This printout has been approved by me, the author. Any mistakes in this printout will not be fixed by the publisher. Here is my signature and the date_____

Phonological Devoicing and Phonetic Voicing in Setswana

One Boyer and Elizabeth Zsiga
Georgetown University

1. Introduction

There have been a lot of debates on the realization of voiced stops in post nasal position in Setswana and some related languages (Coetzee et al. 2010, Gouskova et al. 2011, Hyman 2001, Solé et al. 2010, Zsiga et al. 2006)**.** These studies have described varying realizations such as a lack of alternation in initial position, lenition in intervocalic position and devoicing in post nasal position (henceforth NC position). This devoicing has been reported to result in neutralization of voiced and voiceless stops, as illustrated in examples 1 and 2.

1. underlying voiced stops

bata	look for →	*mpata*	look for me
disa	guard →	*ntisa*	guard me

2. underlying voiceless stops

pata	accompany →	*mpata*	accompany me
tisa	bring →	*ntisa*	bring me

While the lack of alternation in word initial position and intervocalic lenition are not unexpected, post nasal devoicing is surprising because it goes against a purportedly universal tendency for stops following nasals to become voiced. Pater (1999) presents several languages and the techniques that they employ to avoid a voiced stop after a nasal. With this in mind, we aim to address the following questions: Do voiced and voiceless stops neutralize in post nasal position? If so, is this a) categorical, b) productive and c) really devoicing? In answer to these questions we conclude a) yes, b) no, and c) no.

In this paper, we present the results of three studies, two acoustic, and the other perceptual. We argue that in Sengwato, this alternation does not involve a change in voicing at all. We will show, using the results from these studies that, (a) there is neutralization of voiced and voiceless stops in post nasal position, but (b) there is very little or no devoicing in post nasal position, and (c) that listeners rely on some other cues, other than just voicing to distinguish /b/ from /p/ and /d/ from /t/. As a result, an alternative analysis, i.e., one not based on voicing needs to be sort to explain the alternation in post nasal position in Setswana.

2. Language background

Setswana is a Bantu language spoken by approximately 4.5 million people spread across Botswana, South Africa, Zimbabwe and Namibia. There are approximately 10 dialects used across

* This work is part of a study supported by NSF BCS-1052937

Botswana and South Africa (www.ethnologue.com). For this paper, we focus on one dialect, Sengwato, which is spoken mostly in the east central part of Botswana. All data collection was done in a small village named Shoshong. In Sengwato stops are contrastive at three points of articulation as shown in Table 1 below. Aspiration or labialization, or a combination of both, can be used to further contrast the stops. The stops are asymmetrical for voicing. There is no voiced velar stop. The alveolar stop [d] occurs as a variant of /l/, in the context of high vowels [u] and [i] such as in [dusa] 'pregnant, (an animal)' [disa] 'herd,' and [l] occurs after all the other vowels.

Table 1: Sengwato stops

bilabial	dental/alveolar	velar
p p^h b	t t^h t^w t^{hw} (d)	k k^h k^w k^{hw}

3. Acoustic experiment

3.1 Design

For this study, data was collected from 27 speakers. All participants were adult native speakers of Sengwato dialect of Setswana, who were born in Shoshong and had lived there most of their lives. Each of the speakers produced 50 different of real words, in four contexts shown in 3 through 6 below.

3.	phrase-initial	(in isolation)	*bala*	"read"
4.	intervocalic	('Say_____again')	*o re bala xape*	"say read again"
5.	Intervocalic	('To_____again')	*go bala xape*	"to read again"
6.	Post nasal position	('S/he ___ me again').	*o mpala xape*	"s/he reads me again"

Each word was produced with three repetitions in all four contexts. All the words in the study were verbs, which consisted of minimal or near minimal pairs that would allow us to analyze voicing contrasts for stops (See examples 7 and 8 below). All voiced and voiceless stops were included, as well as other consonants distractors. To elicit the words from the speakers, as well as to avoid reading pronunciation, picture art with the action reflecting the verb, was used. In some situations where it was not possible to use picture art, definitions were provided. The speakers were each given a brief trial session to familiarize them with the sentences and the activity required of them.

