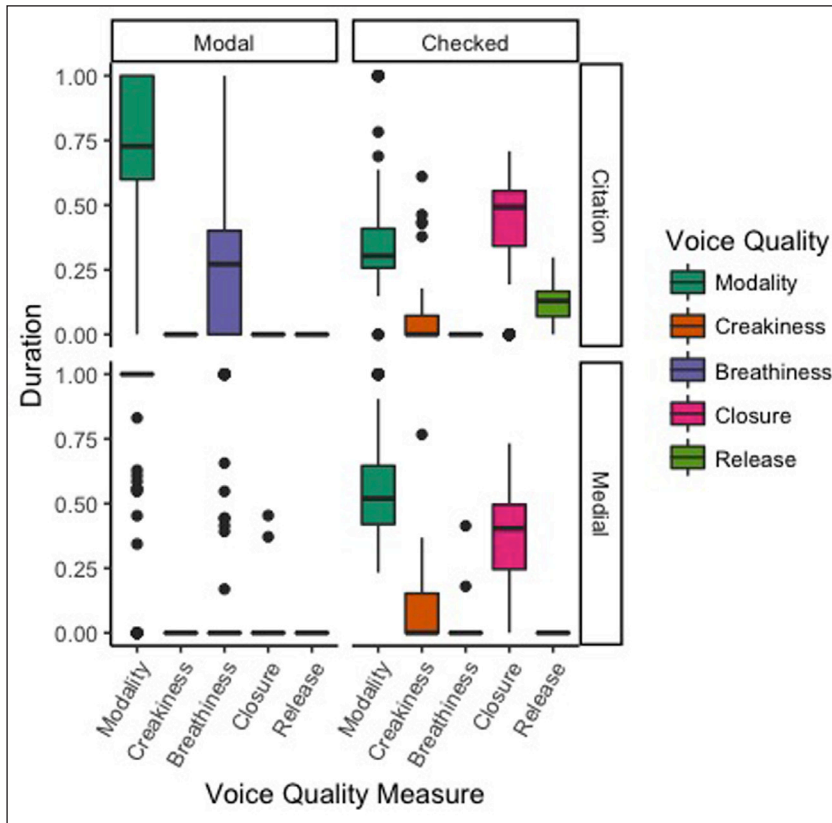


Figure 4 Proportional duration of each voice quality type by phonological vowel type and elicitation context.

5.1.1 Statistical analyses of voice quality (modal vs. checked)

To investigate whether the differences in voice quality proportions shown in [Table 7](#) and [Figure 4](#) were statistically significant, a series of individual mixed-effects linear regression models were fit using the lmer function in the lme4 R package (Bates et al., 2015), each predicting the mean proportional duration of one of the voice quality measures. Original models included as random effects the final lexical tone target¹¹ of the vowel and whether the segment preceding the final vowel was a sonorant, voiced obstruent, or voiceless obstruent. However, neither of these factors improved the fit of any model, so the results of the models are reported here without these factors included.

5.1.1.1 Statistical analysis of modal voicing (modal vs. checked vowels)

[Table 8](#) presents the results of the regression model predicting the proportional duration of modal voicing for phonologically modal and checked vowels in the two production contexts. *Modal* was set as the reference level for phonological phonation type, and *Citation Form* was set as the reference level for context. *Speaker* and *Word* were included as random effects.

¹⁰ Note that all of the proportional durations for each token summed to a total of 1. In other words, none of the duration of any vowel token was excluded from these measurements. The failure of rows in [Table 7](#) to add up to exactly 1 is due to rounding discrepancies.

¹¹ Final tone target was operationalized as the tone target at the end of the vowel, i.e., the pitch target just before the locus of phonation contrast. For example, the final tone target was Low for low and falling tones, Mid for mid and dipping tones, and High for high and rising tones.

FIXED EFFECTS	ESTIMATE	STD. ERROR	T-VALUE	P-VALUE
(Intercept)	0.7534	0.0487	15.4547	0.0213 *
Type				
Checked	-0.3637	0.0313	-11.6223	<0.001 ***
Context				
Phrase-Medial	0.1766	0.0228	7.7404	<0.001 ***
Type * Context				
Checked:Phrase-Medial	-0.0087	0.0384	-0.2255	0.8217

Table 8 Mixed-effects linear regression model: proportional duration of modal voicing.

The results in [Table 8](#) show that the proportional duration of modal voicing is significantly shorter in phonologically checked vowels than it is in phonologically modal vowels ($p < 0.001$), and significantly longer in phrase-medial contexts than in citation form contexts across vowel types ($p < 0.001$). There is no significant interaction in proportional duration of modality between phonological vowel type and utterance context. In other words, the pattern in which phrase-medial vowels have longer proportional durations of modality than vowels in citation form is consistent across the two phonation types.

5.1.1.2 Statistical analysis of creaky voicing (modal vs. checked vowels)

[Table 9](#) presents the results of the regression model predicting the proportional duration of creaky voicing for phonologically modal and checked vowels in the two production contexts. *Modal* was set as the reference level for phonological phonation type, and *Citation Form* was set as the reference level for context. *Speaker* and *Word* were included as random effects.

FIXED EFFECTS	ESTIMATE	STD. ERROR	T-VALUE	P-VALUE
(Intercept)	<0.001	0.0102	0.0097	0.9932
Type				
Checked	<0.001	0.0099	5.4113	<0.001 ***
Context				
Phrase-Medial	<0.002	0.0071	-0.0012	0.9991
Type * Context				
Checked:Phrase-Medial	0.0231	0.0120	1.9260	0.0547

Table 9 Mixed-effects linear regression model: proportional duration of creaky voicing.

The results in [Table 9](#) reveal a significant main effect of phonological vowel type, such that checked vowels have a significantly longer proportional duration of creakiness than phonologically modal vowels ($p < 0.001$). There was no main effect of context on the proportional duration of creakiness, meaning that neither context was likely to have more creaky voicing than the other when abstracting away from vowel type. The interaction between phonological vowel type and utterance context was marginally significant ($p = 0.0501$). Pairwise comparisons showed significant differences between the vowel types in each context ($p < 0.001$), but no significant differences between the contexts for each vowel type; these pairwise results are in line with the main effects in [Table 9](#), showing that creaky voicing is more likely on checked vowels than on modal vowels, and this is true regardless of utterance context.

Note that some cells in [Table 7](#) have 0 values for the creaky voicing column. An average duration of 0 for creaky voicing means that no tokens of these vowels had any period of creaky voicing, as confirmed in [Tables 5](#) and [6](#); in cases such as these, the pairwise comparisons reveal a significantly decreased likelihood of exhibiting a type of voice quality altogether, rather than a shorter proportional duration of this voice quality. This interpretation applies to any model presented here that involves 0 values in [Table 7](#).

5.1.1.3 Statistical analysis of breathy voicing (modal vs. checked vowels)

Table 10 presents the results of the regression model predicting the proportional duration of breathy voicing for phonologically modal and checked vowels in the two production contexts. *Modal* was set as the reference level for phonological phonation type, and *Citation Form* was set as the reference level for context. *Speaker* and *Word* were included as random effects.

FIXED EFFECTS	ESTIMATE	STD. ERROR	T-VALUE	P-VALUE
(Intercept)	0.2453	0.0259	9.4841	0.0148
Type				
Checked	-0.2450	0.0261	-9.3740	<0.001 ***
Context				
Phrase-Medial	-0.1830	0.0181	-10.1222	<0.001 ***
Type * Context				
Checked:Phrase-Medial	0.1881	0.0304	6.1793	<0.001 ***

Table 10 Mixed-effects linear regression model: proportional duration of breathy voicing.

The results in Table 10 reveal a significant main effect of phonological vowel type, such that checked vowels have a significantly shorter proportional duration of breathy voicing than phonologically modal vowels ($p < 0.001$). There was also a significant main effect of speech context ($p < 0.001$); vowels produced phrase-medially have a significantly shorter period of breathy voicing than vowels produced in citation form. Finally, there was a significant interaction between phonological vowel type and utterance context. A pairwise comparison shows that modal vowels produced in phrase-medial contexts, as well as checked vowels in either context, are produced with significantly less breathiness than phonologically modal vowels in citation form ($p < 0.001$). In other words, modal vowels produced in citation form have a higher proportional duration of breathy voicing than any of the other three possible combinations of vowel type and context.

5.1.1.4 Statistical analysis of glottal closure duration (modal vs. checked vowels)

Table 11 presents the results of the regression model predicting the proportional duration of glottal closure for phonologically modal and checked vowels in the two production contexts. *Modal* was set as the reference level for phonological phonation type, and *Citation Form* was set as the reference level for context. *Speaker* and *Word* were included as random effects.

FIXED EFFECTS	ESTIMATE	STD. ERROR	T-VALUE	P-VALUE
(Intercept)	-0.0005	0.0306	-0.0165	0.9892
Type				
Checked	0.4077	0.0153	26.5925	<0.001 ***
Context				
Phrase-Medial	0.0020	0.0126	0.1620	0.8713
Type * Context				
Checked:Phrase-Medial	-0.0526	0.0212	-2.4771	0.0135 *

Table 11 Mixed-effects linear regression model: proportional duration of glottal closure.

The results in Table 11 reveal a significant main effect of phonological vowel type, such that checked vowels are significantly more likely to have some period of glottal closure than modal vowels ($p < 0.001$). There was no significant main effect of utterance context, but there was a significant interaction between phonological vowel type and utterance context ($p = 0.0135$). Pairwise comparisons reveal that though there is no significant difference between modal vowels produced in citation form and those produced phrase-medially ($p = 0.998$), the proportional duration of the glottal closure differed significantly for checked vowels in the different contexts ($p = 0.016$), with checked vowels in citation form exhibiting proportionally longer closures than checked vowels in the medial context.

5.1.1.5 Statistical analysis of glottal release duration (modal vs. checked vowels)

Table 12 presents the results of the regression model predicting the proportional duration of glottal release for phonologically modal and checked vowels in the two production contexts. Modal was set as the reference level for phonological phonation type, and Citation Form was set as the reference level for context. Speaker and Word were included as random effects.

FIXED EFFECTS	ESTIMATE	STD. ERROR	T-VALUE	P-VALUE
(Intercept)	-0.0001	0.0096	-0.0168	0.9891
Type				
Checked	0.0175	0.0040	29.2852	<0.001 ***
Context				
Phrase-Medial	<0.001	0.0033	0.0037	0.9997
Type * Context				
Checked:Phrase-Medial	-0.0117	0.0056	-21.1298	<0.001 ***

Table 12 Mixed-effects linear regression model: proportional duration of glottal release.

The results of the model in Table 12 reveal a significant main effect of phonological vowel type ($p < 0.001$), such that checked vowels are more likely to have a glottal release than modal vowels. There was no significant main effect of utterance context, but the interaction between phonological vowel type and context was significant ($p < 0.001$). There was no portion of any modal vowel in either context produced with a glottal release, and therefore the relevant interaction arises from the difference in checked vowels across contexts. A pairwise comparison showed that checked vowels in citation form had a significantly longer proportional duration of glottal release than checked vowels produced utterance-medially.

