

Wm. G. Bennett, Maxine Diemer, Justine Kerford, Tracy Probert & Tsholofelo Wesi

Rhodes University
w.bennett@ru.ac.za

Setswana (also known as ‘Tswana’ or, more archaically, ‘Chuana’ or ‘Sechuana’) is a Bantu language (group S.30; ISO code tsn) spoken by an estimated four million people in South Africa. There are a further one million or more speakers in Botswana, where it is the dominant national language, and a smaller number of speakers in Namibia. The recordings accompanying this article were mostly produced with a 21-year-old male speaker from the area of Taung, North-West province, South Africa. Some of the accompanying recordings are of a 23-year-old female speaker from Kuruman (approximately 150 km west of Taung). The observations reported here are based on consulting with both these speakers, as well as a third speaker, from Kimberley. All three were speakers of South African Setswana varieties. For discussion of some differences between these varieties and more Northern and Eastern Setswana dialects – including those spoken in Botswana – see (Doke 1954, Cole 1955, University of Botswana 2001).

Previous descriptive work on Setswana includes Jones & Plaatje (1916), Jones (1928), Cole (1955), Chebanne, Creissels & Nkhwa (1997), University of Botswana (2001), as well as a specimen from the 1949 ‘Principles of the IPA’ (IPA 2010: 348–349).

Consonants

The consonants are shown in the table below, followed by example words for each. The transcriptions given here do not indicate effects of predictable vowel length, discussed with suprasegmental features later on. Consonants in parentheses are of marginal status, and are discussed below.

	Bilabial	Labio-dental	Alveolar (apical)	Alveolar lateral	Post-alveolar	Palatal	Velar	Uvular	Glottal
Plosive	p p ^h b		t t ^h d				k k ^h		(ʔ)
Affricate			ts ts ^h	tʃ tʃ ^h	tʃ tʃ ^h dʒ			qχ	
Nasal	m		n			ɲ	ŋ		
Trill			r						
Fricative		(f)	s		ʃ			χ	h
Approximant	w			l		j			

CONSONANT	PHONETIC REPRESENTATION	ORTHOGRAPHIC REPRESENTATION	GLOSS
/p/	pùdì	podì	‘goat’
/t/	táú	tau	‘lion’
/k/	kè	ke	‘I’ ¹
/p ^h /	p ^h ìrì	phiri	‘hyena’
/t ^h /	t ^h átò	thatò	‘love (N)’
/k ^h /	k ^h ùdù	khudu	‘tortoise’
/b/	bàlá	bala	‘count’
/d/	dìdžó	dijò	‘food’
[ʔ]	ʔáχà	aga	‘build’
/ts/	tsèbé	tsèbê	‘ear’
/tʃ/	tʃòmà	tšoma	‘hunt’
/tʰ/	tʰálà	tlala	‘famine’
/ts ^h /	ts ^h éχà	tshêga	‘laugh’
/tʃ ^h /	tʃ ^h ùbà	tšhuba	‘set alight’
/tʰ ^h /	tʰ ^h àpì	tlhapi	‘fish’
/dʒ/	džá	ja	‘eat’
/qχ/	qχósì	kgôsi	‘chief’
/f/	fùlà	fula	‘graze’
/s/	sàlà	sala	‘remain’
/ʃ/	šá	ša	‘burn (INTRANS)’
/χ/	χóχà	gôga	‘pull’
/h/	hùmà	huma	‘become rich’
/m/	ṁmá	mma	‘mother’
/n/	nàmà	nama	‘meat’
/ɲ/	ɲálà	nyala	‘marry’
/ŋ/	ŋàŋà	nganga	‘dispute, argue’
/r/	rátà	rata	‘love’
/l/	lálà	lala	‘sleep’
/w/	wà	wa	‘fall’
/j/	jà	ya	‘go’

Voiceless unaspirated plosives and affricates may be realized as ejectives. One of the speakers consulted for this Illustration systematically produced these sounds as clear ejectives, but the other two speakers did not realize them as ejectives in most tokens. Variable realization as ejectives is also reported in other Setswana data, from previous work, see Coetzee & Pretorius (2010: 406), IPA (2010). The alveolar lateral-release affricate /tʰ/ tends to sound more clearly ejective than other stops and fricatives, and its aspirated counterpart /tʰ^h/ seems to have less aspiration than the other affricates, /ts^h/ and /tʃ^h/.

There is considerable variation in the labial consonants: /f/ varies with /h/ quite generally in most dialects, e.g. [k^hùtsàfàtsà] ~ [k^hùtsàhàtsà] ‘make (oneself) sad’, and [fá] ~ [há] ‘give’.² The voiced stop /b/ may be realized without full closure, as [β]. It also varies phonologically with /ts/ in some words (Cole 1955: 83), e.g. [màbóχò] ~ [màtsóχò] ‘arms’. This variation does not seem to obscure phonemic distinctions, however. In post-nasal context (see discussion below), /h/ that varies with /f/ hardens to [p^h] (e.g. [χòhétà] ‘to finish’ >

¹ The morpheme /kè/ ‘I’ is usually analyzed as a prefix in the Bantuist literature. It is written as a separate word orthographically, however, and our consultant reported no awkwardness about pronouncing it on its own, with no following stem.

² Some sources on other dialects report bilabial [ɸ] rather than [f] (e.g. IPA 2010); our primary consultant’s articulation was labiodental.

