

JOURNAL ARTICLE

# Phonological and phonetic properties of nasal substitution in Sasak and Javanese

Diana Archangeli<sup>1,2</sup>, Jonathan Yip<sup>2</sup>, Lang Qin<sup>2</sup> and Albert Lee<sup>2</sup>

<sup>1</sup> University of Arizona, Tucson, AZ, US

<sup>2</sup> University of Hong Kong, HK

Corresponding author: Diana Archangeli (dba@email.arizona.edu)

Austronesian languages such as Sasak and Javanese have a pattern of morphological *nasal substitution*, where nasals alternate with homorganic oral obstruents—except that [s] is described as alternating with [ɲ], not with [n]. This appears to be an abstract morphophonological relation between [s] and [ɲ] where other parts of the paradigm have a concrete homorganic relation. Articulatory ultrasound data were collected of productions of [t, n, tɕ, ɲ], along with [s] and its nasal counterpart from two languages, from 10 Sasak and 8 Javanese speakers. Comparisons of lingual contours using a root mean square analysis were evaluated with linear mixed-effects regression models, a method that proves reliable for testing questions of phonological neutralization. In both languages, [t, n, s] exhibit a high degree of articulatory similarity, whereas postalveolar [tɕ] and its nasal counterpart [ɲ] exhibited less similarity. The nasal counterpart of [s] was identical in articulation to [ɲ]. This indicates an abstract, rather than concrete, relationship between [s] and its morphophonological nasal counterpart, with the two sounds not sharing articulatory place in either Sasak or Javanese.

**Keywords:** Sasak; Javanese; nasal substitution; ultrasound language research; abstract phonological relations; place of articulation

## 1. Introduction

There is increasing evidence that language sounds which appear to be ‘the same sound’ are in some cases actually distinct from each other, with consequences for phonological analysis. The classification of sound categories is of great importance to understanding the phonological organization in a language, and close attention to small phonetic differences has contributed to the development of both Lexical Phonology (Mohanan, 1982; Kiparsky, 1982) and Articulatory Phonology (Browman & Goldstein, 1986, 1992).

The effect of sameness may arise from categorical perception (Liberman et al., 1957; Harnad, 2003); the fact remains that native speakers do not necessarily attend to all distinctions in speech sounds. For example, sometimes segments with different phonetic properties are classed together as a single sound, such as the different articulations of American English /ɪ/ (Delattre & Freeman, 1968; Mielke et al., 2010; Archangeli et al., 2011; Mielke et al., 2016) and the different acoustics of American English /s/ (Baker et al., 2011). Another class of relevant examples is neutralization, both incomplete and complete. Most striking are instances of incomplete neutralization, where two sounds from different sources were considered to be the same but were later revealed to be slightly different under close phonetic examination (see Port, 1996 and Yu, 2007 for summaries of the issues; final devoicing examples are found in Port & O’Dell, 1985, Warner et al., 2004; Winter & Röttger, 2011 among others; English [l] in Lee-Kim et al., 2013;

tonal near neutralization in Myers & Tsay, 2008; Cheng et al., 2013; incomplete nasal place assimilation in Stephenson & Harrington, 2002; coronal-velar nasal-obstruent heteromorphic sequences in English in Barry, 1991, and Korean heteromorphic palatalization in Sung, 2015). Incomplete neutralizations contrast with complete neutralization, for example complete neutralization of manner for coronals in Korean codas (Kim & Jongman, 1996) and of pre-velar nasals in Italian (Celata et al., 2013). Sometimes in the same language both complete and incomplete neutralization are found, depending on the context: Nasal-obstruent place assimilation in Spanish is complete with a following stop, but it is partial with a following fricative (Kochetov & Colantoni, 2011); /s/ + /j/ in English varies depending on intervening boundary type (Zsiga, 1995).

Such examples raise a question about ‘exceptions’ to otherwise general phonological patterns when a sound appears to have anomalous behavior: Is this truly an exception, inconsistent with the general pattern, or does closer phonetic examination suggest that in fact the sound was misperceived and there is a systematic pattern? To illustrate, the earliest description of Okpe vowels (Hoffmann, 1973) is used as an example of absolute neutralization (Kenstowicz & Kisseberth, 1979): It appears that /ɪ/ and /e/ neutralize to [e], /ʊ/ and /o/ neutralize to [o]. However, Omamor (1973) includes a pilot study suggesting that the high retracted and mid advanced vowel pairs have slightly different formant values—in which case this is not absolute neutralization. Similarly, Mutaka (1995) assumes that Kinande short low vowels are exempt from an otherwise general tongue root advancement harmony, being retracted even in advanced environments; ultrasound examination (Gick et al., 2006) reveals that low vowels have advanced tongue root in advanced contexts. In short, all vowels participate in tongue root harmony and the Kinande pattern is completely symmetric.

Another potential example comes in the phenomenon called nasal substitution, occurring in Austronesian languages, including the Malayo-Polynesian languages Sasak and Javanese. Nasal substitution refers to morphological pairing of obstruent-initial and homorganic nasal-initial forms—except that [s] pairs with a sound described as [ɲ] in both Sasak and Javanese. Our question is whether closer inspection reveals that [s] and the nasal morphologically related to [s] are in fact homorganic sounds, despite the description of them as heterorganic sounds.

Nasal substitution occurs in many Austronesian languages, with similar correspondences across languages. In certain morphological contexts, words with initial voiceless obstruents appear with a homorganic nasal instead of the obstruent—labial with labial, dental/alveolar with dental/alveolar, velar with velar—so-called nasal substitution (De Guzman, 1978; Pater, 1999, 2001; Blust, 2004), as in **Table 1a**. For words with initial voiced obstruents, a homorganic nasal surfaces along with the obstruent, **Table 1b**. Our focus is on the **Table 1a** pattern, the voiceless obstruents and the corresponding homorganic nasals.<sup>1</sup>

What is less straightforward—yet familiar across different languages in this family—is the behavior of word-initial orthographic ‘s.’ Since ‘s’ is typically classed with the other dental/alveolar consonants, we would expect it to pair with [n] in this paradigm. However,

<sup>1</sup> In the orthography of both languages, ‘ng’ is used for [ŋ], ‘ny’ for [ɲ], ‘c’ for [tʃ], and ‘e’ for both [e/ɛ] and [ə]. We use IPA symbols for both languages. For more on nasal substitution, see De Guzman (1978); Archangeli et al. (1998); Pater (1999, 2001) and the excellent summary in Blust (2004) for phonological analysis of Austronesian nasal substitution, Reid (2000) on the historical development of nasal substitution, and Wouk (1999); Austin (2010, 2013) on the syntactic distribution of ‘nasal verbs’ vs. ‘oral verbs’ in Sasak. Here and elsewhere, Sasak items are from fieldnotes by Archangeli and Yip; Javanese items are from Robson & Wibisono (2002) and were verified by a native Javanese-speaking linguist.

**Table 1:** Nasal substitution examples with voiceless and voiced initial obstruents in Sasak (left) and Javanese (right). Non-nasal forms begin with an oral consonant; nasal forms show the corresponding nasal-initial form. (a) Voiceless obstruents pair with single nasal consonants while (b) voiced obstruents pair with nasal-obstruent sequences.

Sasak			Javanese		
<b>a. voiceless initial</b>					
non-nasal	nasal	gloss	non-nasal	nasal	gloss
paku	maku	‘nail’	pariŋi	maririŋi	‘give (respect)’
tambah	nambah	‘hoe’	tuku	nuku	‘buy’
tɛŋjak	ŋɛŋjak	‘grin’	tɛakar	ŋakar	‘scratch’
kadu?	ŋadu?	‘dig’	kirɔ	ŋirɔ	‘estimate’
<b>b. voiced initial</b>					
non-nasal	nasal	gloss	non-nasal	nasal	gloss
bau	mbau	‘pick fruit’	bɔŋɔ	mbɔŋɔ	‘fund’
dɛŋah	ndɛŋah	‘hear’	dɛɔ?	ndɛɔ?	‘see’
dʒait	ŋdʒait	‘sew’	dʒadʒan	ŋdʒadʒan	‘buy snacks’
gita?	ŋgita?	‘see’	gɔwɔ	ŋgɔwɔ	‘carry’

it pairs with orthographic ‘ny’ instead of with ‘n,’ illustrated in **Table 2** where the nasal correspondent of [s] is represented as [N<sub>s</sub>].<sup>2</sup>

The contrast between the behavior of [s] and the behavior of other voiceless obstruents appears to break the ‘homorganic nasal’ pattern. Our question is whether this is indeed the case, arguing in favor of an *abstract pattern*, i.e., a relation between two sounds which is not grounded completely in their phonetic properties. The alternative is that the homorganic nasal pattern is *concrete*, i.e., the relation between the two sounds is fully grounded in their phonetic properties. In an abstract relation, the places of articulation of [s] and of [N<sub>s</sub>] would be quite different from each other while if the relation is concrete, the articulation of [s] and [N<sub>s</sub>] would be relatively similar to each other. In either case, there remains the question of what the place of articulation is for each of these sounds; possibilities are shown in **Table 3**.

Under the *abstract hypothesis*, the two sounds are related morphologically, but the homorganic relation found with other segments does not hold with [s] and its corresponding nasal [N<sub>s</sub>]. We expect that the tongue positions of [s] and [N<sub>s</sub>] would be quite different from each other, and comparable to the way that heterorganic tongue positions differ from each other. The place of the two sounds might correspond to the perceived place, with [s] homorganic to [t, n] (an dental/alveolar [s]; Dart, 1991) and [N<sub>s</sub>] homorganic to [tɕ, ɲ] (**Table 3a**). Alternatively, [s] might be distinct from [t, n] (e.g., [s] is postalveolar in the phonological analysis of Mester, 1986) or [N<sub>s</sub>] might be distinct from [tɕ, ɲ] (**Table 3b, c** respectively).

Under the *concrete hypothesis*, articulatory similarity between [s] and [N<sub>s</sub>] is driving the alternation. Thus, we expect that [s], like other voiceless obstruents, would be homorganic with its nasal counterpart and their tongue positions quite similar to each other. If [s] and [N<sub>s</sub>] are homorganic, they could class with [t, n], with [tɕ, ɲ], or with a third place of articulation (**Table 3d, e, f** respectively). The abstract and concrete hypotheses are summarized in **Table 4**.

<sup>2</sup> We reserve [ɲ] for nasals that are unambiguously postalveolar, such as those related to [tɕ]. Here and in the remainder of this paper, we refer to the nasal which is morphologically related to orthographic ‘s’ as [N<sub>s</sub>].

**Table 2:** Examples with initial [s] in Sasak (left) and Javanese (right). In both languages, non-nasal [s]-initial forms pair with a nasal-initial form, shown here as [N<sub>s</sub>]. The nasal forms are described as beginning with [ɲ], spelled ‘ny,’ and are shown in orthographic representation in the third and seventh columns.

Sasak				Javanese			
non-nasal	nasal	orthog.	gloss	non-nasal	nasal	orthog.	gloss
siduk	N <sub>s</sub> iduk	<i>nyiduk</i>	‘kiss’	siram	N <sub>s</sub> iram	<i>nyiram</i>	‘water’
sean	N <sub>s</sub> ean	<i>nyean</i>	‘divorce’	sɛpa	N <sub>s</sub> ɛpa	<i>nyepa</i>	‘kick’
səboʔ	N <sub>s</sub> əboʔ	<i>nyeboq</i>	‘hide’	səpi	N <sub>s</sub> əpi	<i>nyepi</i>	‘silent’
sapak	N <sub>s</sub> apak	<i>nyapak</i>	‘greet’	sawaŋ	N <sub>s</sub> awaŋ	<i>nyawang</i>	‘see’
soroŋ	N <sub>s</sub> oroŋ	<i>nyorong</i>	‘push’	suruŋ	N <sub>s</sub> uruŋ	<i>nyurung</i>	‘push’
sumpak	N <sub>s</sub> umpak	<i>nyumpak</i>	‘swear’	sənəŋ	N <sub>s</sub> ənəŋ	<i>nyeneng</i>	‘make happy’

**Table 3:** Possible places of articulation for [s] and [N<sub>s</sub>] under the abstract and concrete hypotheses. Outcome d is unlikely given that [N<sub>s</sub>] is consistently described as not being /n/ (and also spelled ‘ny’ as /ɲ/).