5.1.1.6 Summary of statistical analyses of voice quality

The results of voice quality measurements of modal and checked vowels show that, regardless of context (citation form vs. phrase-medial), phonologically modal vowels have a significantly greater proportional duration of modal voicing than checked vowels, while checked vowels are significantly more likely to exhibit creaky voicing. In citation form only, phonologically modal vowels are significantly more likely to exhibit breathy voicing and phonologically checked vowels are significantly more likely to have a glottal closure and audible glottal release.

5.1.2 Statistical analysis of F0 (modal vs. checked)

Table 13 shows the number of tokens of each phonological vowel type and utterance context that were produced with a change in F0, no change, or other F0 trajectory. Modal vowels in citation form tended to be produced with a rapid change or fluctuation in F0 trajectory in the final portion of the vowel, whereas checked vowels in the same context were unlikely to exhibit this change. Vowels in medial position, regardless of phonological vowel type, were also unlikely to exhibit a change in F0 trajectory. It was the checked vowels that were more likely to be coded as “other”; this is likely due to the fact that the onset of the glottalization associated with checked vowels can lead to aperiodicity in the signal and therefore unreliable pitch tracks. We excluded the few “other” tokens from the statistical analysis so that this unreliable data would not interfere with the main analysis of F0 behavior. The same data are visualized in Figure 5 with the “other” tokens omitted.

CONTEXT	VOWEL TYPE	F0 TRAJECTORY		
		CHANGE	NO CHANGE	OTHER
citation	modal	118	44	1
	checked	9	65	13
medial	modal	3	160	1
	checked	3	73	12

Table 13 Number of tokens with each F0 trajectory by phonological phonation type and utterance context.

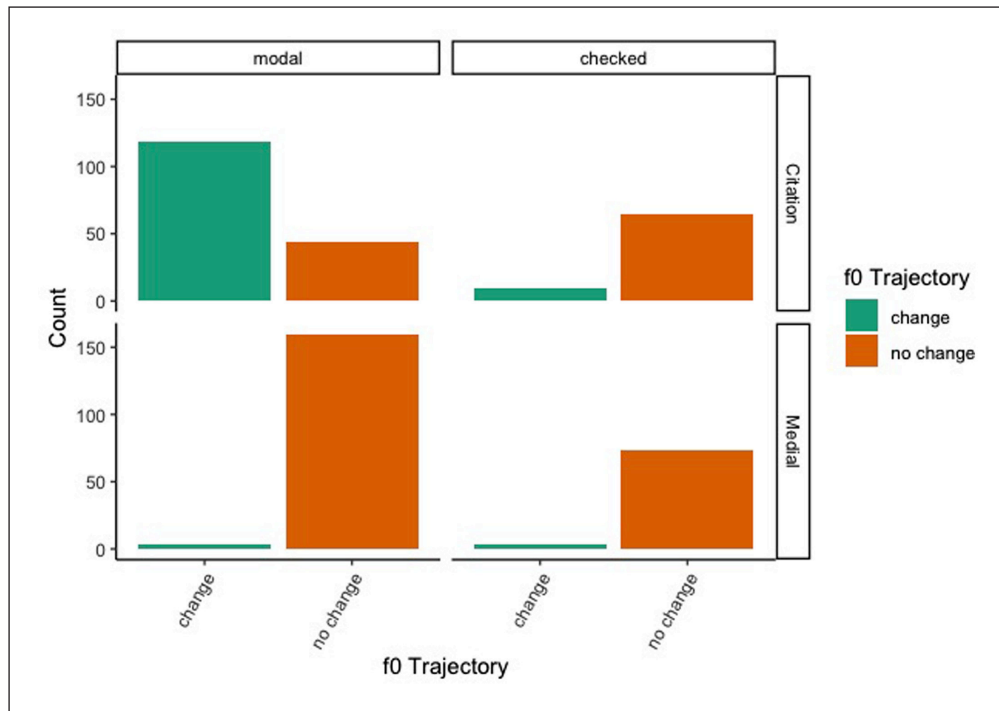


Figure 5 Counts of F0 trajectory types by phonological vowel type and elicitation context.

A mixed-effects logistic regression model was fit using the `glmer` function in the `lme4` R package (Bates et al., 2015) to predict F0 trajectory. **Table 14** presents the results of the regression model predicting “change” or “no change” in F0 for phonologically modal and checked vowels in the two utterance contexts. *Modal* was set as the reference level for phonological phonation type, and *Citation Form* was set as the reference level for context. *Speaker* was included as a random effect.

FIXED EFFECTS	ESTIMATE	STD. ERROR	T-VALUE	P-VALUE
(Intercept)	1.0515	0.4385	2.3982	0.0165
Type				
Checked	-3.1036	0.4158	-7.4641	<0.001 ***
Context				
Phrase-Medial	-5.1777	0.6246	-8.2889	<0.001 ***
Type * Context				
Checked:Phrase-Medial	3.8973	0.9307	4.1873	<0.001 ***

Table 14 Logistic regression model: F0 trajectory.

The results in **Table 14** show a significant main effect of phonological vowel type, such that checked vowels were significantly less likely to exhibit a change in F0 than modal vowels ($p < 0.001$). There was also a significant main effect of utterance context ($p < 0.001$); vowels produced utterance-medially were significantly less likely to exhibit a change in F0 than vowels produced in citation form. The model also revealed a statistically significant interaction between phonological phonation type and utterance context. A pairwise comparison showed that checked vowels in either utterance context, as well as phonologically modal vowels produced utterance-medially, were significantly less likely to be produced with a change in F0 than phonologically modal vowels produced in citation form ($p < 0.001$). In other words, changes in F0 are statistically most likely to occur in modal vowels produced in citation form than in any other vowel type in any other context.

5.1.3 Summary of significant results: modal vs. checked vowels

Table 15 summarizes the results of the statistical analyses presented in sections 5.1.2 and 5.1.3.

CUE		VOWEL TYPE (MODAL VS. CHECKED)	CONTEXT (CITATION VS. MEDIAL)	INTERACTIONS
Voice Quality	Modal	Greater proportion in modal vowels ($p < 0.001$)	Greater proportion in medial contexts ($p < 0.001$)	none
	Creaky	More likely in checked vowels ($p < 0.001$)	No significant difference	Slightly greater proportion for phrase-medial checked vowels, marginally significant ($p = 0.0501$)
	Breathy	Greater proportion in modal vowels ($p < 0.001$)	Greater proportion in citation context ($p < 0.001$)	Greatest proportion for citation form modal vowels ($p < 0.001$)
	Glottal closure	More likely in checked vowels ($p < 0.001$)	No significant difference	Slightly greater proportion for citation form checked vowels than medial ($p = 0.016$); no effect of context for modal vowels ($p = 0.998$)
	Glottal release	More likely in checked vowels ($p < 0.001$)	No significant difference	Most likely for citation form checked vowels ($p < 0.001$)
Change in F0		More likely in modal vowels ($p < 0.0001$)	More likely in citation form ($p < 0.0001$)	Most likely for citation form modal vowels ($p < 0.001$)

Table 15 Summary of statistically significant results (modal vs. checked).

5.2 MODAL, CHECKED, AND REARTICULATED (V1 & V2) VOWELS

Having shown the acoustic patterning of phonologically modal and checked vowels in this data set, we turn now to the acoustics of the two vowel portions that comprise V^2V vowel sequences in MacZ. We do so by conducting the same analyses as in the previous section, but this time examining all four vowel types (modal, checked, V1, and V2), paying special attention to the patterning of V1 and V2 vowels. As explained in section 3.2, the phonological status of the V^2V sequences is ambiguous, with some phonotactic and prosodic evidence in favor of an analysis of rearticulated (V^2V) vowels as a unitary, monosyllabic segment (and thus rearticulation as a third type of phonation, contrasting with modal and checked vowels), and other evidence in favor of analyzing V^2V as a sequence of a checked vowel followed by a modal vowel. To investigate whether there is phonetic support for the latter analysis, we analyze the acoustics of both portions of V^2V here in comparison with the analyses of modal and checked vowels presented above. Similarities between checked vowels and the first portions of rearticulated vowels and between modal vowels and the second portions of rearticulated vowels would provide some support for an analysis in which rearticulated vowels are in fact a sequence of a checked vowel and a modal vowel.

5.2.1 Statistical analyses of voice quality (modal, checked, and rearticulated vowels)

Tables 16 and **17** show the number of tokens for each vowel type and context that are produced with any duration of each of the voice qualities. For ease of comparison, modal and V2 vowels are shown in dark gray while checked vowels and V1 are shown in light gray; the color coding throughout this section is to emphasize the potential similarities between vowels in these two pairs. The proportion of tokens produced with each voice quality measure is similar between the two speakers across voice quality and utterance contexts. In the remainder of this section, we analyze the differences in proportional durations and include *Speaker* as a random effect to account for inter-speaker variation.

Table 18 shows the proportional duration of each voice quality measure for all vowel types in citation form and in utterance-medial context, comparing data from the V1 and V2 portions of rearticulated vowels to the data previously presented in **Table 7**. The same data are visualized as boxplots in **Figure 6**.

V1 portions of rearticulated vowels in both production contexts show an average of 40–50% proportional duration of modal voicing, some period of creakiness, and a relatively long glottal closure. They are produced similarly to checked vowels medial position, in that they do not display the glottal releases typical of citation form checked vowels. It is to be expected that V1 vowels would resemble medial position checked vowels more than citation form checked

CONTEXT	VOWEL TYPE	VOICE QUALITY MEASURES				
		MODALITY	CREAKINESS	BREATHINESS	CLOSURE	RELEASE
citation	modal	84/84 (100%)	0/84 (0%)	56/84 (67%)	0/84 (0%)	0/84 (0%)
	checked	44/47 (94%)	15/47 (32%)	0/47 (0%)	33/47 (70%)	29/47 (62%)
	V1	20/23 (87%)	15/23 (65%)	3/23 (13%)	17/23 (74%)	0/23 (0%)
	V2	23/23 (100%)	0/23 (0%)	18/23 (78%)	0/23 (0%)	0/23 (0%)
medial	modal	90/91 (99%)	0/91 (0%)	2/91 (2%)	0/91 (0%)	0/91 (0%)
	checked	51/51 (100%)	24/51 (47%)	2/51 (4%)	36/51 (71%)	0/51 (0%)
	V1	23/23 (100%)	20/23 (87%)	1/23 (4%)	12/23 (52%)	0/23 (0%)
	V2	23/23 (100%)	2/23 (9%)	1/23 (4%)	0/23 (0%)	0/23 (0%)

Table 16 Number of tokens (proportion of tokens) with each quality measure, Speaker 1.

