



Open Library of Humanities

## Bora's high vowels involve a two-way dental contrast, not a three-way backness contrast

**Jeff Mielke\***, Department of English, North Carolina State University, Raleigh, NC, USA, [jimielke@ncsu.edu](mailto:jimielke@ncsu.edu)

**Steve Parker**, Applied Linguistics Department, Dallas International University, Dallas, TX, USA, [steve\\_parker@diu.edu](mailto:steve_parker@diu.edu)

\*Corresponding author.

---

The Bora language of Peru has six phonemic vowels, conventionally transcribed as /i ε a o i̠ ω/. This inventory is noteworthy since it appears to exhibit a three-way backness distinction among three high unrounded vowels. In this paper we present audiovisual data which confirms two significant facts: (1) /ω/ is high, back, and unrounded; and (2) the only vowel produced with lip rounding is /o/. However, the experiment further reveals that the Bora segment traditionally written /i̠/ is actually not central but rather a possibly novel type of front vowel, articulated with substantial contact between the tongue and both the upper and lower teeth. We use the methodology of Iskarous (2010) to estimate vocal tract area functions for Bora vowels based on their observed formant frequencies and amplitudes. The results of this procedure support our claim that /i̠/ is front and /ω/ is back. Several phonological alternations also motivate this characterization. The opening of the mouth and dental contact visible in the production of /i̠/ are correlated with its acoustic measurements, albeit in a previously undocumented way. We conclude that Bora's /i̠/ should be reclassified as /i̠̚/ – front and dental.

---



## 1. Introduction

Bora is spoken by about 750 persons in the Amazon jungle of Peru and 100 in Colombia. Its phonemic vowels have been described as /i ε a o i u/ (Thiesen & Weber, 2012). A contrast between a central and a back vowel which are otherwise identical (/i/ vs. /u/) is theoretically significant since it implies that the binary feature [ $\pm$ back] is too weak to encode all phonological contrasts along the front/back dimension. A three-way backness contrast would thus challenge phonological models that limit vowel systems to only two degrees of backness, such as Duanmu (2016). Previously, Bora vowels were acoustically pinpointed with measurements of F1-F3 based on audio recordings, indicating that /i/ is indeed farther front than /u/ (Parker, 2001). Furthermore, all segments except /o/ are visibly articulated with unrounded lips. However, if /i i u/ are actually distinguished by a feature other than backness or rounding (as we argue here), that would also be a significant finding. Consequently, two important research questions are: (1) Can the main distinction between /i/ and /u/ in Bora be relegated entirely to a difference in lip configurations or some other feature rather than to a primary contrast of *central* vs. *back*? And (2) if so, what feature is this?

In this paper we discuss an audiovisual experiment focusing on the production of Bora's six phonemic vowels. Ten native speakers were recorded on location in Peru. To facilitate comparison of vowels, particularly lip positions, four brightly colored dots were placed on each speaker's lips, forming a diamond (shown below in **Figure 5**). We analyze the frequencies of F1-F3 as well as the distance between the horizontal and vertical lip dot positions. While the video data is designed to yield information about lip positions, it reveals a crucial difference involving the tongue as well. Specifically, we show that the vowel conventionally described as /i/ is actually distinguished from the other vowels by lingual-dental contact in the mid-sagittal plane, making it a dental and/or lateral vowel. It is therefore not central, but front. We also present evidence from productive morphophonemic alternations and static phonotactic constraints that /i/ patterns as [-back] rather than [+back]. We conclude that Bora's vowel inventory includes three high unrounded vowels, specified as follows: /u/ is back, /i/ is front, and "/i/" is both front and dental. (For the sake of simplicity, in most of this paper we transcribe the latter as /i/, even though our goal is to demonstrate that it is not actually central but front.) This reanalysis thus rejects Parker's (2001) argument that Bora's high vowels require a three-way distinction in the feature [back], a claim based on measurements of F1-F3 only. An additional contribution of this paper is that we examine acoustic features related to potential changes involving frication or vocal tract resonance as a result of this dental constriction, namely, harmonics-to-noise ratio and formant amplitude.

Our instrumental results show that /i/ and /u/ are produced with different lip positions in the vertical dimension only. This is consistent with the presumed difference in jaw height correlated with /u/ being a back vowel and /i/ actually being a front vowel. Lip rounding (indicated by a significantly smaller horizontal distance between the lateral lip dots) is observed only for /o/. This confirms the claim that both /i/ and /u/ are [-round].

Nevertheless, the lingual/dental contact observed in the video data also entails that Bora's "/i/" is articulated differently from what is typically implied by the symbol /i/ used to transcribe analogous vowels in most other languages. This highlights the fact that linguistic descriptions of vowel inventories vary in how much phonetic detail they include. For example, in some cases diacritics are added in order to indicate nuances such as greater lingual constriction than one might expect for an approximant. To illustrate, we briefly mention three other languages reported to have a segment somewhat similar to Bora's /i/. Yet in each of these systems the articulatory features of the sound in question are also distinct, in potentially subtle ways.

First, Urarina (ura) is a language isolate also spoken in Peru, about 250 miles southwest of the Bora area. Elias-Ulloa and Muñoz Aramburú (2021) describe it as having five basic vowels: /i e a i ʉ/. Each of these can also be contrastively lengthened or nasalized except /i/, which curiously is always short and oral. The relative placement of these phonemes in acoustic space is roughly comparable to their Bora counterparts (p. 17). Nevertheless, Elias-Ulloa and Muñoz Aramburú do not mention dental contact for any Urarina vowels, nor do their pictures on page 18 show that the tongue is visible in the articulation of /i/. It could be that the lack of mention of any dental contact in Urarina vowels is due to the authors not focusing on that detail in their description. Nevertheless, we might presume that if dental contact were observed, it would have been noted.

Second, in Banda-Linda (liy), a Ubangian language of the Central African Republic, there is more direct evidence for a canonical (non-dental) /i/. Olson (2019) analyzes this language as having nine vowels, /i ɨ ɯ u e ə o a ɔ/. He uses the diacritic "+" to indicate that the two high, central, unrounded segments exhibit higher F2 values than is typical for /i/ and /ɯ/ in other languages. Nevertheless, the native speaker of Banda-Linda recorded for that study reports that, upon introspection, none of his vowels involve apical contact with the teeth (Ken Olson, personal communication).

Finally, a so-called "apical vowel" has been described in Mandarin. However, this is an allophone of /i/ homorganic with a preceding dental sibilant. Furthermore, based on ultrasound and acoustic data, Lee-Kim (2014) categorizes this segment as a dental approximant ([i̠]), not a true vowel. In contrast to this, Bora's /i/ patterns as an inherently dental vowel, and it does so regardless of surrounding segments.

In summary, this paper has two main, interrelated goals: (1) To empirically establish a revised description of the Bora vowel inventory whereby /i/ is characterized articulatorily as a high, front, dental vowel (with presumed lateral airflow, since there is central contact), rather than as a high central vowel; and (2) to confirm that all Bora vowels except /o/ are [-round]. To our knowledge, every linguist who has analyzed Bora unanimously agrees that /i e a i ʉ/ are produced without lip rounding. However, this description is impressionistic only, based on visual and perceptual observations — watching the speakers' mouths as they talk, and hearing these

sounds pronounced. This study provides instrumental evidence demonstrating that this consensus is correct, and it does so in a way that is independently verifiable and replicable. Consequently, a novel contribution of this work is that we investigate a topic having both theoretical and typological consequences through a unique synergy of techniques – phonological, acoustic, visual, and articulatory modeling. The outcome shows that phonetic fieldwork and laboratory methods can combine in fruitful ways to answer what is basically a phonological question: How to accurately characterize the inventory of Bora vowels, as well as their patterns of behavior.

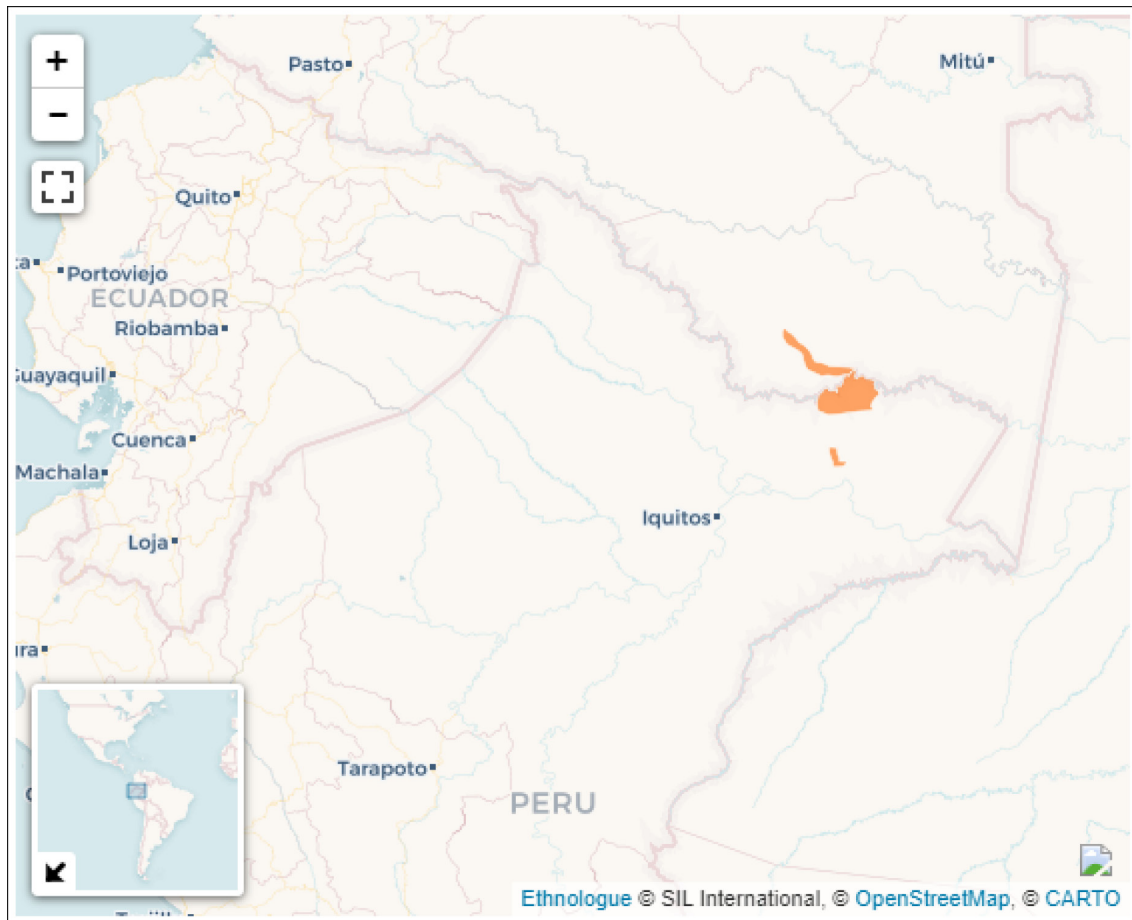
In §2 we give an overview of Bora's current sociolinguistic status and a summary of its phonological system, noting the contrastive segments and their phonotactic distribution. In §3 we describe the design of our audiovisual experiment. The results of that experiment are then presented in two major sections. First, in §4, we analyze lip configurations during the production of each vowel, as well as acoustic measurements that shed further light on their articulatory properties. Afterwards, in §5, we apply Iskarous' (2010) methodology to approximate the vocal tract shapes assumed to correlate with the tongue positions of Bora vowels. We conclude in §6 with a discussion of the theoretical implications of our revised characterization of Bora vowels.

## 2. Overview of the Bora language

### 2.1. Demographics

Since Bora is relatively understudied in the linguistic literature, in this section we provide a brief summary of its current situation. The Ethnologue classifies Bora as one of seven languages in the Witotoan family (Eberhard, Simons, & Fennig, 2017). Its ISO 639-3 code is *boa*. Alternative names for the language include *Miraña* and *Miamunaa* (the autonym). The only other member of the Bora-Muinane branch of Witotoan is *Muinane* (*bmr*), spoken in Colombia. It is closely related to Bora, but not mutually intelligible with it. The Ethnologue classification of Bora as Witotoan follows Aschmann (1993). However, Seifart and Echeverri (2015) consider *Muinane* to be the only language genealogically related to Bora, with an unclear relationship between those two and Witotoan languages.

Recent census data reports an estimated population of 3,000 ethnic Boras (Crevels, 2007, p. 116). Among these, about 750 persons in Peru and 100 in Colombia still speak the language. The varieties of Bora in the two countries are approximately 90–94% similar (Seifart, 2005; Thiesen & Weber, 2012). In this study we focus on Bora as spoken in Peru. There are also several hundred ethnic Boras in Brazil, but no known native speakers. In Peru, Bora is spoken primarily in five villages along the Ampiyacu, Yaguasyacu, and Momón Rivers in Loreto, the northeasternmost region, bordering Colombia and Brazil. The areas where Bora is spoken in Peru (and across the border in Colombia) are shaded in the map in **Figure 1**. Glottolog lists the coordinates for this language as 2°S, 72°15'W, and assigns it the code *bora1263* (Hammarström, Forkel, Haspelmath, & Bank, 2015).



**Figure 1:** A map showing where Bora is spoken (colored areas). Used by permission, © SIL International, 2022. *Ethnologue: Languages of the World*. Twenty-fifth edition. Further redistribution prohibited without permission.

Today there are probably few, if any, monolingual speakers of Bora. Most people over 50 years of age are conversant in Spanish, but still prefer Bora. All those under 50 are fluent in Spanish, and most native speakers of Bora are 20 or older. All children now learn Spanish. Some of them acquire Spanish as their L1 and Bora as L2, while others never learn Bora at all. The Expanded Graded Intergenerational Disruption Scale (EGIDS) level for Bora is 7, corresponding to a shifting language – one that is no longer being transmitted to children (Eberhard et al., 2017; Roe, 2014; Thiesen & Weber, 2012).

## 2.2. Phonology

### 2.2.1. Inventory

Bora exhibits the following basic consonant phonemes: /p t k ʔ p<sup>h</sup> t<sup>h</sup> k<sup>h</sup> t̪s̪ t̪s̪<sup>h</sup> k̪p̪ h β m n r/. Each of these except for /k̪p̪/ also has a palatalized counterpart (see 3b–c). The maximal syllable

is [CVX], where  $X \in \{:, h, \text{?}\}$ . That is, vowel length is contrastive and cannot co-occur with a tautosyllabic coda. Syllable-final consonants are limited to the glottals /h ʔ/. In coda position /h/ is normally pronounced as [x]. Onsets are usually not required. There are two contrastive level tones, high and low. These have a high functional load, with numerous minimal pairs distinguished solely by pitch. The low tone is generally the marked one and high the default. Sequences of heterorganic vowels, as well as long vowels bearing contour tones, are split into distinct syllables phonetically, as in (1)–(2) below (Roe, 2014; Thiesen & Weber, 2012).

As previewed in §1, Bora exhibits six phonemic vowels, traditionally described as shown in **Table 1**. Both the central unrounded /i/ and the back unrounded /u/ are fully productive, canonical segments. Just like the other four vowel phonemes, /i/ and /u/ can occur equally well in open or closed syllables, with high or low tones, as short or long, and in all positions of the prosodic word (initial, medial, and final syllables). Both /i/ and /u/ appear frequently throughout all categories of the Bora lexicon and exhibit numerous minimal pairs. Two examples are presented in (1), contrasting both segments in word-initial and word-final positions. In this paper, we follow most Bora sources in transcribing high tone with an acute accent over vowels ([á]), and leaving low tone unmarked. The page numbers here refer to the dictionary of Thiesen and Thiesen (1998).

|      | front | central | back |
|------|-------|---------|------|
| high | i     | ɨ       | u    |
| mid  | ɛ     |         | o    |
| low  |       | a       |      |

**Table 1:** The traditional vowel inventory of Bora (Thiesen & Weber, 2012).

- (1) /íhi/ [i.í.hi] ‘species of edible caterpillar found on guava fruit’ (p. 154)  
 /íhu/ [i.í.hu] ‘anteater; horse; mule; donkey, burro’ (p. 155)  
 /ihk<sup>h</sup>o/ [ix.k<sup>h</sup>o] ‘to take, grab, or gather fruits’ (p. 153)  
 /u<sup>h</sup>kh<sup>h</sup>o/ [ux.k<sup>h</sup>o] ‘species of giant otter’ (p. 294)

To further illustrate the contrast, the following forms display a lengthened variety of all six vowel phonemes occurring simultaneously in a nearly identical environment:

- (2) /ípa/ [i.í.pa] ‘small ash-colored deer’ (p. 135)  
 /ípa/ [i.í.pa] ‘shad fish species’ (*sábaló*) (p. 148)  
 /uúpa/ [u.ú.pa] ‘species of worm found in meat and fruit in general’ (p. 291)  
 /épa/ [ɛ.é.pa] ‘that one (drum, guava, box, etc.)’ (pp. 127–128)  
 /aápa/ [a.á.pa] ‘to be weak or soft (the voice); to burn weakly (a fire)’ (p. 25)  
 /óópakpa/ [óó.pa.kpa] ‘species of monkey’ (*huapo*) (p. 210)



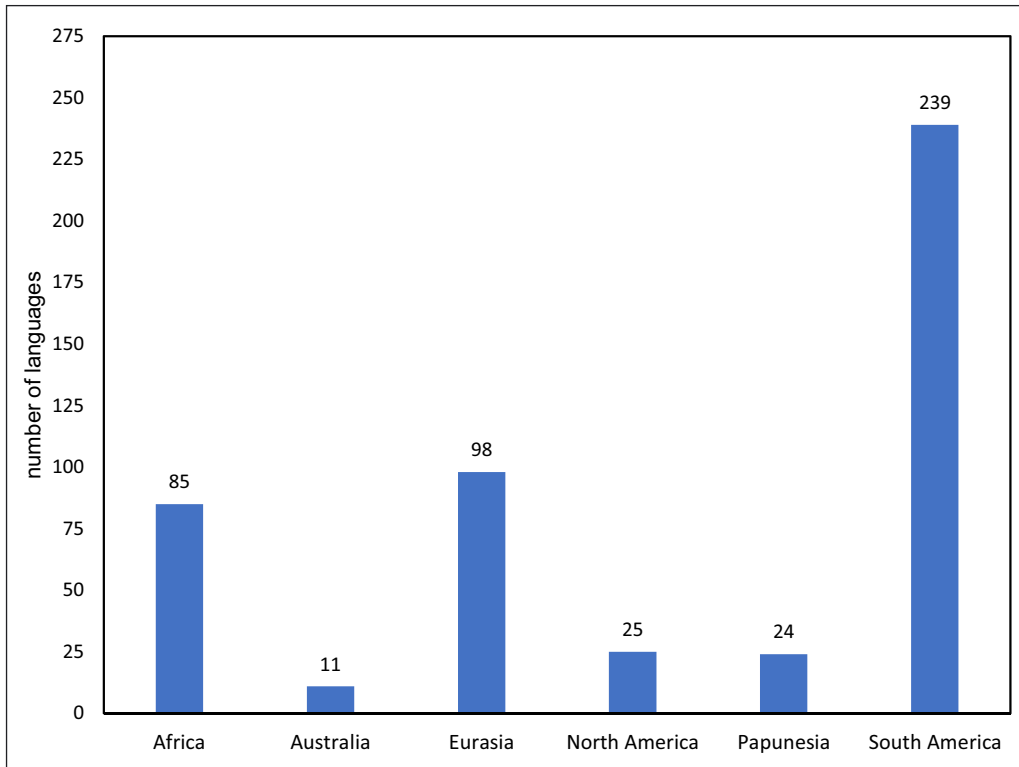
A few minor phonetic details of these segments are also worth mentioning since they shed light on the natural classes we discuss in §2.2.2. Thiesen and Weber (2012) note that “/o/ is the only rounded vowel and is only slightly round” (p. 30). They further assert that /i/ is “tense” while all of the other vowels are “lax” (p. 30). /ɛ/ is raised and pronounced as [e] immediately preceding /i/, and is sometimes lowered to [a] immediately preceding /a/. When /i/ occurs in a syllable closed by the voiceless velar fricative [x] (an allophone of /h/), it is lowered to [ɪ]. Thus /ihk<sup>h</sup>a/ ‘to be, live, exist’ is pronounced [ix.k<sup>h</sup>ɪa]. In this word the vowel /i/ palatalizes the following onset consonant (/k<sup>h</sup>/) across the intervening /h/, as discussed in §2.2.2. In just a few lexical items /a/ is realized as [ɔ] preceding a nasal consonant. The vowel /i/ can sometimes also be pronounced as [ɪ] following labial consonants, especially /m/ (Parker, 2001; Thiesen Kliever & de Thiesen, 1975). The labio-velar stop /k<sup>w</sup>/ is often pronounced as [k<sup>w</sup>] or even [w], especially by younger speakers (Roe, 2014; Thiesen & Weber, 2012).