[m^hétsà] ‘finish me’), while other invariant /h/ hardens to /k^h/ ([χ^húmà] ‘be rich’ > [k^húmà] ‘wealth’).

The glottal stop is non-phonemic, but can be found at the start of underlying vowel-initial morphemes. This may be at the start of a word before a vowel-initial stem (as in [ʔáχà] ‘build’ in the list above), or between a vowel-initial stem and a vowel-final prefix (as in /χ^hè-émà/ ‘to stand’ [χ^hèémà] ~ [χ^hèèmà]). In both contexts, the glottal stop appears to be optional, and our primary consultant normally did not produce glottal stops in the latter context.

The voiced alveolar plosive /d/ occurs only before the phonemic high vowels /i/ and /u/. Previous phonemic analyses generally view it as an allophone of /l/, since [l] does not occur before /i/ and /u/. Before raised mid-high vowels [ɥ i], as well as the glide /w/, we find [l] and not [d], e.g. [mùl]mì ‘cultivator’ and [bòdʒàlwá] ‘beer’.

Intervocally, /h/ is often reduced or elided (e.g. /χ^hè-hùdùwà/ ‘to stir’ is produced as [χ^hèùdùwà]). The trill /r/ normally has 2–3 taps (see section on doubled consonants below for further discussion). The velar nasal has a somewhat restricted distribution: it is most commonly found word-finally, before a velar consonant, or before a /w/ (though with some exceptions, such as [hàhà] ‘dispute’, seen in the list of consonants above).

Nasal consonants may occur before other consonants. Nasals in this context are syllabic, and may undergo penultimate syllable lengthening; see discussion under ‘Doubled consonants’ below for further details. Word-final /ŋ/ commonly assimilates to the place of a following consonant in the next word in connected speech.

The status of the uvular consonants

The phoneme given here as /χ/ normally varies in articulation between a uvular fricative [χ]³ and a voiceless uvular trill [ʀ], though some tokens (especially in connected speech) are closer to [x] or are reduced to [h]. An example of the trilled articulation in the word /nòχà/ ‘snake’ is illustrated in [Figure 1](#). Four periodic lines (which do not match the period of glottal pulses) are clearly distinguishable in the spectrogram during the word-medial /χ/, marked with dashed lines. These correspond to three distinct peaks in the waveform, indicating three separate contacts with the uvula – i.e. a trilled articulation.⁴

Most previous descriptions (generally based on other varieties of Setswana) describe the phoneme /χ/ as a velar fricative /x/ rather than uvular (Doke 1954, Cole 1955, IPA 2010, Gouskova, Zsiga & Tlale Boyer 2011). It is cognate with aspirated velar stops in related languages (compare /áχà/ ‘build’ with Xhosa /ak^ha/). While occasional tokens of /χ/ do sound close to velar [x], the normal realization sounds distinctly post-velar in our data. In post-nasal context, /χ/ hardens to the uvular affricate /qχ/ (e.g. [χá] ‘fetch water’ > [h̄q̄r̄ó] ‘water pot’).

The uvular affricate /qχ/ has been characterized in some previous work as an aspirated uvular plosive /q^h/ (Chebanne et al. 1997, University of Botswana 2001), or as an aspirated velar affricate /kx^h/ (Doke 1954, Cole 1955, IPA 2010, Gouskova et al. 2011).⁵ In our data, the release of this consonant seems fricated rather than just aspirated, though. The quality of the interval between the burst and the following vowel is much more like the fricative /χ/ than the aspiration following stops (and it may involve the same trilling as /χ/; the included recording of [h̄q̄r̄ó] ‘water pot’ illustrates this very clearly). This can be seen in [Figure 2](#), showing both /qχ/ and /k^h/, with the interval from the burst to the start of the following vowel

³ In some tokens the constriction of /χ/ is shortened, which can make it seem closer to [h].

⁴ Similarly trilled articulations of /χ/ are documented in Wolof (Ladefoged & Maddieson 1996: 167). This can also be seen in the recordings of a previous IPA Illustration of Tashlhiyt Berber (Ridouane 2014), and several linguists have impressionistically reported similar articulations of post-velar fricatives in German and Dutch. Thus, while the trilling we notice in Setswana /χ/ is remarkable, it is not necessarily unusual cross-linguistically.

⁵ The previous works that characterize this sound as a uvular plosive do note the affricate as a possible realization; see University of Botswana (2001: 14), Chebanne et al. (1997: 10 fn.).