CONTEXT	VOWEL TYPE	VOICE QUALITY MEASURES				
		MODALITY	CREAKINESS	BREATHINESS	CLOSURE	RELEASE
citation	modal	84/87 (97%)	0/87 (0%)	58/87 (67%)	0/87 (0%)	0/87 (0%)
	checked	46/46 (100%)	17/46 (37%)	0/46 (0%)	43/46 (93%)	45/46 (98%)
	V1	18/20 (90%)	15/20 (75%)	1/20 (5%)	17/20 (85%)	0/20 (0%)
	V2	20/20 (100%)	1/20 (5%)	14/20 (70%)	0/20 (0%)	0/20 (0%)
medial	modal	84/94 (89%)	0/94 (0%)	15/94 (16%)	1/94 (1%)	0/94 (0%)
	checked	51/51 (100%)	11/51 (22%)	0/51 (0%)	49/51 (96%)	0/51 (0%)
	V1	24/25 (96%)	20/25 (80%)	1/25 (4%)	11/25 (44%)	0/25 (0%)
	V2	22/25 (88%)	0/25 (0%)	8/25 (32%)	0/25 (0%)	0/25 (0%)

Table 17 Number of tokens with each quality measure, Speaker 2.

CONTEXT	VOWEL TYPE	VOICE QUALITY MEASURES				
		MODALITY	CREAKINESS	BREATHINESS	CLOSURE	RELEASE
citation	modal	0.75	0	0.25	0	0
	checked	0.39	0.05	0	0.41	0.12
	V1	0.40	0.19	0.05	0.36	0
	V2	0.72	0.003	0.28	0	0
medial	modal	0.92	0	0.07	0.004	0
	checked	0.56	0.08	0.006	0.36	0
	V1	0.48	0.31	0.009	0.20	0
	V2	0.84	0.01	0.12	0	0

Table 18 Mean proportional duration of each voice quality measure for vowels of each phonological type and produced in each context.

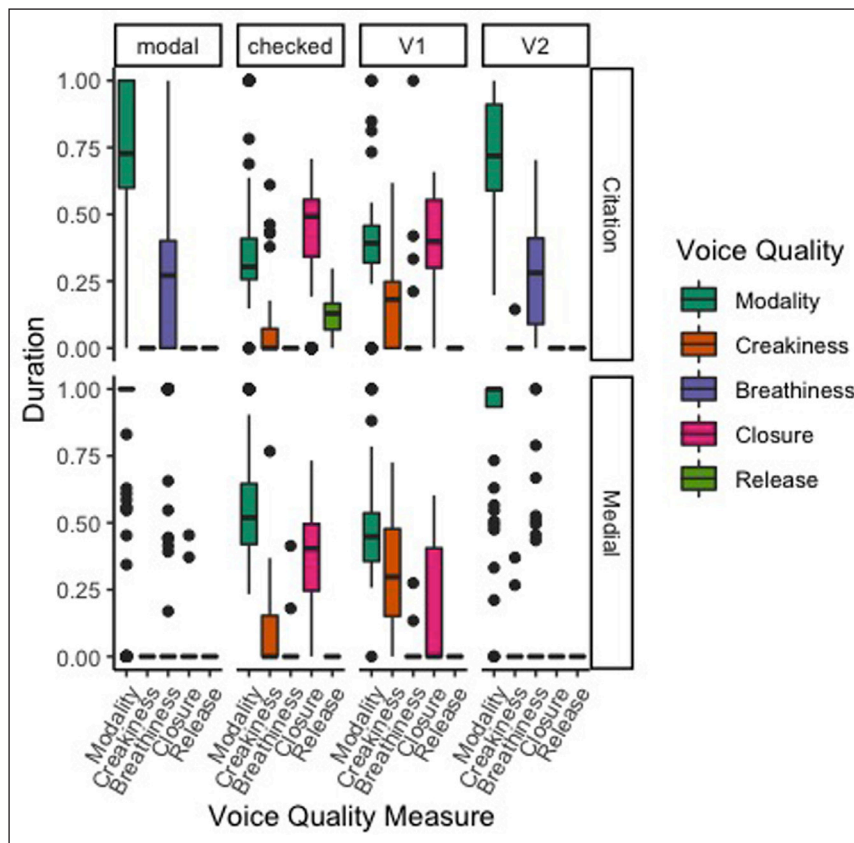


Figure 6 Proportional duration of each voice quality type by phonological vowel type and elicitation context.

vowels, as the citation form checked vowels are in final position and the V1 vowels are not; they are always followed by the V2 portion of the rearticulated vowel. One difference between V1 and checked vowels is that V1 tends to be produced with a greater proportional duration of creakiness, regardless of production context.

On the other hand, V2 vowels tend to behave like modal vowels in both contexts. In citation form, V2s are produced with modal voicing for an average of 75% of their duration, and then with breathy voicing for the remainder of the duration, closely mirroring the measurements for citation form modal vowels. In the phrase-medial context, both phonologically modal and V2 vowels are produced with even more modal voicing, 84% for V2 vowels and 92% for modal vowels on average, and a slightly shorter duration of breathy voicing.

As in section 5.1.1, statistical analyses were conducted using the mean proportional durations for different voice quality types. Individual mixed-effects linear regression models were fit using the `lmer` function in the `lme4` R package (Bates et al., 2015), each predicting the mean proportional duration of one of the voice quality measures. The results of these models are discussed here in turn.

5.2.1.1 Statistical analysis of modal voicing (modal, checked, and rearticulated vowels)

Table 19 presents the results of the regression model predicting the proportional duration of modal voicing for phonologically modal, checked, V1 and V2 vowels in the two production contexts. *Modal* was set as the reference level for phonological phonation type, and *Citation Form* was set as the reference level for context. *Speaker* was included as a random effect.

The results in **Table 19** show that whereas checked and V1 vowels are significantly different from modal vowels with respect to proportional duration of modal voicing ($p < 0.001$), there is no significant difference between modal vowels and V2 vowels. There was a significant main effect of context ($p < 0.001$), such that medial vowels were produced with a significantly longer proportional duration of modal voicing than vowels produced in citation form. A pairwise comparison reveals that the only pairs of vowels that did not show a significant difference in proportional duration of modality were those that compared either modal vowels and V2 vowels or checked vowels and V1 vowels. In other words, the statistics confirm here that with respect to proportional duration of modal voicing, V1 vowels are similar to checked vowels and V2 vowels are similar to phonologically modal vowels.

FIXED EFFECTS	ESTIMATE	STD. ERROR	T-VALUE	P-VALUE
(Intercept)	0.7473	0.0434	17.2348	0.0167 *
Type				
Checked	-0.3571	0.0296	-12.0545	<0.001 ***
V1	-0.3493	0.0392	-8.9039	<0.001 ***
V2	-0.0298	0.0392	-0.7584	0.4485
Context				
Phrase-Medial	0.1753	0.0244	7.1858	<0.001 ***
Type * Context				
Checked: Phrase-Medial	-0.0062	0.0410	-0.1516	0.8795
V1: Phrase-Medial	-0.0904	0.0541	-1.6715	0.0951
V2: Phrase-Medial	-0.04770	0.0541	-0.8818	0.3782

Table 19 Mixed-effects linear regression model: proportional duration of modal voicing.

5.2.1.2 Statistical analysis of creaky voicing (modal, checked, and rearticulated vowels)

Table 20 presents the results of the regression model predicting the proportional duration of creaky voicing for phonologically modal, checked, V1 and V2 vowels in the two production contexts. *Modal* was set as the reference level for phonological phonation type, and *Citation Form* was set as the reference level for context. *Speaker* and *Word* were included as random effects.

FIXED EFFECTS	ESTIMATE	STD. ERROR	T-VALUE	P-VALUE
(Intercept)	0.0001	0.0117	0.0101	0.099301
Type				
Checked	0.0536	0.0126	4.2585	<0.0001 ***
V1	0.1877	0.0168	11.1891	<0.001 ***
V2	0.0039	0.0168	0.2353	0.8141
Context				
Phrase-Medial	<0.001	0.0097	-0.0010	0.9992
Type * Context				
Checked: Phrase-Medial	0.02367	0.0164	1.4454	0.14881
V1: Phrase-Medial	0.1213	0.0216	5.6155	<0.001 ***
V2: Phrase-Medial	0.0107	0.0216	0.4944	0.6212

Table 20 Mixed-effects linear regression model: proportional duration of creaky voicing.

The results in **Table 20** show that while both checked and V1 vowels differ from modal vowels with respect to their proportional duration of creaky voicing ($p < 0.001$), there is no statistical difference here between modal vowels and V2 vowels. Unlike with modal voicing, there is no main effect of context on proportional duration of creaky voicing. However, there was a significant interaction between vowel type and context. Pairwise comparisons show no statistically significant difference between modal and V2 vowels in either context. However, the differences in proportional duration between checked and V1 vowels are significantly different in both contexts ($p < 0.001$). Therefore, while modal and V2 vowels are produced with a similar proportional duration of creaky voicing in both contexts, the difference in proportional duration of creaky voicing between checked vowels and V1 vowels is significant across contexts.

5.2.1.3 Statistical analysis of breathy voicing (modal, checked, and rearticulated vowels)

Table 21 presents the results of the regression model predicting the proportional duration of breathy voicing for phonologically modal, checked, V1 and V2 vowels in the two production

contexts. *Modal* was set as the reference level for phonological phonation type, and *Citation Form* was set as the reference level for context. *Speaker* and *Word* were included as random effects.

FIXED EFFECTS	ESTIMATE	STD. ERROR	T-VALUE	P-VALUE
(Intercept)	0.2460	0.0259	9.5143	0.01060 *
Type				
Checked	-0.2457	0.0259	-9.4806	<0.001 ***
V1	-0.2017	0.0348	-5.7978	<0.001 ***
V2	0.0290	0.0348	0.8323	0.4059
Context				
Phrase-Medial	-0.1828	0.0187	-9.7556	<0.001 ***
Type * Context				
Checked: Phrase-Medial	0.1879	0.0315	5.9598	<0.001 ***
V1: Phrase-Medial	0.1450	0.0416	3.4873	<0.001 ***
V2: Phrase-Medial	0.0281	0.0416	0.6767	0.4988

Table 21 Mixed-effects linear regression model: proportional duration of breathy voicing.

As with previous voice quality measures, the model in [Table 21](#) shows significant differences between modal vowels and both checked and V1 vowels ($p < 0.001$), but no significant difference between modal and V2 vowels ($p = 0.4058$). Here, again, there is a significant main effect of context ($p < 0.001$), such that medial vowels are significantly likely to be produced with less breathy voicing than those in citation form. Pairwise comparisons of the interaction between vowel types show no significant difference between modal vowels and V2 vowels in either utterance context. Similarly, there was no significant difference in proportional duration of breathy voicing between V1 vowels and checked vowels in either context. Therefore, results from this model further confirm that V2 vowels are produced similarly to modal vowels and V1 vowels are most similar to checked vowels.