The inventory of Bora vowels may be unique among the languages of the world (Parker, 2001). To begin with, a contrast between the phonemes /i/ and /ɯ/ is quite rare. The PHOIBLE sample (Moran & McCloy, 2019) contains just 14 out of 2101 distinct languages (0.67%) with both /i/ and /ɯ/. PHOIBLE also lists 461 languages (including Bora) which contrast three high vowel qualities. All but four of these high vowel inventories (99.1%) can be grouped into one of three categories which include a front unrounded vowel such as /i/ and a back rounded vowel like /u/: (1) 358 languages have these two plus a central or back unrounded vowel such as /ɪ/ or /ɯ/. (2) Another 54 languages have these two plus a front rounded vowel such as /y/. (3) Finally, 45 languages have a front unrounded vowel such as /i/ and a back rounded vowel like /u/ plus a third vowel that differs in tenseness, lowering, or centralization from one of these vowels. The other four relevant languages in PHOIBLE have high vowel inventories consisting of /i i ɯ/. One of these is Bora. The other languages with these three high vowels are Huba (Afro-Asiatic; Greive, 1973), Matses (Pano-Tacanan; Fleck, 2003), and Nimboran (Nimboranic; Anceaux, 1965).

An interesting typological detail of the vowel /i/ is that its presence in phoneme inventories appears to be an areal feature of South American languages in particular. For example, Engstrand, Björsten, Lindblom, Bruce, and Eriksson (1998) note that in the UPSID sample of 451 languages, /i/ occurs in 14% of them overall, yet at a rate of 35% among languages in South America specifically. Furthermore, in the much larger PHOIBLE sample of 2186 languages,<sup>1</sup> the geographic distribution of /i/ is even more skewed: Among the 482 languages in PHOIBLE reported to contain /i/, nearly one-half (N = 239) are located in South America alone (see **Figure 2** and **Table 2**). Consequently, of the 333 languages in PHOIBLE pertaining to this continent, 72% of them include /i/. The following histogram and contingency table summarize the statistical distribution of /i/ among the languages in PHOIBLE.

---

<sup>1</sup> This is the same version of PHOIBLE as above. The 2186 languages correspond to 2101 distinct languages in Glottolog.



**Figure 2:** Number of languages in the PHOIBLE sample containing /i/, divided by macroareas.

| Macroarea     | Languages | /i/ Observed | /i/ Expected | $X^2$        |
|---------------|-----------|--------------|--------------|--------------|
| Africa        | 718       | 85           | 158.3        | 34.0         |
| Australia     | 345       | 11           | 76.1         | 55.7         |
| Eurasia       | 455       | 98           | 100.3        | 0.1          |
| North America | 145       | 25           | 32.0         | 1.5          |
| Papunesia     | 190       | 24           | 41.9         | 7.6          |
| South America | 333       | 239          | 73.4         | <b>373.4</b> |
| Total         | 2186      | 482          | 482          | 472.2        |

**Table 2:** Numerical frequency of the phoneme /i/ among the 2186 inventories in the PHOIBLE sample, partitioned by geographic location (macroareas).

A chi-square test of independence on the observed and expected values in **Table 2** is extremely significant:  $X^2(5) = 472.2$ ,  $p < 0.0001$ . We thus conclude that the areal distribution of /i/ is not entirely random, at least in these two samples. As the rightmost column in this table shows, the disproportionate frequency of /i/ among languages spoken in South America is the factor which contributes the most to this outcome ( $373.4 \div 472.2 = 79\%$  of the total  $X^2$  value).



Languages containing high vowel inventories such as /i i u/ pose a challenge to linguistic theory since these three segments cannot be distinguished in models that rely on a single binary feature [ $\pm$ back] for backness distinctions. For example, SPE (Chomsky & Halle, 1968) posits that for classificatory purposes all features should be “strictly binary” (p. 65). Similarly, Duanmu (2016) maintains that binary features are “minimally sufficient to distinguish” all contrastive segments “in the world’s languages” (p. ix). The latter approach explicitly claims that a phonemic opposition between /i/ and /u/ does not exist, and must be captured by a different feature such as [ATR] or [round]. Nevertheless, as noted by Duanmu (2016), many linguists do posit n-ary (non-binary) values for certain features. To illustrate, based on contrasts such as between /y ʉ u/ in Norwegian, Ladefoged and Maddieson (1996, p. 292) conclude, “Consideration of a number of very different cases, such as Nweh and Norwegian, leads us to conclude that it is probably appropriate to recognize a front-back dimension containing three major phonetic categories: [front], [central] and [back].” Citing this precedent, Parker (2001) claims that Bora provides further evidence for a three-way distinction in backness. However, that conclusion is based on formant measurements alone. Given the additional data on lingual-dental contact we present here, it is not necessarily accurate to categorize Bora’s /i/ as [central]. Instead, we argue that it is best described as a front vowel, but one in which the apex and blade of the tongue are implicated in the contrast. The upshot of this situation, then, is that either way Bora does require a new type of feature which the approach of Duanmu (2016) does not countenance, regardless of our experiment. However, if our account here is correct, this feature primarily encodes dental contact, not backness per se. This confirms a prediction made by Björsten and Engstrand (1999) in conjunction with their instrumental study of “damped” /i/ and /y/ in Swedish. They conclude that “these vowels are members of the category of high central unrounded vowels, and that this vowel category is a fairly wide-spread one among the world’s languages. It can be assumed that this vowel type is produced with apicalization in some languages since, according to the above simulation experiment, this would further enhance its damped quality. However, these conclusions remain tentative in view of the limited empirical evidence available so far” (p. 1960). The instrumental data we present here on Bora thus helps to fill this gap.

### 2.2.2. Alternations

In this section we summarize a number of rules and constraints which illustrate the phonological behavior of Bora vowels, in support of the feature specifications just discussed. We focus primarily on processes which change one underlying segment into a different phoneme and/or rules which are triggered mainly in derived environments. In Parker and Mielke (to appear) we show that there are robust facts from Bora phonology establishing a contrast in behavior between the three high vowels /i i u/ in particular. Since numerous data are listed there to document these patterns, in this section we limit the evidence to just one form illustrating each generalization.

These are important in that several morphophonemic alternations converge in confirming the phonological affiliations of the high vowels /i/ and /i/ as front but /u/ as back:

- (3) a. An underlying /i/ cannot surface phonetically preceding /i/, either adjacent or across an intervening consonant: \*[i (C) i]. When morpheme concatenation produces such a sequence, the /i/ changes to [i]. No vowel other than /i/ undergoes this process, and no vowel other than /i/ triggers it.

/i-ʔi/ → [iʔi] ‘this one (bunch or stalk of bananas, etc.)’

this-classifier (Thiesen & Thiesen, 1998, p. 358)

/i-hi/ → [hi] ‘this one (disk, coin, etc.)’

this-classifier (Thiesen & Thiesen, 1998, p. 154)

- b. A process of consonant palatalization is regularly triggered following /i/, and is often triggered by the vowel /a/ in some (but not all) morphemes. These are /a/s which are believed to have historically contained [i] as part of the proto-diphthong /\*ai/. No other vowels systematically correlate with a following palatalized consonant.

/ts<sup>h</sup>i-ʔε/ → [ts<sup>h</sup>iʔ<sup>ʲ</sup>ε] ‘another one (tree, plant, etc.)’

(an)other-classifier (Thiesen & Thiesen, 1998, p. 558)

/i-hi/ → [hi] ‘this one (disk, coin, etc.)’

this-classifier (Thiesen & Thiesen, 1998, p. 154)

- c. Palatalized consonants usually do not occur immediately preceding /i/ or /i/. No other vowels, including /u/, have this inhibitory effect.

/i-ʔi/ → [iʔi] ‘this one (bunch or stalk of bananas, etc.)’

this-classifier (Thiesen & Thiesen, 1998, p. 358)

/i-hu/ → [h<sup>ʲ</sup>u] ‘this one (road, shotgun, etc.)’

this-classifier (Thiesen & Thiesen, 1998, p. 350)

- d. There is a statistical dispreference for [o] to surface phonetically before /u/, at least in verb roots followed by certain suffixes. In such cases, an underlying /o/ changes to [u]. No vowel other than /o/ undergoes this process, and no vowel other than /u/ triggers it. This suggests that /u/ is the segment most articulatorily close (similar) to /o/ in the Bora vowel space.

multiple action singular action gloss of verb root

[oʔt<sup>h</sup>a] [oʔt<sup>h</sup>á-k<sup>h</sup>u] ‘whistle’ (Thiesen & Thiesen, 1998, p. 212)

[poʔto] [poʔtú-k<sup>h</sup>u] ‘row, paddle’ (Thiesen & Thiesen, 1998, p. 57)

[toxt<sup>h</sup>u] [toxt<sup>h</sup>ú-k<sup>h</sup>u] ‘take or carry away; grab’ (Thiesen & Thiesen, 1998, p. 116)

- e. Somewhat analogously, verb roots ending with /u/ exhibit a strong numerical tendency to end with certain suffixes containing /u/ rather than /o/. This is not the case with verbs whose final vowel is /i/.
- f. The generalization shared by (d) and (e) is that, across a morpheme boundary, the vowels /u/ and /o/ tend not to combine with each other phonetically, in either order. Thus there is a mirror-image constraint \*[o C u] and \*[u C o]. The former induces synchronic alternations as in (d) above, whereas the latter is more passive or

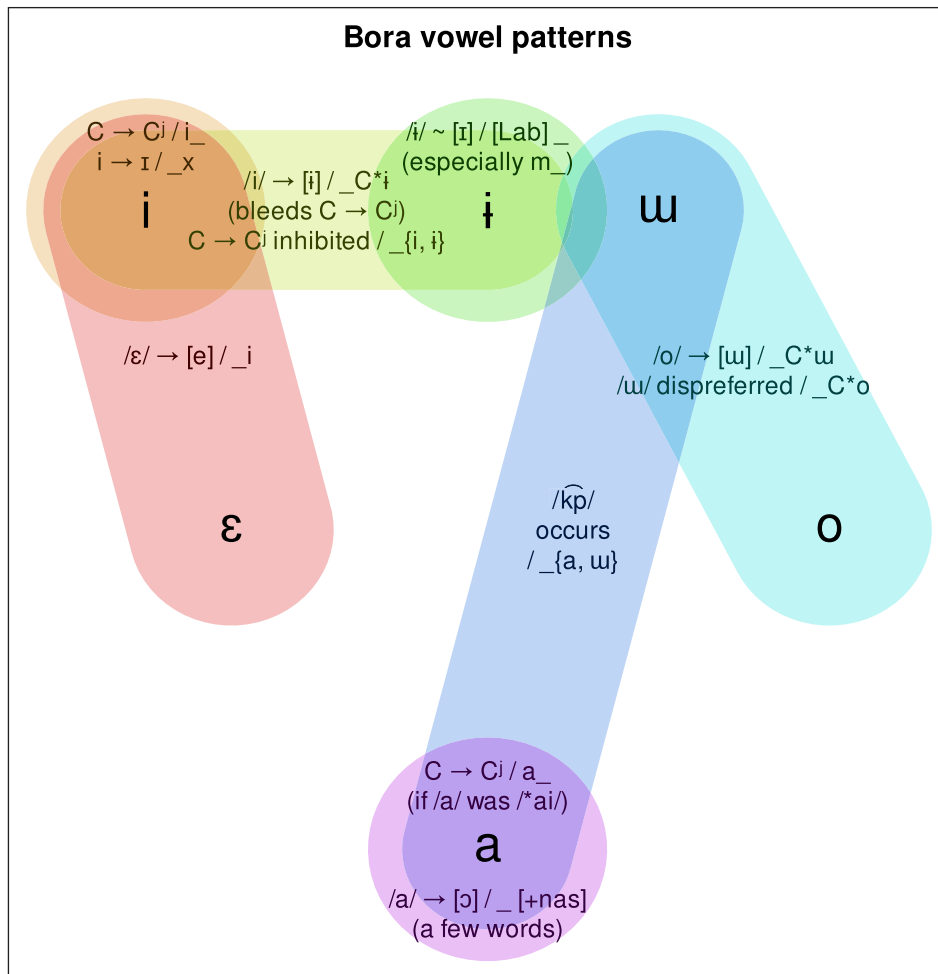
lexicalized (although it can be observed in variants of certain types of verbs). Both of these constraints are violated in Bora, but when they are enforced, the unmarked outcome is invariably [ʷ C ʷ]. That is, [ʷ] emerges at the expense of [o].<sup>2</sup>

- g. The vowel /ɛ/ raises to [e] before /i/ (§2.2.1). This is partially analogous to the raising of /o/ preceding [ʷ]. The vowel /a/ does not undergo either of these processes, nor does the vowel /i/ trigger them.
- h. /ɛ/ is sometimes lowered to [a] before /a/ (§2.2.1), whereas /o/ does not assimilate to /a/.
- i. Finally, the labio-velar stop /kʷ/ appears only before the vowels /a/ and /ʷ/. It is systematically unattested preceding /i i ɛ o/. The motivation for this, we suggest, is that /i i ɛ/ are all [-back], whereas /a/ and /ʷ/ are [+back] (yet [-round]). Since /o/ is [+round], and since /kʷ/ is sometimes pronounced as [k<sup>w</sup>] or [w] (§2.2.1), the prohibition against \*/kʷo/ could be driven by the Obligatory Contour Principle.

The generalizations discussed above are summarized in **Figure 3**. Virtually all of the statements in (3) have exceptions, most of which are sporadic and idiosyncratic. As far as we are aware, however, the generalizations exemplified throughout this section are fully productive and represent the unmarked, default patterns, except where noted. See Parker and Mielke (to appear) for more precise quantification of these tendencies, including additional data. For example, the assimilation of /i/ to [i] is quite systematic in the sense that it takes place with virtually all such sequences derived by morpheme concatenation. These include three prefixes and three word-initial bound adjectival roots. The one regular exception is that it does not apply across the boundary of an enclitic. In fact, the phonotactic condition prohibiting the sequence \*[i C i] holds inside of roots as well. Thus, while forms such as [tsiʔtsi] 'money; coin' are attested, there are no morphemes in which /i/ precedes /i/ (Thiesen & Weber, 2012, p. 31). On the other hand, there are numerous words where the vowel /i/ occurs before a syllable containing /ʷ/, both within and across morphemes.

The distributional facts summarized in (3) and depicted graphically in **Figure 3** provide phonological evidence confirming the specification of /i/ as front and /ʷ/ as back. We see that /i/ is phonologically associated with the front vowel /ɛ/, and /ʷ/ is phonologically associated with the back (and round) vowel /o/, both to the exclusion of the other two high vowels. /i/ shows no evidence of patterning with back vowels, and its phonological patterning in fact groups it with the front vowels. However, the fact that /i/ palatalizes following consonants whereas /i/ does not indicates that /i/ is probably less of a high front vowel than /i/ is. Furthermore, /i/ triggers assimilation of /i/ whereas /ʷ/ does not. The grouping of /i/ and /i/ as a natural class shows that /i/ occupies a position intermediate between /i/ and /ʷ/ in the vowel system, although /i/ could differ from them in some other feature as well. In addition, palatalization of

<sup>2</sup> However, while [o] is avoided in this situation, it is not necessarily the most marked vowel in Bora, numerically speaking. An informal rough count of the entries in Thiesen and Thiesen's (1998) dictionary gives the impression that /o/ occurs in at least the same amount of lexical items as /ɛ/ does. The other four vowels appear to be much more common than these two are.



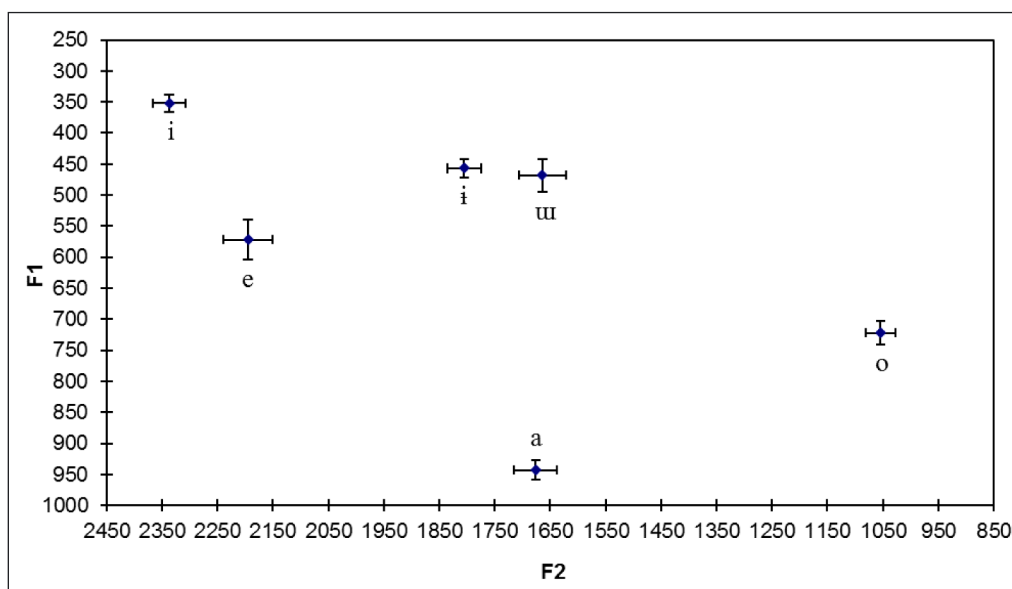
**Figure 3:** Summary of phonological patterns involving Bora vowels.

consonants (after  $/i/$  and  $/a/$ ) is blocked if the targeted consonant precedes  $/i/$  or  $/i/$ , but not if it precedes  $/ʉ/$ . Finally, the bidirectional harmonic pressure exerted on  $/o/$  by the vowel  $/ʉ/$  (but not by  $/i/$ ) also indicates that  $/ʉ/$  is relatively more posterior than  $/i/$  is. The complete assimilation of  $/o/$  to  $[ʉ]$  establishes these two  $[+back]$  segments as adjacent in Bora's vowel space. In short, the phonological behavior sketched in this section strongly demonstrates that the phonemes  $/i/$  and  $/ʉ/$  in Bora need to be distinguished by at least one feature along the front/back dimension. We thus conclude that part of the opposition between these two segments (as well as between them and  $/i/$ ) lies in the horizontal plane, and therefore involves the tongue.

### 2.2.3. Acoustics

An acoustic study by Parker (2001) provides some initial phonetic support for the characterization of Bora vowels sketched above. Using audio data recorded in Peru with 14 adult Bora speakers (eight men and six women), he reports measurements of F1-F3 and segmental duration for all six

vowel phonemes. These were pronounced both in isolation (*V* and *CV* syllables) as well as in 15 specific Bora words, including minimal pairs contrasting /i/ vs. /u/ and /i/ vs. /ɨ/. Most of these lexical items overlap with the word list used in the current experiment (see (4) below). Nevertheless, a crucial advantage — and the focus — of our methodology here is that we also investigate lip and tongue blade positions with video images, which Parker (2001) did not obtain. One of his most important findings is that /i/ consistently exhibits a higher F2 value than /u/ does, for all speakers and all phonological contexts. This difference is always significant, albeit occasionally small; across all tokens the average distance between them ( $[F_2 /i/ - F_2 /u/]$ ) is 166.5 Hz. **Figure 4** illustrates the typical spread between the six Bora vowels, as pronounced in isolated syllables.



**Figure 4:** Scatterplot of mean F1 and F2 frequencies (in Hz) for six female speakers [error bars indicate 95% confidence intervals] (Parker, 2001).

As **Figure 4** exemplifies, the difference in mean F1 between /i/ and /u/ is never significant in Parker's (2001) study. Typically it ranges from 11–16 Hz across conditions (phonological environment and gender). Furthermore, these two vowels fall between /i/ and /ɛ/ in terms of F1. This confirms the informal description of /i/ and /u/ as open (Thiesen & Weber, 2012). On the horizontal dimension, F3 values (not shown here) parallel those of F2, with /i/ about 209 Hz higher on average than /u/. Finally, /i/ is slightly longer in duration than /u/ (99.5 vs. 93.5 ms overall, respectively), and this duration difference is the smallest observed between any pair of vowels. As Parker (2001) concludes, the fact that Bora's /i/ yields higher F2 and F3 values than those of /u/ suggests that /i/ is articulated farther to the front. In order to test this directly, of course, measurements of actual tongue body positions are desirable.