		[t, n]	other	[tɕ, ɲ]
abstract, heterorganic	a.	[s]		[N <sub>s</sub> ]
	b.		[s]	[N <sub>s</sub> ]
	c.	[s]	[N <sub>s</sub> ]	
concrete, homorganic	d.	[s, N <sub>s</sub> ]		
	e.			[s, N <sub>s</sub> ]
	f.		[s, N <sub>s</sub> ]	

**Table 4:** Predictions of the Abstract and Concrete Hypotheses.

	Hypothesis	Prediction	
a.	Abstract Hypothesis	[s] ≠ [N <sub>s</sub> ]	articulatory dissimilarity
b.	Concrete Hypothesis	[s] = [N <sub>s</sub> ]	articulatory similarity

We set out to answer the “Abstract or concrete?” question and to address the related place of articulation issues based on ultrasound data of speakers of both Sasak and Javanese.

## 2. Language background

Sasak is the primary local language of Lombok, Indonesia, with speakers estimated at 2 million (Clynes, 1995) and 2.5 million (Marli, 2015). Sasak, along with Balinese, Sumbawa, Malayic, and Chamic, is within the Malayo-Polynesian sub-group Malayo-Sumbawan (Adelaar, 2005). Sasak is described as having four (Jacq, 1998) or five (Austin, 2003) major dialects. (Austin, 2003 reports that the informal names for the dialects relate to how each group pronounces the deictic words for ‘like this’ and ‘like that’: Ngenó-Ngené (central northeast, central east, and central west coasts of Lombok), Menó-Mené (central Lombok), Ngotó-Ngeté (northeastern Lombok), Ngenó-Mené, also known as Kutó-Kuté (north Lombok), and Meriaq-Meriku (south central Lombok). Jacq, 1998 does not include Ngenó-Mené as a dialect.) The dialects with the widest geographical distribution are Ngenó-Ngené and Menó-Mené, which are the only varieties used by Sasak speakers in this study.

Javanese is the most-spoken regional language of Indonesia and the most-spoken language of the Austronesian language family with approximately 75 million speakers. It is found along the northwest coast of Java (Banten, Krawang, Cirebon) and in the central and eastern areas of this island. Outside of Java, it is used in diasporic communities in neighboring provinces of Sumatra, Kalimantan, and Sulawesi, as well as in Suriname and New Caledonia. Three dialects are usually distinguished (western, central, and eastern) (Ras, 1985), and the western dialect is further divided into seven subdialects (Nothofer, 1980). Our participants are all from the central and eastern dialects, which have not yet been studied in much detail (Nothofer, 2006).

Nasal substitution in Sasak appears on verbs “used when the Patient-like argument is non-referential” according to Austin (2013, p. 41). There are many other factors determining the distribution of oral- or nasal-initial verb forms; see Austin (2013). Similarly, in Javanese, nasal substitution appears on verbs and relates to argument structure: Sato (2008, p. 53) calls nasal substitution the ‘active voice morphology’ and shows that it is necessary in basic transitive clauses, but does not occur in Wh-questions or passives. See also Sato (2015); see Herawati et al. (2016) for discussion of nasal substitution and denominal verbs.

As for the sounds, nasal substitution refers to related pairs of words, typically one which begins with a voiceless oral obstruent and the other with a nasal homorganic to that obstruent. (Voiced obstruents, sonorants, and vowel-initial words have their own patterns; see Clynes, 1995 and Austin, 2013 for Sasak; Dudas, 1976; Robson, 1992; and Lee, 2001 for Javanese, and Pater, 1999, 2001 for the Austronesian pattern in general).

To put nasal substitution in context, the Sasak and Javanese consonant inventories are shown in **Table 5**, based on Clynes (1995); Archangeli et al. (2016) for Sasak<sup>3</sup> and Dudas

**Table 5:** Consonant inventories for Sasak (above) & Javanese (below). The consonants that are our focus here are boxed in these tables.

<b>Sasak:</b>	<b>Bilabial</b>	<b>Dental/Alveolar</b>	<b>Postalveolar/Palatal</b>	<b>Velar</b>	<b>Glottal</b>
Plosive	p b	<span style="border: 1px solid black;">t</span> d		k g	ʔ
Affricate			<span style="border: 1px solid black;">tʃ</span>	dʒ	
Nasal	m	<span style="border: 1px solid black;">n</span>	<span style="border: 1px solid black;">ɲ</span>	ŋ	
Trill/Tap		r			
Fricative		<span style="border: 1px solid black;">s</span>			h
Approximant			j	w	
Lateral approximant		l			

<b>Javanese:</b>	<b>Bilabial</b>	<b>Dental/Alveolar</b>	<b>Postalveolar/Palatal</b>	<b>Velar</b>	<b>Glottal</b>
Plosive	p b	<span style="border: 1px solid black;">t̚</span> d̚ t̚ d̚		k g	ʔ
Affricate			<span style="border: 1px solid black;">tʃ</span>	dʒ	
Nasal	m	<span style="border: 1px solid black;">n̚</span>	<span style="border: 1px solid black;">ɲ</span>	ŋ	
Trill/Tap		r			
Fricative		<span style="border: 1px solid black;">s</span>			h
Approximant			j	w	
Lateral approximant		l			

<sup>3</sup> Clynes (1995) uses ‘laminal’ where we have ‘postalveolar.’

(1976) for Javanese. In these tables, we describe sounds *c* [tʃ], *j* [dʒ], *ny* [ɲ], and *y* [j] as ‘postvelar/palatal’ due to the lack of agreement in the literature as to the precise location of constriction described for these consonantal sounds. Both languages distinguish bilabial, postalveolar, and velar consonants. Sasak has a single dental/alveolar category. Javanese has both dental and alveolar consonants (Dudas, 1976 [citing Horne, 1961; Hayward & Mulijono, 1991]), also described as a dental/retroflex contrast (Suharno, 1982; Robson, 1992; Adisasmito-Smith, 2004; Graff & Jaeger, 2009); ‘s’ is classed with dentals by all sources. Javanese nasal substitution is described as resulting in a dental nasal [ɲ] regardless of whether the corresponding sound is a dental [t] or an alveolar (or retroflex) [t̚]. Because of the challenges of imaging the tip of the tongue with ultrasound, we did not use stimuli with initial alveolar/retroflex stops in the Javanese study.<sup>4</sup>

The languages are similar in that each has only one sibilant, which is one of the sounds targeted in this study. It is possible that there is more variety of articulation for the sibilant because there is no sibilant contrast to be maintained—e.g., there is no [ʃ] alongside the [s]. Clynes (1995) classes Sasak [s] together with [tʃ, dʒ]. Mester (1986) views Javanese /s/ as a postalveolar consonant, and Robson (1992) states that the Javanese ‘s’ is similar to that of English, but “sometimes is heard as approaching *sh*” (p. 12).<sup>5</sup> We did not perceive this fluctuation ourselves, in either language. These different classifications, along with nasal substitution apparently relating ‘s’ and a palatal nasal, raise questions about the phonetic nature of the sibilant in both languages: Could it be articulated somewhere between an dental/alveolar sound and a postalveolar/palatal sound, or is it indeed postalveolar/palatal? On the other hand, because Sasak has only one contrast in the dental/alveolar region while Javanese has two contrasts for the stops, we might expect concomitant differences between the two languages in the nasal counterparts to [s].

### 3. Methods

In order to carry out this study, we collected and analyzed ultrasound tongue imaging data. Sasak data were collected at the Mataram Lingua Franca Institute in Lombok, Indonesia, and Javanese data were collected at the University of Hong Kong.

The procedure for collecting and processing data in the two languages is largely the same. Differences in the methodologies arose because the two studies were carried out independently; we saw the value of putting the two together after the data were collected; analysis methods are as similar as possible given differences in number of stimuli per language and number of repetitions per stimulus for each language.<sup>6</sup> We present the basic methodology here, along with ways in which the procedures for the two data sets differed.

#### 3.1. Participants

For the Sasak part of the study, there were 11 participants who all reported speaking Sasak exclusively until elementary school; all continued to use Sasak on a daily basis throughout their lives. All participants also reported fluency in Bahasa Indonesia and have learned

<sup>4</sup> While there is a somewhat robust literature on the laryngeal contrasts in Javanese, including Brunelle (2010); G. Poedjosoedarmo (1986); G. R. Poedjosoedarmo (1993); Thurgood (2004); Matthews (2015), our focus is on the alternations with voiceless consonants.

<sup>5</sup> Rehg and Sohl (1981) make a similar observation about ‘s’ in Pohnpeian, another Austronesian language, where the ‘s’ sounds somewhere between English ‘s’ and ‘sh,’ with speaker variation about the degree of palatalization.

<sup>6</sup> Archangeli and Yip carried out the Sasak data collection; Lee and Qin collected the Javanese data.

**Table 6:** Demographics of participants for Sasak (top) & Javanese (bottom). M-M refers to the Menó-Mené dialect of Sasak; Ng-Ng refers to the Ngenó-Ngené dialect. CJ refers to the Central Javanese dialect, and EJ to the Eastern Javanese dialect. We found no effect of dialect in either the Sasak or the Javanese analysis.

participant	gender	age	dialect
S2	F	37	M-M
S3	F	30	M-M
S4	M	21	M-M
S6	F	22	Ng-Ng
S7	M	25	M-M
S8	F	25	M-M
S9	M	22	Ng-Ng
S10	M	22	Ng-Ng
S11	M	22	M-M
S12	M	23	M-M
S13	M	19	M-M
J2	F	23	CJ
J3	F	25	CJ
J5	F	29	EJ
J6	F	36	EJ
J7	F	33	EJ
J8	F	30	EJ
J9	F	33	CJ
J10	F	39	EJ

English as a third language. Ages ranged from 19 to 37 (average 24.4); 4 were female and 7 were male. Of these, 8 self-identified as speakers of the Menó-Mené (M-M) dialect while 3 self-identified as Ngenó-Ngené (Ng-Ng) speakers. Data from 2 additional speakers were omitted due to poor quality of the ultrasound images.

For Javanese, 8 female native speakers (and no male speakers) were recorded.<sup>7</sup> All participants reported speaking Javanese (either the eastern or central dialect) on a daily basis until moving to Hong Kong, and all also reported fluency in Bahasa Indonesia. Some learned English as a third language, while others instead learned Cantonese. Ages ranged between 23 and 39 (average 31). The experiment was conducted either in English or in Cantonese, with occasional explanation in Javanese or Bahasa Indonesian by a bilingual Javanese- and Bahasa Indonesian-speaking assistant. (Data from two additional Javanese speakers were omitted due to poor imaging quality in the ultrasound signal).

Information about each Sasak- and Javanese-speaking participant's gender, age, and native dialect appear in **Table 6**.

<sup>7</sup> The Javanese data were collected in Hong Kong while the Sasak data were collected in Lombok, Indonesia. Most Indonesians who come to Hong Kong work as domestic helpers; participants were recruited from this population. However, we were unable to recruit male speakers as there are far fewer male Javanese speakers in Hong Kong.

### 3.2. Stimuli

To examine the relationship between the place of oral and nasal sounds, we identified items with initial coronal voiceless consonants along with their morphologically-related nasal-initial forms. These initial consonants either had a known place of articulation, that is, either dental/alveolar or postalveolar, or they were ambiguous in place ([s] and [N<sub>s</sub>]), shown in **Table 7**, with examples in the rightmost column.

In selecting stimuli, only morphologically-related forms were included, with either [a] or [ə] in the first syllable. The vowels [ə] and [a] were chosen in order to minimize the effects of coarticulation of the consonant from the following vowel. Non-high, central vowel contexts were chosen because high vowels with front or back tongue position (e.g., [i] or [u]) typically show stronger influences on tongue shape and position during consonantal constriction than do other types of vowels (Öhman, 1966; Zharkova & Hewlett, 2009).

For Sasak, one word with each vowel was identified for each of the 6 consonants, resulting in 12 stimuli; an additional [s]-initial pair was included so there are 14 stimuli for Sasak. In one case, the vowel [ɛ] was inadvertently used instead of [ə]: [tɛmbaʔ]/[nɛmbaʔ] ‘shoot.’ Orthographically, the symbol ‘e’ is used for both [ə] and [ɛ/e], leading to confusion about the desired stimuli from the orthographic prompt.

For Javanese, stimuli were selected in a similar way, with either [a] or [ə] present in the first syllable of each target item. For this language, 10 items were identified for each of the 6 target consonants, resulting in a total of 60 Javanese target stimuli. A full list of target Sasak and Javanese stimulus items is presented in the appendix.