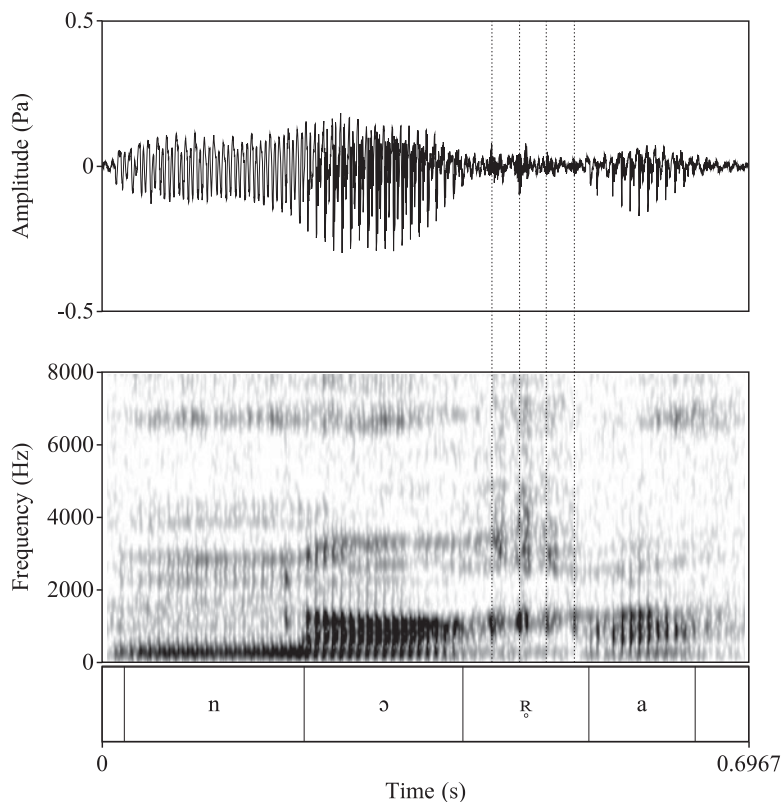


Figure 1 Waveform and spectrogram for /nəχà/ 'snake', showing trilled articulation of /χ/. Dashed lines mark the distinct contacts that comprise the trill.

Table 1 Duration of release intervals for /qχ/ and /k^h/ in [qχʒsɪ] 'chief' and [k^húmɔ] 'wealth'.

Token	qχ burst + affrication	k ^h burst + aspiration
1	97 ms	102 ms
2	114 ms	110 ms
3	80 ms	122 ms
Mean	97 ms	111 ms

indicated by dashed lines. The release of /qχ/, though clearly voiceless, has periodic noise, which closely resembles the trilling normally seen in the uvular fricative /χ/. This indicates that some uvular constriction is retained after release of the closure, unlike the aspiration seen with /k^h/. The duration of the release intervals are fairly close, though measurement of three tokens of each suggests that the release of /qχ/ may be slightly shorter than /k^h/, as shown in Table 1.

Doubled consonants

The liquid and nasal consonants /m n ɲ ŋ r l/ may be 'doubled'; we refer to them in description as 'doubled' and represent them as ⟨mm⟩ rather than ⟨m:⟩ in order to abstract away from issues of phonological representation. Examples are given in (1) below.

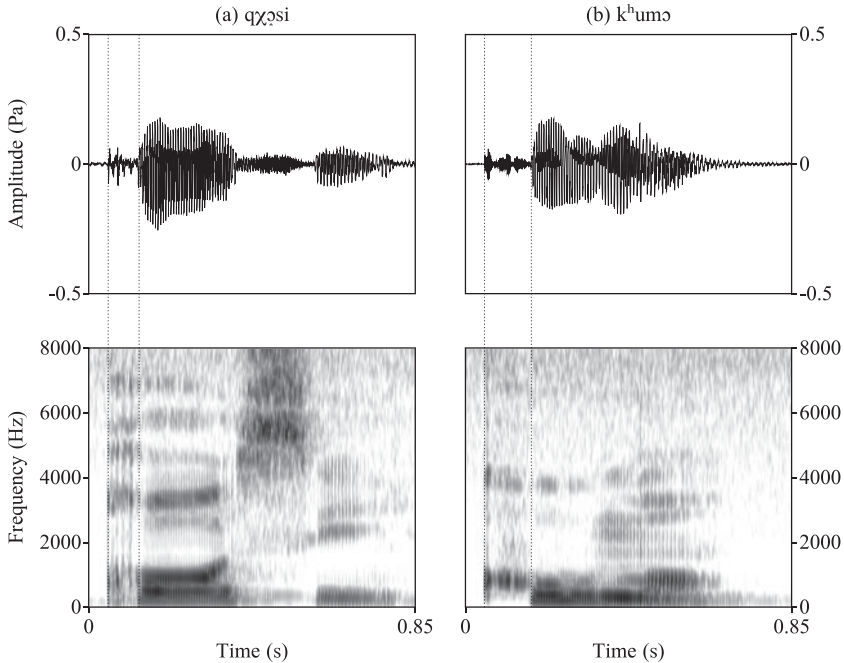


Figure 2 Release of affricate /qχ/ compared to aspirated plosive /kʰ/. Panel (a) shows [qχ] in [qχʒsi] 'chief', with release interval marked by dotted lines. Panel (b) shows release of [kʰ] in [kʰúmò] 'wealth'.

(1) *Examples of doubled sonorant consonants*

- a. ñmá 'mother'
- b. ñnè 'four'
- c. mǒńńá 'man'
- d. ńńwè 'one'
- e. ñré 'father'

The distinction between doubled and singleton consonants may be phonologically contrastive for at least the nasals, as in [mǒńńá] 'man' vs. [mòńńà] 'jealousy'. The typical origin of these doubled consonants appears to be historical loss of a penultimate vowel between two identical sonorants (University of Botswana 2001).

Doubled nasals typically show a noticeable dip or drop in intensity in the middle of the segment. This is illustrated in Figure 3, with the dip marked by the arrow.