A striking detail of **Figure 4** is the relative proximity of /i/ and /u/ on the F2 axis. This explains in part why they rarely contrast in the same language: They are perceptually very

similar. Also noteworthy is the large distance in F2 between /ɯ/ and /o/. However, since lip rounding is known to lower F2, this is expected if /o/ is rounded and /ɯ/ is not. The relative positioning of these six segments in Bora's vowel space is consistent with measurements of analogous phonemes in various languages, including other studies involving /i/ and/or /ɯ/ (Elias-Ulloa & Muñoz Aramburú, 2021; Olson, 2019; Parker, 2001; Schwartz, Boë, Vallée, & Abry, 1997; Thomas, 2017). Furthermore, Bora's formant values confirm a generalization noted in Becker-Kristal's (2010) acoustic summary of 230 languages: Non-peripheral vowels (such as /i/) exhibit a tendency to lie more to the front than to the back, and this trend is especially strong among higher vowels.

#### 2.2.4. Summary of the problem

To summarize thus far, Bora has three high vowels, conventionally transcribed as /i i ɯ/. What is the best way to account for their contrast and divergent behavior? Three main possibilities might be considered. First, based on the phonetic and phonological evidence presented above, we could simply continue to categorize these vowels as front unrounded, central unrounded, and back unrounded, respectively. This is consistent with the impressions of Thiesen Kliever and de Thiesen (1975) and Parker (2001) that none of these vowels are rounded. However, this description would contradict the claim that languages do not make use of more than two degrees of backness (Duanmu, 2016). A second possibility is that /i ɯ/ are produced with similar tongue positions but are distinguished by means of a difference in lip posture which was not documented in previous work on Bora. Selkirk (1993) in fact argues that a contrast between vowels such as /i/ and /ɯ/ must be captured in this way, due to the architecture of the Labial Node in the feature geometry tree she proposes. This hypothesis is also consistent with the phonological facts, as well as the assumption that phonological alternations typically involve phonetically-related sounds. A third possibility is that the primary phonetic difference between /i ɯ/ is due to some feature other than backness or rounding. We turn now to instrumental data that allows us to evaluate these possibilities. We will show that while /i/ and /ɯ/ are both unrounded, /i/ is actually produced with the tongue protruded toward the front of the mouth and with a small mouth opening caused by a lingual-dental constriction.

## 3. Methodology

### 3.1. Speakers

The audio/video recordings discussed in this section were collected by Amy Roe in 2011, as part of a larger study on Bora tone (Roe, 2014).<sup>3</sup> Ten native speakers of Bora participated (five males

---

<sup>3</sup> These procedures were approved by the IRB of the University of North Dakota, in conjunction with Roe (2014). Thanks to Amy Roe for sharing her recordings with us.

and five females). Three additional speakers were recorded, but their data is not included in our analysis due to technical problems with the recordings.

Roe (2014) personally interviewed each speaker prior to inviting them to participate. She reports that they ranged from 23 to 68 years of age. None of them presented any noticeable impairments that would affect their speech. All of them grew up in Bora-speaking villages and acquired the language as L1. Furthermore, each of them continues to use Bora on a regular basis. Each Bora speaker was also fluent in Spanish, the metalanguage used to obtain the speech samples, and can read fluently in both languages. Recordings were made in three locations in Peru: (1) Iquitos, a major Spanish-speaking city on the Amazon River; (2) San Andrés, a Bora village near Iquitos, on the Momón River; and (3) Brillo Nuevo, a Bora village more distant from Iquitos, on the Yaguasyacu River. Each speaker was paid for their efforts. In Iquitos the recordings were made in the home of one of the speakers. In San Andrés and Brillo Nuevo the recordings were made in empty houses in order to minimize background noise. See Roe (2014) for further details concerning the selection of Bora participants and other logistical considerations.

### 3.2. Materials

The stimuli were presented on written sheets, using the practical Bora orthography already familiar to each participant. Thirteen Bora words were elicited, including several minimal pairs. These are listed in (4). The word list was pronounced five times by each speaker, with the order of presentation of the items counterbalanced across the five sheets. A couple of participants pronounced a few items less than five times due to technical problems.

Each target word was embedded in the frame ‘[tʰɛ]\_\_’ ‘Say \_\_’. This consists of the imperative singular prefix /ti-/ plus the verb root /neé/ ‘say’. When these are combined, the /n/ undergoes palatalization (see §2.2.2). As Roe (2014, p. 39) explains, the context ‘[tʰɛ]\_\_’ was selected for elicitation not because of a preference for the target word (in the blank) to be sentence final. Rather, the speakers she worked with were unable to think of an appropriate Bora word to come right after the substitution item (i.e., in position *X* in the hypothetical sentence ‘Say \_\_ *X*’).

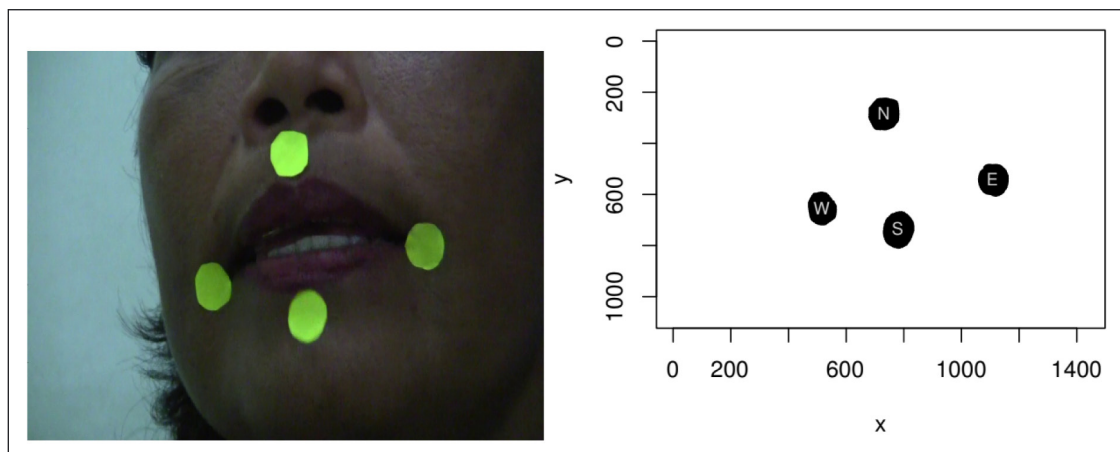
- (4) Bora stimuli, including minimal pairs contrasting /i/ vs. /u/ and /i/ vs. /i/.
- |    |                 |   |
|----|-----------------|---|
| a. | [tʰa-hu]        | ‘one pistol, etc.’                            |
| b. | [tʰa-hi]        | ‘one coin, etc.’                              |
| c. | [tʰa-mu]        | ‘one drum, etc.’                              |
| d. | [tʰa-mi]        | ‘one canoe, etc.’                             |
| e. | [tʰá-hú-ʔkʰopa] | ‘one big pistol’                              |
| f. | [tʰá-hí-ʔkʰopa] | ‘one big coin’                                |
| g. | [uxkʰu]         | ‘to get, obtain’                              |
| h. | [ixkʰu]         | ‘species of bird’ ( <i>paucar</i> in Spanish) |



- |    |                     |  |
|----|---------------------|--|
| i. | [ípa]               | ‘small ash-colored deer’                                 |
| j. | [ípaá] <sup>4</sup> | ‘shad fish species’ ( <i>sábalo</i> in Spanish)          |
| k. | [í-kpaá]            | ‘this machete, etc.’                                     |
| l. | [íkpaá-βε]          | ‘action of opening and closing the mouth’ (onomatopoeic) |
| m. | [miʔé-kpa]          | ‘dam closure of river to catch fish’                     |

In the stimuli in (4), the contrast between the three high vowels (particularly /i/ and /u/) is controlled for in terms of prosodic factors. For example, in items (a–d) these phonemes occur word-finally with low tone. In items (e–f) they are both suffixed and thus shift to high tone. In items (g–h) they occur in initial, closed syllables. The vowel /i/ also occurs in both open and closed syllables in this list, as well as with high and low tones. The list also contains at least two instances of each of the other three vowel phonemes.

Each speaker was filmed with a video camera fixed on a tripod approximately two feet in front and slightly to the side of the body, at the height of the chin. Video files were initially saved in mpg format at a rate of 29.97 frames per second. The video’s audio tracks were recorded in stereo and sampled at 48 kHz. The left channel was extracted and used for acoustic analysis. Four yellow dots were affixed to each person’s mouth: One at each corner (the horizontal axis), one on or just below the middle of the lower lip, and one on or just above the middle of the upper lip, at the base of the philtrum (see **Figure 5**). The two dots in the middle of the upper and lower lips form the vertical axis.



**Figure 5:** Quantification of lip image data. Left: Yellow dots placed on a speaker’s lips to measure lip position. Right: Yellow pixels from the same image, labeled as North, South, East, and West, at dot centers. x and y are horizontal and vertical position, respectively, in pixels.

<sup>4</sup> This word is the same lexical item which appears in (2). In Bora, citation forms are sometimes modified prosodically when placed in context, for grammatical reasons.

The goal of this method of lip measurement is to provide a straightforward way to extract a low-dimensional representation of lip postures that can be used to recognize phonetically important labial gestures. An alternative approach is to apply blue lipstick, threshold the video images, and then extract measures such as the area inside the blue lip region as an indication of lip rounding (Lallouache, 1991; Ménard et al., 2013; Noiray, Cathiard, Ménard, & Abry, 2011).<sup>5</sup> For images of lips unadorned with markers or lipstick, lip landmarks may also be tagged manually (Mielke, 2012). Manually tagged lips are used to train neural networks to recognize analogous landmarks on other speakers (Wrench & Balch-Tomes, 2022). This may become a standard technique in the future. In our experiment, the vertical markers were not placed at the very edge of the lips. Consequently, the dots do not meet when the lips are closed and hence they do not directly represent the area of the lip opening. Nevertheless, the size of the lip opening can still be estimated from their relative positions when open by comparing this measurement with the distance between the dots when the lips are closed.

Video images are not typically designed to measure tongue positions, but this is possible when the tongue tip approaches the mouth opening. For example, in studying an interdental approximant in Kagayanen, Mielke, Olson, Baker, and Archangeli (2011) manually tagged the tongue tip and then measured the degree of tongue protrusion based on lateral images of the face. Since the experiment here involves frontal view images, it is difficult to precisely quantify the degree of tongue tip protrusion. Nevertheless, at the very least we are able to categorically label images according to whether there is visible lingual-dental contact, and if so, whether this contact is central and/or lateral.

### 3.3. Data analysis

We analyzed all of the recordings in which the speaker's lips remained in view of the camera and in which the acoustic signal was of sufficient quality for formant analysis. These criteria yielded 963 usable vowel tokens from the target words listed in (4). Token counts by speaker and phone are displayed in **Table 3**.

#### 3.3.1. Speech segmentation

Audio was segmented at the word and phone levels using an adaptation of the Penn Phonetics Lab Forced Aligner (P2FA; Yuan & Liberman, 2008). P2FA was adapted following the procedure used by Milne (2011) for French and Dicanio et al. (2013) for Yoloxóchitl Mixtec. Specifically, we first assigned an English phone to each phone of the language under investigation (as shown in **Table 4**), then we force-aligned the speech from that language as if it were English,

---

<sup>5</sup> In the laboratory, lips may be tracked directly using Electromagnetic Articulometry (Schönle et al., 1987) or Opto-Trak (e.g., Campbell, 2004; Noiray et al., 2011).

| Speaker: | 1   | 2   | 3  | 4  | 5  | 6   | 7   | 8  | 9  | 10 | TOTAL |
|----------|-----|-----|----|----|----|-----|-----|----|----|----|-------|
| /a/      | 58  | 45  | 36 | 23 | 14 | 54  | 39  | 33 | 22 | 37 | 361   |
| /ε/      | 9   | 8   | 4  | 3  | 3  | 9   | 8   | 2  | 2  | 3  | 51    |
| /i/      | 17  | 12  | 9  | 8  | 5  | 17  | 11  | 12 | 8  | 9  | 108   |
| /i/      | 35  | 24  | 24 | 10 | 8  | 32  | 18  | 13 | 13 | 15 | 192   |
| /ɯ/      | 28  | 25  | 20 | 12 | 10 | 32  | 26  | 12 | 16 | 18 | 199   |
| /o/      | 10  | 7   | 3  | 4  | 1  | 6   | 4   | 6  | 4  | 7  | 52    |
| TOTAL    | 157 | 121 | 96 | 60 | 41 | 150 | 106 | 78 | 65 | 89 | 963   |

**Table 3:** Token counts by speaker and vowel (speakers 1–5 are female and 6–10 are male).

| Bora phones | P2FA model used (Arpabet, IPA) |      |
|-------------|--------------------------------|------|
| [p], [ʔ]    | P                              | [p]  |
| [tʰ], [t̪]  | CH                             | [t̪] |
| [kʰ], [k̪]  | K                              | [k]  |
| [h, x]      | HH                             | [h]  |
| [β]         | V                              | [v]  |
| [m]         | M                              | [m]  |
| [ɲ]         | N                              | [n]  |
| [i]         | IY1                            | [i]  |
| [i], [ɯ]    | AH1                            | [ʌ]  |
| [ε]         | EH1                            | [ε]  |
| [o]         | OW1                            | [o̯] |
| [a]         | AA1                            | [ɑ]  |

**Table 4:** English phone substitutions used for forced alignment of Bora words, using the Penn Phonetics Lab Forced Aligner.

and finally we replaced the English phone labels with the correct labels for the target language (Bora in this case). This works reasonably well because most of what the aligner does is place boundaries between consonants and vowels using acoustic models that are somewhat similar to the actual phones being segmented. For alignment, we created a lexicon of Bora words occurring in the recordings, using Arpabet labels. After alignment we replaced these labels in our annotations with the appropriate IPA symbols. To complete the process, the resulting boundaries were visually inspected in Praat (Boersma & Weenink, 2007) and hand-corrected when necessary.

### 3.3.2. Image analysis

Using `avconv` from `Libav` (Newmarch, 2017), we extracted the video frames during vowel intervals, as well as all /m/ and / $\widehat{kp}$ / consonant intervals. In order to identify the yellow dots, these images were analyzed in `R` by thresholding according to red, green, and blue pixel color values. In general, the yellow dots had a high ratio of green to blue, and more red than blue. Thresholds were set independently for each recording file, as needed. In some cases, a yellowish background object required us to manually specify regions to exclude for a particular recording. Contiguous pixels identified as “yellow” were automatically grouped into blobs, and the four largest blobs were interpreted as the four dots. These dots were designated as “North”, “South”, “East”, and “West” according to their position in the image, and the (x,y) coordinates of their centers were analyzed further. We analyzed the vertical distance between the two central dots (N and S in **Figure 5**), and the horizontal distance between the two lateral dots (W and E in **Figure 5**).

Since the absolute distance between lip dots varies from speaker to speaker, the distances were scaled between the speaker mean and a realistic minimum value. Given that the dots never completely reach each other (§3.2), the realistic minimum distance is not zero but rather the most extreme lip closure or the most extreme lip rounding. In the case of vertical lip distance, the minimum value is taken as the average minimum distance for each speaker during their production of [m], which involves complete labial closure. Horizontal lip distance is more complicated because the minimum value is expected to occur during the most extreme rounding, which is normally with /o/. However, not all Bora speakers produce /o/ with particularly rounded lips (as will be seen in the results below). This makes /o/ a poor reference point. Consequently, as an alternative we tried using / $\widehat{kp}$ / as a reference point for lip rounding since it is frequently pronounced as [k<sup>w</sup>] or even [w], but we found that this was not a stable reference point, either. We therefore established a minimum horizontal lip dot distance reference point by taking into account the fact that speakers who do consistently round /o/ typically produce it with 85–90% of their mean lip distance. Accordingly, we set the minimum horizontal lip dot value as 85% of the mean horizontal lip dot distance for each speaker's vowel midpoint frames.

### 3.3.3. Acoustic analysis

After segmentation, vowel formant frequencies were extracted automatically in Praat. We used a script which measures each vowel with several different combinations of numbers of LPC coefficients and maximum formant frequencies, and selects the candidate whose formant frequencies and log bandwidths are closest to a prototype. This procedure essentially follows the technique which Evanini (2009) and Labov, Rosenfelder, and Fruehwald (2013) apply to English. Formant frequency was measured for all vowels at their midpoints, and was normalized with the Lobanov method (Lobanov, 1971). Formant frequency outliers were hand-corrected by visually inspecting spectrograms.

To explore potential differences other than formant frequency, the harmonics-to-noise ratio was measured for the middle 50% of each vowel, using Praat. Furthermore, multitaper spectra were calculated for the middle 50% of each vowel using the `spectRum` package for R (Reidy, 2013). Multitaper spectra reduce variance and increase temporal precision relative to Discrete Fourier Transform-based spectra, making it easier to estimate the properties of spectral peaks (Reidy, 2015). We used these spectra to measure formant amplitudes and identify aperiodic noise that might be caused by frication.

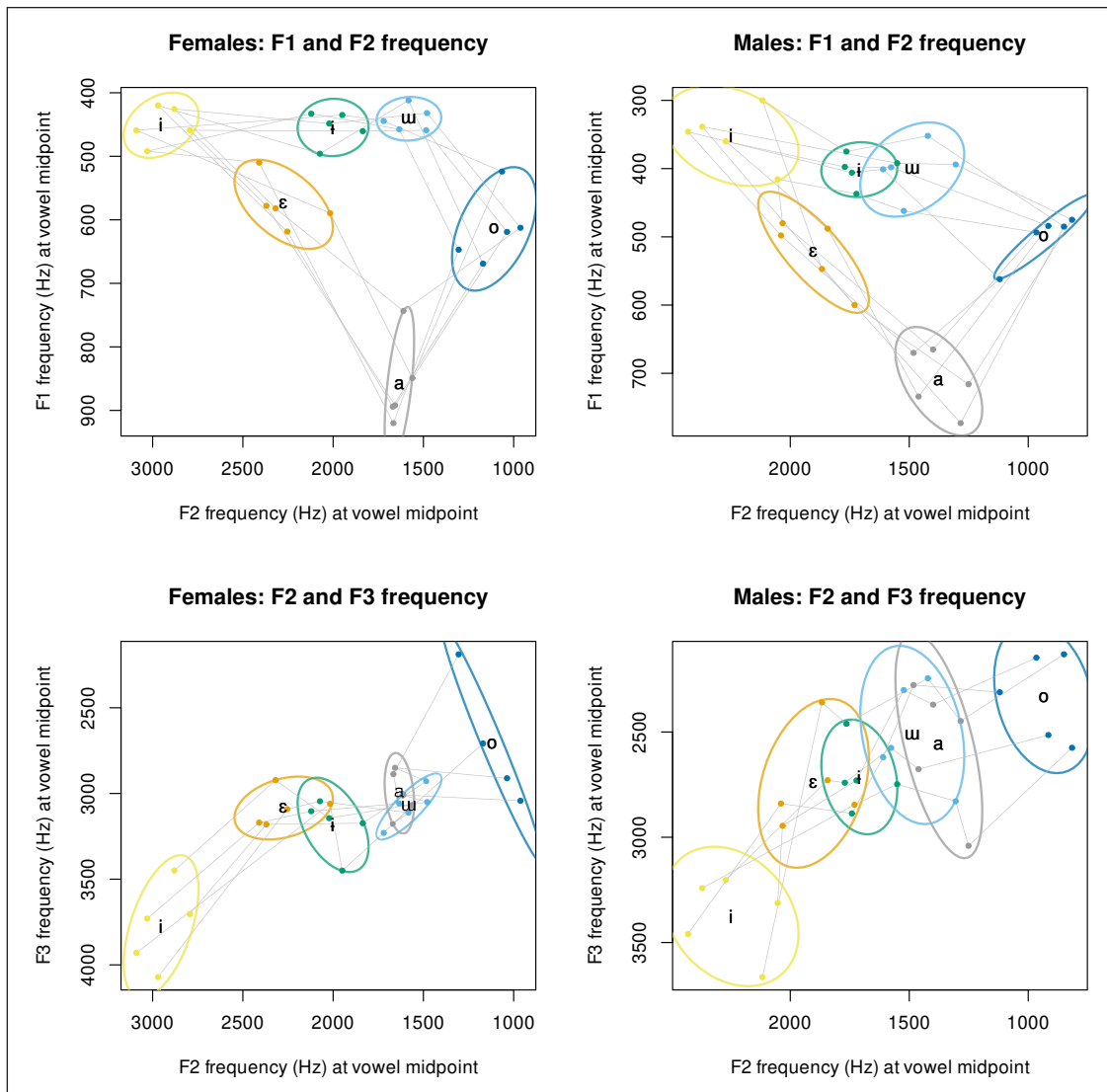
## 4. Results

### 4.1. Formant frequencies

**Figure 6** shows the median F1, F2, and F3 frequencies for each vowel phoneme for each of the 10 speakers. The positions of the vowels are generally consistent with Parker's vowel plot in **Figure 4** above. We also replicate Parker's (2001) finding that /i/ and /u/ are distinguished by F2 and F3. We include F3 in the analysis in order to look for distinctions not captured by F1 and F2. Among other things, F3 may indicate differences in lip rounding or rhoticity. To confirm significant differences in formant values among Bora's high vowels, linear mixed effects regressions were performed on the normalized formant frequency measurements of /i i u/ using the `lme4` package (Bates, Maechler, Bolker, & Walker, 2014) in R (R Core Team, 2015). Vowel and speaker gender were specified as fixed effects, with /i/ as the reference level for vowel. We include gender because we observe some apparent differences between males and females in F1 frequency and other measurements. The interaction of vowel and gender was included, as well as random intercepts for speaker and word. Here we compare only the three high vowels, since other pairwise comparisons are obviously quite different, and the token counts for the two mid vowels are somewhat low. We refer readers to Parker (2001) for more details and discussion of Bora vowel formant frequencies. In the statistical results below we treat differences as significant when  $|t| > 2$ .