7.	bilabials:	*baka* "learn lesson from"	~	*paka* "carry food for trip"
		bala "read"	~	*pala*, "refuse"
		bata "look for"	~	*pata* "accompany"
8.	dental alveolars:			
		dika "gang up on"	~	*tika* "throw at"
		dila "decorate floor"	~	*tila* "avoid"
		disa "herd"	~	*tisa* "bring"
		duba "knead"	~	*tuba* "wish harm on"

The data was analyzed using Praat (Boersma & Weenink 2010). The acoustic measurements of the data collected were based on spectrograms and waveforms. We measured closure, voicing during closure, burst and Voice Onset Time (VOT) durations. Closure duration was measured from the endpoint of the preceding sonorant to the point of initiation of the burst noise. These measurements were taken for all the four different contexts.

3.2. Acoustic results

For this paper we present the results for 16 out of the 27speakers involved in the study.[1] We compare the measurements for voiced versus voiceless stops in word initial, intervocalic, and post nasal positions.

3.2.1. Word initial position

As expected, we noted that in word initial position, underlying voiced stops have some degree of voicing (average 92 ms) while the voiceless stops do not.[2] We found VOT for [p] to have a longer duration (average 10ms), compared to that for [b](average 3ms). We did not find much of a difference for burst durations: 5ms for [b] and [6] ms for [p].

3.2 2. Intervocalic positions

We compared the voiced and voiceless stops in intervocalic position in two context, (Say____ again) and (To_____ again). The results for the two contexts were similar. In both these positions all the stops were produced with more that 50% voicing during closure, with [b] and [d] having slightly higher percentage of voicing (88% and 96%)[3] than their voiceless counterparts. On the other hand [p] had slightly longer VOT duration(8ms) than [b] (3ms), but about the same burst durations(5ms). In terms of closure durations, the voiced stops had shorter durations(average 96ms) than their voiceless counterparts (average 135 ms). Again, the results for these contexts were not unexpected: voiced stops tend to be shorter in intervocalic position; the percentage of closure is longer for voiced stops because their closure duration is comparatively shorter.

3.2.3 Post nasal position

In this position we note that there is neutralization of closure, voicing during closure, burst and VOT durations for the underlying voicing contrasts (as can be seen in Figure 1). Both voiced and voiceless stops are at least 72 to 80% voiced.

Figure 1: Post nasal stops durations (ms)

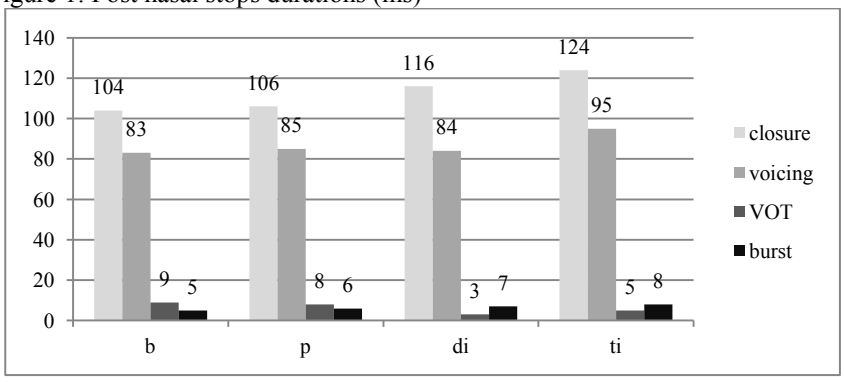


Figure 2 is a spectrogram and waveform of the phrase [ʊmbala] 'he/she reads me', with [b] in post nasal position. Note the fully voiced [b] for the entire duration of the closure, with a voicing break only

[1] The results for the remaining 11 speakers are still being analyzed.

[2] Word initial closure duration could not be accurately measured on Praat since there is no waveform or spectrogram corresponding to closure. Also note that the results for [d] and [t] presented here were in the environment of the vowel [i] as [d] only occurs before [u] and [i] in Setswana while for [b] and [p] they were in the context of [a].