5.2.1.4 Statistical analysis of duration of glottal closure (modal, checked, and rearticulated vowels)

[Table 22](#) presents the results of the regression model predicting the proportional duration of glottal closure for phonologically modal, checked, V1 and V2 vowels in the two production contexts. *Modal* was set as the reference level for phonological phonation type, and *Citation Form* was set as the reference level for context. *Speaker* and *Word* were included as random effects.

FIXED EFFECTS	ESTIMATE	STD. ERROR	T-VALUE	P-VALUE
(Intercept)	-0.0004	0.0258	-0.0158	0.9895
Type				
Checked	0.4080	0.0169	24.0922	<0.001 ***
V2	0.3654	0.0225	16.2295	<0.001 ***
V2	0.0034	0.0225	0.1524	0.8789
Context				
Phrase-Medial	0.0020	0.0135	0.1491	0.8815
Type * Context				
Checked: Phrase-Medial	-0.0522	0.0227	-2.2976	0.0219 *
V1: Phrase-Medial	-0.1653	0.0300	-5.5144	<0.001 ***
V2: Phrase-Medial	-0.0043	0.0300	-0.1418	0.8873

Table 22 Mixed-effects linear regression model: proportional duration of glottal closure.

Similarly to previous models, the results from this model show that there is a significant difference in proportional duration of the glottal closure between modal vowels and both checked and V1 vowels ($p < 0.001$), but no significant difference between modal and V2 vowels ($p = 0.8789$). While the main effect of context was not significant ($p = 0.8815$), the interaction between vowel type and context did achieve significance ($p < 0.001$). Pairwise comparisons show no significant difference between modal vowels and V2 vowels in either context; as supported by the previous models, it seems that these two vowel types behave relatively similarly in both contexts. While there was no significant difference between V1 vowels and checked vowels in citation form, the two vowels did differ significantly with respect to their proportional duration of the glottal closure in medial position.

5.2.1.5 Statistical analysis of duration of glottal release (modal, checked, and rearticulated vowels)

Table 23 presents the results of the regression model predicting the proportional duration of glottal release for phonologically modal, checked, V1 and V2 vowels in the two production contexts. *Modal* was set as the reference level for phonological phonation type, and *Citation Form* was set as the reference level for context. *Speaker* was included as a random effect.

FIXED EFFECTS	ESTIMATE	STD. ERROR	T-VALUE	P-VALUE
(Intercept)	-0.0001	0.0073	-0.0166	0.9892
Type				
Checked	0.1174	0.0035	33.3302	<0.001 ***
V1	0.0006	0.0047	0.1291	0.8973
V2	0.0006	0.0047	0.1291	0.8973
Context				
Medial	<0.001	0.0029	0.0032	0.9975
Type * Context				
Checked:Medial	-0.1173	0.0049	-24.0557	<0.001 ***
V1:Medial	-0.0007	0.0064	-0.1209	0.9038
V2:Medial	-0.0007	0.0064	-0.1209	0.9038

Table 23 Mixed-effects linear regression model: proportional duration of glottal release.

Unlike with previous voice quality measures, the results in **Table 23** show a significant main effect of phonological vowel type only between modal and checked vowels. There was no significant main effect of context ($p = 0.9974$), but the interaction between phonological vowel type and context did achieve significance ($p < 0.001$). Pairwise comparisons here showed significant differences only between checked vowels in citation form and all other vowel/context combinations. In other words, checked vowels in citation form had a significantly longer proportional duration of a glottal release than any other vowel produced in any other context.

5.2.1.6 Summary of statistical analyses of voice quality

To summarize the voice quality results across all four phonological vowel types, there is strong statistical evidence that modal vowels and V2 portions of rearticulated vowels are produced similarly with respect to all voice quality measures in all contexts. While V1 vowels and checked vowels also show many such similarities, they differ with respect to some voice quality measures in medial form, specifically the proportional duration of a glottal closure and release burst. These differences are likely due to the fact that the checked vowels in this data set are all word-final, but V1 vowels are inherently non-final in all cases, as they are always followed by the corresponding V2 portion of the rearticulated vowel.

5.2.2 Statistical analysis of F0 (modal, checked, and rearticulated vowels)

Table 24 shows the number of tokens in each F0 trajectory category by phonological vowel types and elicitation context. The data for modal and checked vowels is the same as previously reported in **Table 13**; data from the V1 and V2 portions of phonologically rearticulated vowels is added here for comparison.

		F0 TRAJECTORY		
CONTEXT	VOWEL TYPE	CHANGE	NO CHANGE	OTHER
citation	modal	118/163 (73%)	44/163 (27%)	1/163 (0.6%)
	checked	9/87 (10%)	65/87 (75%)	13/87 (15%)
	V1	13/43 (30%)	28/43 (65%)	2/43 (5%)
	V2	24/34 (71%)	9/34 (26%)	1/34 (3%)
medial	modal	3/164 (2%)	160/164 (98%)	1/164 (0.6%)
	checked	3/88 (3%)	73/88 (82%)	12/88 (14%)
	V1	16/48 (33%)	28/48 (58%)	4/48 (8%)
	V2	2/35 (6%)	32/35 (91%)	1/35 (3%)

Table 24 Number of tokens with each F0 trajectory by phonological vowel types and elicitation context.

Table 24 shows the number of tokens for each vowel type and elicitation context that exhibited each type of F0 trajectory. Though the overall number of tokens coded as “other” is small, the checked vowels had more of these tokens than any of the other vowels. As reported above, we assume that this pattern is due to the fact that the onset of the glottalization required to create a checked vowel (that is, one that ends with a glottal closure) sometimes interferes with the periodicity required for pitch tracking. The same data are visualized in Figure 7 with the “other” tokens omitted.

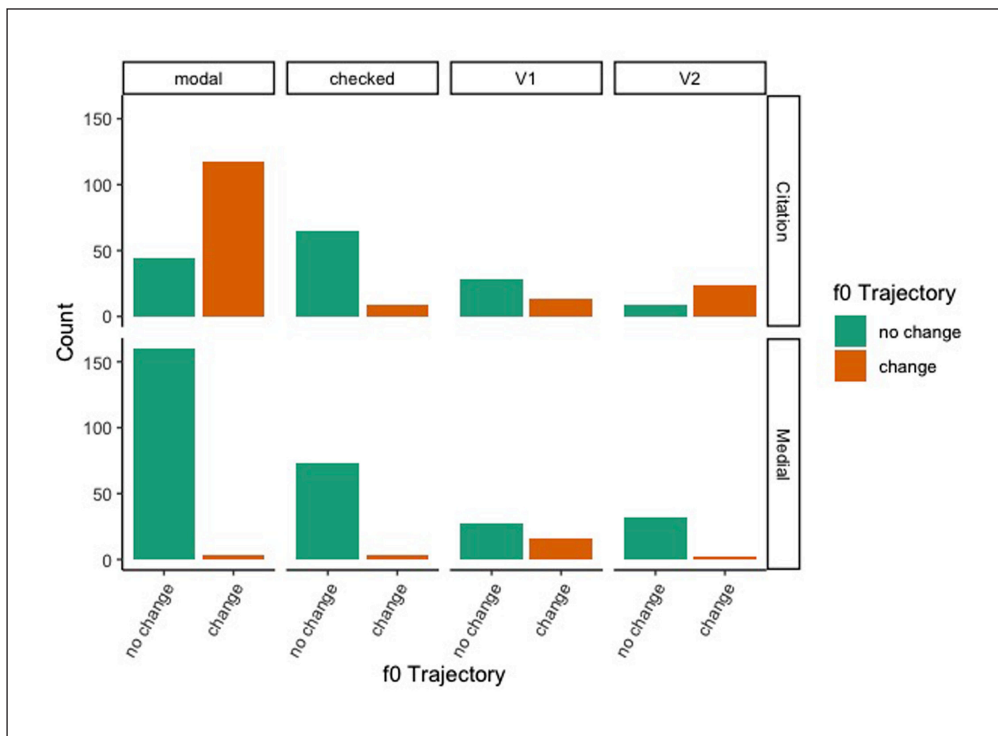


Figure 7 Counts of F0 trajectory types by phonological vowel type and elicitation context.

Figure 7 shows that modal vowels in citation form had more than twice as many tokens with a change in F0 trajectory than it did tokens without one. Though there were fewer tokens of rearticulated vowels, the V2 portions of these vowels in citation form show a similar pattern; more than twice as many tokens with F0 changes than without. For checked and V1 vowels in all contexts, and for modal and V2 vowels in medial contexts, a majority of the tokens were produced with no change in the F0 trajectory.

A mixed-effects logistic regression model was fit using the `glmer` function in the `lme4` R package (Bates et al., 2015) to predict F0 trajectory. We excluded the few “other” tokens from

the statistical analysis so that this unreliable data would not interfere with the main analysis of F0 behavior. *Table 25* presents the results of the regression model predicting “change” or “no change” in F0 for phonologically modal and checked vowels in the two utterance contexts. *Modal* was set as the reference level for phonological phonation type, and *Citation Form* was set as the reference level for context. *Speaker* was included as a random effect.

FIXED EFFECTS	ESTIMATE	STD. ERROR	T-VALUE	P-VALUE
(Intercept)	1.0777	0.5023	2.1453	0.0319 *
Type				
Checked	-3.1643	0.4171	-7.5861	<0.001 ***
V1	-1.9729	0.4036	-4.8879	<0.001 ***
V2	-0.0180	0.4467	-0.0402	0.9679
Context				
Medial	-5.2556	0.6230	-8.4357	<0.001 ***
Type * Context				
Checke:Medial	3.9601	0.9315	4.2512	<0.001 ***
V1:Medial	5.5009	0.7882	6.9790	<0.001 ***
V2:Medial	1.2062	1.0401	1.1597	0.2462

Table 25 Logistic regression model: F0 trajectory.

The results show a significant main effect of vowel type ($p < 0.001$), such that checked vowels and V1 vowels were significantly less likely to exhibit a change in F0 than modal vowels; the difference between V2 and modal vowels was not significant ($p = 0.9680$). There was also a main effect of context; vowels produced in utterance-medial contexts were significantly less likely to be produced with a change in F0 trajectory than vowels that were produced in citation form.

Pairwise comparisons of the interaction between vowel type and context show no significant difference between modal vowels and V2 vowels in either context, supporting the notion that modal and V2 vowels are produced similarly with respect to F0 trajectory. Similarly, there was no significant difference between checked vowels and V1 vowels in citation form. However, in medial contexts, the difference between checked vowels and V1 vowels was significantly different ($p = 0.0013$), with V1 vowels more likely to show a change in F0 trajectory than checked vowels.

5.3 OVERALL SUMMARY OF RESULTS

The phonetic characteristics of the MacZ phonation contrast presented in this section show that checked vowels are frequently produced with a glottal closure, while modal vowels are frequently produced without such a closure. In the citation form context, Speaker 1 produced 70% of the checked vowel tokens with a glottal closure, and Speaker 2 produced 98% of the checked vowel tokens with a glottal closure. In the phrase-medial context, Speaker 1 produced 75% of the checked vowel tokens with a glottal closure, and Speaker 2 produced 96% of the checked vowel tokens with a glottal closure. Modal vowels were almost never produced with a glottal closure for either speaker in either context. A glottal closure, or lack thereof, was thus the most consistent cue to the phonation contrast across the data examined, which is in line with the way the contrast between modal and checked vowels is generally described for Zapotec languages (see section 3.2).