Starting with the F1 model in **Table 5**, none of the vowel main effects are significant. However, there are significant interactions between gender and both vowels (/i/ and /u/), indicating that only the males have an F1 difference between /i/ and /u/, as seen in **Figure 6**. The expected F2 differences between the three high vowels are significant, although smaller for the males, even in these normalized values. Similar differences are observed in F3.

Recall that Thiesen and Weber (2012) characterize all Bora vowels other than /i/ as being lax (§2.2.1). Despite this claim, we observe a difference in F1 between the three high vowels only in males, who also show a reduced F2 difference between these three vowels compared to the females. The mean F1 frequencies (in Hz) for the high vowels are as follows: females: /i/ = 450, /e/ = 465, /u/ = 446; males: /i/ = 346, /e/ = 391, /u/ = 398. The smaller F1 differences among high vowels in our data vs. Parker's (2001) results is mostly accounted for by



**Figure 6:** Median formant frequencies by speaker. The top row shows  $F1 \times F2$  plots with females on left and males on right. The bottom row shows  $F2 \times F3$  plots. Each speaker's data points are connected by gray lines. Ellipses represent two standard deviations around mean values.

relatively high F1s for /i/ in the present data. F1 is expected to be slightly higher for high back vowels compared to high front vowels, since the more posterior constriction location causes the back cavity to be smaller (see, e.g., de Boer, 2011; Ladefoged, 1996, 127–128). A small part of the explanation for both the different F1 comparisons between genders in our results, and between the present data vs. Parker's (2001) data, is that our sample size is smaller than his was. Notwithstanding this detail, higher F1 in central and back vowels does not necessarily indicate a tense/lax difference since it is an expected acoustic correlate of backness. Clearer evidence for a tense/lax difference would come from pairs of vowels where the apparently tense one is lower in

| <b>F3 Frequency</b>   | <b>Estimate</b> | <b>Std. Error</b> | <b>t value</b> |
|-----------------------|-----------------|-------------------|----------------|
| (Intercept)           | 0.1349          | 0.1498            | 0.900          |
| vowel u               | -0.4663         | 0.1075            | -4.338         |
| vowel i               | 1.6172          | 0.1276            | 12.679         |
| gender male           | -0.1446         | 0.2001            | -0.722         |
| vowel u × gender male | 0.0896          | 0.1156            | 0.775          |
| vowel i × gender male | 0.3191          | 0.1305            | 2.446          |
| <b>F2 Frequency</b>   | <b>Estimate</b> | <b>Std. Error</b> | <b>t value</b> |
| (Intercept)           | 0.2397          | 0.0575            | 4.169          |
| vowel u               | -0.8177         | 0.0758            | -10.787        |
| vowel i               | 2.0185          | 0.0906            | 22.284         |
| gender male           | 0.1656          | 0.0569            | 2.909          |
| vowel u × gender male | 0.1726          | 0.0676            | 2.552          |
| vowel i × gender male | -0.5127         | 0.0764            | -6.715         |
| <b>F1 Frequency</b>   | <b>Estimate</b> | <b>Std. Error</b> | <b>t value</b> |
| (Intercept)           | -0.7780         | 0.0494            | -15.748        |
| vowel u               | -0.0989         | 0.0579            | -1.707         |
| vowel i               | -0.0962         | 0.0691            | -1.392         |
| gender male           | -0.0097         | 0.0552            | -0.175         |
| vowel u × gender male | 0.1236          | 0.0544            | 2.273          |
| vowel i × gender male | -0.2254         | 0.0614            | -3.674         |

**Table 5:** Linear mixed effects regressions for normalized formant frequencies of /i i u/. (/i/ is the reference level.)

F1 and at the same time slightly more peripheral in F2. The fact that /i i u/ all differ from each other primarily in F2 indicates that their differences are due to something other than tenseness or [ATR].

## 4.2. Lip positions

Lip rounding involves constriction of the orbicularis oris muscle encircling the mouth. This has the effect of pulling the four colored dots closer together, especially along the horizontal axis. Since other speech gestures (such as jaw raising and lowering) more directly affect the relative position of the two central dots, we expect the horizontal distance between the two lateral dots to be the clearest indicator of lip rounding. **Figure 7** displays sample images of all six vowel phonemes for speaker 3 (a female), showing the rounded /o/ (with lateral dots relatively close



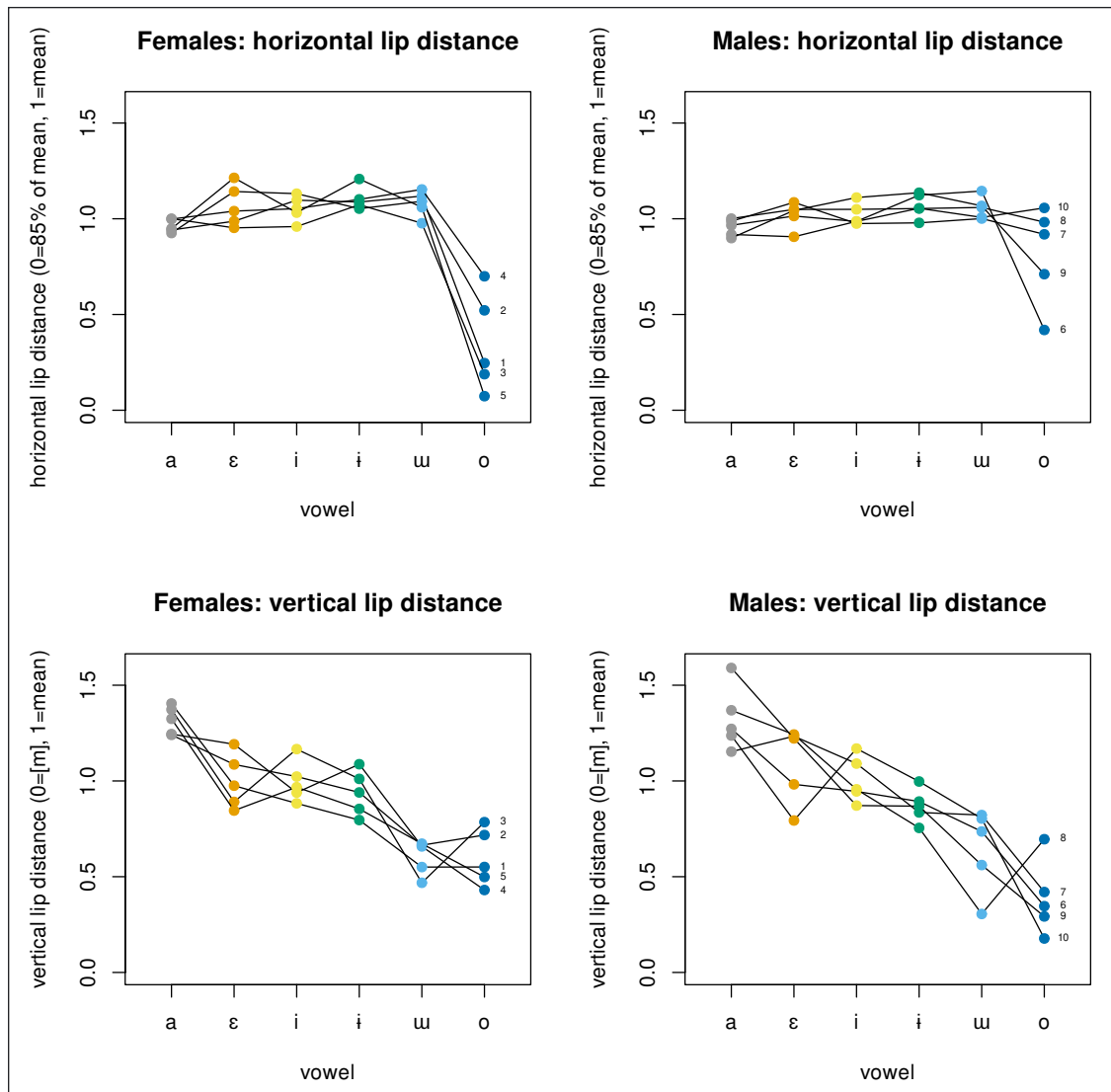


**Figure 7:** Representative lip/jaw/tongue positions for speaker 3 (a female). Top row (left to right): [i] in [í-kpáá], [ɨ] in [ixk<sup>h</sup>ʉ], and the first [ʉ] in [uxk<sup>h</sup>ʉ]; Bottom row: [ɛ] in [miʔtʃé-kpa], [a] in [ikpáá-βɛ], and [o] in [ts<sup>h</sup>á-hí-ʔk<sup>h</sup>opa].

together) and lack of rounding in the other vowels (with more distance between the lateral dots). The central dots are naturally farther apart in lower vowels, which are produced with more jaw opening. Furthermore, no lip compression (inrounding) is apparent in this image of /ʉ/. Rather, the observed vertical difference between /i/ and /ʉ/ (i.e., the distance between their central dots) appears to be associated with jaw height.

**Figure 8** summarizes the horizontal and vertical lip dot distance measures, which were normalized as explained in §3.3.2. Speaker numerical codes appear at the far right side of each panel. Speakers 1–5 are female and speakers 6–10 are male. The top panels of **Figure 8** show that in the horizontal dimension /o/ involves lip rounding for most speakers, especially the females. Furthermore, there is very little difference between the other five vowels, whose values are all close to 1 (the mean for all of the vowel tokens). Crucially, there is no consistent difference in lip rounding between /i/ and /ʉ/. The bottom panels in **Figure 8** show that there is more variation in vertical lip dot distance, compared with the horizontal axis. The low vowel /a/ is the most open, and /ʉ/ and /o/ are the most closed.

**Tables 6–7** display the summaries of linear mixed effects regressions performed on the normalized lip measurements, using the `lme4` package in R. Unlike in the formant frequency models above, here we are specifically interested in comparing the high vowels to all other vowels, in particular /o/, which is expected to be rounded. Consequently, all six vowels are included in the models. In this case /ʉ/ is designated the reference level vowel in order to facilitate comparison of /ʉ/ with /i/ and /o/. Factors included were gender and its interaction with vowel, and whether the vowel is preceded or followed by a labial consonant (which could be



**Figure 8:** Speaker mean horizontal and vertical lip distances by vowel (females on left and males on right). The top panels show the normalized distances between the lateral (W and E) dots. The bottom panels show the normalized distances between the central (N and S) dots.

any of /p kβ m/). This is important since a labial consonant affects lip shape during the vowel. The fit for these lip measurements (already normalized within speaker) was singular when all six vowels and a random intercept for speaker was included. Therefore, only a random intercept for word was included in these two models since our priority is comparing vowel categories.

**Table 6** gives the results for horizontal lip dots. Using  $|t| > 2$  as a criterion for significance once again, we see that /a/ and /o/ have significantly less distance between the two lateral lip dots than /u/ does. On the other hand, /u/ is not different from the other three vowels (/ε i i/). Furthermore, the difference in lip rounding between /u/ and /o/ in particular is quite large,

| Horizontal            | Estimate | Std. Error | t value |
|-----------------------|----------|------------|---------|
| (Intercept)           | 1.0926   | 0.0201     | 54.472  |
| vowel i               | 0.0444   | 0.0287     | 1.544   |
| vowel a               | -0.0965  | 0.0265     | -3.639  |
| vowel ε               | 0.0147   | 0.0440     | 0.334   |
| vowel i               | -0.0261  | 0.0374     | -0.698  |
| vowel o               | -0.6565  | 0.0469     | -14.002 |
| gender male           | -0.0219  | 0.0248     | -0.886  |
| preceded by labial    | -0.0411  | 0.0176     | -2.336  |
| followed by labial    | -0.0607  | 0.0194     | -3.135  |
| vowel i × gender male | -0.0278  | 0.0353     | -0.786  |
| vowel a × gender male | 0.0008   | 0.0308     | 0.026   |
| vowel ε × gender male | -0.0529  | 0.0551     | -0.961  |
| vowel i × gender male | 0.0595   | 0.0419     | 1.420   |
| vowel o × gender male | 0.4604   | 0.0546     | 8.437   |

**Table 6:** Linear mixed effects regression for normalized horizontal lip dot distances of all vowels. (/ʊ/ is the reference level.)

indicating that /o/ is a rounded vowel while /ʊ/ is unrounded, just like all of the other vowels. The smaller difference between /ʊ/ and /a/ is consistent with the correlation between vowel height and horizontal lip distance that is apparent in vowels other than /o/ in **Figure 8**. Preceding and following labial consonants also reduce horizontal lip distance. There is a significant positive interaction between /o/ and gender, which is explained by the fact that three of the five males do not exhibit lip rounding in /o/ (as seen in **Figure 8**). We take this as confirmation of Thiesen and Weber's (2012, 30) observation that /o/ is only slightly rounded in Bora (§2.2.1).

The analysis of vertical lip dot distances is presented in **Table 7**. Here we see that /a/ is considerably more open than /ʊ/, while /ε i i/ are all slightly more open than /ʊ/. However, we observe no significant difference between /ʊ/ and /o/. There is also a main effect for gender, as well as a complicated set of interactions between gender and most vowels. Nevertheless, none of these negate the small but significant vertical difference between /i/ and /ʊ/.

Interpreting vertical lip distance is more complicated than interpreting horizontal lip distance since vertical lip distance can be influenced by jaw height. In terms of height, /a/ is at the open end of the scale. With respect to lip compression or rounding, /o/ is at the closed end of the scale. Since /ʊ/ and /o/ differ in rounding, their similarity in vertical lip distance could be superficial. Consequently, we explore this further with a visual inspection of the images. Furthermore, we see quite a bit of inter-speaker variation in the direction of difference. Speakers 7, 8, and 10 (all

| Vertical              | Estimate | Std. Error | t value |
|-----------------------|----------|------------|---------|
| (Intercept)           | 0.6414   | 0.0627     | 10.225  |
| vowel i               | 0.2861   | 0.0906     | 3.158   |
| vowel a               | 0.6775   | 0.0830     | 8.163   |
| vowel ε               | 0.4065   | 0.1340     | 3.035   |
| vowel i               | 0.3990   | 0.1198     | 3.331   |
| vowel o               | 0.0176   | 0.1431     | 0.123   |
| gender male           | 0.1102   | 0.0306     | 3.598   |
| preceded by labial    | -0.0236  | 0.0655     | -0.361  |
| followed by labial    | -0.0475  | 0.0728     | -0.652  |
| vowel i × gender male | -0.1284  | 0.0437     | -2.937  |
| vowel a × gender male | -0.1382  | 0.0382     | -3.621  |
| vowel ε × gender male | -0.0522  | 0.0685     | -0.762  |
| vowel i × gender male | -0.1039  | 0.0519     | -2.001  |
| vowel o × gender male | -0.3284  | 0.0680     | -4.826  |

**Table 7:** Linear mixed effects regression for normalized vertical lip dot distances of all vowels. (/ʊ/ is the reference level.)

males) are the three speakers whose horizontal lip distances reveal no rounding in /o/ (**Figure 8**). One explanation for this is they are using vertical lip compression instead of rounding, a gestural property that could be shared with /ʊ/. Speakers 7 and 10 indeed have vertical lip distance for /o/ that is less than for all the other vowels, including /ʊ/. Speaker 8, on the other hand, shows more vertical lip distance in /o/ than in /ʊ/. The two speakers with the smallest vertical distance for /ʊ/ (speakers 3 and 8), appear to have minimal lip compression for /o/. All 10 speakers show at least a very small vertical lip dot difference between /i/ and /ʊ/, with the distance for /i/ being invariably bigger. In **Figure 7**, speaker 3 has a particularly large vertical difference between /i/ and /ʊ/, and considerable rounding for /o/, across all tokens in general. The apparent difference in jaw height between /i/ and /ʊ/ potentially facilitates a lingual difference between these two vowels. In confirmation of this hypothesis, the tongue tip is clearly in contact with the upper and lower teeth in the [i] image in **Figure 7**. It is also clearly visible in the [i] image there, although in a somewhat different position. Taken together, these facts indicate that there are no speakers who use lip compression in a similar way in /o/ and /ʊ/. On the other hand, there are several reasons for assuming that /i/ and /ʊ/ are produced with rather different tongue positions.

One purpose for examining the lip data is to determine whether a distinction in lip postures can account for the small F2 difference observed between /i/ and /ʊ/. Otherwise (if this is

not the case), a tongue position difference must account for it instead, since the lips cannot. Presumably this gestural difference involves tongue backness. In conclusion, we have found that the labial configuration does not reliably account for the contrast between /i/ and /u/ in Bora. Nevertheless, there is an apparent lingual difference visible in the images. In the following subsections we explore this articulatory difference and its possible acoustic consequences. Specifically, we examine how lingual-dental contact is distributed across the vowel categories and what role it plays in the production of /i/ (and possibly other vowels).

### 4.3. Tongue blade position

Although the video images were not collected for the purpose of examining the position of the tongue tip, videos for six of the 10 speakers proved suitable for this (speakers 1, 2, 3, 6, 9, and 10: three females and three males). The videos of the other four speakers resulted in either an infelicitous angle for seeing inside the mouth, or insufficient dynamic range. For the images produced by the six speakers listed above, tongue blade position was coded as follows. While blinded to vowel phoneme labels, the first author examined all 3344 images of the 678 tokens produced by these six speakers (on average about five frames per vowel), and determined that the tongue was (at least remotely) visible in 724 of them. Some of these frames are indicative of a tongue gesture as seen for /i/ in **Figure 7** that probably peaks near the middle of the vowel's duration. In certain other frames the tongue is visible due to a gesture relating to a neighboring consonant. In the latter cases it is expected that the contact appears to be greatest at the edge of the vowel (its beginning or end). Finally, some of the frames are expected simply because the mouth is wide open. The images where the tongue is visible to any degree whatsoever mainly involve /i/ ( $n = 399$ ), /i/ ( $n = 161$ ) and /a/ ( $n = 137$ ). The remaining 27 frames (out of 724) correspond to /ɛ/ (26 times) and /u/ (just once). None of the frames involving /o/ ever show the tongue visible at all.