### 3.3. Procedure

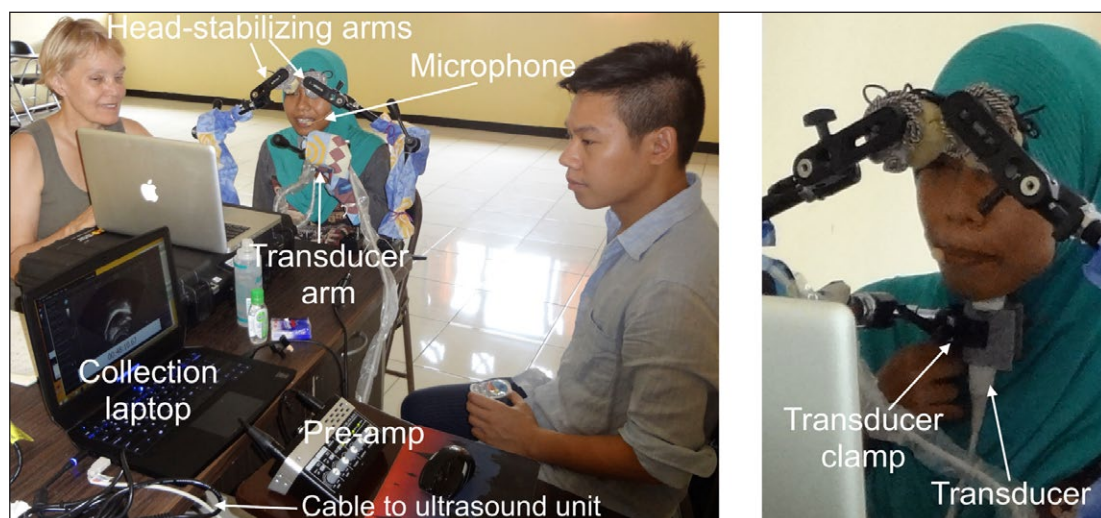
Each data collection session began with an explanation of the study and the data collection methods. The participants were seated in front of a display laptop, which was used to present visual prompts. Two posable camera arms (Manfrotto 143 Magic Arm) were adjusted to provide a stable headrest for the participants and to minimize head movement throughout the entire duration of the collection session. A third fixed the position of the ultrasound sensor along the centerline of the lower jaw at a location where the full midsagittal contour of the tongue imaged most clearly. This setup is shown in **Figure 1**, where a close-up of head-to-probe stabilization method is shown in the image to the right.

When stabilization arm adjustments were complete, participants were asked whether they were willing to continue with the study. On agreement, each was asked to sip water slowly through a straw, in order to create an image of the palate. Once a good palate image was obtained, participants were instructed to produce each target stimulus from

**Table 7:** Dental/alveolar and postalveolar test comparisons. Non-nasal forms (lefthand column) are those which do not have the nasal prefix; nasal forms (righthand column) are related forms that appear with the nasal prefix. Sounds viewed as postalveolar are highlighted in gray; sounds viewed as dental/alveolar are in white cells. Sasak and Javanese stimuli are found in the appendix.

	Non-nasal forms	Related nasal forms	Examples (Sasak)		
Dental/Alveolar	t	n	tambah	nambah	‘hoe (v.)’
Postalveolar	ʈɕ	ɲ	ʈɕapɔk	ɲapɔk	‘take s.o.’s property’
Ambiguous	s	N <sub>s</sub>	sapak	N <sub>s</sub> apak	‘greet’





**Figure 1:** Sasak data collection set up in Lombok. The possible camera arms are cloth-covered to make them more friendly to participants. Two arms stabilize the forehead while a third holds the probe in a fixed position. One laptop is used for presenting stimuli while the other is used to monitor and collect data. The ultrasonic scan unit, which was placed underneath the desk in order to mitigate fan noise, is not shown here.

the dedicated display laptop's screen, and prompts were advanced for each participant by an experimenter. Throughout the task, another experimenter monitored the imaging quality during collection to ensure that good ultrasound images were obtained.

Participants were asked to read each of the target word items, which appeared on the screen of the display laptop, one at a time. In the case of Sasak words, target items were produced in isolation, whereas for Javanese, target words were presented in the carrier phrase *Kata* \_\_\_ '(the) word (is)\_\_\_,' in order to make speakers produce a preceding [a] vowel immediately before the target initial-consonant sounds.

Stimuli were presented in a randomized order that was unique to each participant. For Sasak speakers, randomization was performed within each of six stimulus blocks, with each block containing one iteration of each target word. This resulted in six productions of each item and a total of up to 84 token productions per session for Sasak. For Javanese speakers, randomization was performed within one large block, in which each target item appeared two times, resulting in two productions of each target item and a total of up to 120 token productions per session for Javanese. The target number of prompts and repetitions for each sound in each language is summarized in **Table 8**.

Because the collection procedures for Sasak and Javanese were designed independently of each other and occurred separately, the number of target items for Javanese was much larger than that for Sasak in this study. On the other hand, the number of iterations of each item in Sasak was larger than that for each item in Javanese. The consequence is that there are 50% more token productions in the Javanese data than in the Sasak data.

For all recording sessions, the ultrasound images were collected using a 2–4 MHz convex ultrasound sensor (Teleded MC4-2R20N) coupled with a Teleded ClarUs-EXT portable, ultrasonic beam-former connected to a high-performance laptop that functioned as a machine dedicated to audio- and video-data collection. The ultrasound images were constructed and displayed using Echo Wave II ultrasound imaging software (Teleded 2015) at approximately 60 frames per second. On-screen renderings of these images were collected using a combination of desktop-display software (XSplit Broadcaster: SplitmediaLabs,

**Table 8:** Number of tokens for each place category for each language. ‘Total’ refers to the maximum number of analyzable tokens produced by each participant for the relevant CV sequence. Recall that in Sasak, a [t/n]-initial form with [ɛ] was used instead of a form with [ə], hence ‘ $\epsilon_{SAS}$ ’ and ‘ $\epsilon_{JAV}$ ’ in the top lefthand cell.

initial C	V context	Sasak				Javanese			
		sounds	items	reps	total	sounds	items	reps	total
t/n	$\epsilon_{SAS}$ $\epsilon_{JAV}$	2	1	6	24	2	5	2	40
	a	2	1	6		2	5	2	
tɕ/ɲ	ə	2	1	6	24	2	5	2	40
	a	2	1	6		2	5	2	
s/N <sub>s</sub>	ə	2	1	6	36	2	5	2	40
	a	2	2	6		2	5	2	
<i>total tokens per participant</i>					84	<i>total tokens per participant</i>			120

2015) and real-time on-screen video-capture software (Fraps: Beepa, 2015) at a stable rate of 60 frames per second.<sup>8</sup> Each session involved under 30 minutes of recording.

Audio was captured with an over-the-ear condenser microphone. In order to synchronize audio and ultrasound video, for Sasak, one experimenter produced a series of 6–14 tokens of the voiceless post-alveolar click [k̠] at the end of each speaker’s video and audio recordings immediately before stopping those recordings. The Javanese recordings lacked these click productions, and thus, in order to synchronize each ultrasound video to its corresponding audio signal, 10 instances of the release of the unaspirated voiceless velar stop [k], present at the beginning of each item’s carrier phrase *Kata \_\_\_\_*, were selected at random throughout each recording and analyzed instead of clicks. Video-to-audio synchronization was achieved by determining the mean temporal lag between the onset time of the acoustic release burst of each stop (voiceless post-alveolar click [k̠] for Sasak; velar plosive [k] for Javanese) and the time of the ultrasound frame immediately prior to visible articulatory release.<sup>9</sup> In most cases, the ultrasound video signal had a consistent lag of between 1 and 3 seconds after the audio, and for each ultrasound video, frame times were subsequently readjusted in order to align them temporally with corresponding acoustic events in the audio signal. Post-collection and post-alignment, single ultrasound image frames were extracted from the video recordings as PNG-format image file sequences using digital video playback software (QuickTime Pro: Apple 2010). The recordings for 4 Javanese speakers (J2, J5, J6, and J7) were inadvertently halted mid-collection and thus their productions were recorded in multiple video files. However, these speakers did not move out of position when the recordings were halted, and in these cases, video-to-audio synchronization simply required the calculation of lag for each individual video file relative to its corresponding audio signal. (See manual alignment techniques in Miller & Finch, 2011.)

<sup>8</sup> There is a slight mismatch, of at most a 16.7-ms lag in the frame rate of the Echo Wave software’s ultrasonic image construction and that of the video-capture software. This is not considered problematic for temporal synchronization: For nasals and fricatives, whose duration is much longer than 16.7 ms, the medial frame is selected so the selected frame is always within the segment; for stops, we took the frame *preceding* the release burst (see section 3.4) so if there is an effect of the latency, the frame selected is slightly earlier than desired but certainly during the stop-constriction interval, and the frame does not come after the release of the stop, when the tongue has moved away from occlusion. This was verified during visual inspection of each extracted frame.

<sup>9</sup> Multiple samples of click/stop release bursts for synchronization were measured in order to improve calculations of lag between audio and video within a given recording. On average, the standard deviation between multiple lag measures in the same recording was  $11.3 \pm 4.3$  ms across participants.

For the Sasak recordings, collected in Lombok, there were environmental factors that increased the noise-to-signal level in each audio recording, such as roosters crowing at random times, mosque calls for prayer, a pre-school promotion ceremony, motor-scooters passing by, and echo-y rooms to record in. The condenser microphone, positioned close to participants' mouths and set to record at a low-level setting for gain during recording, served to improve the signal-to-noise ratio in the audio signal. These disruptions have had at most a minor impact on this study since the primary focus is the articulatory gestures associated with the sounds of interest, not their acoustic properties, and were not severe enough to prevent aural identification of items. The Javanese recordings were collected in a quiet room at the University of Hong Kong and did not have such issues.

### 3.4. Analysis

Analysis involved four steps: (i) identifying frames to analyze, (ii) assigning coordinates to tongue contours, (iii) quantifying the distance between contours, (iv) statistically modeling the distribution of distance values across conditions.

Frames were identified through the audio recording, using Praat (Boersma & Weenink, 2015). Details follow about how frames were identified for different sound types.

- a. *Oral stops and affricates* [t, tʃ]. For oral stops and affricates, the frame of interest was defined as the last frame before the release of the oral stop constriction, which was identified from the corresponding waveform and spectrogram. The extracted ultrasound frame was assumed to represent a full stop constriction (i.e., the achievement of a full postalveolar constriction gesture) because the temporal distance between ultrasound frames (16.67 ms) was shorter than the time it would take for the tongue to maintain a stop constriction prior to release. The frames of interest for affricates were determined in the same manner because of the oral-stop portion of such sounds contained a complete oral constriction at the location for the affricates.
- b. *Nasal stops* [n, ŋ, N̥]. For nasal stops, the extracted frame was the frame closest to the acoustic midpoint of the interval of nasal-stop constriction, as determined from the waveform and spectrogram. In Sasak, where words were collected in isolation, the nasal stop was preceded by silence and the onset of nasalization was identified as the onset of the vocal fold vibration in the acoustic waveform. For the end of the nasal, and in Javanese where the nasal stop was preceded by a vowel due to the carrier phrase, the nasal stop boundaries were identified by the loss of vowel formant structure and a significant decrease in acoustic intensity in the acoustic waveform and spectrogram.
- c. *Fricative* [s]. For fricatives, the frame of interest was identified as the frame closest to the acoustic midpoint of the frication associated with the [s] articulation. The onset and offset of frication was determined by the presence of aperiodic noise in the waveform. The frame closest to the midpoint of the frication interval was assumed to best represent the [s] articulation because the corresponding acoustic pattern in the spectrogram was most characteristic of [s] at the center of the fricative.

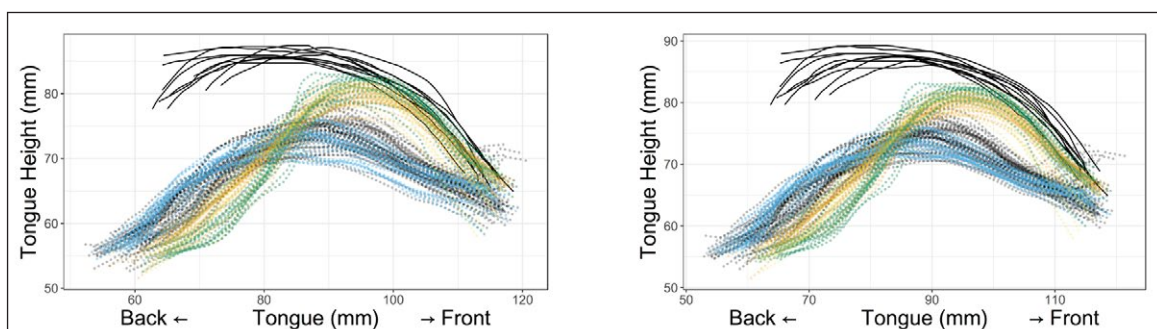
The next step was to convert the images into coordinates corresponding to the tongue surface as shown in the image. EdgeTrak software (Li et al., 2005) was used to determine the boundary edges (corresponding to the surface of the tongue) and fit a smoothed graphical spline curves onto the boundaries; edges were hand-corrected as needed. The data for each spline were exported as a set of 100 equidistant coordinate points.