The duration of doubled sonorants in penultimate-syllable position, like those in (1), normally seems to be approximately three times that of singletons. Comparing the recordings of [mòńńà] 'jealousy' and [mǒńńá] 'man' included with this article, we find that the singleton [n] has a duration of 99 ms, while the doubled [nn] is 310 ms. Comparing the singleton [n] in [χǒńwá] 'to drink' to the doubled [nn] in [χǒńńá] 'to live', we find a similar ratio: 137 ms in the singleton nasal, 437 ms in the doubled one. Additionally, while the singleton trill [r] normally involves 2–3 contacts, the doubled trills involve 6–8 contacts (rather than 4–6).⁶ This may be a reflection of penultimate syllable lengthening (discussed later on).

⁶ We asked the speakers we consulted for their intuitions about syllables; for words with doubled consonants, they were not consistent. One speaker segmented *rré* 'father' as [r.re], in line with the syllabic geminate interpretation noted here. Two other speakers characterized it as a single syllable. We take this to be a task effect, and infer that asking speakers is not a viable way to probe subconscious knowledge about syllable boundaries and syllabification.

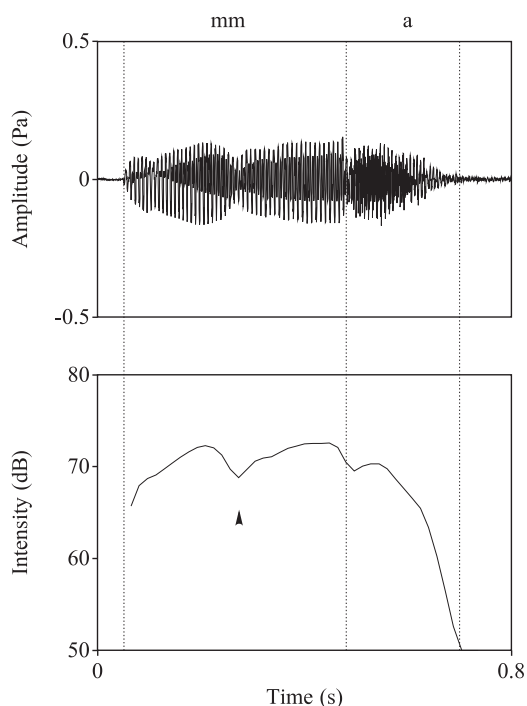


Figure 3 Intensity discontinuity seen in doubled /mm/ from [m̩.má] ‘mother’. Dashed lines mark segment boundaries; arrow marks intensity dip.

Though previous descriptions imply that doubling of /l/ is quite common (Cole 1955: 29, University of Botswana 2001: 25), it was rare in the data we recorded from our consultants. One token of /m̩-l̩l̩/ ‘fire’ as [m̩l̩.l̩] was collected from a sentence from our primary consultant. The duration of the doubled liquid in this form is 156 ms, substantially longer than singleton [l]s. (When the same word was pronounced in isolation, as [m̩l̩l̩], the singleton [l]s had durations of approximately 80 ms to 110 ms). A clear dip in intensity can be seen midway through the doubled [ll], comparable to that observed in the doubled nasals. This is shown in Figure 4.

Some previous work on Setswana describes the doubled consonants as syllabic (Cole 1955: 28, Chebanne et al. 1997, IPA 2010). This description alone is not sufficient, however: the doubled liquids are not observed before another consonant – a restriction that does not follow from simply interpreting them as syllable nuclei. Coetzee (2001) and Gouskova et al. (2011: 2123) analyze the doubled consonants as geminates, comprising the nucleus of a syllable and the onset of the following one, e.g. [m̩.má] ‘mother’. This interpretation seems consistent with our observations. The dip in intensity in doubled nasals and [ll] suggests that these segments are not simply lengthened, but consist on some level of multiple consonantal units. This also seems to be corroborated by the observation that doubled consonants in penultimate position are three times as long as their singleton counterparts. Penultimate syllables are normally lengthened (discussion below). If the doubled consonants are geminates in which the first half of the geminate is syllabic, then only half of the doubled segment should undergo penultimate lengthening. Thus, we might expect lengthening of a doubled consonant to produce a sequence with three abstract units of length: one from each half of the geminate, plus one more unit of length from further lengthening of just the half of the geminate that is syllabic (i.e. /m.ma/ → [m̩.ma]).

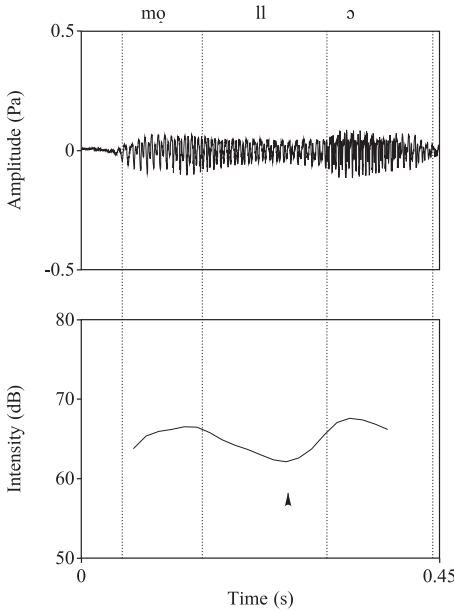


Figure 4 Intensity dip (marked by an arrow) in doubled [ll], from /mo-lele/ ‘fire’ in connected speech.