Subsequently, these 724 images with a visible tongue were coded for lingual-dental contact by the first author and another trained phonetician, independently of each other, and both blinded to phone labeling. For the purpose of coding, lingual-dental contact was defined as whether the tongue visibly contacted any of the upper incisors (as seen for [i] in **Figure 7**). There was 76.1% agreement between the two coders on this initial step. Images where there was disagreement were then re-coded by both coders together (still blinded to phone labeling as well as the original coding), to reach a consensus. After this consultation, 479 of the 724 images were coded as having contact between the tongue and the upper incisors. These images were then coded again by both phoneticians according to whether the contact appeared to be central and laminal (as for a sound produced with lateral airflow) vs. primarily along the sides of the tongue (as for a typical vowel with mid-sagittal grooving and lateral bracing). There was 72.0% agreement between

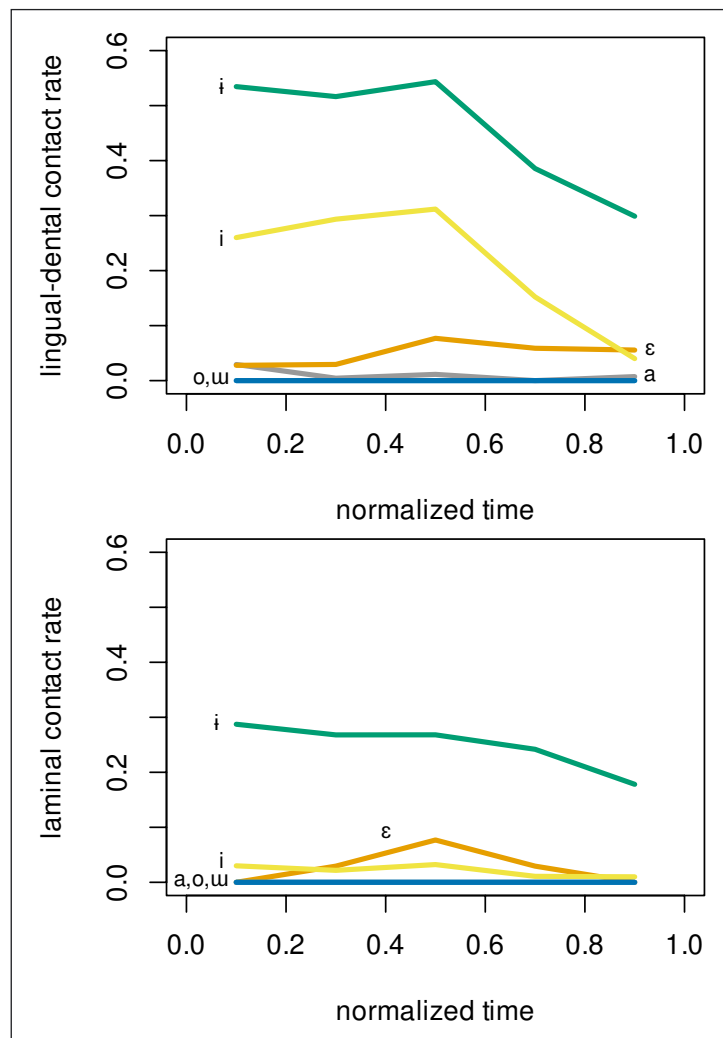
the two coders at this point. Images involving disagreement were then re-coded by both coders together (blinded to phone labeling and the original coding), to reach a consensus. Ultimately, 210 of the 479 images with visible lingual-dental contact had central laminal contact, and nearly all of these were images of /i/. The results are summarized in **Table 8**.

| Vowel | Images | Tongue visible |        | Lingual-dental contact |       | Central laminal contact |       |
|-------|--------|----------------|--------|------------------------|-------|-------------------------|-------|
| /a/   | 1178   | 137            | (12%)  | 13                     | (1%)  | 0                       | (0%)  |
| /ɛ/   | 166    | 26             | (16%)  | 8                      | (5%)  | 4                       | (2%)  |
| /i/   | 477    | 161            | (34%)  | 100                    | (21%) | 10                      | (2%)  |
| /i̥/  | 792    | 399            | (50%)  | 358                    | (45%) | 196                     | (25%) |
| /o/   | 134    | 0              | (0%)   | 0                      | (0%)  | 0                       | (0%)  |
| /ʊ/   | 597    | 1              | (0.2%) | 0                      | (0%)  | 0                       | (0%)  |
| TOTAL | 3344   | 724            |        | 479                    |       | 210                     |       |
| Vowel | Tokens | Tongue visible |        | Lingual-dental contact |       | Central laminal contact |       |
| /a/   | 259    | 54             | (21%)  | 10                     | (4%)  | 0                       | (0%)  |
| /ɛ/   | 38     | 10             | (26%)  | 4                      | (11%) | 1                       | (3%)  |
| /i/   | 77     | 36             | (47%)  | 24                     | (31%) | 2                       | (3%)  |
| /i̥/  | 147    | 87             | (59%)  | 78                     | (53%) | 43                      | (29%) |
| /o/   | 38     | 0              | (0%)   | 0                      | (0%)  | 0                       | (0%)  |
| /ʊ/   | 141    | 1              | (1%)   | 0                      | (0%)  | 0                       | (0%)  |
| TOTAL | 700    | 188            |        | 116                    |       | 46                      |       |

**Table 8:** Coding of visible lingual-dental contact by two phoneticians: Number of images (top) and number of tokens (bottom) exhibiting any tongue visibility or dental contact.

There were no images (frames) of /o/ or /ʊ/ where either coder reported lingual-dental contact. In all ten tokens where lingual-dental contact was observed in conjunction with /a/, this vowel is immediately preceded by /tʰ/ (items a–f in (4)). This indicates that the coronal consonant accounts for the lingual-dental contact in those cases, rather than the vowel /a/ itself. Among the remaining three vowels, 53% of /i̥/ tokens had visible lingual-dental contact, compared to 31% for /i/ and 11% for /ɛ/. Nearly all of the tokens of /i̥/ with visible lingual-dental contact had contact primarily involving the sides of the tongue, consistent with lateral bracing (typical for nonlow vowels and visible from the front of the mouth in extremely front vowels). 29% of the tokens of /i̥/ were produced with central laminal contact that was visible through the mouth opening.

**Figure 9** shows the rates of observed lingual-dental contact (of any type) and central laminal contact specifically, within five time bins. Lingual-dental contact for /a/ peaks (mildly) at the start of the vowel, consistent with it being caused by the preceding  $\widehat{ts}^h$ . For /i i ε/, lingual-dental contact peaks in the middle of the vowel interval. Furthermore, it is much higher overall in the first half of the vowel for both /i/ and /i/, before dropping steadily thereafter. Crucially, the increased lingual-dental contact in the first half of the vowel /i/ cannot be due to coarticulation with a preceding consonant since none of the /i/s in the words we elicited are adjacent to a consonant articulated with the tongue blade (4). In that list, one of the words with /i/ does have a nearby coronal consonant (/miʔtʃékpa/), yet lingual-dental contact is less frequent in this word

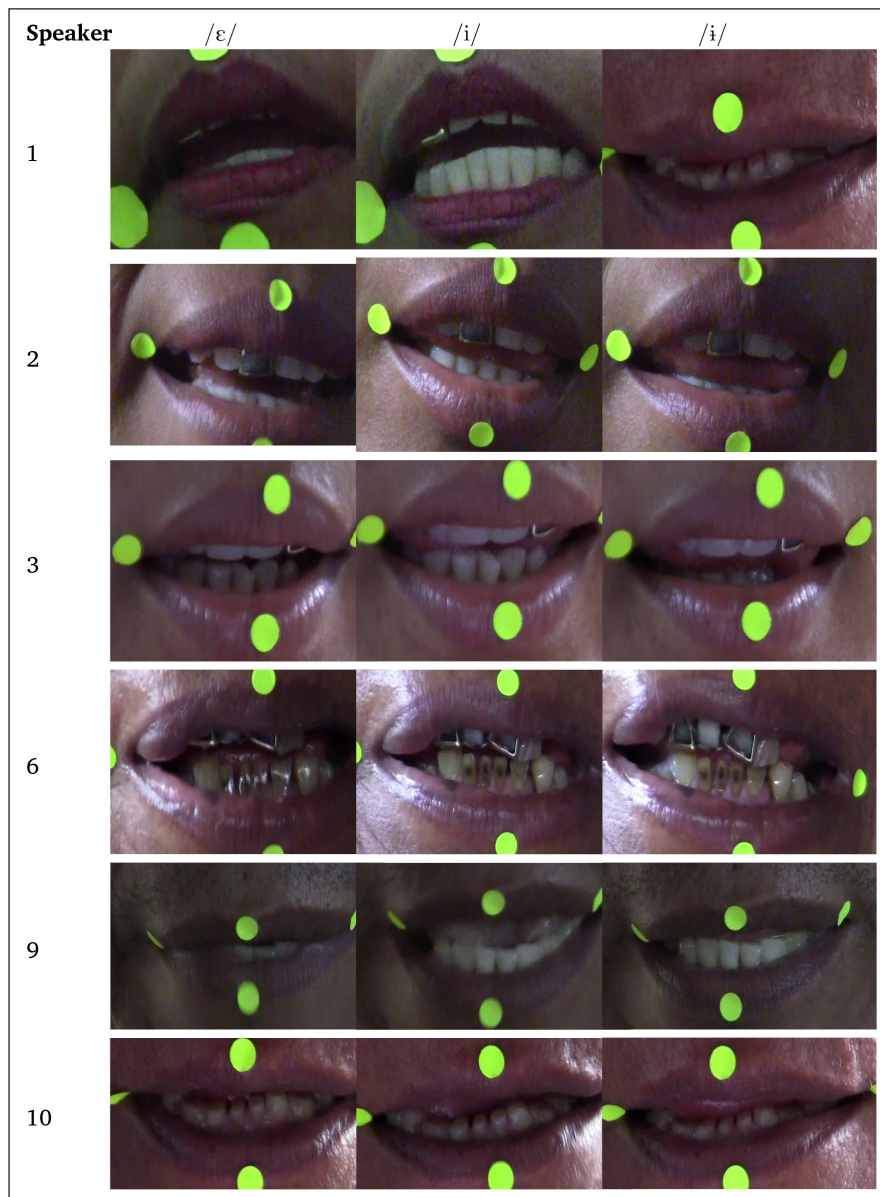


**Figure 9:** Coding of five speakers' vowel images by overall lingual-dental contact rate (top) and central laminal lingual-dental contact rate (bottom). Normalized time indicates the proportion of the vowel's total duration.



than in the others, where the /i/ is word-initial before a labial consonant (/ípa/ and /íkpáá/). Consequently the lingual-dental contact observed in /i/ and /i/ is inherent to the production of these vowels rather than being an effect of coarticulation with an adjacent segment. Only /i/ has appreciable central laminal contact between the tongue blade and teeth, and this is fairly evenly distributed throughout the vowel's duration, as seen in the bottom panel of **Figure 9**.

**Figure 10** shows the best exemplars of lingual-dental contact for /i i ε/ for each of the six coded speakers (i.e., the frames with the most contact or the closest to achieving contact). All



**Figure 10:** The most extreme lingual-dental contact frames involving the vowels /ε i i/ for each of the six visually-coded speakers.

of the /i/ images in **Figure 10** were coded as having lingual-dental contact. The /ε/s of speaker 6 and 9 were considered to have lingual-dental contact but in these two cases it appears to be incidental since the word is /miʔtʃékpa/. The token of speaker 6's /ε/ shown here is the one token of /ε/ that was coded as having central laminal contact (bottom half of **Table 8**). Among the six images of /i/ in **Figure 10**, all except speaker 1's were coded as involving lingual-dental contact. The tokens of /i/ shown here for speakers 2 and 3 are the two tokens of [i] that were coded as having central laminal contact. Both of these tokens have less visible tongue bulging and more apparent mid-sagittal grooving than the same speakers' tokens of /i/.

In summary, whereas /a/ and /ε/ are occasionally produced with some contact between the tongue tip and the lower teeth, the vowels /i/ and /i/ are frequently produced with considerable contact between the tongue and teeth, including the upper teeth. Such an anterior tongue position is not expected if /i/ is a canonical central vowel, as the traditional IPA symbol suggests. The dental contact observed in these images of Bora /i/ is typically central laminal contact with the upper incisors, often accompanied by bulging of the tongue blade between the upper and lower teeth. Such a constriction would cause complete obstruction of central (mid-sagittal) airflow, making /i/ a lateral vowel. In contrast to this, the lingual contact observed for /i/ does not appear to be laminal. Instead, the front of the tongue seems to be braced on either side of the mid-sagittal groove. This is typical of nonlow vowels, and is normally visible through the mouth opening in front vowels. Consequently, we posit that central laminal contact is a defining characteristic of Bora /i/ since it is the only vowel regularly coded as having this feature. The fact that it was observed in only 29% of the tokens is likely due in large part to the difficulty of seeing the tongue in some of the frames that were coded. This again highlights the need for more direct and precise data on tongue positions.

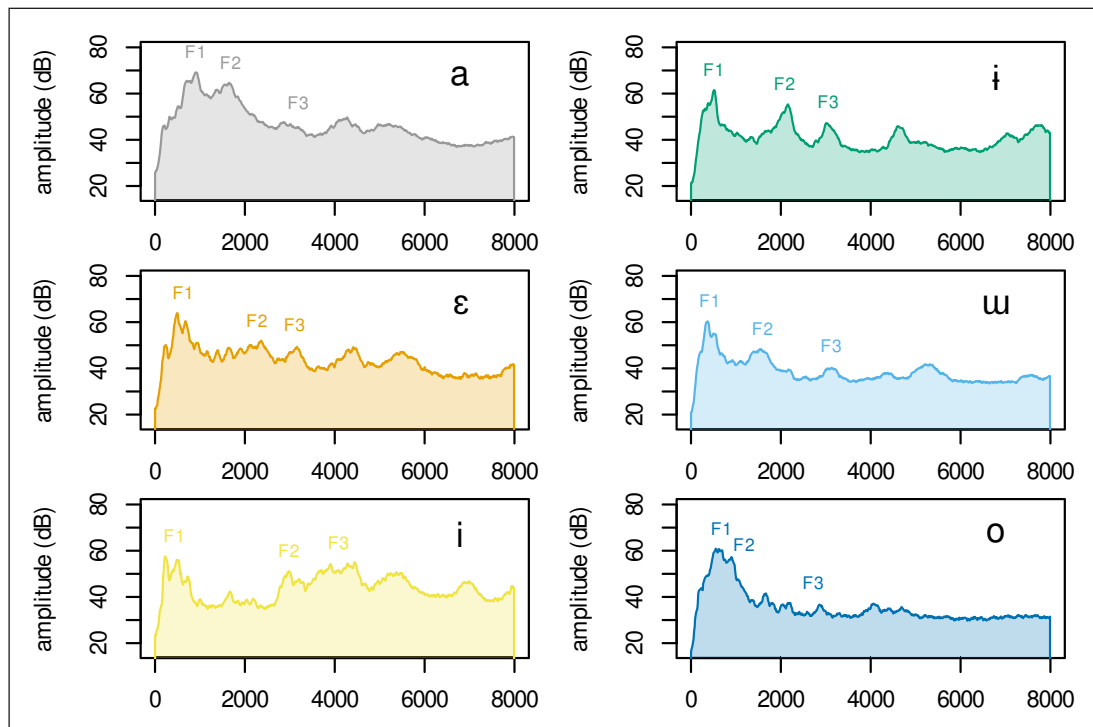
With respect to the production of /u/, its closed mouth configuration makes it difficult to observe where the tongue is located. Nevertheless, there is no positive evidence that the tongue blade is near the front of the mouth for this vowel. Furthermore, the tongue is never seen to be pressed out between the teeth, as frequently occurs in the production of /i/. Nor is there any reason to suspect that this would happen since all of the phonetic and phonological data converges on the fact that /u/ is a back vowel. Accordingly, the articulatory differences that can be visually observed between /i/ and /u/ is that /i/ is normally produced with the tongue pressed against the back of the upper and lower teeth, whereas for /u/ the jaw is raised and the upper and lower teeth are close together, with no apparent lingual-dental contact.

In summary, the difference between the phonemes /i/ and /i/ in Bora is basically the presence of central laminal lingual-dental contact for /i/, with concomitant lateral airflow. The next subsection further explores the acoustic consequences of these articulatory features. The involvement of the tongue blade in the production of this segment could entail that it is a fricative vowel (see, e.g., Faytak, 2018; Lee-Kim, 2014). If so, we would expect /i/ to exhibit more high frequency noise and a lower harmonics-to-noise ratio than /u/ does. In addition,

the apparent lateral airflow in /i/ could cause it to show the presence of antiformants or other acoustic correlates of laterality. Nevertheless, as we will see, none of these features in fact turn out to be present.

#### 4.4. Spectral shape

Multitaper spectra were calculated for the middle 50% of each vowel token for all 10 speakers, as described in §3.3.3. **Figure 11** shows the mean spectrum for one female speaker's vowels. The F1, F2, and F3 labels represent the median formant frequencies for each vowel shown, and are centered over the spectral peaks corresponding to these formants. The shapes of the spectral peaks clearly vary across vowels.



**Figure 11:** Mean spectra of each vowel for speaker 1 (a female). The x axis in each panel displays frequency in Hz.

Since lingual-dental contact could produce frication in /i/ or /i/, we measured harmonics-to-noise ratio for all of the vowels, using the default settings in Praat. If /i/ is a fricative vowel, we expect this ratio to be lower for /i/ than for /u/, since /u/'s energy is basically periodic whereas /i/ is (hypothetically) a mixture of periodic signal and noise. **Table 9** reports a linear mixed effects regression comparing the three high vowels of all ten speakers, with the same random

| harmonics-to-noise ratio              | Estimate | Std. Error | t value |
|---------------------------------------|----------|------------|---------|
| (Intercept)                           | 6.0233   | 1.5471     | 3.893   |
| vowel $\text{u}$                      | -0.5742  | 1.2813     | -0.448  |
| vowel $\text{i}$                      | -0.9018  | 1.5619     | -0.577  |
| gender male                           | -2.0793  | 1.8592     | -1.118  |
| vowel $\text{u}$ $\times$ gender male | 0.4540   | 0.7716     | 0.588   |
| vowel $\text{i}$ $\times$ gender male | 4.1053   | 0.9240     | 4.443   |

**Table 9:** Linear mixed effects regression for harmonics-to-noise ratio among high vowels. (/i/ is the reference level.)

effect structure as the formant frequency models above. However, for the regression models here (involving just the high vowels), /i/ is used as the reference level in order to facilitate comparing it with its two immediate neighbors. There are no significant main effects for vowel or gender. Furthermore, the small estimate of a difference between /i/ and /u/ is in the opposite direction of what we would expect if /i/ is a particularly noisy vowel. There is a significant interaction between /i/ and gender, driven by the fact that /i/ has particularly high harmonics-to-noise ratio for some males. However, there is no indication that /i/ is noisier than /u/. This lack of noise coincides with the impression that /i/ does not sound to us as fricated. In addition, the spectra in **Figure 11** do not in general exhibit greater high-frequency energy in /i/ than in /u/. Nor do we observe any signs of antiformants in /i/'s spectra.

Although aperiodic noise does not appear to distinguish /i/ from /u/, many of the spectra, including those in **Figure 11**, indicate that /i/ has more prominent spectral peaks than /u/ does. By this we mean that they are higher relative to the troughs between the formants. Formant amplitude was measured as the maximum amplitude (in dB) in the vicinity of each measured formant frequency. For these purposes *vicinity* was defined as a 200-Hz window centered on the measured formant frequency. In the event that two formants appear within 200 Hz of each other, the boundary was set at the midpoint between those two formant frequencies in order to keep the amplitude measurement windows from overlapping.

As shown in **Figure 12**, there are many differences in formant amplitude between the six vowels. Among these, it is especially noteworthy that /i/ and /u/ have a consistent difference in the amplitude of F2 and F3, even though the two segments are so similar in terms of their formant frequencies.<sup>6</sup>

---

<sup>6</sup> The recordings of speaker 3 contain only one usable token of /o/, so her values for that vowel are omitted from these plots.