Although continuous attempts were made to ensure that the collected ultrasonic data did not contain any head movement relative to the transducer, comparisons of 9–10 successive traces of each speaker's palate collected throughout the ultrasound recordings indicated that head movement occurred during the scanning for 5 talkers (S9, S10, J6, J7, and J10). Based on the palate data, the approximate moment of each significant shift

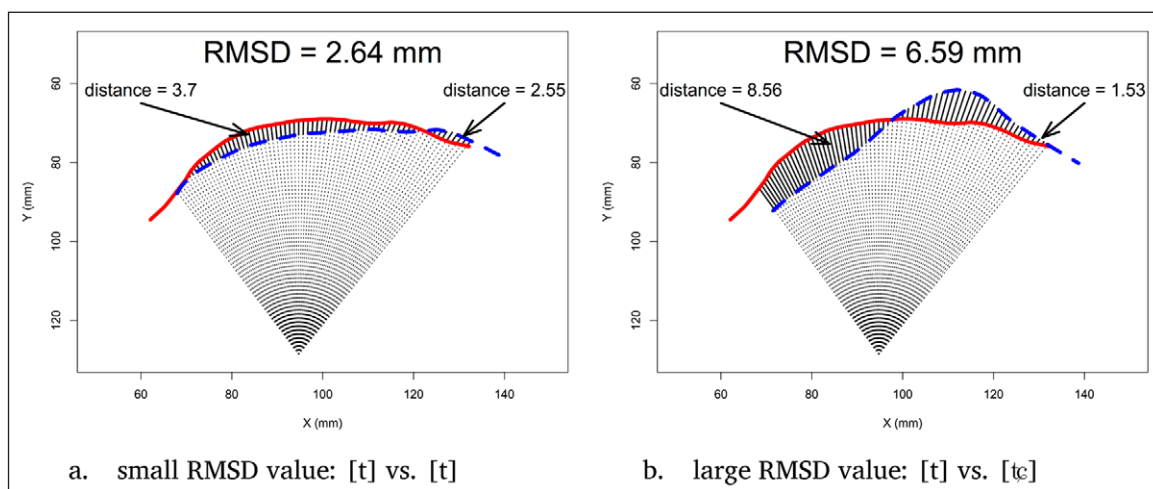
in head position during production was identified, and palate contours affected by the movement were adjusted via spatial translation until the anterior portion of these affected palate contours were situated in the same region as those of the other palate traces. Then all sound contours of interest for these participants that were also affected by the head movement were adjusted in the same manner as the adjusted palate traces in order to correct for head movement in the spline data. This adjustment resulted in an overall reduction in variation in spatial position of the tongue splines in the data for the 5 talkers with head movement. An example of all palate and tongue contour data for participant S9 before adjustment and after adjustment is given in **Figure 2**.

In order to derive measures of similarity/dissimilarity between production tokens, root mean squared distances (RMSDs) between tongue contour pairs were calculated. Mean squared distance and RMSD values are used for various purposes in ultrasound tongue position research, e.g., comparing native and non-native speakers' articulations (Li et al., 2005; Davidson, 2005; Berry et al., 2012), understanding coarticulation (Irfana & Sreedevi, 2016), and evaluating the accuracy of edge-detection algorithms (Roussos et al., 2009; Fasel & Berry, 2010; Csapó & Lulich, 2015). For our purposes, low RMSD values indicate that two sound tokens were articulated with high similarity, whereas high RMSDs indicate that the tokens were articulated with quite distinct lingual contours. RMSDs were calculated from the spatial distances between the contours from each token pair along each angle with an integer value shared between the two contours with respect to the location of the origin. The origin was defined as the point of intersection between the lines representing the leftmost and rightmost boundaries of the ultrasonic image for each talker, and the origin's location depended solely on the scan settings (scan frequency, scan depth, and field of view) used in the EchoWave software. The RMSD calculation procedure is depicted in **Figure 3**. Distances at each angle were squared individually, then summed together, and divided by the total number of angles, and the square root of this value was taken as the RMSD measure for the token pairing.

For each spline pair, no measures were taken at angles ( $\theta$ ) that intersected with only a single trace, as shown at both ends of the images in **Figure 3**. This method contrasts with the mean Euclidean distance algorithm used in (Zharkova & Hewlett, 2009), which calculates the arithmetic mean of the shortest distances of all points along each spline to all points along the paired spline. Where there was a mismatch in spline



**Figure 2:** Aggregated tongue contours for Sasak participant S9 before adjustment (left) and after adjustment (right), based on palate traces (black solid lines). Contours are shown as dotted lines, and sound groups are coded by place (dental/alveolar = black, postalveolar = orange, ambiguous [s] = blue, ambiguous [N<sub>s</sub>] = green). Here and henceforth, the anterior part of the oral cavity is oriented toward the righthand side of each figure.



**Figure 3:** Illustration of the calculation of RMSD values between contours. Distances at each angle between the pair of contours are shown as red/solid and blue/dashed line segments, with distance values at example angles indicated to the sides. Rays at the relevant angles from the origin are indicated with dotted lines extending up to the two tongue contours. Measures were taken at each angle.

length along the  $x$ -dimension (the  $\theta$ -dimension in our method), as seen in **Figure 3**, the Zharkova and Hewlett approach could overestimate mean distance between splines because points at the extreme ends of each spline would be measured as having longer distances to their nearest point along the paired spline. Our method, on the other hand, essentially ignores measures at angles ( $\theta$ ) where both contours were not present in the spline data, and the potential overestimation of distance between spline pairs at the extreme ends is less likely here than in Zharkova and Hewlett. Since our aim is to determine the degree of similarity/difference (or degree of homorganicity/heterorganicity) between contours, we chose a method that would not necessarily overestimate mean distances between contours, i.e., not exaggerate their differences, in cases where contours simply differed in length along the  $\theta$ -dimension rather than in articulatory place of constriction.

RMSD data for each language were submitted to linear mixed-effects regression (LMER) models using the `lmer()` function in the `lmerTest` package (Kuznetsova et al., 2012) in R statistical software (R Core Team, 2016). Both languages' LMER models contained fixed effects of Place (homorganic, heterorganic, ambiguous) and Nasality (shared, contrastive) and random slopes and intercepts for factor levels within Subject. Word was omitted as a random effect because the number of tokens per item in the Javanese data set was small (2 iterations per word), and the addition of this factor did not improve the fit of the model for either language. Estimates of RMSD values from these LMER models allowed for comparisons of magnitude of difference between sound pairs according to Place and Nasality conditions. The number of RMSD values per speaker and per language and condition are reported in **Tables 9** and **10**.

#### 4. Results

In order to determine the significance of the above observations, we carried out LMERS on RMSD differences for various classes of sounds: The RMSD between two tongue splines serves as a measure of similarity of articulatory tongue position. We divided the relevant sound pairs into six categories, shown in **Table 11**. 'Homorganic' refers to sounds with the same place of articulation, dental/alveolar or postalveolar, while 'heterorganic'

**Table 9:** Total number of extracted splines and mean number of degrees (corresponding to mean number of distance values per comparison) involved in calculating RMSD values for each speaker of Sasak and Javanese. The number of extracted splines corresponds to the number of stimuli for each language as shown in Table 8.

Sasak			Javanese		
participant	no. of splines	mean degrees	participant	no. of splines	mean degrees
S2	84	71.4	J2	118	59.5
S3	84	54.6	J3	117	67.6
S4	84	66.3	J5	118	59.7
S6	83	57.0	J6	118	57.9
S7	84	62.0	J7	118	46.3
S8	84	80.2	J8	118	53.4
S9	84	60.1	J9	118	57.4
S10	84	55.3	J10	118	54.2
S11	84	56.5			
S12	84	37.8			
S13	84	74.1			

**Table 10:** The total number of contour-pair comparisons across participants per condition for Sasak (left) and Javanese (right). These numbers are identical to the count of RMSD values calculated per condition.

	Place	Nasality			Place	Nasality	
		same	contrastive			same	contrastive
Sasak	homorganic	2,893	3,156	Javanese	homorganic	6,061	6,380
	heterorganic	3,156	3,156		heterorganic	6,380	6,380
	ambiguous	3,366	3,564		ambiguous	2,744	2,800

**Table 11:** Six classes of sound pairings based on place and nasality. Here and elsewhere, ‘s’ indicates that two sounds have the same nasality regardless of place of articulation—that is, both are oral as in [t]–[t] or [t]–[t̥], or both are nasal as in [n]–[n] or [n]–[ŋ], while ‘c’ indicates that two sounds have contrastive nasality—one is oral while the other is nasal, [t]–[n] and [t̥]–[n] for example.

		Place		
		homorganic	heterorganic	ambiguous
Nasality	s(ame)	[t]–[t], [n]–[n], [t̥]–[t̥], [ŋ]–[ŋ]	[t]–[t̥], [n]–[ŋ]	[s]–[s], [N <sub>s</sub> ]–[N <sub>s</sub> ]
	c(ontrastive)	[t]–[n], [t̥]–[ŋ]	[t]–[ŋ], [t̥]–[n]	[s]–[N <sub>s</sub> ]
		unambiguous		

is a cross between these two. These comparisons give a measure for RMSD for homorganic and heterorganic sounds. The ‘ambiguous’ category contains the sounds tested by this study, [s] and [N<sub>s</sub>]. ‘Nasality-same’ (nasality<sub>s</sub>) means either both sounds are oral or both

sounds are nasal; ‘nasality-contrastive’ (nasality<sub>c</sub>) means that the two sounds in a pair disagree for nasal/oral articulation.<sup>10</sup>

Using the categories from **Table 11**, we are able to make our predictions explicit. Our first focus is on the categories homorganic and heterorganic, whether oral or nasal. We use the homorganic class to establish reasonable RMSDs for sounds made with the same place of articulation (same or similar articulatory position of the tongue). RMSDs are predicted to be small in this case because homorganic sounds are made with the same place of articulation by definition. In contrast, we expect large RMSDs when the two sounds have different places of articulation, the heterorganic condition. Putting these together, we expect that the  $RMSD_{\text{homorganic}}$  is smaller than the  $RMSD_{\text{heterorganic}}$ .

When nasality is added in, we expect to find that homorganic sounds with contrastive nasality still have a small RMSD, but it is slightly larger than when the two sounds are truly identical due to differences introduced by the nasal/oral contrast because of the different manners of articulation. Our four expectations are summarized in **Table 12**.<sup>11</sup>

The box plot in **Figure 4** shows the distribution of RMSDs for both the unambiguous sounds and the ambiguous sounds. Focusing on the unambiguous sounds, we see that the RMSDs for homorganic sounds are small, well below 5 mm, with a lower mean when sounds are identical (i.e., the nasality<sub>s</sub> case; Sasak: 2.5 mm, Javanese: 1.6 mm) than when nasality differs (nasality<sub>c</sub>; Sasak: 3.4 mm, Javanese: 3.9 mm), exactly as expected:  $RMSD_{\text{Homorganic-s}} < RMSD_{\text{Homorganic-c}}$ . (This difference is significant, as seen in **Table 14**.) In contrast, the means for heterorganic sounds are above 5 mm for both Sasak and Javanese, whether there is a nasality contrast or not, again as expected:  $RMSD_{\text{Homorganic-s,c}} \ll RMSD_{\text{Heterorganic}}$ . RMSDs pattern as expected; we now have a measure to use in quantifying comparisons involving ambiguous sounds.