Post-nasal consonant hardening

Previous work on Tswana observes a series of changes that apply to root-initial consonants in ‘post-nasal’ context, conditioned by several prefixes.⁷ These changes include: devoicing and/or lenition of voiced plosives; fortition of /h/ and /f/⁸ to aspirated plosives [k^h] and [p^h], respectively; affrication of other fricatives /χ/, /s/ and /ʃ/ to [qχ], [ts^h], and [tʃ^h]; fortition of /r/ to [t^h] and of /l/ to [t]; and insertion of [k] before vowel-initial roots. Examples are given in (2).

(2) *Consonant fortition examples*

a.	sɛχɔ̃fàtsà	‘bless’	ts ^h ɛχɔ̃fàtʃɔ̃	‘blessing’	s ~ ts ^h
b.	rátà	‘love (v)’	ì ^h átà	‘love oneself’	r ~ t ^h
c.	lwà	‘fight (v)’	ɲtwá	‘fight (N)’	l ~ t
d.	χàlàlèlè	‘shine (v)’	qχàlàlèlè	‘shine (N)’	χ ~ qχ
e.	χɔ̃ítsì	‘to know’	kítʃɔ̃	‘knowledge’	Ø ~ k
			wàŋkítsì	‘he knows me’	
			kèàikítsì	‘I know myself’	

We did not observe any consistent phonetic differences between the plosives and affricates derived in this way, as compared to their underlying counterparts in root-initial contexts. VOT values for the stops derived from liquids (/r/ → [t^h] and /l/ → [t]) seem comparable to underlying /t^h/ and /t/, suggesting that the hardening is a categorical shift (though more rigorous quantitative examination would be needed to confirm this). For further discussion

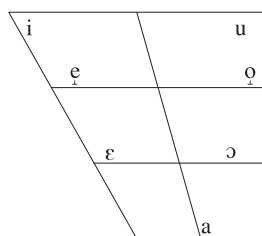
⁷ These consist of the reflexive prefix /i(N)-/, the first person singular object marker [iN-], and the prefix for nouns of class 9, /N-/. All three of these prefixes historically had a homorganic nasal consonant (represented as ⟨N⟩ here). This nasal is retained in the first-person singular object marker, but is systematically absent in the reflexive prefix, which always surfaces as [i-]. The prefix for nouns of class 9 surfaces as a homorganic nasal if the root is monosyllabic, like with /lwà/ ‘fight’, but otherwise has no overt segmental realization, as in ‘blessing’ in (2a).

⁸ Recall from the earlier section on consonants that /f/ varies with /h/, but /f/ still hardens to [p^h], whereas /h/ hardens to [k^h].

of post-nasal consonant changes, see Hyman (2001), Zsiga, Gouskova & Tlale (2006), Coetzee, Lin & Pretorius (2007), Coetzee & Pretorius (2010), Solé, Hyman & Monaka (2010), Gouskova et al. (2011), Boyer & Zsiga (2013).

Vowels

There are seven contrastive vowels, /i ɛ̞ ɛ a ɔ̞ ɔ u/. Previous descriptions of Setswana characterize the high-mid vowels as high or semi-open, e.g. as /i/ and /u/ (Cole 1955, Chebanne et al. 1997; see also Dichabe 1997, le Roux & le Roux 2008, le Roux 2012). In our data, the majority of tokens seem to be lower than this notation suggests, but still higher than [e] and [o], and we therefore transcribe them as raised close-mid [ɛ̞] and [ɔ̞] (a finding similar to that of le Roux & le Roux 2008). In the linguistics literature and some older materials, the orthographic convention is to use *e* and *o* for the close-mid vowels [ɛ̞] and [ɔ̞], while *ê* and *ô* are used to represent open-mid [ɛ] and [ɔ]. In the modern standard orthography, however, the diacritic is normally left out, and both sets of vowels are represented as *e* and *o* (see Chebanne et al. 2003 for more details of the orthographic conventions).



Vowel	PHONETIC REPRESENTATION	ORTHOGRAPHIC REPRESENTATION	GLOSS
/i/	ts ^h ípì	tshipi	'iron'
/ɛ̞/	séχà	sega	'cut'
/ɛ/	ts ^h éχà	tsêga	'laugh'
/a/	ts ^h ábà	tshaba	'fear'
/ɔ̞/	ʃókà	šôka	'twist'
/ɔ/	tʃòmà	tšoma	'hunt'
/u/	k ^h úmò	tšuba	'set alight'

A formant-scaled plot of the vowels is given below. Measurements were taken at the midpoint of each vowel, and averaged across three tokens of each of the example words listed above.

Both pairs of mid vowels (open-mid /ɛ ɔ/ and close-mid /ɛ̞ ɔ̞/) have raised allophones [ɛ̞ ɔ̞] and [i̞ u̞], respectively. Previous work reports that these raised allophones are conditioned by a following higher vowel, or by the consonants /s ʃ ts^h tʃ tʃ^h tʃ^h tʃ^h ɲ/, and by the locative suffix [-ŋ] (Cole 1955). For our primary consultant, this raising was subject to variation, happening clearly in some lexical items, but not others, and to varying degrees. Some previous sources report the raised allophones of the open-mid vowels in contexts where they are not derived, thus leading to an analysis with nine phonemic vowels (Chebanne et al. 1997, Dichabe 1997; see also Khabanyane 1991 on the same issue in the closely related vowel system of Sesotho).