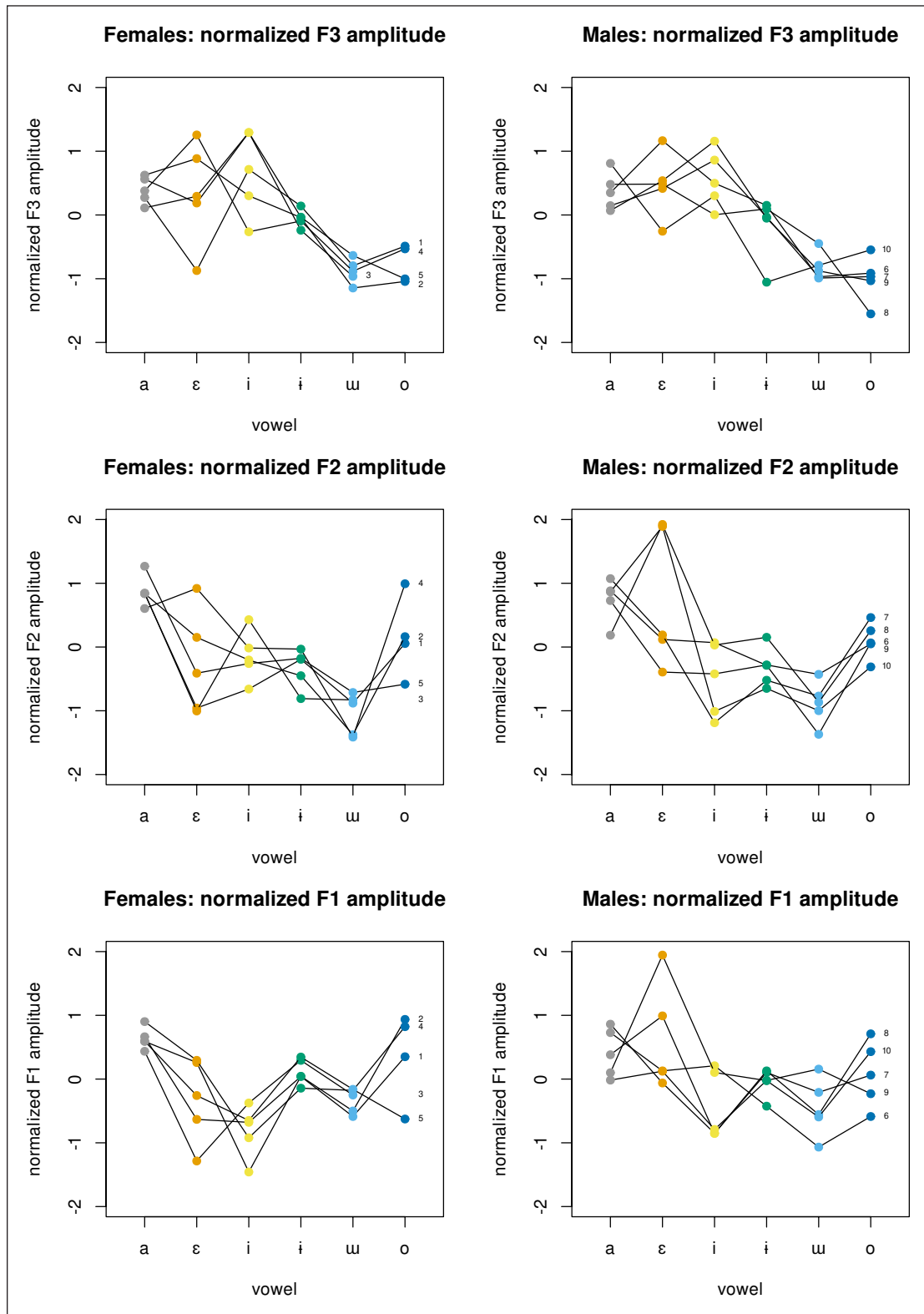


Figure 12: Median formant amplitudes of each vowel for all speakers.

**Table 10** summarizes three linear mixed effects regressions for formant amplitude among the three high vowels. We include gender as a factor and use the same random effect structure as the models of formant frequency and harmonics-to-noise ratio in previous tables. /ʊ/ has significantly lower F2 and F3 amplitude than /i/. Also, /i/ has a significantly higher F3 amplitude than /i/. There are no gender effects involving the contrast between /i/ and /ʊ/. However, there are significant interactions involving /i/ for F1 and F2, with males having rather higher-amplitude F1 and somewhat lower-amplitude F2 than females.

| <b>F3 Amplitude</b>   | <b>Estimate</b> | <b>Std. Error</b> | <b>t value</b> |
|-----------------------|-----------------|-------------------|----------------|
| (Intercept)           | -0.1037         | 0.1880            | -0.551         |
| vowel ʊ               | -0.7376         | 0.2398            | -3.076         |
| vowel i               | 0.7039          | 0.2893            | 2.433          |
| gender male           | -0.1109         | 0.1630            | -0.680         |
| vowel ʊ × gender male | 0.1594          | 0.1599            | 0.997          |
| vowel i × gender male | 0.0610          | 0.1804            | 0.338          |
| <b>F2 Amplitude</b>   | <b>Estimate</b> | <b>Std. Error</b> | <b>t value</b> |
| (Intercept)           | -0.3897         | 0.1704            | -2.287         |
| vowel ʊ               | -0.6038         | 0.2331            | -2.591         |
| vowel i               | 0.1439          | 0.2816            | 0.511          |
| gender male           | 0.0832          | 0.1218            | 0.683          |
| vowel ʊ × gender male | -0.0660         | 0.1447            | -0.456         |
| vowel i × gender male | -0.3310         | 0.1633            | -2.027         |
| <b>F1 Amplitude</b>   | <b>Estimate</b> | <b>Std. Error</b> | <b>t value</b> |
| (Intercept)           | -0.1095         | 0.2969            | -0.369         |
| vowel ʊ               | -0.3197         | 0.4055            | -0.788         |
| vowel i               | -0.7823         | 0.4938            | -1.584         |
| gender male           | -0.2467         | 0.1586            | -1.556         |
| vowel ʊ × gender male | 0.1347          | 0.1578            | 0.853          |
| vowel i × gender male | 0.8525          | 0.1779            | 4.792          |

**Table 10:** Linear mixed effects regressions for formant amplitude among high vowels. (/i/ is the reference level.)

In conclusion, in these spectra we do not observe noise consistent with frication in any of the vowels. The presence of visible lingual contact in /i/, but not in /ʊ/, combined with greater formant amplitudes in /i/ relative to /ʊ/, indicates that the two vowels are probably articulated quite differently, despite their close proximity in the F1-F2 space.

The relationship between formant amplitude and vocal tract configuration is fleshed out in the next section, and further supports our claim that Bora /i/ should be classified as front rather than central.

## 5. Articulatory estimates

Obviously it would be extremely helpful to have direct articulatory images in order to pinpoint the tongue positions used to produce Bora vowels, beyond what is visible through the mouth opening. Accordingly, we have made multiple attempts to collect ultrasound data on location in Peru, but these have been unsuccessful due to logistical factors beyond our control. Furthermore, a work-around solution may not be feasible for a long time.<sup>7</sup> Nevertheless, what we can do at the moment is apply available quantitative methods to estimate vocal tract shapes and compare these with the articulatory evidence from video images.

The perturbation theory of vowel acoustics (Chiba & Kajiyama, 1941) models the relationship between changes in formant frequency and constrictions at particular locations in the vocal tract. Rice and Öhman (1976) also show that constrictions at a different set of vocal tract locations are associated with changes in formant bandwidth. Building on these facts, Iskarous (2010) proposes a method for estimating vocal tract shape parameters from formant frequencies and amplitudes, using amplitude as a proxy for bandwidth. In this section we apply Iskarous's model to estimate the vocal tract configurations of Bora vowels from their formant frequency and amplitude measurements. See Appendix A for more details about this technique, including an illustration of how it relates the acoustic and articulatory parameters that are of particular interest for Bora's high vowels.

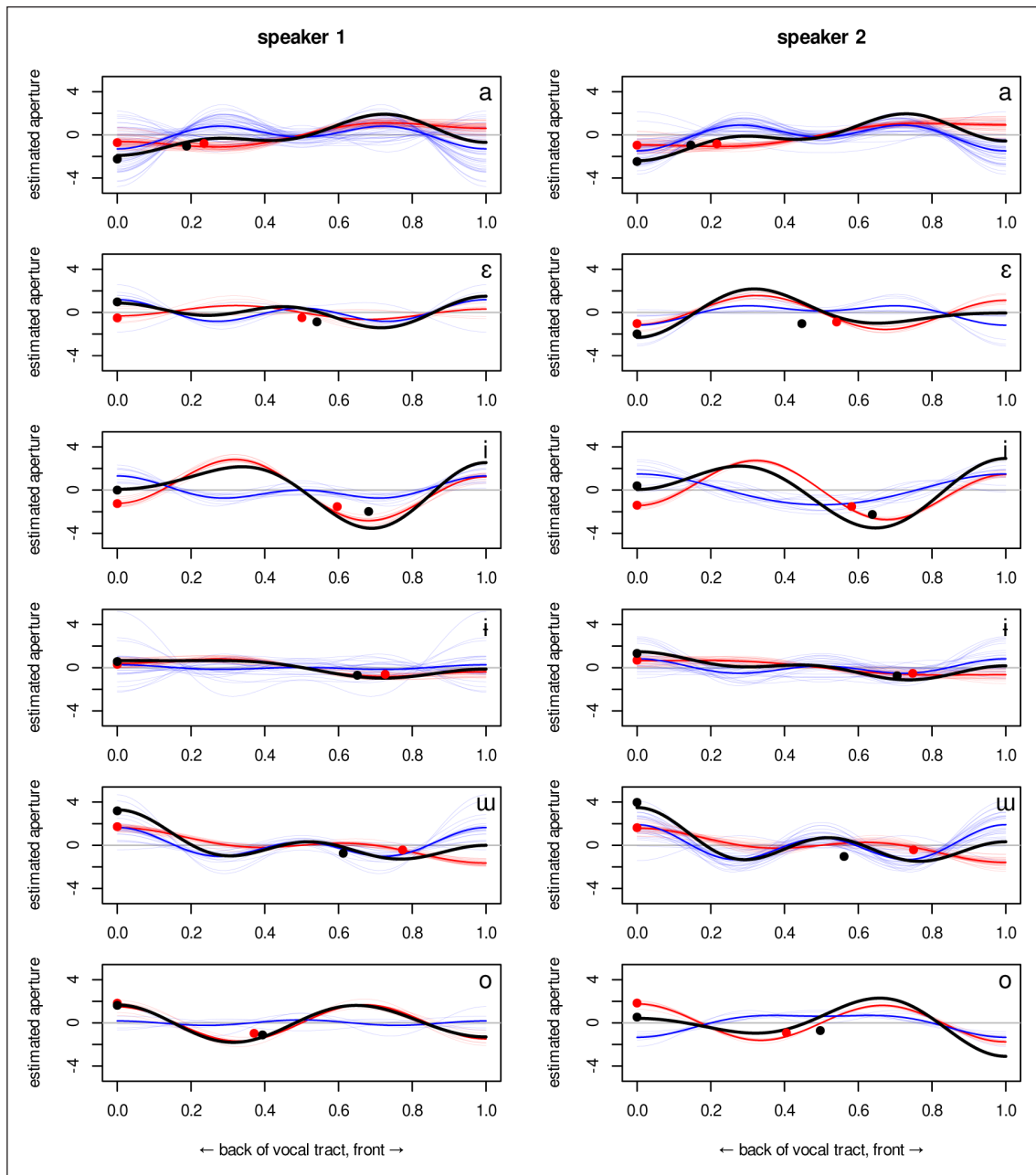
Both the frequencies and amplitudes of F1 and F2 were z-scored and then used to estimate area functions for all Bora vowel tokens. Each estimated area function is expressed by a plot in which  $x$  values represent positions along the vocal tract and  $y$  values represent how wide the vocal tract is at each  $x$  value. For each area function, the estimated lip aperture was defined as the  $y$ -value at the lips ( $x = 1$ ). The lingual constriction location is defined as the weighted average of the non-labial vocal tract locations with constrictions, and the lingual constriction degree is the mean deviation from a neutral vocal tract for these constrictions.

**Figure 13** shows the estimated vocal tract shapes for the six vowel phonemes for two female speakers. The dark black contour is the average vocal tract shape calculated by averaging across all the estimated vocal tracts for that vowel category. The thick red contour is the part of the vocal tract shape contributed by formant frequencies. The thick blue contour is the part of the vocal tract shape contributed by formant amplitudes. The thin red and blue contours represent

---

<sup>7</sup> This was the situation when we submitted this paper. Recently we obtained articulatory data which we are just beginning to analyze.



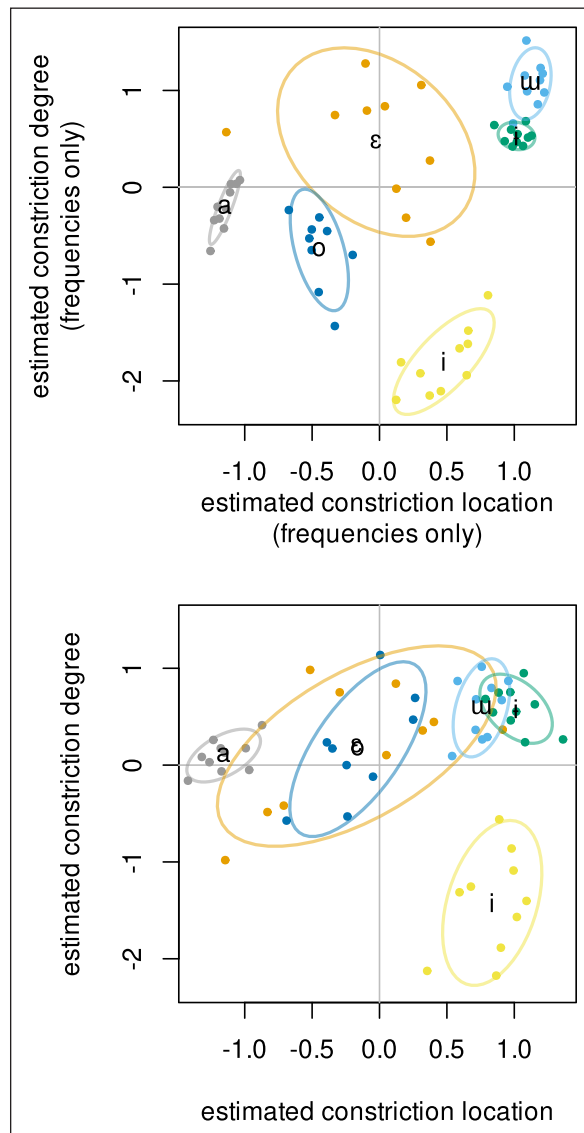


**Figure 13:** Mean estimated vocal tract area functions for six vowel categories for two female Bora speakers. The  $x$ -axis ranges from 0 (the glottis) to 1 (the lips). The  $y$ -axis values can be interpreted as  $z$ -scored deviations from a uniform tube.

these components of individual tokens. The estimated aperture values are to be interpreted as  $z$ -scored deviations from a uniform tube, where negative values indicate constriction and positive values indicate expansion. The  $x$ -axis values are to be interpreted as ranging from the glottis ( $x = 0$ ) to the lips ( $x = 1$ ). The two black dots in each panel indicate (respectively) the estimated

average (1) lip aperture and (2) constriction location and degree. The red dots indicate the estimates for constriction location and degree based on formant frequencies only. The difference between the red dots and the corresponding black dots is due to the contribution of formant amplitudes to the computation.

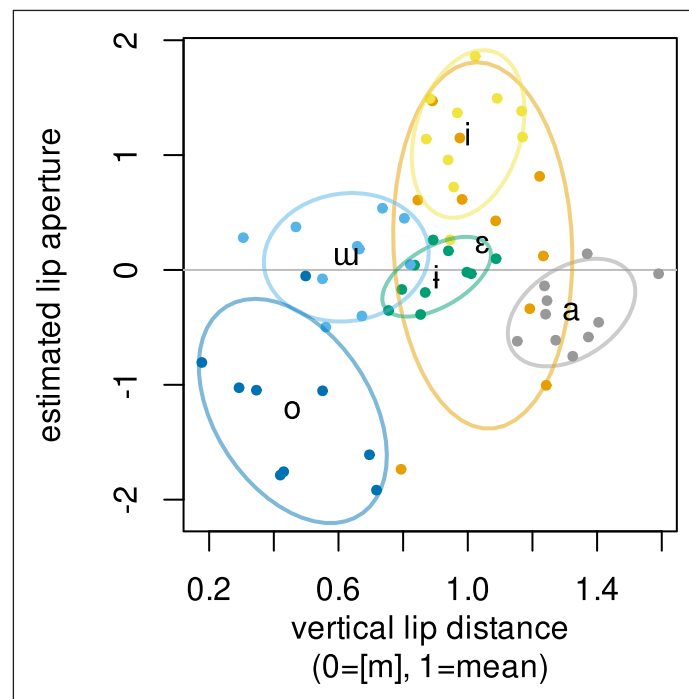
**Figure 14** shows the median estimated constriction location and degree for each vowel for each speaker. /a/ has its estimated constriction primarily in the posterior part of the vocal tract,



**Figure 14:** Estimated constriction location and degree for all ten speakers, by vowel. Values are z-scores. The estimates on the top are based only on formant frequency. The estimates on the bottom are based on formant frequency and amplitude. Dots represent speaker medians. Ellipses represent two standard deviations around mean values.

and /i/ has its estimated constriction location toward the front. /o/ and /ε/ are intermediate in location. The fact that /a/ and /o/ differ primarily in constriction location and /i/ and /ε/ differ primarily in constriction degree is typical of how back and front vowels of different heights are articulated (Wood, 1979). The vowels /i/ and /u/ are estimated to have the most open vocal tracts, with relatively anterior constriction locations. When only formant frequency is taken into account (the top panel in **Figure 14**), /i/ and /u/ have similar (anterior) constriction locations. However, when formant amplitude is added to the estimates (the bottom panel), /u/'s constriction location is more posterior, yet still anterior to /o/. This is typical of how vowels characterized as /u/ (or its unrounded counterpart) are articulated (Wood, 1979). That is, they involve a constriction location at the soft palate, unlike the nonhigh back vowels, which are articulated in the pharynx. This estimate of /u/'s constriction location is driven by the low amplitude of its F2, which moves /u/'s constriction farther back.

**Figure 15** shows the median estimated lip aperture for each vowel for each speaker. In §4.2 we observed that only /o/ involves lip rounding in the horizontal dimension; apart from that, there is very little difference between the other vowels with respect to the horizontal dots. In the dimension of vertical lip distance, however, the vowels in **Figure 15** are relatively spread out,



**Figure 15:** Estimated lip aperture compared to measured vertical lip distance for all ten speakers, by vowel. y-axis values are z-scores. Measured lip distances are normalized as described in §3.3.2. Dots represent speaker medians. Ellipses represent two standard deviations around mean values.

a fact which appears to be related to jaw height. We include measured vertical lip distance on the *x*-axis of this figure and note that the lip aperture values estimated from formant frequency (*y*-axis) are similar to the observed vertical lip distances (see **Figure 8** and **Table 7**). Nevertheless, /ʊ/ is estimated to have a somewhat larger lip opening than we observed, and /a/ is estimated to have a somewhat smaller lip opening relative to the other vowels.

In conclusion, the application of Iskarous' procedure to Bora vowels has yielded two significant results: (1) it corroborates our claim that /i/ involves a more anterior tongue position than /ʊ/ does; and (2) it shows that this difference in vocal tract shape can account for the higher amplitude of /i/'s formants.

## 6. Discussion

In this paper we have sought to account for the three-way contrast among high vowels in Bora. We have replicated Parker's (2001) observation that /i/, /i/, and /ʊ/ yield similar F1 values but are consistently distinguished by F2 and, to a lesser extent, F3. Furthermore, Parker (2001) shows that the duration differences between these vowels are minimal, so the phonological contrast between them cannot be ascribed to length.

A priori, the most likely explanation for the different F2 frequencies among the three high vowels is that they differ in tongue backness and/or lip rounding. With respect to the latter, we observe systematic labial differences across certain vowels in the vertical plane only (the distance between the central dots placed on the upper and lower lips). Specifically, the low vowel /a/ has the largest vertical lip opening and the rounded vowel /o/ has the smallest. Furthermore, among the high vowels, /ʊ/ is produced with consistently more closed lips than /i/ and /i/. In the horizontal dimension, however, we have shown that none of the high vowels is produced with pursing of the lips, so /ʊ/ is [–round] in a categorical sense. /ʊ/ is also generally produced with the upper and lower teeth close together.

In terms of tongue body position, the lingual-dental contact observed in /i/ (and to a lesser extent in /i/) indicates that the tongue is very far front in the mouth for /i/. This in turn entails very different tongue postures for /i/ and /ʊ/. This is corroborated by the estimated vocal tract parameters derived from F1 and F2 frequency and amplitude, using Iskarous' (2010) method. This procedure estimates a front vowel-like constriction location for /i/ and a central vowel-like constriction location for /ʊ/. We hypothesize that the observed vertical lip differences among high vowels are related to the jaw height necessary to produce a dorsal constriction in /ʊ/, but a more anterior constriction in /i/ and /i/. Nevertheless, the parameter of vertical lip distance is correlated with estimated lip aperture, suggesting that the constriction in the front of the mouth for /i/ has an acoustic effect similar to lip rounding, namely, lowering of F2 and F3.