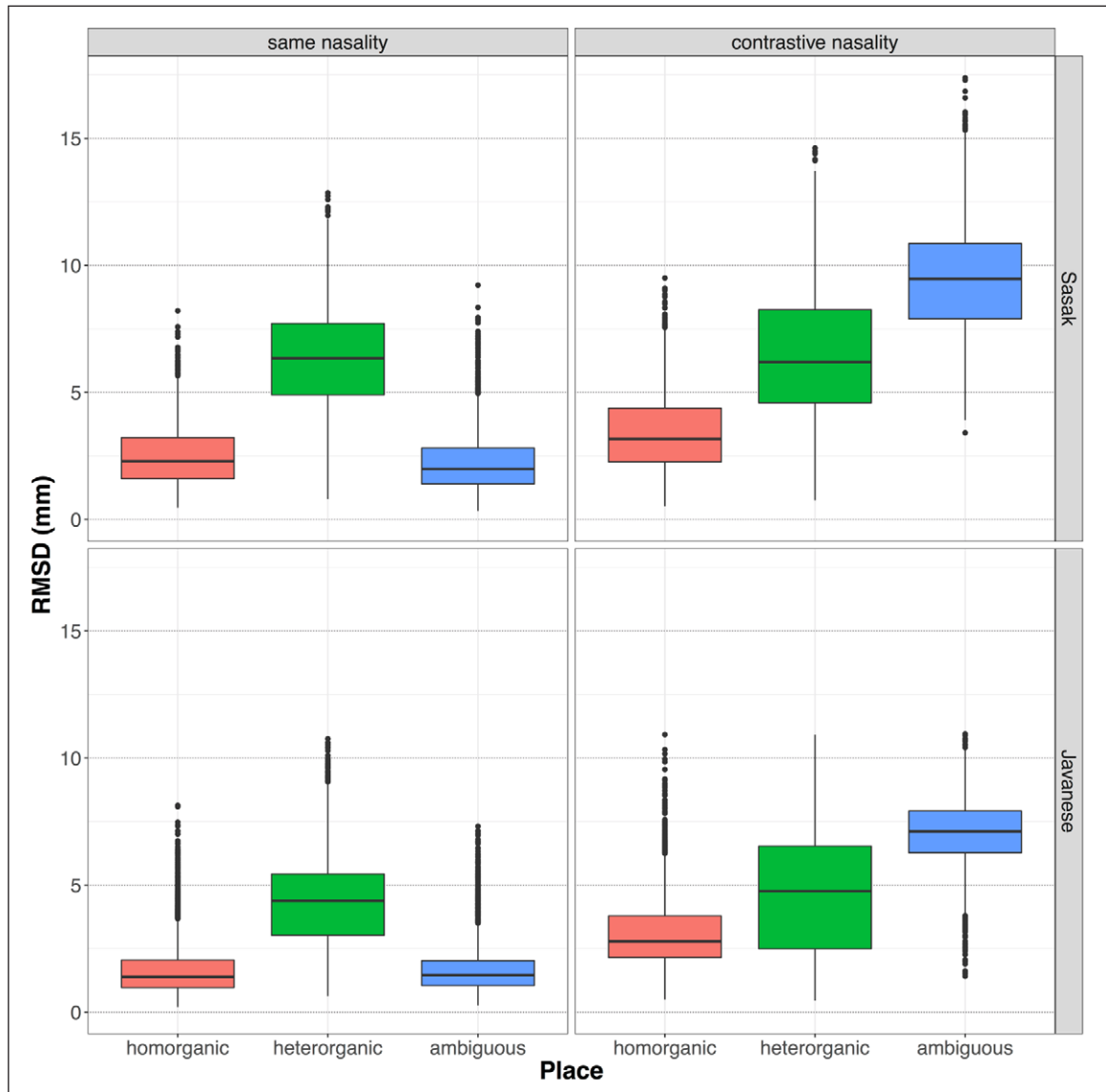
In particular, we now can understand the abstract/concrete hypotheses’ predictions in terms of RMSD. Under the abstract hypothesis, [s] is heterorganic to [N<sub>s</sub>], shown by a relatively large RMSD. Under the concrete hypothesis, [s] is homorganic with [N<sub>s</sub>], shown by a RMSD similar to that of homorganic-contrastive pairs since [s] and [N<sub>s</sub>] have

**Table 12:** Predictions for RMSDs in mm for impact of Place and Nasality.

	Place	Nasality	
a.	Homorganic	same	The estimated RMSD is relatively small.
		contrastive	The estimated RMSD is relatively small, but larger than in the homorganic-same case.
b.	Heterorganic	same	The estimated RMSD is relatively large.
		contrastive	The estimated RMSD is relatively large.
c.	Summary	$RMSD_{\text{Homorganic-s}} < RMSD_{\text{Homorganic-c}} \ll RMSD_{\text{Heterorganic}}$	

<sup>10</sup> A caveat is in order: Ultrasound images capture only part of the tongue yet there are multiple other dimensions for comparing articulation that do not appear in these images (such as the larynx, lips, posterior pharyngeal wall), and the images are two dimensional while the vocal tract is three dimensional.

<sup>11</sup> Differences between homorganic oral and nasal consonants would be consistent with results in Gibbon et al. (2007) and Shosted et al. (2012). Gibbon et al. (2007) show that [t] and [d] have more contact than [n] in normal adult speakers of English, using electropalatography, while Shosted et al. (2012) show that degree of contact also varies between different languages: [ɲ] in Peninsular Spanish involves a fair degree of occlusion in the alveopalatal region, whereas in Brazilian Portuguese the [ɲ] has a degree of closure more like an approximant, with an articulatory target that is neither occluded nor anterior in the oral cavity. Thus we do not have clear predictions about whether same or contrastive nasality affects the RMSD of heterorganic sounds.



**Figure 4:** Box plots showing the distribution of RMSD values for unambiguous homorganic and heterorganic pairs and for the ambiguous sounds, whether they have the same or contrasting nasality (same, contrastive respectively). Homorganic sounds with the same oral/nasal value are identical in this table, [t]–[t], [n]–[n], etc.

different nasality values; see **Table 13**. Under both hypotheses, the self-comparison values ( $RMSD_{\text{Ambiguous-s}}$ )—i.e., [s] compared with [s] and  $[N_s]$  compared with  $[N_s]$ —are expected to be small.

**Table 14** evaluates the predictions from **Table 13** in terms of estimated RMSD values for different comparisons, using LMERS to make those comparisons. The Place effects show that homorganic sounds have a significantly smaller estimated RMSD than do heterorganic sounds (Sasak same nasality: 2.5 mm < 6.4 mm,  $p < 0.0001$ ; Sasak contrastive nasality: 3.4 mm < 6.5 mm,  $p < 0.0001$ ; Javanese same nasality: 1.6 mm < 4.3 mm,  $p < 0.0001$ ; Javanese contrastive nasality: 3.9 mm < 4.5 mm,  $p < 0.0001$ ), with a general difference of approximately 3 mm. This is consistent with the prediction for place differences. In contrast, according to the LMER results for both languages, estimated RMSD values for homorganic sounds with same nasality (i.e., [t] compared with [t], [n] with [n], [tɕ] with [tɕ], [ɲ] with [ɲ]) and for ambiguous sounds with same nasality (i.e., [s] with [s],  $[N_s]$  with  $[N_s]$ ) differ by a much smaller margin (Sasak: homorganic<sub>s</sub> 2.5 mm > ambiguous<sub>s</sub>



**Table 13:** Predictions about RMSD value for the ambiguous sounds [s]-[N<sub>s</sub>] under the abstract and concrete hypotheses.

	comparison	prediction	measure
Abstract	[s]-[s], [N <sub>s</sub> ]-[N <sub>s</sub> ]	homorganic	The estimated RMSD is relatively small.
	[s]-[N <sub>s</sub> ]	heterorganic	The estimated RMSD is relatively large.
	Summary	$RMSD_{\text{ambiguous-same}} = RMSD_{\text{homorganic-same}}$ $RMSD_{\text{ambiguous-same}} < RMSD_{\text{ambiguous-contrastive}}$	
Concrete	[s-s], [N <sub>s</sub> ]-[N <sub>s</sub> ]	homorganic	The estimated RMSD is relatively small.
	[s]-[N <sub>s</sub> ]	homorganic	The estimated RMSD is relatively small.
	Summary	$RMSD_{\text{ambiguous-same}} = RMSD_{\text{homorganic-same}}$ $RMSD_{\text{ambiguous-contrastive}} = RMSD_{\text{homorganic-contrastive}}$	

**Table 14:** LMER estimates of RMSD values (in mm) for unambiguous and ambiguous sound classes in Sasak (top) and Javanese (bottom). “Est. 1” shows RMSD estimates for the lefthand side of each prediction; “Est. 2” shows RMSD estimates for the righthand side of each comparison. The box highlights the *p*-value that is not significant. Sounds are categorized by place and nasality as in Table 11.

		Sasak					
	Prediction	Est. 1	Est. 2	S.E.	t-value	p-value	
Place	homorganic <sub>s</sub> << heterorganic <sub>s</sub>	2.50	6.36	0.04	87.55	<0.0001	
	homorganic <sub>c</sub> << heterorganic <sub>c</sub>	3.42	6.47	0.04	70.89	<0.0001	
	homorganic <sub>s</sub> = ambiguous <sub>s</sub>	2.50	2.24	0.04	6.04	<0.0001	
Nasality	homorganic <sub>s</sub> < homorganic <sub>c</sub>	2.50	3.42	0.04	20.80	<0.0001	
	heterorganic <sub>s</sub> < heterorganic <sub>c</sub>	6.36	6.47	0.04	2.644	0.0082	
	ambiguous <sub>s</sub> << ambiguous <sub>c</sub>	2.24	9.49	0.04	176.262	<0.0001	
		Javanese					
	Prediction	Est. 1	Est. 2	S.E.	t-value	p-value	
Place	homorganic <sub>s</sub> << heterorganic <sub>s</sub>	1.63	4.29	0.03	104.43	<0.0001	
	homorganic <sub>c</sub> << heterorganic <sub>c</sub>	2.97	4.54	0.03	62.38	<0.0001	
	homorganic <sub>s</sub> = ambiguous <sub>s</sub>	1.63	1.60	0.03	0.93	0.354	
Nasality	homorganic <sub>s</sub> < homorganic <sub>c</sub>	1.63	2.97	0.03	52.48	<0.0001	
	heterorganic <sub>s</sub> < heterorganic <sub>c</sub>	4.29	4.54	0.03	9.75	<0.0001	
	ambiguous <sub>s</sub> << ambiguous <sub>c</sub>	1.60	6.91	0.04	140.05	<0.0001	

2.2 mm, *p* < 0.0001; Javanese: homorganic<sub>s</sub> 1.6 mm = ambiguous<sub>s</sub> 1.6 mm, *p* = 0.354), consistent with the predictions pertaining to ambiguous<sub>same</sub> sounds.

Turning to the effect of nasality, we see that the RMSD estimates for homorganic comparisons with the same nasality are smaller than those for homorganic comparisons with contrastive nasality (Sasak: 2.5 mm < 3.4 mm, *p* < 0.0001; Javanese: 1.6 mm < 3.9 mm, *p* < 0.0001), although the RMSD value for homorganic<sub>contrastive</sub> pairings is still relatively

small. Heterorganic<sub>same</sub> and heterorganic<sub>contrastive</sub> pairings were similar to each other in RMSD estimates but still differed significantly (Sasak: 6.4 mm < 6.5 mm,  $p = 0.0082$ ; Javanese: 4.3 mm < 4.5 mm,  $p < 0.0001$ ), and corresponding RMSD values in these conditions were relatively large. Importantly, RMSD estimates for ambiguous sounds with contrastive nasality (ambiguous<sub>c</sub>) were large (Sasak: 9.5 mm; Javanese: 6.9 mm) and differed drastically from RMSD estimates for ambiguous sounds with matching nasality (Sasak: 9.5 mm >> 2.2 mm,  $p < 0.0001$ ; Javanese: 6.9 mm >> 1.6 mm,  $p < 0.0001$ ). We conclude that [s] and [N<sub>s</sub>] exhibit the greatest distinctness in lingual contour shape and place of articulation in both languages, exactly as predicted under the abstract hypothesis.

The answer to the abstract/concrete question raises the issue of whether [s] is dental/alveolar sound and [N<sub>s</sub>] is postalveolar, or whether one or the other is something else. In order to test whether [s] and [N<sub>s</sub>] are homorganic with other dental/alveolar and postalveolar sounds respectively, we revised our earlier LMER models to target these two ambiguous sounds, including pairwise comparisons between each of the ambiguous sounds [s, N<sub>s</sub>] and each of the unambiguous sounds [t, n, tɕ, ɲ] (a separate model was generated for each ambiguous sound, with only relevant sound comparisons included in the data set). The revised models provide RMSD estimates for each sound-sound pair as well as  $p$ -values for comparisons with corresponding unambiguous homorganic and unambiguous heterorganic categories given in **Table 11**, homorganic<sub>s</sub>, homorganic<sub>c</sub>, heterorganic<sub>s</sub>, and heterorganic<sub>c</sub>. RMSD estimates and  $p$ -values from the revised LMER models are reported in **Table 15**. The closer the RMSD estimate for a given sound pair is to the RMSD of one of the unambiguous categories, the more likely that that pair is a member of that category. If a  $p$ -value indicates that the RMSD comparison for the ambiguous sound does not differ from the RMSD for the corresponding unambiguous homorganic comparison, then the place-ambiguous sound in the pair is homorganic with the other sound in the pair.