It is not clear whether the raised allophones of the close-mid vowels /ɛ̞/ and /ɔ̞/ are consistently distinct from the underlying high vowels /i/ and /u/ for our speakers. Based on F1

⁹ The prefix /ɔ̞/ indicates agreement with a second person singular subject; /ó/ indicates agreement with a third person subject of noun class 1. Class 1 nouns are canonically human, so we gloss this as (s)he. Note that some previous literature on Setswana spoken in Botswana reports that the second person prefix may be high toned in certain morphosyntactic contexts, such as in negative or subjunctive forms (Chebanne et al. 1997: 31ff.).

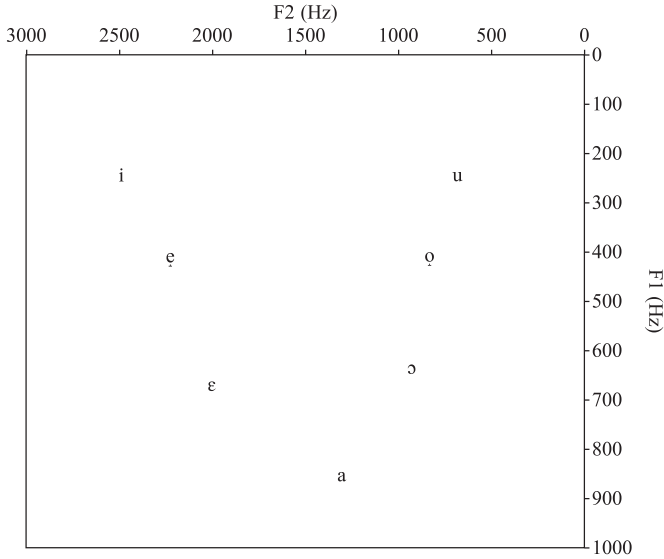


Figure 5 Formant-scaled vowel plot.

measurements of an assortment of tokens, as well as impressionistic transcriptions, it seems that raising of the close-mid vowels /e/ and /o/ can result in vowels that sound very close to [i] and [u] respectively. Raising of open-mid /ε/ and /ɔ/ to [ɛ] and [ɔ̄] respectively very often produces vowels that are slightly lower than [e] and [o], but the F1 and F2 ranges of both vowels overlap to a very high degree. For more detailed consideration of the vowel raising patterns, see Dichabe (1997), le Roux & le Roux (2008).

While the raised vowels [ī ū] (from /e o/) are phonetically high, they do not behave phonologically like underlying high vowels with respect to the distribution of [l] and [d]. They occur with [l], systematically; [d] is found only with underlying /i u/.

Vowels in word-final position are often devoiced, e.g. /χḁ-hétsà/ → [χḁhét͡sḁ̀] ‘to finish’, particularly for words at the end of a phrase. They may also be centralized or otherwise reduced.

Suprasegmental features

Tone

Setswana has two phonemic tones, high and low. The tonal phonology of Setswana is quite complex, and tonal processes can cause high and low tones to manifest on the same syllable, resulting in a surface rising or falling tone (see Jones 1928 and Chebanne et al. 1997 for much more detailed discussion of tone patterns). Some examples of lexical tone distinctions are given in (3).

(3) *Tonal contrast examples*

a.	χḁhùlà	‘to shoot’	χḁhùlà	‘to graze’
b.	χḁbùà	‘to speak’	χḁbùà	‘to skin’
c.	dʒá	‘eat’	jà	‘go’
d.	bàlá	‘count’	bónà	‘see’
e.	tàú	‘lion’	tʰóù	‘elephant’
f.	m̄pá	‘belly’	ḁtá	‘louse’
g.	ḁbíná	‘(s)he dances’	ḁbíná	‘you dance’ ⁹

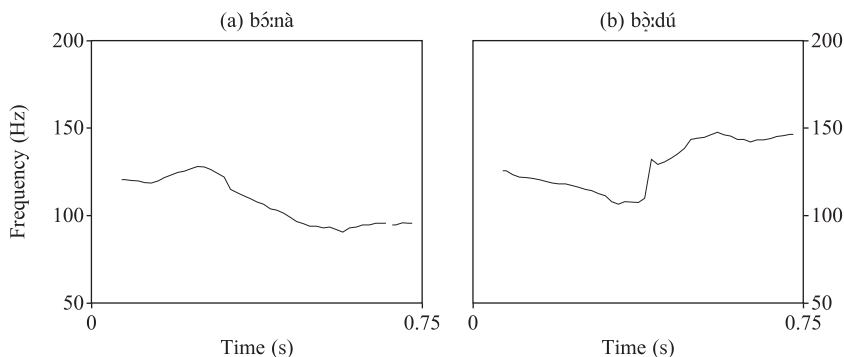


Figure 6 Example of tonal contrast with little absolute pitch difference. Panel (a) shows H-L sequence in [bó:nà] 'see', and panel (b) shows L-H sequence in [bò:dú] 'rotten'.

In utterance-initial contexts, high and low tones may start from approximately the same pitch. They are still distinguished by contours, however: low tones involve a fall, while high tones maintain the same pitch, or involve a rise. This is illustrated in Figure 6, with pitch tracks of [bò:dú] 'rotten' and [bó:nà] 'see', both said in isolation. These words have different tones on each syllable. The starting pitch is approximately the same in both words, however: the tonal difference on the first syllable is realized only towards the end of the syllable. (In the second syllable of each word, the tonal distinction does come with a clear difference in absolute pitch over the full duration of the syllable, though.)