The analysis of the mouth videos indicates that lingual-dental contact in /i/ is central, suggesting that this vowel is lateral. For /i/, however, the lingual-dental contact appears to be more lateral, with central airflow. The type of contact observed in /i/ further implies that it should differ acoustically from /u/ in ways other than formant frequency. However, the lack of any difference in harmonics-to-noise ratio, together with the lack of any fricative perceptual quality, indicates that /i/ is not a fricative vowel. Nevertheless, we do observe that /i/ and /i/ have higher F2 and F3 amplitudes than /u/ does. The fact that /i/ and /u/ differ only slightly in the *frequency* of their formants, yet differ consistently in the *shape* of their formants, suggests that they yield similar formant frequencies via different articulatory routes. This is consistent with the observation that /u/ is produced with a raised jaw (evidenced by the overlapping upper and lower teeth), which can facilitate raising and backing of the tongue body to produce a back unrounded vowel. This conclusion is also confirmed by the estimated vocal tract parameters.

In summary, /i/ can be understood acoustically as being similar to an extremely front rounded vowel such as [y], wherein the first three formants are resonances of the long back cavity and the narrow tube formed by the tongue passage and the lips (Fant, 1970). In this case, however, the narrow opening of the vocal tract is due to the lingual-dental contact rather than extreme lip rounding. The F2 of /i/ is more likely to be a front cavity resonance, with F1 and F3 due to the back cavity. As such, the similar formant frequencies of /i/ and /u/ are not necessarily caused by analogous vocal tract configurations.

Since the lingual-dental contact in /i/ is central and the airstream (presumably) flows around the sides of the constriction, /i/ can be described articulatorily as a lateral vowel. If laterality is a key feature of this vowel, we expect to see antiformants caused by the side cavity formed by the air passing over the tongue. Acoustic-articulatory modeling by Zhou (2009) demonstrates that in order for a lateral sound to have an antiformant in the region where F3, F4, and F5 typically occur, the lateral channels need to be 3–6 cm long and asymmetrical. The constriction we have observed for /i/ is too anterior for lateral channels of this length, so it is not surprising that we have not observed antiformants in /i/. The upshot of this situation is that this vowel may be articulatorily lateral without manifesting the normal acoustic feature of laterality, just as it is unrounded while appearing to share the articulatory basis of its formants with a front rounded vowel. In fact, some English-speaking listeners remark that Bora /i/ (but not /u/) sounds like it has a [ø] or [l] quality, which is consistent with its dental contact and apparent lateral airflow.

We conclude that the most accurate way to characterize the Bora vowel previously described as high central unrounded (/i/) is as a high front dental vowel. Once this middle high vowel is reinterpreted as front, the main research question is no longer how to distinguish the two

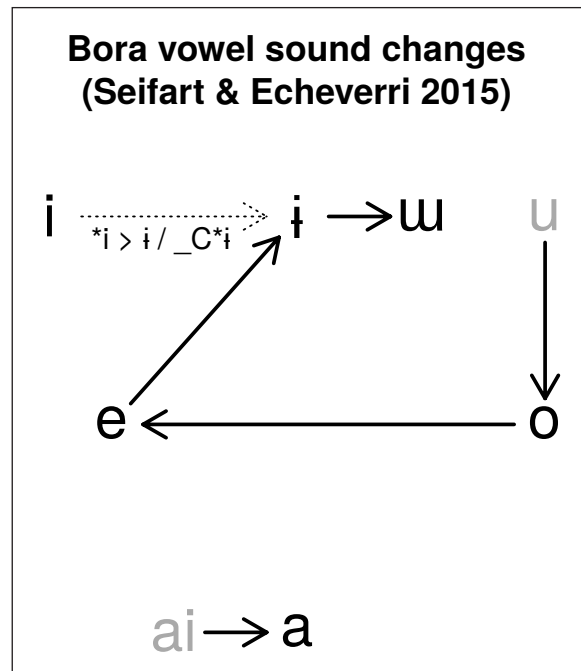
high non-front vowels (/i u/), but how to distinguish the two high front vowels (/i i/). What actually distinguishes the vowel traditionally transcribed as /i/ from /i/ articulatorily is its tight lingual-dental constriction, which in turn probably drives its lower F2 frequency. We propose to transcribe this vowel as /i̠/, highlighting the dental contact by using the IPA diacritic for dental. This is typically invoked only for consonants, but seems like the most appropriate option in this case. If more vowels with lateral airflow are discovered in other languages (i.e., with central constrictions at places other than the teeth), it might be appropriate to group together all such vowels as lateral, rather than focus on the dental place of articulation specifically in Bora. Alternatively, since lingual-dental contact is observed in both /i/ and /i/ in Bora, another possibility is to specify /i/ as dental and /i/ as both dental and lateral. However, since lateral bracing is observed in most speech sounds (Gick, Allen, Roewer-Després, & Stavness, 2017) yet central contact is rare in vowels, we choose for the moment to focus on the unusual lingual-dental contact in /i̠/.

**Table 11** summarizes our proposed reclassification of the Bora vowel inventory (compare with the traditional **Table 1** above). One implication of our account is that Bora no longer provides evidence supporting a three-way distinction in the feature [back], contra Parker (2001). This is a partial vindication of Duanmu's (2016) approach. Nevertheless, we still require a novel binary feature to distinguish between /i/ and /i̠/ since neither [high], [round], nor [ATR] is appropriate. As suggested above, one possibility is [lateral]. Another option is to specify /i̠/ as [+distributed], as is done for dental consonants. We leave this topic for a subsequent paper.

|      | front<br>non-dental | front<br>dental | back<br>unrounded | back<br>rounded |
|------|---------------------|-----------------|-------------------|-----------------|
| high | i                   | i̠              | u                 |                 |
| mid  | ɛ                   |                 |                   | o               |
| low  |                     |                 | a                 |                 |

**Table 11:** Revised inventory of Bora vowels, reanalyzing /i/ as front and dental (/i̠/) rather than central, and confirming /u/ as back unrounded.

Some clues about the phonological relationships among Bora vowels are found in the reconstructed Proto-Bora-Muinane vowel system. **Figure 16** illustrates the vowel sound changes posited by Seifart and Echeverri (2015) to have occurred between Proto-Bora-Muinane and Bora. According to their analysis, the present-day non-back high vowel set /i i/ were historically the front vowels /\*i \*e/, respectively. Aschmann (1993), on the other hand, claims that Bora's vowel system is conservative relative to Muinane, with only one sound change in Bora (\*ai > a), and a major vowel rotation (in the opposite direction) in Muinane. Either of these two diachronic analyses sheds some light on why /i/ favors palatalization more than the other member of this



**Figure 16:** Historical sound changes between Proto-Bora-Muinane and Bora, as posited by Seifart and Echeverri (2015). The dotted arrow indicates a context-sensitive change and the solid arrows indicate context-free changes. Gray vowel symbols indicate Proto-Bora-Muinane vowels not found in Bora today.

vowel class does (/i/ or /\*e/). In Seifart and Echeverri's reconstruction, the present-day front vowel set /i ε/ was /\*i \*o/, and the current back vowel set /u o/ was historically the non-front high vowels /\*i \*u/, respectively. Since /i/ (= /i/) evolved from /\*e/, and since palatalization apparently dates back to before the /\*ai/ > /a/ sound change, it is possible that the palatalization blockers (present-day /i/ and /i/) were established at a stage of the language in which they were the front vowel class /\*i \*e/. Furthermore, the instrumental data we have reported here indicates that the sound change which Seifart and Echeverri (2015) describe as /\*e/ > /i/ (raising and centralization) can now more accurately be described as raising and dentalization (//\*e/ > /i/). Thus, part of the diachronic motivation for the robust lingual-dental contact in Bora /i/ (erstwhile /i/) is to prevent it from merging completely with historical /\*i/, given that its predecessor (//\*e/) was pressured to raise from mid to high due to the encroachment of proto /\*o/ into the front mid space. In other words, the current situation may have resulted from a rather complicated push chain effect. To the degree that this hypothesis is correct, it supports Vaux and Samuels' (2015) claim that vowel inventories which are non-optimally dispersed, such as Bora's, may be better accounted for by evolutionary forces rather than synchronic constraints on maximal perceptibility.

Languages described as having both /i/ and /u/ are rare, but not vanishingly rare. As noted in §2.2.1, 14 out of 2101 languages (0.67%) in PHOIBLE (Moran & McCloy, 2019) contain both



/i/ and /ɯ/. We have also seen that /i/ is particularly frequent among inventories in South America. Engstrand et al. (1998) and Björsten and Engstrand (1999) even suggest that the high rate of occurrence of /i/ in South American languages may indicate that it is actually some other kind of vowel rather than a prototypical [i]. Specifically, they speculate that in some cases it may be an apical or fricativized vowel, or a segment similar to Swedish Viby-colored vowels. Swedish Viby-colored /i: y:/ are described as having a damped or buzzy quality and a lower-than-expected F2 compared to cardinal high front vowels. Some of the articulatory details of Swedish Viby-colored /i: y:/ are still unclear, including whether the main point of constriction is more anterior or more posterior than in typical high front vowels (Engstrand et al., 1998; Schötz, Frid, & Löfqvist, 2011). Ultimately, these segments may be similar to Bora's /i/, which clearly involves the tongue blade and a much more anterior tongue posture than would normally correspond to a vowel with such a low F2.

## 7. Conclusion

The lip position data analyzed in this study confirms that Bora /i/ and /ɯ/ are both [–round] vowels. We also maintain that /ɯ/ is a high back vowel. However, we have posited that the previously-observed differences in F2 and F3 between them are best accounted for by analyzing /i/ as a front vowel (rather than a central vowel) with concomitant lingual-dental contact that results in an acoustic effect similar to lip rounding. The conclusion that these two vowels are produced with rather different articulatory configurations is somewhat surprising given their close proximity in the F1-F2-F3 space. Nevertheless, it is consistent with the observation that they differ in the amplitude of their formants. This in turn was shown to coincide with the vocal tract parameters for Bora's vowels, as estimated from their observed formant frequencies and amplitudes using Iskarous' (2010) method. /ɯ/ is produced with a raised jaw, which affects the lip positions in a way that is distinct from the lip rounding observed for /o/. This fact is also consistent with the assumption that /ɯ/, but not /i/, is made with a dorsal constriction in the back of the oral cavity. In light of these observations, we have proposed that a more phonetically accurate transcription of /i/ is /i̟/. This vowel is produced with lateral airflow due to the contact between the tongue blade and the central incisors. The vertical lip/jaw difference between /i/ and /ɯ/ is comparable to that between /a/ and /i/. Precise measurements of lingual positions (beyond what can be observed through the teeth) could confirm these claims. This will be an important next step, but has not been feasible to date.

These phonetic findings are consistent with the phonological observations documented in Parker and Mielke (to appear), and summarized in (3) above. In the vowel system of Bora, /ɯ/ patterns with the back vowels /a/ and /o/, while /i/ patterns with the front vowel /ɛ/. /i/ does not pattern with back vowels at all, but it does pattern with the front vowel /i/ in blocking

palatalization. As we noted, this might be attributed to its historical status as /\*e/ in Proto-Bora-Muinane. Furthermore, /i/ harmonizes with /i/ to the exclusion of /u/. /i/ also overlaps with the front vowel /i/ in that both of them can be realized as [i], although in different contexts.

Among the world's languages, the Bora vowel system is most likely unique in that it includes the dental vowel /i/, and its /i/ also involves some lingual-dental contact. Several phonological patterns crucially require the contrast between /i/ and /u/ to be referred to primarily by a difference in tongue backness. In future work we plan to more explicitly address the featural specifications of these segments as part of a formal explanation for their phonotactic behavior. A principal contribution of the present study is the convergence of new phonetic facts with previously observed phonological patterns to confirm that the three segments /i i u/ are all [+high] and [−round], while differing primarily in the position of the tongue body and tongue blade. Some of the other 13 languages in PHOIBLE described as containing both /i/ and /u/ might involve a similar tongue blade difference. Nevertheless, since the Bora segment previously transcribed as /i/ has not been documented previously (as far as we are aware), most of these other languages probably exhibit a more straightforward central-back contrast. These findings thus highlight the need to carry out instrumental studies to pinpoint what are impressionistically transcribed as analogous vowels in various languages.

---

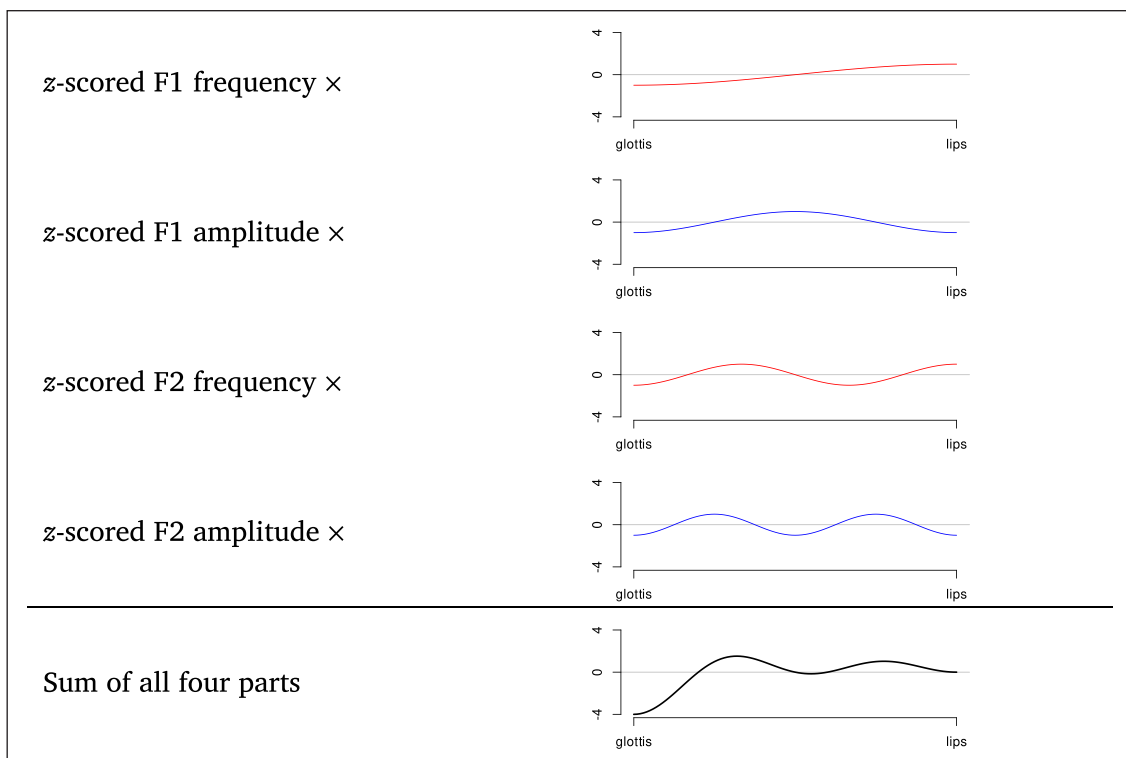
## Appendix A. A more detailed explanation of vocal tract estimation

Iskarous' (2010) technique is based on the observation that a spatial discrete Fourier transform of the logarithm of the vocal tract area function yields components which are linearly related to formant frequencies (Mermelstein, 1967; Schroeder, 1967). Specifically, the antisymmetric components are related to the resonant frequencies of a vocal tract that is closed at one end and open at the other. The more the vocal tract area function resembles a half period of a cosine wave with a constriction near the glottis and an expansion in the lip opening, the higher the frequency of F1 (as in [a]; see the illustration in Schroeder 1967). Conversely, the more it resembles the opposite of this, the lower the frequency of F1 (as in [i]). This is the basis of acoustic perturbation theory (Chiba & Kajiyama, 1941). Chiba and Kajiyama's classic figure showing the first four standing waves of the vocal tract (not pictured here) illustrates the relationship between the first four formant frequencies and the first four antisymmetric sinusoid components. Constricting the tube at a velocity antinode decreases the frequency of the corresponding formant, while constricting the tube at a velocity node increases the frequency of that formant. Enlarging a uniform tube at one of these places has an effect in the opposite direction, but much smaller in magnitude (McGowan, 2020). Given these relationships, it is possible to estimate the vocal tract shape on the basis of formant frequencies, but the antisymmetric components alone do not provide a unique solution (for illustrations of the nonuniqueness of solutions based solely on antisymmetric components, see Mermelstein 1967).

Schroeder (1967) observes that formant frequencies of a tube open at one end are most helpful in modeling vowels (like [i] and [a]) which have a constriction close to one end of the vocal tract but worst for vowels (like [u], or [ɪ] and [ʊ]) that have a constriction near the middle. A more precise estimate of vocal tract area function is achieved by including the symmetric components as well. The symmetric components are associated with the resonances of a tube closed at both ends, but they are also associated with the vocal tract impedance function measured at the lips (Schroeder, 1967), and with formant *bandwidths* (Rice & Öhman, 1976). Iskarous' (2010) procedure makes use of formant bandwidths because, unlike the resonances of a closed tube and the vocal tract impedance function, bandwidths are accessible to listeners during vowel production. Although Schroeder (1967) shows that symmetrically narrowing the middle of the vocal tract and widening it at the ends has negligible effect on formant frequencies, Rice and Öhman (1976) demonstrate that such constrictions do affect bandwidths. Expanding the middle of the vocal tract, as shown in the second curve in **Figure 17**, decreases the bandwidth of F1 (increasing its amplitude), while the opposite gesture (constricting the middle of the vocal tract) increases the bandwidth of F1 (and decreases its amplitude). A symmetrical widening of the vocal tract on either side of the midpoint and narrowing elsewhere, as shown in the fourth curve of **Figure 17**, reduces the bandwidth of F2 (and increases amplitude), while the opposite modification increases the bandwidth of F2 and decreases its amplitude. Iskarous (2010) demonstrates that the bandwidths of F1 and F2 (but not that of F3) are inversely related to their amplitudes. Consequently, since formant amplitudes are easier to measure than formant bandwidths, he concludes that formant amplitudes are the most

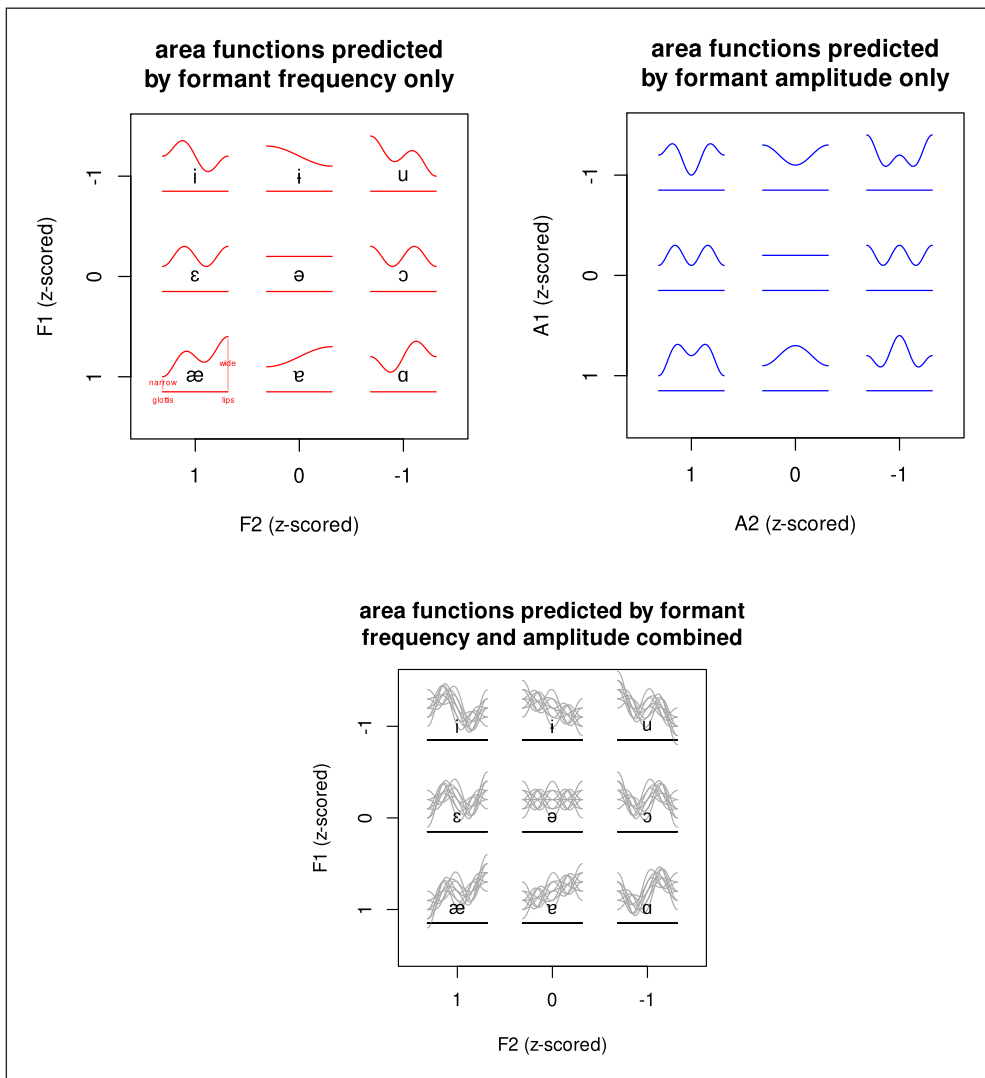
practical way of estimating the symmetric components of vocal tract area functions from recorded speech. This supplies the additional information that is crucial to refine area function estimates, compared with just formant frequencies alone.