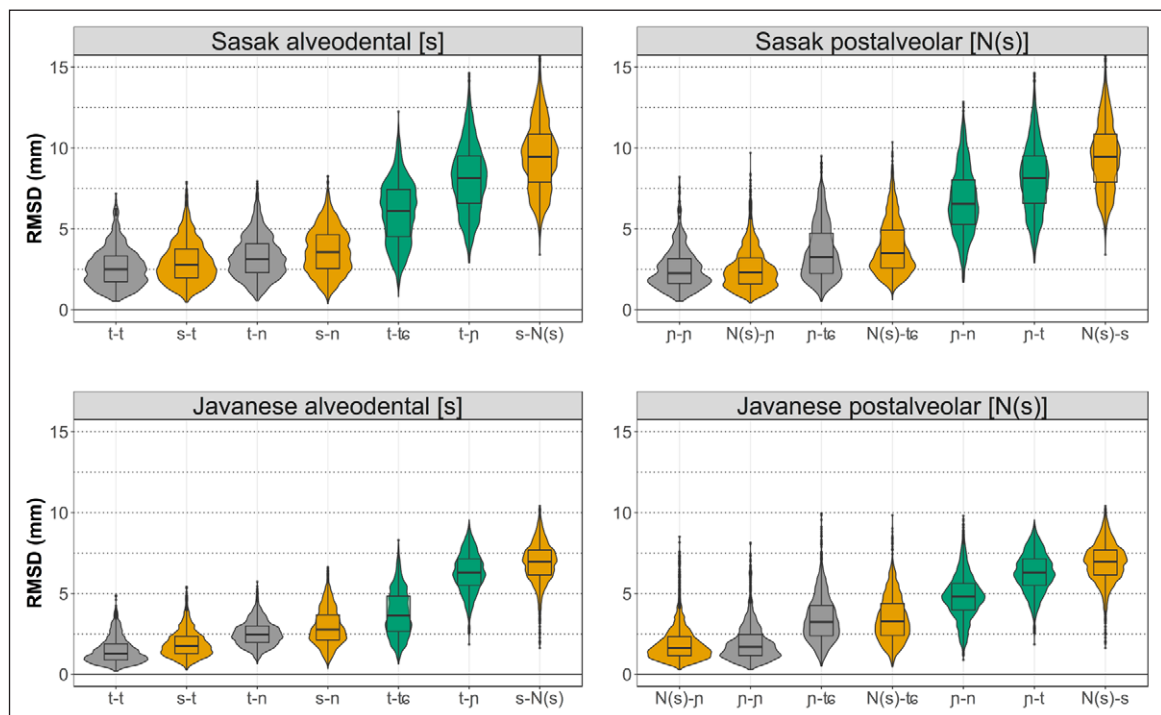
In all comparisons shown in **Table 15**, the RMSD estimates for [s] compared with either [t] or [n], and for [N<sub>s</sub>] compared with either [tɕ] or [ɲ] are similar to the RMSD estimates for corresponding homorganic comparisons (columns 4–6, **Table 15**) and much smaller than the estimates for corresponding heterorganic comparisons (columns 7–9, **Table 15**).

**Table 15:** RMSD estimates in mm (values highlighted in gray) and  $p$ -values for comparisons of ambiguous sounds [s, N<sub>s</sub>] with unambiguous dental/alveolar and postalveolar sounds. Boxed  $p$ -values indicate RMSD differences that are not significant.

	Homorganic comparison					Heterorganic comparison		
	Pair	RMSD	Pair	RMSD	$p$ -value	Pair	RMSD	$p$ -value
Sasak	[s]-[t]	2.95	[t]-[t]	2.62	<0.0001	[t]-[tɕ]	6.02	<0.0001
	[s]-[n]	3.64	[t]-[n]	3.25	<0.0001	[t]-[ɲ]	8.13	<0.0001
	[N <sub>s</sub> ]-[ɲ]	2.55	[ɲ]-[ɲ]	2.52	0.618	[ɲ]-[n]	6.70	<0.0001
	[N <sub>s</sub> ]-[tɕ]	3.86	[ɲ]-[tɕ]	3.59	<0.0001	[ɲ]-[t]	8.13	<0.0001
	[s]-[N <sub>s</sub> ]	9.49	[t]-[n]	3.31	<0.0001	[t]-[ɲ]	8.13	<0.0001
	Javanese	[s]-[t]	1.89	[t]-[t]	1.48	<0.0001	[t]-[tɕ]	3.77
[s]-[n]		2.93	[t]-[n]	2.53	<0.0001	[t]-[ɲ]	6.30	<0.0001
[N <sub>s</sub> ]-[ɲ]		1.90	[ɲ]-[ɲ]	1.97	0.0468	[ɲ]-[n]	4.81	<0.0001
[N <sub>s</sub> ]-[tɕ]		3.45	[ɲ]-[tɕ]	3.41	0.190	[ɲ]-[t]	6.30	<0.0001
[s]-[N <sub>s</sub> ]		6.91	[t]-[n]	2.52	<0.0001	[t]-[ɲ]	6.30	<0.0001

In general, RMSD estimates for comparisons of [s] to both [t] and [n] and of [N<sub>s</sub>] to both [ŋ] and [tɕ] are within the range of homorganic contrasts (a RMSD near 3-mm), rather than in the range of heterorganic contrasts (a RMSD at or above 6-mm). In two cases, comparisons between Sasak [N<sub>s</sub>]-[ŋ] and Javanese [N<sub>s</sub>]-[tɕ] with their corresponding homorganic pairs [ŋ]-[ŋ] and [ŋ]-[tɕ] result in large *p*-values (*p* = 0.618 and *p* = 0.190, respectively), indicating no significant difference between RMSDs for these specific sound pairs and their unambiguous homorganic analogues. In Javanese, RMSDs for [N<sub>s</sub>]-[ŋ] differ significantly from those for the known homorganic pair [ŋ]-[ŋ] (*p* = 0.0468), though to a lesser extent than for other presumably homorganic comparisons, and the RMSD estimate (1.9 mm) is even smaller than most other homorganic RMSDs. These relations are shown visually in **Figure 5**, in which presumed homorganic pairings, [s]-[t], [s]-[n], [N<sub>s</sub>]-[ŋ], and [N<sub>s</sub>]-[tɕ] (middle and leftmost orange shapes in each plot), are more similar to corresponding homorganic-same and -contrastive groups (gray shapes) than their heterorganic-same or -contrastive counterparts (green shapes) in terms of mean RMSD between contours. In each case for both languages, the comparison between [s] and [N<sub>s</sub>] (rightmost orange shapes) has a larger mean RMSD value than any other homorganic or heterorganic comparison.

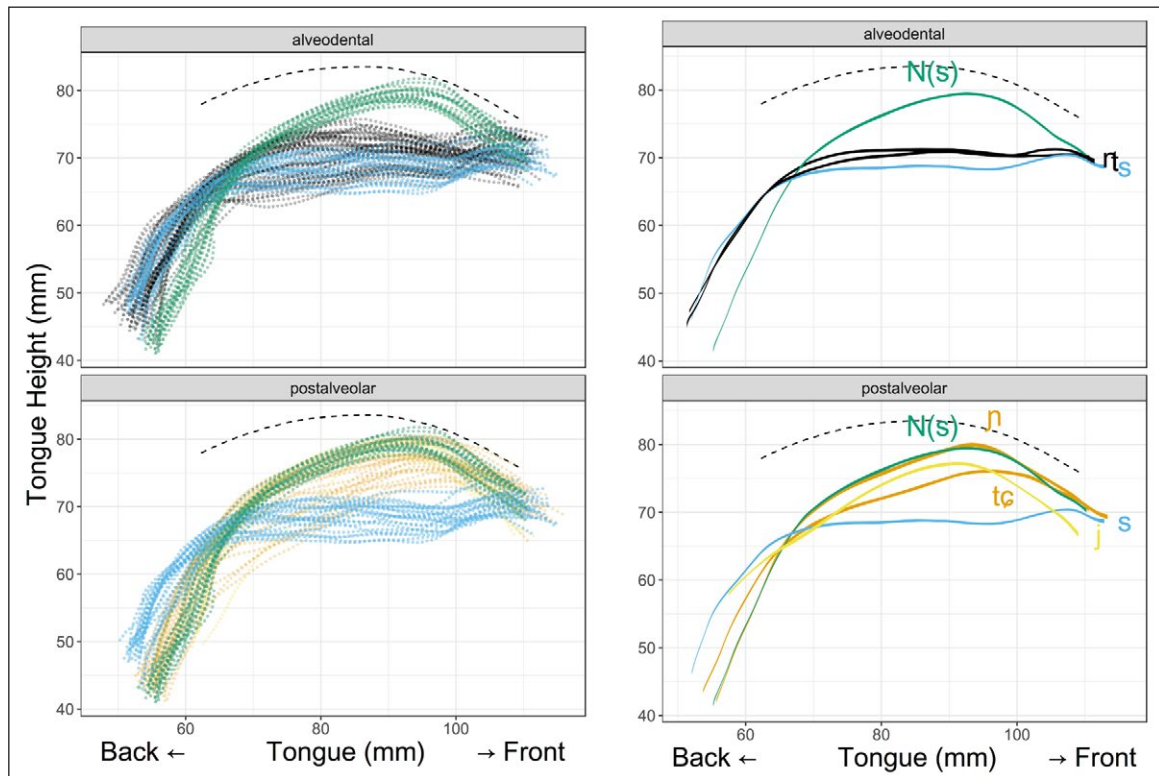
These results are consistent with the dental/alveolar/postalveolar hypothesis, that, in both languages, [s] is a member of the known dental/alveolar class along with [t] and [n] and that [N<sub>s</sub>] is a member of the postalveolar/palatal class along with [tɕ] and [ŋ].



**Figure 5:** Violin plots with overlaid box plots for RMSD values (in mm) for each group in the revised LMER models in Table 15. RMSDs for [s]- and [N<sub>s</sub>]-containing pairs (column 2 in Table 15) are highlighted in orange, whereas homorganic pairs (column 5, Table 15) are indicated in gray and heterorganic pairs (column 8, Table 15) are shown in green. In each plot, probability density of a given RMSD is indicated by the cross-sectional width of the outer shape at that value, interquartile ranges are indicated with box plots within each shape, and median values are indicated by the horizontal line in each box plot.

### 5. Discussion

Comparison of lingual contours in ultrasound images for the relevant sounds shows that in these languages, the nasal substitution pattern relating a voiceless obstruent with its homorganic nasal stop breaks down with [s] and [N<sub>s</sub>]. As illustrated by the contour plots and Smoothing Spline ANOVAs (SS-ANOVAs, Davidson 2006)<sup>12</sup> for S2 in **Figure 6**, these two sounds have different articulations: [s] is an anterior, dental/alveolar sound in both Sasak and Javanese, with an articulation similar to that of [t] and [n], while [ɲ]—whether related to [s] or to [tɕ]—is a postalveolar sound like [tɕ]. This pattern is borne out across all talkers in both languages. Given the morphological pairing of [s] with [ɲ] and the attested variation among possible articulations for [s] sounds (Dart, 1991), we would not have been surprised to find a postalveolar [s] in these languages, with an articulation more similar to [ɲ]. However, this is not the case, nor is it the case that [s] and [N<sub>s</sub>] share a single articulatory configuration that contrasts with both dental/alveolar and postalveolar consonants. Instead, there is an abstract morphophonological relation between an dental/alveolar [s] and derived postalveolar/palatal nasal [ɲ], just as impressionistic accounts lead us to expect.



**Figure 6:** Aggregated tongue splines (left) and SS-ANOVA by sound (right) for Sasak participant S2. The width of lines in the SS-ANOVA plot corresponds to a 95% Bayesian confidence interval for radius *r* at each value of *θ* for each sound’s lingual contours. In the upper figures, unambiguous dental/alveolar splines are black; in the lower figure, unambiguous postalveolar/palatal splines are orange. In all figures, ambiguous splines are blue for [s] and green for [N<sub>s</sub>]. Note that the SS-ANOVA contour for [ɲ] for this speaker lies completely underneath the contour for ambiguous nasal [N<sub>s</sub>].

<sup>12</sup> SS-ANOVA plots were generated in polar-coordinate parameters *r* (radial distance) and *θ* (angle), using the same origin as that used in the RMSD analysis and the `ssanova()` function in the `gss` package (Gu, 2014) in R. Rather than creating SS-ANOVAs by token, we created them by sound. For example, all [t]s were grouped together regardless of the word each came from.