Penultimate lengthening

The penultimate syllable of a prosodic phrase is normally lengthened (Cole 1955: 55, University of Botswana 2001: 31, IPA 2010: 349). These lengthened syllables are also sometimes described as stressed (University of Botswana 2001: 32). Suffixation can produce alternations, affecting which syllable is lengthened, as in [ɲà:ɲà] 'argue' vs. [ɲàɲè:là] 'argue for' (seen in the list below); the first vowel [a] in these two words shows a difference in duration of approximately 70 ms. Penultimate lengthening can apply to syllabic sonorant consonants as well as vowels, as in [ɲ:tʰá] 'point'. The degree of lengthening depends on position in the utterance: words in isolation or at the end of an utterance exhibit more lengthening than phrase-final words in the middle of an utterance. We follow the convention in previous work (Cole 1955: 55) of transcribing the former as full length, and the latter as half length (i.e. [a:] vs. [a]); examples can be seen in the recorded passage at the end of this illustration.

(4) Examples of length differences

- | | | |
|----|-----------|---------------------|
| a. | ɲàɲè:là | 'argue for' |
| | ɲà:ɲà | 'argue' |
| b. | ɲ:tʰá | 'point' |
| | tʰà:pì | 'fish' |
| c. | màbóχó:ɲ | 'arms (LOC) |
| | màbó:χò | 'arms' |
| d. | kókó:ɲ | 'fowl (LOC)' |
| | kò:kò | 'fowl' |
| e. | tʰàtsí:sà | 'fill (CAUS)' |
| | χòtʰá:tsà | 'fill' |
| f. | dìtsí:sà | 'look after (CAUS)' |
| | dí:sà | 'look after' |

Syllable structure

Syllables in Setswana are most commonly V or CV in shape. Syllabic sonorants also occur, but such syllables may not have a distinct consonantal onset (i.e. *ŋ* occurs, but not *Cŋ*). Syllable-internal consonant clusters are limited to a consonant followed by [w], such as [χð:nwá] ‘to drink’, [bɔ̌.dzà.lwá] ‘beer’, [ɲ.ŋwè] ‘one’, and [sè.tʃʰwá.ntʃʰ] ‘picture’. Sonorants in such clusters can be identified as non-syllabic, because they do not exhibit penultimate lengthening.

Vowel–vowel sequences are permitted morpheme-internally, as in [tʰó.ù] ‘elephant’, [tà.ú] ‘lion’, and [pò.ó] ‘bull’. Hiatus between a root and a prefix may optionally show insertion of a glottal stop, though, as noted earlier, and elision of /h/ can also lead to hiatus. In connected speech, these sequences are very often reduced to, and do not show glottal stops. Intervocalic consonants may also be lenited or reduced.

Transcription of a recorded passage

This is a translation of ‘The North Wind and the Sun’, the standard text used for Illustrations of the IPA. The transcription below is based on a recording from our primary consultant; a second recording of the same story from another, female, speaker is also included with the accompanying recordings. The repetitions of [la] and [ʔɛŋɛ] in lines 2 and 3 were identified by the speaker as errors made while telling the story, and not deliberate repetitions. The orthographic version follows the modern convention in leaving out diacritics on mid vowels.

Broad phonetic transcription

p^héó já bókò̀̀ni lè lètsà̀̀tsi bánè̀̀báχà̀̀nè̀̀tsà̀̀nà χò̀̀rè | kémá hò̀̀:nà nà̀̀lè̀̀mà̀̀tʰá· χà̀̀isà̀̀ŋ
 || bábóná mò̀̀è̀̀tì m̀̀mé lètsà̀̀tsí là rá: p^héó lá? là̀̀rè || jò ɔ̌tʰá pól·ám mò̀̀è̀̀tì dzásé |
 ʔɛŋɛ: | ʔɛŋɛ:ʔó: || ké̀̀nè̀̀ ʔónà̀̀lɪŋ má̀̀tʰá:mà ʔáχà̀̀isà̀̀ŋ || lètsà̀̀tsí lɪnɛ̀̀là itʃʰú:bá: || mó:
 m̀̀rə̀̀hó χà̀̀lɪrú || ké̀̀fá· p^héó· já bókò̀̀ni já b̀̀tʃwè̀̀là ká: m̀̀tʰá || mò̀̀è̀̀tì ɔ̌n·é à̀̀ χò̀̀χá
 dzá·sè̀̀ à̀̀ itʃʰè̀̀rélɛ̀̀:tsè̀̀ || p^héó já bókò̀̀ni já bájá itʰ̀̀ð̀̀bò̀̀χà || lètsà̀̀tsí lè né là:tʃʰwá
 làhí·sá || k'ámóχó̀̀tè̀̀ m̀̀mé mò̀̀è̀̀tì à̀̀béá:pó̀̀llà dzásè̀̀ || lètsà̀̀tsí lè p^héó· já bókò̀̀ni
 bánè̀̀bádú̀̀mà̀̀lànà χò̀̀rè lètsà̀̀tsí kél̀̀ǹ̀lè̀̀ lè nà̀̀lè̀̀m m̀̀tʰá || áχà̀̀isà̀̀ŋ

Orthographic version

Pheho ya bokoni le letsatsi ba ne ba ganetsana gore, ke mang o ne a na le matla a a gaisang. Ba bona moeti mme letsatsi la raya pheho la re yo o tla a apolang moeti jase e ne . . . e ne o . . . ke ene o na leng matla a a gaisang. Letsatsi le ne la itšhuba mo morago ga leru. Ke fa pheho ya bokoni ya butšwela ka matla. Moeti o ne a goga jase a itšhereletsa, pheho ya bokoni ya ba ya itlhoboga. Letsatsi le ne la tšhwa la hisa ka mogote mme moeti a be a apolola jase. Letsatsi le pheho ya bokoni ba ne ba dumalana gore letsatsi ke lone le le naleng matla a a gaisang.