However, the realization of additional phonetic cues (beyond the existence or non-existence of a glottal closure) differed depending on the speech context (citation form words in isolation versus words produced within a carrier phrase). *Table 26* summarizes the results presented in this section with regards to the voice quality characteristics of phonologically modal and checked vowels as well as the V1 and V2 portions of rearticulated vowels in both the citation form and phrase-medial speech contexts.

For the citation form data, it is clear that phonologically modal vowels are likely to surface as modal followed by a period of breathiness, as was found for Betaza Zapotec by Crowhurst, Kelly,

	MODAL	V2 REARTICULATED	CHECKED	V1 REARTICULATED
Citation form	Vowels tend to be produced with some final breathiness, and often with a rapid change in F0 towards the final portion of the vowel.		Most vowels are produced with a closure and audible release; sometimes creakiness is also present.	Similar to checked vowels, but with less likelihood of a closure and no audible release.
Phrase-medial	Vowels may be produced with some breathiness and occasional F0 fluctuation, but these properties are significantly less frequent than they were in the citation form context.		No audible releases when glottal closures are present; creakiness is sometimes present.	Similar to checked vowels, but tend to have a longer period of creakiness and are less likely to have a glottal closure.

Table 26 Overall summary of results.

and Teodocio (2016). We additionally found that modal vowels tend to exhibit a fluctuation in F0 that is independent of final breathiness. Phonologically checked vowels exhibit phonetic glottalization, usually in the form of a full glottal closure with an audible and visible release burst. Modal vowels exhibited a significantly greater proportional duration of modal voicing than creaky vowels regardless of production context and a significantly greater proportional duration of breathy voicing than creaky vowels in citation form. Checked vowels exhibited a significantly greater proportional duration of glottal closure and burst than modal vowels in citation form, and a significantly greater proportional duration of creaky voicing than modal vowels in the phrase-medial context. This finding also mirrors that of Crowhurst, Kelly, and Teodocio (2016) for Betaza Zapotec, who found fewer full glottal closures and more creakiness in phrase-medial positions.

Overall, the phrase-medial data differs from the citation form data primarily in the fact that we see fewer phonetic cues to the contrast. Very few of the modal vowels exhibited breathiness or F0 fluctuations as they did in the citation form data. While phonologically checked vowels tend to surface in citation form contexts as modal throughout the majority of the vowel with a glottal closure in the final portion followed by an audible burst, we see no such bursts in the phrase-medial data.

The statistical analyses of phonetic cues to V1 and V2 vowels largely confirm the prediction that V1 vowels will show phonetic cues similar to those of checked vowels, and that that V2 vowels will show phonetic cues similar to those of modal vowels. This prediction, however, does not always hold for the V1 vowels, especially in citation form contexts. Phonologically checked vowels have a shorter period of creaky voicing than V1 vowels, and have a longer glottal closure and release on average than V1 vowels. We assume that this difference is due the fact that whereas V1 vowels are inherently non-word-final, as they are always followed by a corresponding V2, checked vowels here are always word-final and, in the citation form contexts, utterance-final.

6 DISCUSSION

The data presented in the previous section reveal that there are diverse phonetic cues to the phonation contrast in MacZ. We show that the nature and distribution of phonetic cues to the phonation contrast are different in the two contexts: citation form words in a word list versus the same words in a carrier phrase-medial position. In the citation form context, both voice quality and F0 provide phonetic cues, while fewer cues surface in the phrase-medial context. In this section, we focus primarily on the cues to the contrast between checked and modal vowels. We describe how the patterns that emerge from the citation form data are straightforwardly accounted for by contrast and enhancement theory (Hall 2011). We argue that the Functional Phonology approach suggested by Boersma (1998) as an alternative to phonetic enhancement theories does not fully capture the data presented here, as it is not clear how this approach can account for redundant phonetic cues along multiple, distinct acoustic dimensions. However, we suggest that incorporating principles from Functional Phonology (Boersma 1998) into Hall's (2011) contrast and enhancement theory can help to explain the variation in phonetic realizations based on phrase position, which is not accounted for by the contrast and enhancement approach alone. In particular, we suggest that a probabilistic approach to Optimality-theoretic constraint ranking as proposed by Boersma (1998) and others (e.g., Goldwater & Johnson 2003) can help to account for our overall results. We turn to the question of rearticulated vowels and their phonological status in section 6.5.

6.1 DEFINING THE PHONATION CONTRAST

Phonetically-driven approaches to phonology differ in terms of how they characterize the realization of redundant phonetic cues associated with a given phonological contrast like those examined in this paper. As discussed above in our review of literature on phonetic enhancement (Section 2), Hall’s (2011) contrast and enhancement theory proposed that active contrastive features, as identified through a hierarchy of ordered phonological features (Dresher 2009; see [Figure 1](#) in Section 2), are eligible for enhancement via any phonetic configuration that contributes to the perceptual distinctiveness of that feature. Boersma’s Functional Phonology (1998) argued against characterizing this process as one of phonetic enhancement, instead proposing that the set of phonetic cues that contribute to a contrast should be thought of as the gestural implementation of “Max” and “Min” Optimality-theoretic constraints on distinctive features. Boersma (1998) therefore rejects the notion of phonetic cues enhancing abstract phonological features (such as [back]), and instead proposes that all phonetic implementations correspond to perceptual features, which in turn correspond to a given acoustic-perceptual continuum (such as F2).

Both of these approaches assume that phonological contrasts rely on features whose elements should be maximally perceptually distinct. For Hall (2011), this is a binary [+/-] specification of traditional abstract phonological features. Under such an approach, we could assume that the phonation contrast in MacZ makes use of a [glottalized] feature, such as the *laringización* feature proposed by Arellanes Arellanes (2009), such that modal vowels are specified as [-glottalized] while checked vowels are specified as [+glottalized]. For Hall (2011), the glottalization distinction could be enhanced with a range of phonetic configurations, ranging from directly to indirectly related to the core acoustic and articulatory targets of the feature, as discussed in Section 2.

Boersma (1998) on the other hand, argues against the use of such binary features (p. 355), and prefers to understand features as corresponding to a continuous perceptual correlate (see also Boersma 2009). For example, the traditional [back] feature for vowels corresponds to the continuum of F2 values and the traditional [high] feature corresponds to the continuum of F1 values. Languages make use of these formant spaces in different ways through language-specific gestural implementations of these perceptual features. In the case of phonation contrasts, such an approach may be informed by existing proposals regarding voice quality continua, such as the proposal of Gordon and Ladefoged (2001), which associates breathy voice with a less constricted glottis (greater open quotient) and creaky voice with a more constricted glottis (smaller open quotient). Of course, for Boersma, the continuum should associate with a perceptual rather than articulatory feature. The open quotient gesture corresponds acoustically to the difference between the first two harmonics (H1 minus H2¹²) of the spectrum, with breathier sounds showing a greater difference between H1 and H2 than creakier sounds (e.g., Kuang 2017). This type of continuum is represented in [Figure 8](#).

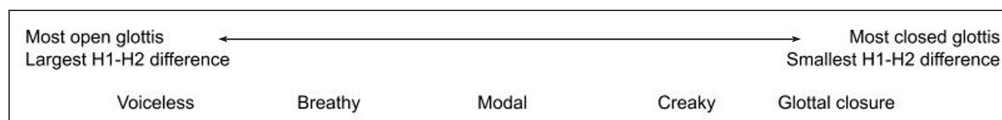


Figure 8 A simple articulatory and perceptual voice quality continuum.

Taking a Functional Phonology (Boersma, 1998) approach to analyzing the MacZ data, and drawing an analogy with how Boersma explains lip rounding of back vowels as part of the gestural implementation of MaxF2, we might wish to say that checked vowels are a gestural implementation of a “Min” constraint on the perceptual feature H1-H2, while modal vowels are the gestural implementation of a “Max” constraint on the same feature. Surface realizations of the former need not always be a complete glottal closure, and surface realizations of the latter need not always be voiceless; this variation can be captured by the interaction of the Min and Max constraints on H1-H2 with other markedness and faithfulness constraints. For example, there could be another constraint, Max-SNR, which prefers modal vowels that optimize signal-to-noise ratio over breathy or creaky voice, both of which can reduce recoverability for other important

¹² There are several possible phonetic correlates to these glottal configurations, including multiple spectral tilt measures, periodicity measures, etc. (e.g., Garellek 2015). We use H1-H2 as the perceptual cue to the glottal contrast as an illustrative example, but do not claim that this measure is more accurate or better-suited to the language-specific phonetics of this contrast than other possible cues.

contrasts such as tone and vowel quality. A high ranking for such a constraint would result in winning candidates having modal voicing, despite the fact that modal voicing is in the middle of the continuum in [Figure 8](#) and therefore not preferred by either the Min or Max H1-H2 constraints.

Having established how these two approaches might characterize the MacZ phonation contrast that is central to this paper, we move in the next few sections to discussing how they explain or do not explain the phonetic realizations revealed through our analysis of the MacZ data, addressing the advantages and disadvantages of each approach.

6.2 THEORETICAL ACCOUNTS OF THE CITATION FORM DATA

The phonetic characteristics of the MacZ phonation contrast for words in citation form are largely consistent with a contrast and enhancement approach (Hall 2011) in which each contrastive feature specification, in this case [+/-glottalized], is produced with additional phonetic cues that make the contrast sufficiently perceptible. Under Hall's (2011) contrast and enhancement approach, redundant phonetic cues may enhance the auditory impression of the [+/-glottalized] feature in multiple ways and along different acoustic dimensions. Some enhancements work to push a perceptual cue to the contrastive feature further along a given acoustic continuum, while others involve separate acoustic effects that increase the overall salience of the distinction between contrastive elements. Such an approach is helpful for explaining the breathiness and F0 cues present on the majority of the phonologically modal vowels in citation form. Most of the modal vowels for both speakers have a breathy voice cue in the final portion of the vowel or throughout the entire duration of the vowel, and over half of the tokens show a non-steady F0 trajectory.

The breathiness cue in this case could be interpreted under Hall's (2011) contrast and enhancement approach as a type of enhancement in which "a feature can be enhanced by the amplification of its articulatory and acoustic/auditory correlates" (p. 20). Analogous to producing a contrastively [-back] vowel in the front of the vocal tract rather than the center, breathy voice pushes the H1-H2 difference further to the left along the voice quality continuum presented in [Figure 8](#). A Functional Phonology approach (Boersma 1998) would account for the breathiness cue in a similar way. Analogous to the way lip rounding on back vowels emerges from Boersma's MaxF2 constraint, breathy voicing on modal vowels in MacZ could be seen to emerge from a MaxH1-H2 constraint that pushes the gestural implementation of modal vowels in citation form further to the left of the continuum presented in [Figure 8](#).