Further examination of the splines and SS-ANOVAs provides a better understanding of the articulation of these consonants in the two languages. Reviewing these images shows that nearly all talkers (10 out of 11 Sasak, 8 out of 8 Javanese) articulated postalveolar sounds [tɕ] and [ɲ] as seen in **Figure 6**—with somewhat dissimilar articulatory positions: [tɕ] was articulated with a tongue-blade constriction posterior to dental/alveolar [t, s, n] but without full raising of the tongue body to the hard palate, whereas [ɲ], whether derived from [tɕ] or [s], was articulated similarly with the tongue blade but with a complete palatal occlusion between the tongue body and hard palate, at or slightly anterior to the constriction locations for palatal sounds [j] and [i] in the two languages. Thus, [ɲ] in Sasak and in Javanese is characterized as having a longer region of palatal constriction than [tɕ], and this articulation is consistent with that observed in electropalatographic data for Peninsular Spanish alveolo-palatal ñ [ɲ] (Martínez Celdrán & Fernández Planas, 2007; Fernández Planas, 2009; Shosted et al., 2012).

These articulator relations align most closely with the first abstract, heterorganic hypothesis in **Table 3a**, and are summarized as in **Table 16**. Furthermore, we see that not only is the relation between [s] and ‘N<sub>s</sub>’ abstract, so too is the relation between [tɕ] and [ɲ]: They are not entirely homorganic sounds.

That these articulatory patterns obtain in a pair of related languages suggests that the anomalous pairing of [s] and [ɲ] with respect to nasal substitution is stable, despite being an abstract linguistic relation. The pairing appears to have resisted pressure to move towards a concrete relation that would, over time, serve to regularize the pattern. Archangeli et al. (2012) find the same sort of stability in Bantu vowel harmony: Bantu height harmony occurs primarily in verbs, where sequences of mid vowels are preferred to mid-high sequences, with the exception of the permitted sequence [e...u]. The same pattern is found in nouns to a lesser degree—with the exception that even [e...u] occurs less often than expected. The difference is that in verbs, the [e...u] sequence arises across a morpheme boundary, and so any changes in this sequence disrupt an entire morphological paradigm while with the nouns, gradual item-by-item change is possible. To recast Archangeli et al. (2012) in terms of the current discussion, every nasal substitution form relating [s] and [ɲ] constitutes pressure against reanalyzing any single instance relating [s] and [ɲ]. Such pressure predicts that anomalous patterns that enter a morphophonological relation (through whatever means) are likely to remain in a language.

In addition to finding a lack of within-paradigm regularity in our study (the concrete-homorganic hypothesis in **Table 3d–f**), we also find little evidence of even partial assimilation. Although [s] is clearly articulated in the dental/alveolar region with the tongue tip/blade, we do not observe articulations of its nasal counterpart ([N<sub>s</sub>]) that are more similar to that of [s] than would be expected from a palatal nasal sound. Comparisons between [N<sub>s</sub>] and [ɲ] result in similar RMSD values (or smaller, in the

**Table 16:** Place relations between Sasak and Javanese sounds that were investigated in this study. [ɲ] corresponds to /ɲ/ as well as ‘N<sub>s</sub>.’

	Dental/Alveolar	Postalveolar	Palatalized postalveolar/ Alveolo-palatal
Plosive	t		
Affricate		tɕ	
Nasal		n	ɲ
Fricative	s		

case of Javanese) than comparisons within [ɲ] contours alone, indicating that [N<sub>s</sub>] is no different from [ɲ] than other productions of [ɲ]. Thus, it is not the case that [s] as a derivational source of [N<sub>s</sub>] influences that articulation of its nasal substitution counterpart with its anterior, apical/laminal articulation, and in fact the RMSD values between [s] and its nasal counterpart [N<sub>s</sub>] are among the largest values observed in the entire data set for each language.

This work uses RMSD calculations to determine whether two sounds are homorganic or heterorganic in articulatory place, giving a quantified comparison of tongue-contour data from ultrasound images. RMSDs provide individual numerical values to represent the magnitudes of spatial difference within each contour comparison, and these values can be used for further comparison. Moreover, when coupled with a mixed-effects linear regression model, RMSD values from multiple talkers can be compared within a single analysis by treating Speaker as a random effect. This kind of comparison enables us to identify patterns that hold generally across speakers of a given language. Statistical comparisons of RMSD values across test conditions of interest allow for an evaluation of similarity or difference between those conditions, adding to the similarity tools proposed in Mielke (2012).

## 6. Conclusion

To conclude, this study has shown that there is an abstract morphophonological relation in the nasal substitution paradigm in both Sasak and Javanese. Our analysis of ultrasound images supports [s/ɲ] nasal substitution as an abstract relation within what appeared to be an otherwise general and concrete homorganic pattern: We show that in both languages, [s] is articulated quite similarly to the dental/alveolar sounds and that the morphologically-related [N<sub>s</sub>] is indistinguishable from [ɲ]. Further examination of the data shows that the voiceless affricate is alveolo-palatal [tʃ], which corresponds to palatal [ɲ] in the same paradigm, again an abstract relation. Thus, the pattern is not as homorganic as impressionistic analysis suggests. We suggest that the pattern has resisted regularization because it is a robust morphophonological relation, held in place by morphological paradigm pressure.

Finally, we introduce the RMSD/LMER method for comparing tongue contours, pooling comparisons from multiple subjects in order to better understand general patterns within each language.

## Additional File

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

- **Appendix.** The appendix contains a complete list of the stimuli for both languages. DOI: <https://doi.org/10.5334/labphon.46.s1>

## Acknowledgements

For the collection of Sasak data, we are grateful to Issyatul Mardiah and Silvana Sandi for their help in identifying stimuli and with our logistics, and to Untung Waluyo, Husni Mu'adz, and the Mataram Lingua Franca Institute (MaLFI) for sponsoring our visit to Lombok. We also thank the University of Hong Kong, students in Spring 2015 LING 3003 field trip class, and Suki Yiu for their support in working with Sasak speakers. For the collection of the Javanese data, we are grateful to Dalan Peranginangin for his help with recruiting and interacting with our speakers. Thank you also to Eric Yadong Liu for his careful work in identifying frames and extracting contours.

## Competing Interests

The authors have no competing interests to declare.

## References

- Adelaar, A. 2005. Malayo-Sumbawan. *Oceanic Linguistics*, 44, 357–388. DOI: <https://doi.org/10.1353/ol.2005.0027>
- Adisasmito-Smith, N. 2004. Phonetic and phonological influences of Javanese on Indonesian (Unpublished doctoral dissertation). Cornell University, Ithaca, New York.
- Apple. 2010. QuickTime Player [Computer program]. Version 7.7.3 (2890). (accessed 12-Nov-14).
- Archangeli, D., Baker, A. and Mielke, J. 2011. Categorization and features: Evidence from American English /r/. In: Ridouane, R. and Clements, G. (Eds.), *Where Do Phonological Contrasts Come from?* 173–196. Amsterdam: John Benjamins.
- Archangeli, D., Mielke, J. and Pulleyblank, D. 2012. Greater than noise: Frequency effects in Bantu height harmony. In: Botma, B. and Noske, R. (Eds.), *Phonological Explorations: Empirical, Theoretical and Diachronic Issues*, 191–222. Berlin: Mouton de Gruyter.
- Archangeli, D., Moll, L. and Ohno, K. 1998. Why not \*nɔ̄. In: *Proceedings from the Main Session of the Chicago Linguistics Society's Thirty-fourth Meeting*, 1–26.
- Archangeli, D., Tanashur, P. and Yip, J. 2016. Illustration of the IPA: Sasak, Meno-Mené dialect. University of Arizona and University of Hong Kong. (unpublished manuscript, University of Arizona and University of Hong Kong).
- Austin, P. 2003. The linguistic ecology of Lombok, eastern Indonesia. (unpublished manuscript, SOAS, University of London).
- Austin, P. 2010. Dialect variation in the voice system of Sasak: when is a nasal-verb not a nasal-verb? In: Indonesian Syntax Symposium, TUFs. [to appear in proceedings].
- Austin, P. 2013. Too many nasal verbs: dialect variation in the voice system of Sasak. *Voice Variation in Austronesian Languages of Indonesia. NUSA: Linguistic Studies of Languages in and Around Indonesia*, 54, 29–46.
- Baker, A., Archangeli, D. and Mielke, J. 2011. Variability in American English s-retraction suggests a solution to the actuation problem. *Language Variation and Change*, 23, 347–374. DOI: <https://doi.org/10.1017/S0954394511000135>
- Barry, M. 1991. Temporal modelling of gestures in articulatory assimilation. In: *Proceedings of the 12th International Congress of Phonetic Sciences*, 4, 14–17.
- Beepa. 2015. Fraps [computer program]. Version 3.5.99. <http://www.fraps.com>. (accessed 29-May-15).
- Berry, J., Fasel, I., Fadiga, L. and Archangeli, D. 2012. Training deep nets with imbalanced and unlabeled data. In: *Interspeech*, 1756–1759.
- Blust, R. 2004. Austronesian nasal substitution: a survey. *Oceanic Linguistics*, 43, 73–148. DOI: <https://doi.org/10.1353/ol.2004.0004>
- Boersma, P. and Weenink, D. 2015. Praat: Doing phonetics by computer [Computer program]. Version 5.4.06. <http://www.praat.org>. (accessed 21-Feb-15).
- Browman, C. and Goldstein, L. 1986. Towards an articulatory phonology. *Phonology*, 3(21), 219–252.
- Browman, C. and Goldstein, L. 1992. Articulatory phonology: An overview. *Phonetica*, 49, 155–180. DOI: <https://doi.org/10.1159/000261913>
- Brunelle, M. 2010. The role of larynx height in the Javanese. *Austronesian and Theoretical Linguistics*, 167, 7.
- Celata, C., Calamai, S., Ricci, I. and Bertini, C. 2013. Nasal place assimilation between phonetics and phonology: An EPG study of Italian nasal-to-velar clusters. *Journal of Phonetics*, 41, 88–100. DOI: <https://doi.org/10.1016/j.wocn.2012.10.002>
- Cheng, C., Chen, J.-Y. and Gubian, M. 2013. Are Mandarin sandhi tone 3 and tone 2 the same or different? The results of functional data analysis. *Pacific Asia Conference on Language, Information, and Computation*, 27, 296–301.

- Clynes, A. 1995. Sasak. In: Tryon, D. (Ed.), *Comparative Austronesian Dictionary: An Introduction to Austronesian Studies*, 511–520. Berlin: Mouton de Gruyter.
- Csapó, T. G. and Lulich, S. M. 2015. Error analysis of extracted tongue contours from 2d ultrasound images. In: *Sixteenth Annual Conference of the International Speech Communication Association*.
- Dart, S. N. 1991. Articulatory and acoustic properties of apical and laminal articulations. *UCLA Working Papers in Linguistics*, 79, 1–155.
- Davidson, L. 2005. Addressing phonological questions with ultrasound. *Clinical Linguistics & Phonetics*, 19, 619–633. DOI: <https://doi.org/10.1080/02699200500114077>
- Davidson, L. 2006. Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance. *Journal of the Acoustical Society of America*, 120(1), 407–415. DOI: <https://doi.org/10.1121/1.2205133>
- De Guzman, V. P. 1978. A case for nonphonological constraints on nasal substitution. *Oceanic Linguistics*, 17, 87–106. DOI: <https://doi.org/10.2307/3622906>
- Delattre, P. C. and Freeman, D. C. 1968. A dialect study of American r's by x-ray motion picture. *Linguistics*, 44, 29–68. DOI: <https://doi.org/10.1515/ling.1968.6.44.29>
- Dudas, K. 1976. The phonology and morphology of Modern Javanese (Unpublished doctoral dissertation). University of Illinois, Urbana-Champaign.
- Fasel, I. and Berry, J. 2010. Deep belief networks for real-time extraction of tongue contours from ultrasound during speech. In: *20th International Conference on Pattern Recognition (ICPR)*, 1493–1496. DOI: <https://doi.org/10.1109/ICPR.2010.369>
- Fernández Planas, A. M. 2009. Características linguopalatales de la nasal palatalizada en español [Linguopalatal characteristics of the palatalized nasal in Spanish]. *Estudios de Fonética Experimental*, 18, 161–174.
- Gibbon, F. E., Yuen, I., Lee, A. and Adams, L. 2007. Normal adult speakers' tongue palate contact patterns for alveolar oral and nasal stops. *Advances in Speech Language Pathology*, 9(1), 82–89. DOI: <https://doi.org/10.1080/14417040600954824>
- Gick, B., Pulleyblank, D., Campbell, F. and Mutaka, N. 2006. Low vowels and transparency in Kinande vowel harmony. *Phonology*, 23, 1–20. DOI: <https://doi.org/10.1017/S0952675706000741>
- Graff, P. and Jaeger, T. F. 2009. Locality and feature specificity in OCP effects: Evidence from Aymara, Dutch, and Javanese. In: *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, 45, 127–141.
- Gu, C. 2014. GSS: General smoothing splines. [R Package]. Version 2.1–4. <http://cran.r-project.org>. (last accessed 26-Jan-15).
- Harnad, S. 2003. Categorical perception. In: *Encyclopedia of Cognitive Science*, 67. MacMillan: Nature Publishing Group.
- Hayward, K. and Mulijono. 1991. The dental/alveolar contrast in Javanese. *Bulletin of the School of Oriental and African Studies*, 54(01), 126–144. DOI: <https://doi.org/10.1017/S0041977X00009642>
- Herawati, N., Rustono, M. H. and Poedjosoedarmo, S. 2016. Prefixes of denominal verb formation in Javanese language. *Journal of Education and Social Sciences*, 5, 151–156.
- Hoffmann, C. 1973. The vowel harmony system of the Okpe monosyllabic verb. *Department of Linguistics and Nigerian Languages Research Notes*, 6, 79–112.
- Horne, E. 1961. *Beginning Javanese*. New Haven: Yale University Press.
- Irfana, M. and Sreedevi, N. 2016. Coarticulatory aggression and direction of coarticulation: An ultrasound study. *International Journal of Innovative Research and Development*, 5(5), 24–29.
- Jacq, P. 1998. How many dialects are there? In: *Working Papers in Sasak*, 67–90. Melbourne: University of Melbourne.