Acknowledgements

This work was supported by a grant from the Rhodes University Research Council. Authors’ names are listed in alphabetical order. We owe great thanks to our Setswana consultants for all their help; for other helpful discussion and input, we also thank Iyad Issa, Seunghun Lee, Aaron Braver, and Andries Coetzee, and an anonymous reviewer.

Supplementary material

To view supplementary material for this Illustration, please visit <http://dx.doi.org/10.1017/S0025100316000050>.

References

- Boyer, One & Elizabeth Zsiga. 2013. Phonological devoicing and phonetic voicing in Setswana. *Selected Proceedings of the Annual Conference on African Linguistics (ACAL)* 43, 82–89.
- Chebanne, A[ndy] M., D. Creissels & H. W. Nkhwa. 1997. *Tonal morphology of the Setswana verb*. München: LINCOM Europa.
- Chebanne, Andy [M.], Makali Mokitimi, Lits'episo Matlosa, Rosalia Nakin, Mildred Wakimelo Nkolola, Sekgothe Mokgoatšana & Mats'epo Machobane. 2003. *The standard unified orthography for Sotho/Tswana languages: Botswana, Lesotho, Namibia, South Africa, Zambia and Zimbabwe*. Cape Town: The Centre for Advanced Studies of African Society (CASAS).
- Coetzee, Andries W. 2001. Nasal-only syllables in Tswana. In Graham Horwood, & Se-Kyung Kim (eds.), *RuLing Papers II*, 1–20. New Brunswick, NJ: Rutgers University.
- Coetzee, Andries W., Susan Lin & Rigardt Pretorius. 2007. Post-nasal devoicing in Tswana. In Jürgen Trouvain & William J. Barry (eds.), *Proceedings of the International Congress of Phonetic Sciences 16 (ICPhS XVI)*, 861–864.
- Coetzee, Andries W. & Rigardt Pretorius. 2010. Phonetically grounded phonology and sound change: The case of Tswana labial plosives. *Journal of Phonetics* 38, 404–421.
- Cole, Desmond T. 1955. *An introduction to Tswana grammar*. Cape Town: Longman.
- Dichabe, Seipati Bernice. 1997. Advanced tongue root harmony in Setswana. M. A. thesis, University of Ottawa.
- Doke, Clement M. 1954. *The Southern Bantu languages*. London: Oxford University Press.
- Gouskova, Maria, Elizabeth Zsiga & One Tlale Boyer. 2011. Grounded constraints and the consonants of Setswana. *Lingua* 121(15), 2120–2152.
- Hyman, Larry M. 2001. On the limits of phonetic determinism in phonology: *NC revisited. In Elizabeth Hume & Keith Johnson (eds.), *The role of speech perception in phonology*, 141–185. New York: Academic Press.
- IPA [International Phonetic Association]. 2010. The principles of the International Phonetic Association (1949). *Journal of the International Phonetic Association* 40(3), 299–358. [Reprint of 1949 *Principles of the International Phonetic Association*, ed. by Daniel Jones.]
- Jones, Daniel. 1928. *The tones of Sechuana nouns*. London: International Institute of African Languages and Cultures. [Reprinted 1970, Gregg International Publishers, Farnborough, England.]
- Jones, Daniel & Solomon Tshekisho Plaatje. 1916. *A Sechuana reader*. London: University of London Press. [Reprinted 1970, Gregg International Publishers, Farnborough, England.]
- Khabanyane, Khathatso Evelyn. 1991. The five phonemic heights of Southern Sotho: An acoustic and phonological analysis. *Working Papers of the Cornell Phonetics Laboratory* 5, 1–36.
- Ladefoged, Peter & Ian Maddieson. 1996. *The sounds of the world's languages*. Cambridge, MA: Blackwell.
- Le Roux, Mia. 2012. An acoustic comparison: Setswana vowels versus the cardinal vowels. *South African Journal of African Languages* 32(2), 175–180.
- Le Roux, Mia & Jurie Le Roux. 2008. An acoustic assessment of Setswana vowels. *South African Journal of African Languages* 28, 156–171.
- Ridouane, Rachid. 2014. Tashlhiyt Berber. *Journal of the International Phonetic Association* 44(2), 207–221.
- Solé, Maria-Josep, Larry M. Hyman & Kemmony C. Monaka. 2010. More on post-nasal devoicing: The case of Shekgalagari. *Journal of Phonetics* 38, 604–615.
- University of Botswana, Department of African Languages and Literature. 2001. *The Sound system of Setswana*, 2nd edn. Gaborone: Lightbooks.
- Zsiga, Elizabeth, Maria Gouskova & One Tlale. 2006. On the status of voiced obstruents in Tswana: Against *ND. In Christopher Davis, Amy Rose Deal & Yuri Zabbal (eds.), *Proceedings of the 36th Annual Meeting of the North East Linguistic Society (NELS 36)*, 721–734.