However, the F0 fluctuations cannot easily be understood as pushing the H1-H2 feature further along its perceptual continuum in the same way, as there is not a direct acoustic relationship between these F0 fluctuations and the H1-H2 difference. We therefore understand the F0 fluctuation to be an instance of one of the more indirect relationships Hall (2011) describes, in which "a feature with a particular acoustic/auditory correlate can be enhanced by a *separate* acoustic/auditory effect that increases the relative salience of that correlate" (p. 20; emphasis added).

This indirect relationship has to do with the fact that the [-glottalized] feature in our data is also enhanced by breathiness, as discussed above. It has been shown in previous work that the switch from modal to breathy voicing can result in F0 perturbations (Garellek & Keating 2011). We therefore argue that MacZ speakers produce changes in F0 to increase the relative salience of the breathy correlate to the [-glottalization] feature, one of the types of phonetic enhancement that Hall's (2011) contrast and enhancement approach accounts for.¹³ Importantly, given that our data show that the breathy voicing and F0 fluctuations do not always co-occur, our claim is that the F0 fluctuation is being used to optionally further enhance the phonation contrast, and not that it surfaces as an automatic mechanical byproduct of the breathy voicing enhancement

¹³ It is important to note that the F0 fluctuations observed do not seem to simply be an artifact of the transition from modal to breathy voicing, as there is not a one-to-one relationship between occurrences of the two cues. Of the 159 vowel tokens that exhibited breathy voicing after a period of modal voicing, only 105 (66%) also surfaced with a fluctuation in F0. Conversely, of the 186 tokens that surfaced with a fluctuation in F0, only 105 (56%) were also produced with a period of breathy voicing following the modal voicing. Additionally, in cases where both breathiness and an F0 fluctuation were present, it was not the case that the two occurred at the same time in the production of the vowel. In other words, though the onset of breathy voicing tends to lead to F0 fluctuations, it is not the case that every F0 fluctuation observed in our data was the result of a transition from modal to breathy voicing (e.g., [Figure 3](#)). For this reason, we see the F0 movements as an additional phonetic cue to modal phonation.

described above. Functional Phonology (Boersma 1998) offers no clear mechanism to account for phonetic cues like these F0 fluctuations, which are not an obvious gestural implementation of MaxH1-H2 (or any unidimensional perceptual continuum) and instead add salience to the overall phonation contrast by drawing on an additional acoustic dimension.

The presence of creaky voicing and glottal bursts on checked vowels in the citation form data is also straightforwardly captured by Hall's (2011) contrast and enhancement approach. Another of Hall's (2011) possible phonetic enhancements occurs when "a feature with a particular articulatory correlate can be enhanced by the amplification of a natural mechanical by-product of that gesture" (p. 20). Creaky voicing and glottal bursts are natural mechanical by-products of the glottal closure gesture (e.g., Chong & Garellek 2018), and creaky voicing in particular has been found to be a phonetic correlate of the [+glottalized] feature across Zapotec languages (e.g., Arellanes Arellanes 2010; Chávez Peón 2010; Esposito, 2010; Crowhurst, Kelly, & Teodocio 2016). Therefore this is a predictable enhancement of the [+glottalized] feature of checked vowels according to Hall's (2011) approach.

Here again, however, it is not clear how Functional Phonology (Boersma 1998) accounts for these additional cues. If we assume that checked vowels are a gestural implementation of the constraint MinH1-H2, the creaky voicing cue in fact moves us the wrong way along the continuum in [Figure 8](#), as creaky voicing actually creates a greater H1-H2 difference than a glottal closure. In addition, the glottal bursts which are present on so many of the checked vowels in citation form do not seem to have a place along the continuum in [Figure 8](#) whatsoever, as they have no direct relationship to the perceptual H1-H2 difference. The glottal closure that is at the right end of this continuum should be the most faithful implementation of a constraint like MinH1-H2, and indeed, a glottal closure is reliably produced in nearly all of the checked tokens in our data. However, the additional creaky voicing and burst cues are not easily accounted for following Boersma's (1998) approach.

In sum, our results suggest that the checked/modal phonation contrast in MacZ exhibits phonetic enhancement following the predictions put forth by the contrast and enhancement theory (Hall 2011). A contrast and enhancement approach predicts that phonetic enhancement may occur along multiple dimensions to increase the salience of the overall contrast, which is precisely what we see in our data. This occurs in the citation form data via breathiness and F0 fluctuations for [-glottalized] and glottal bursts and/or creakiness for [+glottalized]. Contrast and enhancement neatly accounts for these particular phonetic characteristics, because it explicitly predicts how phonetic enhancements will be related to the phonological feature active in the contrast. Phonetic enhancement of the [+/- glottalized] feature in MacZ does not just take advantage of one unidimensional percept to glottalization, the way that lip rounding enhances the continuum of the vowel backness (F2) percept. Rather, cues may be a natural mechanical byproduct of an associated gesture, or they may involve an auditory effect that increases the salience of a separate auditory cue.

It is not clear how Functional Phonology (Boersma 1998) accounts for additional phonetic cues beyond those associated with a single perceptual continuum that underlies the main contrast. A constraint like MinH1-H2 only predicts gestural implementations that have a direct effect on the H1-H2 difference. Boersma requires that contrasts be explained using distinctive perceptual features (like F2), rather than abstract features with articulatory labels (like [back]), so we could consider other perceptual correlates of glottalization such as low periodicity and low overall spectral tilt (e.g., Garellek 2015). Still, it is not clear how to combine these into a single faithfulness constraint that explains cues such as F0 fluctuations and glottal bursts. It seems that these are only captured by appealing to a more abstract phonological feature, such as [glottal], and allowing for more indirect enhancements to this feature.

6.3 THEORETICAL ACCOUNTS OF THE PHRASE-MEDIAL DATA

The most consistently occurring cue to checked phonation in MacZ, the glottal closure, is realized only on the final portion of the vowel after the tonal cue has been produced. This word-final placement of the cue could make the contrast particularly difficult to perceive in a phrase-medial context. According to the contrast and enhancement theory, this lack of perceptibility should trigger phonetic enhancements. However, the phonetic enhancements seen in the citation form data (breathiness, F0 fluctuations, creaky voicing, and glottal bursts) are not all

present in the phrase-medial data. While the contrast and enhancement theory (Hall, 2011) neatly accounts for the enhancements of the overall contrast between checked and modal vowels, it does not straightforwardly explain why we see more phonetic enhancement in our citation form data than in the phrase-medial data. Although Hall (2011) is explicit that “phonetic enhancement is variable across languages, speakers and contexts” (p. 19), he does not provide mechanisms to predict or account for this variability.

Functional Phonology (Boersma 1998), on the other hand, explicitly accounts for such optionality across different production contexts by allowing for a stochastic element in constraint evaluation (Chapter 15: 329–346). As Boersma (1998) explains:

The account of optionality presented here naturally encapsulates pragmatics-based reranking. For instance, if you want to speak more clearly, you may raise all your faithfulness constraints by, say, 5 along the continuous ranking scale. In this way, an 80%–20% preference for place assimilation will turn into a 18%–82% preference against. Depending on whether the faithfulness constraint is ranked above or below its rival, slight variation may turn into obligations or the reverse. If the ranking difference is large to begin with, however, nothing happens; so we see that discrete properties of surface rerankability are compatible with, and may well follow from, a general continuous rerankability of all constraints (p. 346).

To account for the variability in our results across production contexts, we can appeal to variable rankings of interacting faithfulness constraints as Boersma (1998) suggests. However, as we argued above, constraints like MinH1-H2 that correspond to a single perceptual continuum as Boersma suggests are not in line with the range of phonetic cues found in MacZ. Rather, our constraints must be something like Max[glottal] and Min[glottal], faithfulness constraints that are gradiently satisfied by candidates with any subset of phonetic cues to the abstract [glottal] feature. Constraints of this more abstract nature better capture the range of phonetic cues shown in our data (breathy voicing, F0 fluctuations, creaky voicing, glottal bursts). We propose that the pressure to maximize perceptual distinctiveness is stronger in the citation form context than in the phrase-medial context. On the other hand, we suggest that the pressure to minimize articulatory effort (e.g., Lindblom & Maddieson 1988; Boersma 1998; Flemming 2002; 2004) and to maintain the perceptibility of other contrasts, such as vowel quality and tone, are stronger in phrase-medial contexts than in citation form contexts, where hyperarticulation is common. In a Functional Phonology approach, which draws its basis from OT (Prince and Smolensky 1993), this interplay is represented through the interaction of faithfulness constraints, which represent the pressure to maximize perceptual distinctiveness, and markedness constraints, which represent the pressure to minimize articulatory effort.

We present an analysis of this type, adopting a maximum entropy (MaxEnt) approach (Goldwater and Johnson 2003; Wilson 2006; Hayes et al. 2009), a more recent probabilistic version of OT similar to Boersma’s (1998) Gradual Learning Algorithm. In a MaxEnt model, constraints are weighted for their relative importance rather than strictly ranked with respect to each other. Constraint violations are multiplied by the weight of that constraint, and the winning candidate is that which incurs the lowest weighted sum of violations. Unlike the Gradual Learning Algorithm, MaxEnt requires only one parameter to set, and is more generalizable to processes of learning in other domains (Goldwater and Johnson, 2003). For these reasons, we use MaxEnt as a tool to present an analysis in which the context-dependent phonetic enhancement patterns present in the data emerge from the relative weighting of perceptual faithfulness and markedness constraints. However, we do not claim that other stochastic grammar modeling methods are necessarily less suited to the data we present.

Tables 27 and **28** show how Max[glottal] and Min[glottal] constraints interact with other markedness and faithfulness constraints to account for the MacZ data. **Table 27** captures the citation form data, in which checked vowels are produced as creaky with a final closure and burst, and phonologically modal vowels are produced with modal voicing throughout and a fluctuation in F0.¹⁴ Here, the markedness constraint *Effort, which militates against candidates

¹⁴ Though MaxEnt was developed in part to model surface variation in speech, we assume that the winning candidates in these tableaux are ideal productions, and leave a more thorough analysis of the free variation seen in our data to future work.

that require articulatory effort, has a weight of 0; the winning candidate may be articulatory effortful without incurring any violations. The Max[glottal] and Min[glottal] constraints interact with Max-SNR, a markedness constraint that prefers candidates produced with modal voicing, as these increase the perceptibility of contrasts related to vowel quality and contrastive tone. Though candidate b. as the surface form of an input checked vowel incurs a violation of the relatively high-weighted Max-SNR constraint, it receives fewer total violation points than candidate a., which also violates Max[glottal], and c., which incurs no violations of Max-SNR but violates Max[glottal] twice. When the input vowel is phonologically modal, candidates that are produced with non-modal incur a violation of Max-SNR. Candidate d., which does not violate Max-SNR, incurs more violations of Min[glottal] than the optimal candidate, as it is produced neither with breathy voicing, which farthest to the left of the glottal continuum, nor the F0 fluctuation which we argue phonetically enhances the [-glottal] feature specification.