Iskarous (2010) introduces a simplified technique for using these insights to estimate vocal tract shape parameters from formant frequencies and amplitudes, as illustrated in **Figure 17**.  $z$ -scores for the frequencies of the first two formants are used as coefficients for the first two antisymmetric components, and  $z$ -scores for the amplitudes are used as coefficients for the first two symmetric components. These four components are then summed to estimate properties of the area function (in the form of deviations from a uniform tube). Articulatory parameters are then estimated as follows: the lingual constriction degree is the absolute value of the mean of all the negative  $z$ -scores (i.e. constrictions) in the posterior 90% of the vocal tract (i.e., excluding the lips). The constriction location is defined as the weighted average of the vocal tract positions with negative  $z$ -scores, i.e. the center of gravity of the non-labial constrictions. Note that we define constriction location and constriction degree differently from Iskarous (2010) because we are focusing on non-peripheral vowels that are not well characterized by the location of the maximum constriction degree. The lip aperture is the area function value at the lips ( $x = 1$ ). Iskarous (2010) uses American English vowels from the Wisconsin X-ray microbeam database (Westbury, 1994) to demonstrate that contrastive differences in constriction location, constriction degree, and lip aperture can be recovered from the frequencies and amplitudes of the first two formants.



**Figure 17:** Estimating vocal tract shape from speech. (Iskarous, 2010).

**Figure 18** illustrates the vocal tract shapes that are predicted using Iskarous' technique. The top-left panel shows the effect of a change of 1 standard deviation in F1 or F2 frequency while holding normalized amplitude constant at 0. All of these shapes are antisymmetric and all of them have the same value at the very middle of the vocal tract. Nine hypothetical vowels are represented by all of the possible combinations of z-scored F1 and F2 values of -1, 0, or 1. The curve above each IPA symbol represents the estimated deviation from a straight tube. The straight line underneath each vowel serves to give the visual impression of a tube, but the distance between the straight line and the curve above it is arbitrary. The quantities produced by this method are the deviations from a straight tube, e.g., the straight upper curve shown for [ə],



**Figure 18:** Illustration of area functions estimated from 1 SD differences in formant frequency, amplitude, or both combined.

which is the result of applying the technique to a vowel with normalized formant frequencies of 0. A vowel with normalized F1 of  $-1$  (such as [i] in this example) is estimated to have an expansion at the back of the vocal tract and a narrowing at the front, while a vowel with a normalized F1 of 1 ([e]) is estimated to have the exact opposite shape.

The top-right panel of **Figure 18** shows the effect of a change of 1 standard deviation in F1 or F2 amplitude while holding normalized frequency constant at 0. All of these shapes are symmetrical. Low F2 amplitude (e.g., normalized F2 amplitude of  $-1$ ) is associated with narrowing on either side of the middle of the vocal tract and expansion at the middle and the ends, while high F2 amplitude is associated with the opposite perturbation: narrowing at the middle and ends but not in between. We note that symmetric perturbation can be modeled with different basis functions. Rice and Öhman use sine basis functions, which yield somewhat different predictions from Iskarous's cosine basis functions. Among other differences, sines have values of zero at the vocal tract endpoints and predict no influence of lip constriction on formant bandwidth. This deserves further exploration, but here we have followed Iskarous's method and used cosines.

The bottom panel of **Figure 18** shows the 81 possible combinations of the normalized formant frequency and amplitude values shown in the other two panels. For example, the area function estimate for a vowel such as [i] (or [u]) which has low F1 and moderate F2 would be the result of combining the [i] shape from the top-left panel with any of the amplitude-derived curves shown in the top-right panel. A [i]- or [u]-type vowel has low F1 and moderate F2, and on this basis it is estimated to resemble [i]'s vocal tract shape shown in the top left of **Figure 18**. If this vowel additionally has low F2 amplitude (and wider bandwidth), it is estimated to differ from that shape in being relatively more constricted on either side of the vocal tract midpoint and less constricted elsewhere, as depicted in the  $A_1 = 0, A_2 = -1$  shape in the top right panel of **Figure 18**. With this model we can now approximate a more precise estimate of the articulatory features of Bora vowels from their observed acoustic values, even in the absence of direct images of tongue positions. The results of this procedure are discussed in §5.

## Acknowledgements

The following Bora speakers participated in this research and gave permission to publicly acknowledge their contributions: Álida Estela Soria Arirama, Clever José Panduro Miveco, Elena Carla Soria Arirama, Eli Soria Vega, Estefanía Rodríguez Mibeco, Etelevina Meléndez Tilley, Hernán López Rodríguez, Inés Chichaco Aquirina de Pérez, Julia Ruiz Mibeco, Lucio Roque Meléndez, Marcelina Chichaco Nepiré, Manuel Ruiz Mibeco, Nélica Estefanía López Rodríguez, Oscar López Flores, René Rodríguez Torres, Roger Gonzalo Cupay Flores, and Sonia Vega Torres.

We also thank the following people for their contributions. Ian Maddieson suggested to us the technique of pasting colored dots on speakers' lips to facilitate measurements. Amy Roe collected the recordings and shared them with us. Ruth Thiesen Kerr put us in contact with

many Bora speakers. Madeleine Oakley coded images for lingual-dental contact. Our interns and lab assistants have included Katherine Mudd at DIU and Frankie Pennington, Cali Powell, and Ryan Deklerk at NCSU. We received helpful input from Paul de Lacy, Chris Golston, Khalil Iskarous, audiences at Acoustics '17, Boston (especially Matt Faytak and Richard Wright), the 2018 Annual Meeting of the Linguistic Society of America in Salt Lake City, Dallas International University, North Carolina State University, the UNC Linguistics Spring Colloquium, Cornell University, the University of Georgia, Brown University, Duke University, and California State University, Fresno. Steve Walter helped with the statistical analysis. Frank Seifart and David Weber kindly answered questions about Bora. Two anonymous referees and the associate editor provided valuable comments on an earlier draft.

This work benefited from funding from the National Science Foundation grant BCS-1562134 “Phonetic and Phonological Documentation of Kalasha, an endangered Indo-Aryan language”.

## Competing interests

The authors have no competing interests to declare.

## Author contributions

Steve Parker designed and oversaw the speech data collection and analyzed the phonological data. Jeff Mielke took the lead on processing and analyzing the acoustic and image data and conducted the vocal tract estimation procedure. Both authors analyzed the results and wrote the paper.

---

## References

- Anceaux, J. C. (1965). *The Nimboran language: Phonology and morphology*. The Hague: Martinus Nijhoff. DOI: <https://doi.org/10.1007/978-94-017-5934-2>
- Aschmann, R. P. (1993). *Proto-Witotoan*. *SIL International Publications in Linguistics 114*. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4*. (R package version, 1(7))
- Becker-Kristal, R. (2010). *Acoustic typology of vowel inventories and Dispersion Theory: Insights from a large cross-linguistic corpus* (Unpublished doctoral dissertation). University of California, Los Angeles.
- Björsten, S., & Engstrand, O. (1999). Swedish ‘damped’ /i/ and /y/: Experimental and typological observations. In *Proceedings of ICPhS XIV* (pp. 1957–1960).
- Boersma, P., & Weenink, D. (2007). Praat: doing phonetics by computer [Computer software manual]. Retrieved from <http://www.praat.org> ([Computer program]).
- Campbell, F. M. (2004). *The gestural organization of North American English /r/: A study of timing and magnitude*. (MA thesis, University of British Columbia).



- Chiba, T., & Kajiyama, M. (1941). *The vowel: its nature and structure*. Tokyo: Kaiseikan. DOI: <https://doi.org/10.11501/1677006>
- Chomsky, N., & Halle, M. (1968). *The Sound Pattern of English*. New York: Harper & Row.
- Crevels, M. (2007). South America. In C. Moseley (Ed.), *Encyclopedia of the world's endangered languages* (pp. 103–196). London: Routledge.
- de Boer, B. (2011). First formant difference for /i/ and /u/: A cross-linguistic study and an explanation. *Journal of Phonetics*, 39, 110–114. DOI: <https://doi.org/10.1016/j.wocn.2010.12.005>
- DiCanio, C., Nam, H., Whalen, D. H., Timothy Bunnell, H., Amith, J. D., & García, R. C. (2013). Using automatic alignment to analyze endangered language data: Testing the viability of untrained alignment. *The Journal of the Acoustical Society of America*, 134(3), 2235–2246. DOI: <https://doi.org/10.1121/1.4816491>
- Duanmu, S. (2016). *A theory of phonology features*. Oxford: Oxford University Press. DOI: <https://doi.org/10.1017/S0022226716000372>
- Eberhard, D. M., Simons, G. F., & Fennig, C. D. (Eds.). (2017). *Ethnologue: Languages of the World. Twenty-fifth edition*. Dallas, Tex.: SIL International. Retrieved from <https://www.ethnologue.com/language/boa>
- Elias-Ulloa, J., & Muñoz Aramburú, R. (2021). Upper-Chambira Urarina. *Journal of the International Phonetic Association*, 51(1), 137–169. DOI: <https://doi.org/10.1017/S0025100319000136>
- Engstrand, O., Björsten, S., Lindblom, B., Bruce, G., & Eriksson, A. (1998). *Hur udda är Viby-i? Experimentella och typologiska observationer*. (Folkmålsstudier 39:83–95).
- Evanini, K. (2009). *The permeability of dialect boundaries: A case study of the region surrounding Erie, Pennsylvania* (Unpublished doctoral dissertation). University of Pennsylvania.
- Fant, G. (1970). *Acoustic theory of speech production: with calculations based on X-ray studies of Russian articulations*. The Hague: Walter de Gruyter. DOI: <https://doi.org/10.1515/9783110873429>
- Faytak, M. D. (2018). *Articulatory uniformity through articulatory reuse: Insights from an ultrasound study of Sūzhōu Chinese* (Doctoral dissertation, University of California Berkeley). Retrieved from <https://escholarship.org/uc/item/47j8969j>. DOI: <https://doi.org/10.5070/P7141042486>
- Fleck, D. W. (2003). *A grammar of Matses*. (Unpublished doctoral dissertation). Rice University.
- Gick, B., Allen, B., Roewer-Després, F., & Stavness, I. (2017). Speaking tongues are actively braced. *Journal of Speech, Language, and Hearing Research*, 60(3), 494–506. DOI: [https://doi.org/10.1044/2016\\_JSLHR-S-15-0141](https://doi.org/10.1044/2016_JSLHR-S-15-0141)
- Greive, J. A. (1973). Kilba. In M. E. Kropp-Dakubu (Ed.), *West African language data sheets*, 1, 326–334. West African Linguistics Society.
- Hammarström, H., Forkel, R., Haspelmath, M., & Bank, S. (2015). *Glottolog 4.5*. Retrieved from <http://glottolog.org> (Leipzig: Max Planck Institute for Evolutionary Anthropology, Accessed on 2022-04-29) DOI: <https://doi.org/10.5281/zenodo.5772642>
- Iskarous, K. (2010). Vowel constrictions are recoverable from formants. *Journal of Phonetics*, 38, 375–387. DOI: <https://doi.org/10.1016/j.wocn.2010.03.002>

- Labov, W., Rosenfelder, I., & Fruehwald, J. (2013). One hundred years of sound change in Philadelphia: Linear incrementation, reversal, and reanalysis. *Language*, 89(1), 30–65. DOI: <https://doi.org/10.1353/lan.2013.0015>
- Ladefoged, P. (1996). *Elements of acoustic phonetics, 2nd edition*. Chicago: University of Chicago Press. DOI: <https://doi.org/10.7208/chicago/9780226191010.001.0001>
- Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Oxford: Blackwell.
- Lallouache, M. T. (1991). *Un poste "visage-parole" couleur: Acquisition et traitement automatique des contours des lèvres* (Unpublished doctoral dissertation). Grenoble INPG.
- Lee-Kim, S.-I. (2014). Revisiting Mandarin 'apical vowels': An articulatory and acoustic study. *Journal of the International Phonetic Association*, 44(3), 261–282. DOI: <https://doi.org/10.1017/S0025100314000267>
- Lobanov, B. M. (1971). Classification of Russian vowels spoken by different speakers. *Journal of the Acoustical Society of America*, 49(2B), 606–608. DOI: <https://doi.org/10.1121/1.1912396>
- McGowan, R. S. (2020). *On formants*. Lexington, MA: CReSS Books.
- Ménard, L., Toupin, C., Baum, S. R., Drouin, S., Aubin, J., & Tiede, M. (2013). Acoustic and articulatory analysis of French vowels produced by congenitally blind adults and sighted adults. *Journal of the Acoustical Society of America*, 134(4), 2975–2987. DOI: <https://doi.org/10.1121/1.4818740>
- Mermelstein, P. (1967). Determination of the vocal-tract shape from measured formant frequencies. *The Journal of the Acoustical Society of America*, 41(5), 1283–1294. DOI: <https://doi.org/10.1121/1.1910470>
- Mielke, J. (2012). A phonetically based metric of sound similarity. *Lingua*, 122, 145–163. DOI: <https://doi.org/10.1016/j.lingua.2011.04.006>
- Mielke, J., Olson, K. S., Baker, A., & Archangeli, D. (2011). Articulation of the Kagayanen interdental approximant: An ultrasound study. *Journal of Phonetics*, 39(3), 403–412. DOI: <https://doi.org/10.1016/j.wocn.2011.02.008>
- Milne, P. M. (2011). Finding schwa: Comparing the results of an automatic aligner with human judgments when identifying schwa in a corpus of spoken French. *Canadian Acoustics*, 39(3), 190–191. DOI: <https://doi.org/10.1121/1.3654646>
- Moran, S., & McCloy, D. (2019). *Phoible 2.0*. Retrieved from <http://phoible.org> (Jena: Max Planck Institute for the Science of Human History, Accessed on 2022-02-16) DOI: <https://doi.org/10.5281/zenodo.2593234>
- Newmarch, J. (2017). Ffmpeg/Libav. In *Linux sound programming* (pp. 227–234). Berkeley, CA: Apress. DOI: <https://doi.org/10.1007/978-1-4842-2496-0>
- Noiray, A., Cathiard, M.-A., Ménard, L., & Abry, C. (2011). Test of the movement expansion model: Anticipatory vowel lip protrusion and constriction in French and English speakers. *The Journal of the Acoustical Society of America*, 129(1), 340–349. DOI: <https://doi.org/10.1121/1.3518452>
- Olson, K. S. (2019). Reanalyzing the Banda-Linda vowel system. In *Proceedings of ICPhS 2019* (p. 2061–2064).

- Parker, S. (2001). The acoustic qualities of Bora vowels. *Phonetica*, 58(3), 179–195. DOI: <https://doi.org/10.1159/000056198>
- Parker, S., & Mielke, J. (To appear). *Phonological data for studying vowel patterns in Bora*. Retrieved from <https://diu.edu/academics/special-electronic-publications/> (Special Electronic Publications. Dallas International University).
- R Core Team. (2015). *R language definition*. Available from CRAN sites.
- Reidy, P. F. (2013). *spectRum*. Retrieved from <https://github.com/patrickreidy/spectRum> (R package).
- Reidy, P. F. (2015). A comparison of spectral estimation methods for the analysis of sibilant fricatives. *JASA Express Letters*, 137(4), EL248–EL254. DOI: <https://doi.org/10.1121/1.4915064>
- Rice, L., & Öhman, S. (1976). On the relationship between formant bandwidths and vocal tract shape features. In *UCLA Working Papers in Phonetics 31* (pp. 27–31). <https://escholarship.org/uc/item/31f5j8m7>
- Roe, A. (2014). *The phonetics and phonology of Bora tone*. (MA thesis). University of North Dakota.
- Schönle, P. W., Gräbe, K., Wenig, P., Höhne, J., Schrader, J., & Conrad, B. (1987). Electromagnetic articulography: Use of alternating magnetic fields for tracking movements of multiple points inside and outside the vocal tract. *Brain and Language*, 31(1), 26–35. DOI: [https://doi.org/10.1016/0093-934X\(87\)90058-7](https://doi.org/10.1016/0093-934X(87)90058-7)
- Schötz, S., Frid, J., & Löfqvist, A. (2011). Exotic vowels in Swedish: A project description and an articulographic and acoustic pilot study of /i:/. In *Proceedings of Fonetik 2011, Speech, Music, and Hearing 51*.
- Schroeder, M. R. (1967). Determination of the geometry of the human vocal tract by acoustic measurements. *The Journal of the Acoustical Society of America*, 41(4B), 1002–1010. DOI: <https://doi.org/10.1121/1.1910429>
- Schwartz, J.-L., Boë, L.-J., Vallée, N., & Abry, C. (1997). The dispersion-focalization theory of vowel systems. *Journal of Phonetics*, 25(3), 255–286. DOI: <https://doi.org/10.1006/jpho.1997.0043>
- Seifart, F. (2005). *The structure and use of shape-based noun classes in Miraña (North West Amazon)*. (Unpublished doctoral dissertation). Radboud University.
- Seifart, F., & Echeverri, J. A. (2015). Proto Bora-Muinane. *LIAMES*, 15(2), 279–311. DOI: <https://doi.org/10.20396/liames.v15i2.8642303>
- Selkirk, E. (1993). *[Labial] relations*. (University of Massachusetts Amherst ms.).
- Thiesen, W., & Thiesen, E. (1998). *Diccionario bora-castellano, castellano-bora*. Serie Lingüística Peruana no. 46. Lima: Instituto Lingüístico de Verano.
- Thiesen, W., & Weber, D. (2012). *A grammar of Bora: with special attention to tone*. SIL International Publications in Linguistics no. 148. Dallas: SIL International.
- Thiesen Kliewer, W., & de Thiesen, E. A. (1975). Fonemas del bora. In *Datos etno-lingüísticos no. 1* (pp. 1–12). Lima, Peru: Instituto Lingüístico de Verano.

- Thomas, P. N. (2017). The central vowel of Kawaiisu. *International Journal of American Linguistics*, 83, 539–559. DOI: <https://doi.org/10.1086/691588>
- Vaux, B., & Samuels, B. (2015). Explaining vowel systems: Dispersion theory vs natural selection. *The Linguistic Review*, 32(3), 573–599. DOI: <https://doi.org/10.1515/tlr-2014-0028>
- Westbury, J. R. (1994). *X-ray microbeam speech production database user's handbook*. (University of Wisconsin, Madison, WI).
- Wood, S. (1979). A radiographic analysis of constriction locations for vowels. *Journal of Phonetics*, 7, 25–43. DOI: [https://doi.org/10.1016/S0095-4470\(19\)31031-9](https://doi.org/10.1016/S0095-4470(19)31031-9)
- Wrench, A., & Balch-Tomes, J. (2022). Beyond the edge: Markerless pose estimation of speech articulators from ultrasound and camera images using DeepLabCut. *Sensors*, 22(3), 1133. DOI: <https://doi.org/10.3390/s22031133>
- Yuan, J., & Liberman, M. (2008). Speaker identification on the SCOTUS corpus. In *Proceedings of Acoustics 2008* (pp. 5687–5690). DOI: <https://doi.org/10.1121/1.2935783>
- Zhou, X. (2009). *An MRI-based articulatory and acoustic study of American English liquid sounds /r/ and /l/* (Unpublished doctoral dissertation). University of Maryland, College Park.