- Kenstowicz, M. and Kisseberth, C. 1979. *Generative phonology: Description and theory*. London: Academic Press.
- Kim, H. and Jongman, A. 1996. Acoustic and perceptual evidence for complete neutralization of manner of articulation in Korean. *Journal of Phonetics*, 24(3), 295–312. DOI: <https://doi.org/10.1006/jpho.1996.0016>
- Kiparsky, P. 1982. Lexical morphology and phonology. In: Yang, I.-S. (Ed.), *Linguistics in the Morning Calm*, 3–91. Seoul: Hanshin.
- Kochetov, A. and Colantoni, L. 2011. Spanish nasal assimilation revisited: A cross-dialect electropalatographic study. *Laboratory Phonology*, 2, 487–523. DOI: <https://doi.org/10.1515/labphon.2011.018>
- Kuznetsova, A., Brockhoff, P. and Christensen, R. 2012. lmerTest: Tests in linear mixed effect models. [R Package]. Version 2.0–20. <http://cran.r-project.org>. (last accessed 25-Jan-15).
- Lee, A. 2001. Nasal substitution in Malayo-Javanic: A new analysis. In: *Austronesian Formal Linguistics Association*, 7.
- Lee-Kim, S.-I., Davidson, L. and Hwang, S. 2013. Morphological effects on the darkness of English intervocalic /ɪ/. *Laboratory Phonology*, 4(2), 475–511. DOI: <https://doi.org/10.1515/lp-2013-0015>
- Li, M., Kambhamettu, C. and Stone, M. 2005. Automatic contour tracking in ultrasound images. *Clinical Linguistics & Phonetics*, 19(6–7), 545–554. <http://www.pcultrasound.com>. (accessed 13-Mar-15).
- Lieberman, A. M., Harris, K. S., Hoffman, H. S. and Griffith, B. C. 1957. The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, 54(5), 358–368. DOI: <https://doi.org/10.1037/h0044417>
- Marli, Y. 2015. Welcome to NTB, the province of West Nusa Tenggara. Mataram, Lombok. (unpublished manuscript, Mataram Lingua Franca Institute).
- Martínez Celdrán, E. and Fernández Planas, A. M. 2007. *Manual de fonética española: Articulaciones y sonidos del español [Handbook of spanish phonetics: Spanish articulations and sounds]*. Barcelona, Spain: Ariel.
- Matthews, M. 2015. An acoustic investigation of Javanese stop consonant clusters. In: Camp, A., Otsuka, Y., Stabile, C. and Tanaka, N. (Eds.), *AFLA 21: The Proceedings of the 21st Meeting of the Austronesian Formal Linguistics Association*, 201–218.
- Mester, R. A. 1986. Studies in tier structure (Unpublished doctoral dissertation). University of Massachusetts.
- Mielke, J. 2012. A phonetically based metric of sound similarity. *Lingua*, 122(2), 145–163. DOI: <https://doi.org/10.1016/j.lingua.2011.04.006>
- Mielke, J., Baker, A. and Archangeli, D. 2010. Variability and homogeneity in American English /ɪ/ allophony and /s/ retraction. In: Fougéron, C., Kuhnert, B., D’Imperio, M. and Vallée, N. (Eds.), *Variation, Detail, and Representation: Labphon*, 10, 699–730. Berlin: Mouton de Gruyter.
- Mielke, J., Baker, A. and Archangeli, D. 2016. Individual-level contact reduces phonological complexity: evidence from bunched and retroflex /ɪ/. *Language*, 92, 101–140. DOI: <https://doi.org/10.1353/lan.2016.0019>
- Miller, A. and Finch, K. 2011. Corrected high-frame rate anchored ultrasound with software alignment. *Journal of Speech, Language, and Hearing Research*, 54(2), 471–486. DOI: [https://doi.org/10.1044/1092-4388\(2010/09-0103\)](https://doi.org/10.1044/1092-4388(2010/09-0103))
- Mohanan, K. P. 1982. Lexical phonology (Unpublished doctoral dissertation). Massachusetts Institute of Technology, Cambridge, MA.
- Mutaka, N. 1995. Vowel harmony in Kinande. *Journal of West African Languages*, 25, 41–55.


- Myers, J. and Tsay, J. 2008. Neutralization in Taiwan Southern Min tone sandhi. In: Hsiao, Y.-C., Hsu, H.-C., Wee, L.-H. and Ho, D.-A. (Eds.), *Interfaces in Chinese Phonology: Festschrift in Honor of Matthew Y. Chen on his 70th Birthday*, 47–78. Taiwan: Academia Sinica.
- Nothofer, B. 1980. *Dialektgeographische Untersuchungen in West-Java und im Westlichen Zentral-Java*. Harrassowitz.
- Nothofer, B. 2006. Javanese. In: Keith, B. (Ed.), *Encyclopedia of Language and Linguistics* (Second Edition). 113–115. Oxford: Elsevier.
- Öhman, S. E. G. 1966. Coarticulation in VCV utterances: Spectrographic measurement. *Journal of the Acoustical Society of America*, 39, 151–168. DOI: <https://doi.org/10.1121/1.1909864>
- Omar, A. P. 1973. Uvwię – a case of vowels merging. *Research Notes from the Department of Linguistics and Nigerian Languages, University of Ibadan*, 6, 113–143.
- Pater, J. 1999. Austronesian nasal substitution and other NÇ effects. In: Kager, R. van der Hulst, H. and Zonneveld, W. (Eds.), *The Prosody-Morphology Interface*, 310–343. Cambridge: Cambridge University Press.
- Pater, J. 2001. Austronesian nasal substitution revisited: What's wrong with \*NC (and what's not). In: Lombardi, L. (Ed.), *Segmental Phonology in Optimality Theory: Constraints and Representations*, 159–182. Cambridge: Cambridge University Press.
- Poedjosoedarmo, G. 1986. The symbolic significance of pharyngeal configuration in Javanese speech: Some preliminary notes. *Miscellaneous Studies of Indonesian and Other Languages in Indonesia. NUSA*, 8.
- Poedjosoedarmo, G. R. 1993. Uses of phonation type in Javanese. *Oceanic Linguistics Special Publications*, 24, 123–131.
- Port, R. F. 1996. The discreteness of phonetic elements and formal linguistics: Response to A. Manaster Ramer. *Journal of Phonetics*, 24, 491–511. DOI: <https://doi.org/10.1006/jpho.1996.0027>
- Port, R. F. and O'Dell, M. L. 1985. Neutralization of syllable-final voicing in German. *Journal of Phonetics*, 13, 455–471.
- Ras, J. J. 1985. *Inleiding tot het modern javaans*, 1. Foris Publications.
- R Core Team. 2016. R: A language and environment for statistical computing. Vienna, Austria. Retrieved from: <https://www.R-project.org/> (Version 3.3.2 (last accessed 26-Jan-15)).
- Rehg, K. and Sohl, D. 1981. *Ponapean Reference Grammar*. Honolulu: University Press of Hawai'i.
- Reid, L. A. 2000. Sources of proto-oceanic initial prenasalization: The view from outside Oceanic. *Oceanic Linguistics Special Publications*, 29, 30–45.
- Robson, S. 1992. *Javanese Grammar for Students*. Monash University: Centre of Southeast Asian Studies.
- Robson, S. and Wibisono, S. 2002. *Javanese English Dictionary*. Tuttle Publishing.
- Roussos, A., Katsamanis, A. and Maragos, P. 2009. Tongue tracking in ultrasound images with active appearance models. In: *2009 16th IEEE International Conference on Image Processing (ICIP)*, 1733–1736. DOI: <https://doi.org/10.1109/ICIP.2009.5414520>
- Sato, Y. 2008. Minimalist interfaces: Selected issues in Indonesian and Javanese (Unpublished doctoral dissertation). University of Arizona, Tucson, Arizona.
- Sato, Y. 2015. Argument ellipsis in Javanese and voice agreement. *Studia Linguistica*, 69, 58–85. DOI: <https://doi.org/10.1111/stul.12029>
- Shosted, R., Hualde, J. I. and Scarpace, D. 2012. Palatal complexity revisited: An electropalatographic analysis of /ɲ/ in Brazilian Portuguese with comparison

- to Peninsular Spanish. *Language and Speech*, 55(4), 477–502. DOI: <https://doi.org/10.1177/0023830911434120>
- SplitmediaLabs. 2015. XSplit Broadcaster [computer program]. Version 1.3.1403.1202. <https://www.xsplit.com>. (accessed 29-May-15).
- Stephenson, L. and Harrington, J. 2002. Assimilation of place of articulation: Evidence from English and Japanese. In: *Proceedings of the 9th Australian International Conference on Speech Science and Technology*, 592–597.
- Suharno, I. 1982. *A Descriptive Study of Javanese*, 45. Australian National University: Department of Linguistics, Research School of Pacific Studies.
- Sung, J.-H. 2015. *A Cross-linguistic Articulatory Analysis of Palatalization in Korean, English, and Scottish Gaelic*. (Unpublished doctoral dissertation). University of Arizona.
- Teleded. 2015. Echo Wave II [computer program]. Version 3.4.0. <http://www.pcultrasound.com>. (accessed 13-Mar-15).
- Thurgood, E. 2004. Phonation types in Javanese. *Oceanic Linguistics*, 43(2), 277–295. DOI: <https://doi.org/10.1353/ol.2005.0013>
- Warner, N., Jongman, A., Sereno, J. and Kems, R. R. 2004. Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from Dutch. *Journal of Phonetics*, 32, 251–274. DOI: [https://doi.org/10.1016/S0095-4470\(03\)00032-9](https://doi.org/10.1016/S0095-4470(03)00032-9)
- Winter, B. and Röttger, T. 2011. The nature of incomplete neutralization in German: Implications for laboratory phonology. *Grazer Linguistische Studien*, 76, 55–74.
- Wouk, F. 1999. Sasak is different: A discourse perspective on voice. *Oceanic Linguistics*, 38, 91–114. DOI: <https://doi.org/10.2307/3623394>
- Yu, A. 2007. Understanding near mergers: The case of morphological tone in Cantonese. *Phonology*, 24, 187–214. DOI: <https://doi.org/10.1017/S0952675707001157>
- Zharkova, N. and Hewlett, N. 2009. Measuring lingual coarticulation from midsagittal tongue contours: Description and example calculations using English /t/ and /a/. *Journal of Phonetics*, 37(2), 248–256. DOI: <https://doi.org/10.1016/j.wocn.2008.10.005>
- Zsiga, E. C. 1995. An acoustic and electropalatographic study of lexical and postlexical palatalization in American English. In: Connell, B. and Arvaniti, A. (Eds.), *Papers in Laboratory Phonology*, 4, 282–302. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511554315.020>

**How to cite this article:** Archangeli, D., Yip, J., Qin, L. and Lee, A. 2017 Phonological and phonetic properties of nasal substitution in Sasak and Javanese. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 8(1): 21, pp. 1–37, DOI: <https://doi.org/10.5334/labphon.46>

**Submitted:** 25 August 2016 **Accepted:** 24 April 2017 **Published:** 04 September 2017

**Copyright:** © 2017 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

 *Laboratory Phonology: Journal of the Association for Laboratory Phonology* is a peer-reviewed open access journal published by Ubiquity Press.

**OPEN ACCESS** 