	MAX[GLOTTAL] /CHECKED/	MAX-SNR	MIN[GLOTTAL] /MODAL/	*EFFORT	
	12.1	15.2	7.5	0	
<i>/checked/</i>					
a. creaky voicing, no burst	1	1	0	0	27.3
☞ b. creaky voicing, glottal burst	0	1	0	1	15.2
c. modal throughout	2	0	0	1	24.2
<i>/modal/</i>					
d. modal throughout	0	0	2	1	15
e. creaky voicing	0	1	2	0	30.2
f. breathy voicing	0	1	1	0	22.7
☞ g. modal throughout, F0 fluctuation	0	0	1	2	7.5
h. breathy voicing, F0 fluctuation	0	1	0	1	15.2

Table 27 A MaxEnt model of the citation form data.

	MAX[GLOTTAL] /CHECKED/	MAX-SNR	MIN[GLOTTAL] /MODAL/	*EFFORT	
	13.7	27.7	0	20.6	
<i>/checked/</i>					
☞ a. creaky voicing, no burst	1	1	0	0	41.1
b. creaky voicing, glottal burst	0	1	0	1	48.3
c. modal throughout	2	0	0	1	48
<i>/modal/</i>					
☞ d. modal throughout	0	0	2	1	20.6
e. creaky voicing	0	1	2	0	27.7
f. breathy voicing	0	1	1	0	27.7
g. modal throughout, F0 fluctuation	0	0	1	2	41.2
h. breathy voicing, F0 fluctuation	0	1	0	1	48.3

Table 28 A MaxEnt model of the phrase-medial form data.

The differences in phonetic cues to the glottalization contrast across contexts is modeled here as a simple reweighting of the active constraints. Unlike in citation form (*Table 27*), *Effort in the modeling of phrase-medial data (*Table 28*) has a relatively high weighting, preventing articulatorily effortful candidates from surfacing. This high weighting captures the generalization that in medial form, articulatorily difficult phonetic enhancements do not surface, and modal voicing is more prominent. On the other hand, the weight of Min[glottal] in this context is 0, far outweighed here by Max-SNR. As a result of this relative weighting, optimal candidates are those that are produced with modal voicing, and therefore that allow for perceptibility of vowel quality and tone contrasts, at the expense of minimizing or maximizing the [glottal] percept.

We would emphasize, though, that theories which explain surface representations primarily through tradeoffs between maximizing perceptual distinctiveness and minimizing articulatory effort (e.g., Boersma; 1998) do not on their own offer the same explanatory power that a contrast and enhancement approach (Hall; 2011) offers for the data presented here. In a system that highly prioritizes the pressure to maximize perceptual distinctiveness, as we see for our citation form data, contrast and enhancement helps to predict what types of phonetic correlates we would expect to see (articulatory and perceptual cues directly and indirectly associated with enhancing the glottalization contrast) and not expect to see (random ways of adding in phonetic distinctiveness). We therefore see the principle of minimizing articulatory effort as working in tandem with the contrast and enhancement approach. Because the type of context predicts the extent to which articulatory ease may be prioritized by the speaker, it therefore also both constrains and predicts the types of phonetic enhancements of phonological contrasts that are likely to surface.

6.4 F0 AS A CUE TO CONTRASTIVE PHONATION IN A LARYNGEALLY COMPLEX LANGUAGE

Though the F0 fluctuation as a phonetic enhancement to a [glottal] contrast can be easily captured by theories such as Hall's (2011), we note here that it is somewhat surprising in the context of MacZ's phonological system. As shown in Section 3.1, MacZ exhibits both contrastive tone and contrastive phonation. It would be reasonable to imagine that the presence of two laryngeal contrasts in a given system would limit the ways in which each of these contrasts is enhanced. In other words, the existence of contrastive tone in this language could be expected to eliminate the possibility of an F0 enhancement of the phonation contrast. However, our data show that the presence of linguistic tone in this system does not preclude F0 from being employed as a phonetic enhancement to the phonation contrast. Rather, the trajectory of F0 is a cue that enhances the phonation contrast while, apparently, not obscuring the perceptibility of the tone contrast, also cued by F0. We assume that this is made possible by the timing of laryngeal contrasts in MacZ, discussed above, in which contrastive tone is conveyed in the first part of the vowel and phonation contrasts only in the final portion. Though we leave a discussion of how temporal configurations fit into abstract phonological contrasts to future work, we note here that the F0 cue surfaces as a phonetic enhancement of the phonation contrast only after the tone gesture is complete.

The F0 cues to tone and phonation, respectively, are also perceptually distinguished by the nature of their pitch trajectories. The F0 cues to phonation appear to lack specific pitch targets and manifest primarily as a deviation from the pitch trajectory. The lack of an apparent pitch target, combined with the relatively short duration of the F0 fluctuations observed here, distinguish the cues to tone from the cues to phonation. We assume that this distinction makes for a situation in which F0 is able to simultaneously provide the primary cue to tonal contrasts and a phonetic enhancement of the phonation contrast. As a result, our data provide an instance of one laryngeal cue, F0, simultaneously contributing to two laryngeal contrasts, tone and phonation.

6.5 THE PHONOLOGICAL STATUS OF REARTICULATED VOWELS

The focus of this discussion has been the phonetic enhancement of the phonation contrast between checked ($V^?$) and modal (V) vowels in MacZ. A secondary goal of this paper was to consider phonetic evidence related to the phonological status of rearticulated vowels, and whether they are best analyzed as single vowels with a glottal interruption ($V^?V$) or as disyllabic sequences of a checked vowel followed by a modal vowel ($V^?.V$). A summary of the phonotactic arguments for and against each analysis was presented in Section 3.2. The current study presented an opportunity to add phonetic evidence to our understanding of the status of rearticulated vowels by comparing their acoustic characteristics to those of checked and modal vowels.

Under existing analyses of other Zapotec languages, rearticulated vowels are considered single vowels with a glottal interruption ($V^?V$) that differ from checked vowels primarily in that their [+glottalized] feature is anchored to the center portion of the vowel rather than the final portion of the vowel (Arellanes Arellanes 2014; López Nicolas 2014). Under such an analysis, the vowel

portion following the glottalization is simply an echo or continuation of the vowel preceding the glottalization, and it does not have the phonological status of a modal vowel. Therefore if this is the correct analysis, we might expect to see acoustic cues to glottalization similar to those of checked vowels near the central portion of the vowel rather than at the end, but we would not necessarily expect phonetic enhancements like those of modal vowels to occur in the final portion of the vowel.

In Section 3.2, we suggested another possible interpretation of these vowels based on the phonotactic characteristics of MacZ. What have been assumed to be individual vowels with a glottal closure produced in the middle ($V^?V$), based largely on the syllable structures of other Zapotec varieties, could in MacZ be better analyzed as a sequence of two vowels with the same quality, the first checked and the second phonologically modal ($V^?.V$). If this is the case, we would expect to see phonetic enhancements similar to those of checked vowels in the V1 portion of the vowel sequence and phonetic enhancements similar to those of modal vowels in the V2 portion of the sequence.

Our data showed that the V2 vowels showed the same phonetic enhancements as modal vowels (F0 fluctuations, breathiness) and patterned like modal vowels according to context (citation form versus phrase-medial). The V1 vowels showed phonetic enhancements similar to those of phrase-medial checked vowels, but with less likelihood of a full glottal closure and greater likelihood of creakiness. As discussed above, we posit that these differences may be due to the fact that these vowels necessarily have different positions in the word in our data, with our checked vowels always appearing word-finally and our V1 vowels necessarily appearing word-medially before the corresponding V2. In other words, the phonetic differences between V1 and checked vowels in our data do not preclude an analysis in which V1 vowels are in fact medial checked vowels. Though further evidence is needed to make a more conclusive claim about the phonological status of $V^?V$ sequences, our data lends phonetic support to the idea that initial portions of these sequences could in fact be checked vowels and the final portions of these sequences could in fact be modal vowels, thereby supporting an analysis of purported rearticulated vowels as disyllabic vowel sequences.

7 CONCLUSION

This paper has provided data illustrating the phonetic cues to a phonation contrast in Macuilianguis Zapotec, an under-documented variety of Sierra Juárez Zapotec. We analyze the phonetic patterning using Hall's (2011) contrast and enhancement theory of phonetic enhancements, incorporating principles of maximizing and minimizing perceptual features from Boersma (1998, 2008). We show that the phonetic enhancements of the contrast can be modeled as resulting from the relative weights of Max/Min[glottal] constraints and other markedness and faithfulness constraints. The context-dependent variation in cues is modeled as a reweighting of the relevant constraints in different contexts. To our knowledge, ours is the first analysis that uses a phonetic enhancement approach to account for a phonation contrast.

We also make two smaller contributions that should be further explored in future work. First, the data suggest that F0 can provide a phonetic cue to two laryngeal contrasts, phonation and tone, simultaneously. While many studies have analyzed voice quality as an additional cue to tone contrast (e.g., Morén & Zsiga 2006; Nguyen & Macken 2008; Yu and Lam 2014; Uchihara 2016; Kuang 2017), we are aware of no other work suggesting that F0 provides an additional cue to contrastive phonation. Laryngeally complex languages like MacZ offer the promise of further exploration of this finding. We also address the phonological status of $V^?V$ sequences, showing that our data support a possible phonological analysis in which these do not form a third member of the [glottal] contrast, as traditionally analyzed for MacZ and other Zapotec languages, but rather are sequences of a checked vowel followed by a modal vowel. We note that future work should further investigate the phonetic and phonological patterning of these vowels to clarify their phonological status.

ABBREVIATIONS

MacZ San Pablo Macuilianguis Zapotec

The additional file for this article can be found as follows:

- **Appendix A.** Word List. DOI: <https://doi.org/10.5334/gjgl.959.s1>

ACKNOWLEDGEMENTS

The MacZ speakers we recorded for this study are part of the Grupo Cultural Tagayu' and are dedicated language activists in their communities; we are grateful for their participation and acknowledge their work *Xarutekuali*. Thank you to the editors and anonymous reviewers for your comments, which significantly strengthened the final version of this paper. Thanks also to the audiences at the 42nd Annual PLC, the 26th mfm, and LabPhon 17 for their feedback on previous versions of this work.

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR AFFILIATIONS

Maya L. Barzilai

Georgetown University, US

Katherine J. Riestenberg  orcid.org/0000-0003-4865-9237

Haverford College, US

REFERENCES

- Arellanes Arellanes, F. 2010. Dos “grados” de laringización con pertinencia fonológica en el zapoteco de San Pablo Güilá. In E. Herrera Zendejas (ed.), *Entre cuerdas y velo: Estudios fonológicos de lenguas otomangués*, 85–122. Mexico City: Colegio de México. DOI: <https://doi.org/10.2307/j.ctv6jmx5b.7>
- Arellanes Arellanes, F. 2014. El anclaje temporal de los rasgos laringeos en el zapoteco de San Pablo Güilá y una nueva escala de laringización: Análisis bajo el marco de la Teoría de la Optimidad. In R. Gutiérrez Bravo, F. Arellanes Arellanes and Peónio Peón Peón (eds.), *Nuevos estudios de teoría de la optimidad*. Mexico City: Colegio de México.
- Avelino Becerra, H. 2004. Topics in Yalalág Zapotec with particular reference to its phonetics. University of California Los Angeles, Ph.D. Dissertation.
- Bates, D, M. Mächler, B. Bolker and S. Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67(1), 1–48. DOI: <https://doi.org/10.18637/jss.v067.i01>
- Boersma, P. 1998. *Functional phonology: Formalizing the interactions between articulatory and perceptual drives*. The Hague: Holland Academic Graphics.
- Boersma, P. 2009. Cue constraints and their interactions in phonological perception and production. *Phonology in perception* 15. 55–110.
- Boersma, P. and D. Weenink. 2018. Praat: Doing phonetics by computer. Retrieved 15 August 2018 from <http://www.praat.org/>.
- Campbell, Eric. 2017. Otomanguan historical linguistics: Exploring the subgroups. *Language and Linguistics Compass* 11(4). 1–23. DOI: <https://doi.org/10.1111/lnc3.12240>
- Chávez-Peón, M. E. 2010. The interaction of metrical structure, tone, and phonation types in Quiavini Zapotec. University of British Columbia, Ph.D. Dissertation.
- Chong, A. J. and M. Garellek. 2018. Online perception of glottalized coda stops in American English. *Laboratory Phonology* 9(1). 1–24. DOI: <https://doi.org/10.1109/TMTT.2014.2364584>
- Crowhurst, M. J., N. E. Kelly and A. Teodocio. 2016. The influence of vowel laryngealisation and duration on the rhythmic grouping preferences of Zapotec speakers. *Journal of Phonetics* 58. 48–70. DOI: <https://doi.org/10.1016/j.wocn.2016.06.001>
- Denes, P. 1955. Effect of duration on the perception of voicing. *The Journal of the Acoustical Society of America* 27(4). 761–764. DOI: <https://doi.org/10.1121/1.1908020>
- Dreher, J. J. and P. C. E. Lee. 1968. Instrumental investigation of single and paired Mandarin tonemes. *Monumenta serica* 27(1). 343–373. DOI: <https://doi.org/10.1080/02549948.1968.11731059>
- Dresher, B. Elan. 2009. *The contrastive hierarchy in phonology*. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511642005>
- Esposito, C. M. 2010. Variation in contrastive phonation in Santa Ana Del Valle Zapotec. *Journal of the International Phonetic Association* 40(2). 181–198. DOI: <https://doi.org/10.1017/S0025100310000046>

- Flemming, E. 2002. *Auditory representations in phonology*. New York: Routledge.
- Flemming, E. 2004. Contrast and perceptual distinctiveness. In B. Hayes, R. Kirchner and D. Steriade (eds.), *Phonetically based phonology*, 232–276. Cambridge, UK: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511486401.008>
- Foreman, J. 2006. The morphosyntax of subjects in Macuilianguis Zapotec. University of California Los Angeles. Ph.D. Dissertation.
- Gandour, J. 1977. On the interaction between tone and vowel length: Evidence from Thai dialects. *Phonetica* 34(1). 54–65. DOI: <https://doi.org/10.1159/000259869>
- Garellek M. 2015. Perception of glottalization and phrase-final creak. *The Journal of the Acoustical Society of America* 137(2). 822–831. DOI: <https://doi.org/10.1121/1.4906155>
- Garellek, M. and P. Keating. 2011. The acoustic consequences of phonation and tone interactions in Jalapa Mazatec. *Journal of the International Phonetic Association* 41(2). 185–205. DOI: <https://doi.org/10.1017/S0025100311000193>
- Goldwater, S. and M. Johnson. 2003, April. Learning OT constraint rankings using a maximum entropy model. In *Proceedings of the Stockholm workshop on variation within Optimality Theory* (Vol. 111120).
- Gordon, M. and P. Ladefoged. 2001. Phonation types: A cross-linguistic overview. *Journal of Phonetics* 29. 386–406. DOI: <https://doi.org/10.1006/jpho.2001.0147>
- Hall, D. C. 2011. Phonological contrast and its phonetic enhancement: Dispersedness without dispersion. *Phonology* 28(1). 1–54. DOI: <https://doi.org/10.1017/S0952675711000029>
- Hayes, B., C. Wilson and B. George. 2009. Manual for Maxent grammar tool. <http://linguistics.ucla.edu/people/hayes/MaxentGrammarTool/ManualForMaxentGrammarTool.pdf>.
- INALI. 2008. Catálogo de las lenguas indígenas nacionales: Variantes lingüísticas de México con sus autodenominaciones y referencias geoestadísticas. Retrieved from http://www.inali.gob.mx/pdf/CLIN_completo.pdf.
- Keyser, S. J. and K. N. Stevens. 2006. Enhancement and overlap in the speech chain. *Language* 82(1). 33–63. DOI: <https://doi.org/10.1353/lan.2006.0051>
- Kingston, J. 2011. Tonogenesis. In M. van Oostendorp, C. J. Ewen, E. Hume and K. Rice (eds.), *The Blackwell Companion to Phonology* 4. 2304–2334. Oxford, UK: Blackwell. DOI: <https://doi.org/10.1002/9781444335262.wbctp0097>
- Kuang, J. 2017. Covariation between voice quality and pitch: Revisiting the case of Mandarin creaky voice. *The Journal of the Acoustical Society of America* 142(3). 1693–1706. DOI: <https://doi.org/10.1121/1.5003649>
- Liljencrants, J. and B. Lindblom. 1972. Numerical simulation of vowel quality systems: The role of perceptual contrast. *Linguistic Society of America* 48(4). 839–862. DOI: <https://doi.org/10.2307/411991>
- Lindblom, B. and I. Maddieson. 1988. Phonetic universals in consonant systems. In C. Li and L. M. Hyman (eds.), *Language, Speech and Mind*, 62–78. London: Routledge.
- Lisker, L. 1986. “Voicing” in English: A catalogue of acoustic features signaling /b/ versus /p/ in trochees. *Language and Speech* 29(1). 3–11. DOI: <https://doi.org/10.1177/002383098602900102>
- López Nicolas, O. 2014. El zapoteco de Zochina: Tópicos en fonología y morfosintaxis. CIESAS (Mexico City, Mexico). Ph.D. Dissertation.
- Marlett, S. A. and V. B. Pickett. 1987. The syllable structure and aspect morphology of Isthmus Zapotec. *International Journal of American Linguistics* 53(4). 398–422. DOI: <https://doi.org/10.1086/466066>
- Morén, B. and E. C. Zsiga. 2006. The lexical and post-lexical phonology of Thai tones. *Natural Language and Linguistic Theory* 24. 113–178. DOI: <https://doi.org/10.1007/s11049-004-5454-y>
- Nguyen, H. T. and Macken, M. A. 2008. Factors affecting the production of Vietnamese tone: A study of American learners. *Studies in Second Language Acquisition* 30. 49–77. DOI: <https://doi.org/10.1017/S0272263108080030>
- Pérez Báez, G. 2015. Morphological valence-changing processes in Juchitán Zapotec. In N. Operstein and A. H. Sonnenschein (eds.), *Valence Changes in Zapotec. Synchrony, diachrony, typology*, 93–115. Philadelphia: John Benjamins. DOI: <https://doi.org/10.1075/tsl.110.06per>
- Pickett, V. B. 1953. Isthmus Zapotec Verb Analysis I. *International Journal of American Linguistics* 19(4). 292–296. DOI: <https://doi.org/10.1086/464235>
- Pickett, V. B. 1955. Isthmus Zapotec Verb Analysis II. *International Journal of American Linguistics* 21(3). 217–232. DOI: <https://doi.org/10.1086/464336>
- Prince, A. and P. Smolensky. 1993. Optimality Theory: Constraint interaction in generative grammar. *Optimality Theory in phonology* 3.
- Riestenberg, K. J. 2017. Acoustic salience and input frequency in L2 lexical tone learning: Evidence from a Zapotec revitalization program in San Pablo Macuilianguis. Georgetown University. Ph.D. Dissertation.
- Silverman, D. 1997. Laryngeal complexity in Otomanguean vowels. *Phonology* 14. 235–261. DOI: <https://doi.org/10.1017/S0952675797003412>
- Simons, G. F. and C. D. Fennig (eds.) 2018. *Ethnologue: Languages of the world*. Dallas, TX: SIL International. Retrieved from <http://www.ethnologue.com>.

- Smith-Stark, T. C. 2007. Algunas isoglosas zapotecas. In C. Buenrostro, J. J. Rendón, L. Valiñas, M. A. Vargas Monro, O. Schumann, S. Herrera Castro and Y. Lastra (eds.), *Clasificación de las lenguas indígenas de México: Memorias del III Coloquio Internacional de Lingüística Mauricio Swadesh*, 69–133. Mexico City: UNAM-INALI.
- Stevens, K. N. and S. J. Keyser. 1989. Primary features and their enhancement in consonants. *Language* 65(1). 81–106. DOI: <https://doi.org/10.2307/414843>
- Stevens, K. N., S. J. Keyser and H. Kawasaki. 1986. Toward a phonetic and phonological theory of redundant features. In J. S. Perkell and D. H. Klatt (eds.), *Invariance and variability in speech processes*, 426–449. Hillsdale, NJ: Lawrence Erlbaum.
- Tejada, L. 2012. Tone gestures and constraint interaction in Sierra Juarez Zapotec. University of Southern California. Ph.D. Dissertation.
- Uchihara, H. 2016. Tone and registrogenesis in Quiavini Zapotec. *Diachronica* 33(2). 220–254. DOI: <https://doi.org/10.1075/dia.33.2.03uch>
- Wilson, C. 2006. Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive science* 30(5). 945–982. DOI: https://doi.org/10.1207/s15516709cog0000_89
- Yu, K. M. and H. W. Lam. 2014. The role of creaky voice in Cantonese tonal perception. *The Journal of the Acoustical Society of America* 136(3). 1320–1333. DOI: <https://doi.org/10.1121/1.4887462>

TO CITE THIS ARTICLE:

Barzilai, Maya L. and Katherine J. Riestenberg. 2021. Context-dependent phonetic enhancement of a phonation contrast in San Pablo Macuilianguis Zapotec. *Glossa: a journal of general linguistics* 6(1): 18. 1–36. DOI: <https://doi.org/10.5334/gjgl.959>

Submitted: 03 April 2019

Accepted: 25 October 2020

Published: 09 February 2021

COPYRIGHT:

© 2021 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

Glossa: a journal of general linguistics is a peer-reviewed open access journal published by Ubiquity Press.