[3] This is expected as in previous studies we have found that in Setswana, when in intervocalic position, [b] tends to lenite. We will be analyzing the data further to find evidence of lenition in this position for voiced stops.

at the point of release of a strong burst. There is a period of voicelessness (VOT) before the onset of voicing for the following vowel.

Figure 2: Spectrogram & waveform for N_bala

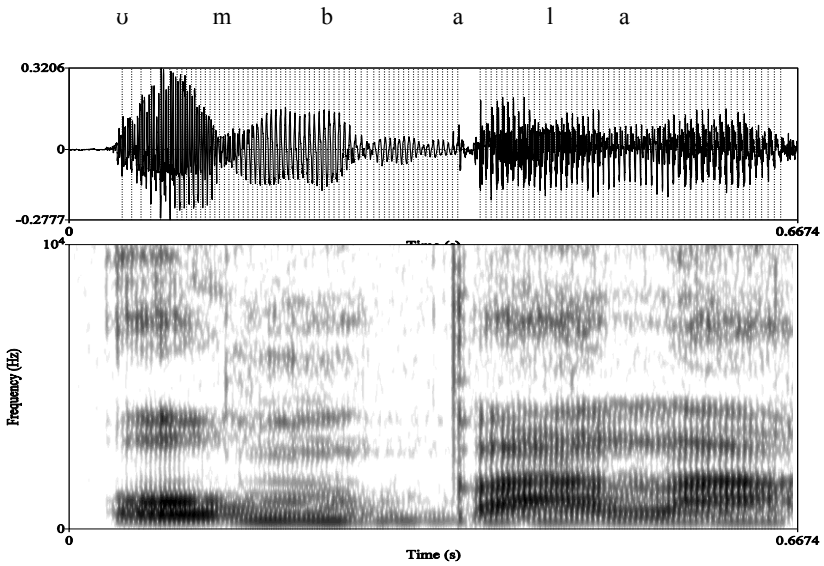
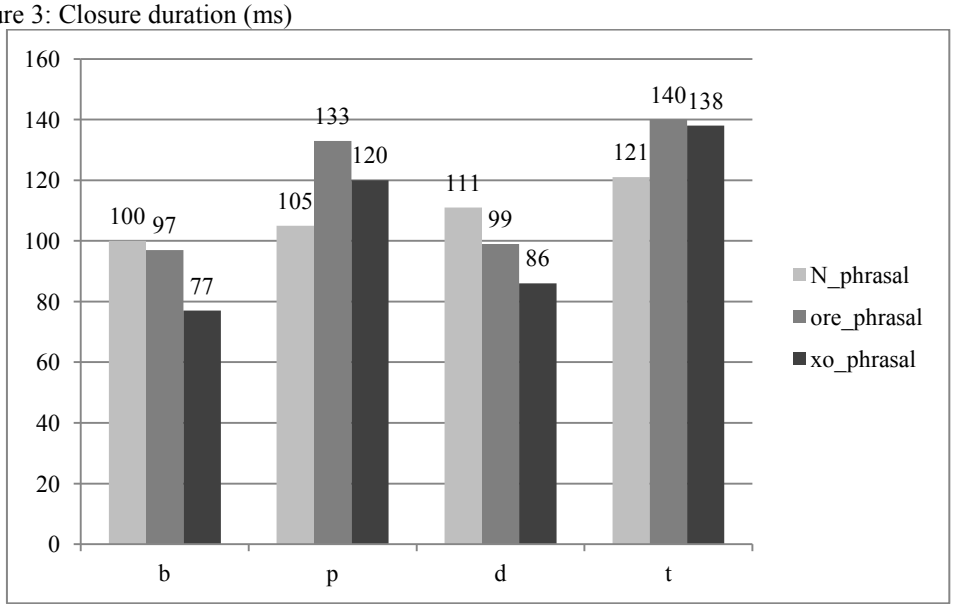


Figure 3 shows a comparison of the closure duration of the stops in intervocalic and in post nasal position. We observe that in the former position, the voiceless stops occur with a longer closure duration, which is expected. On the other hand, the closure duration for voiced stops remains almost the same for both positions. If indeed the stops were devoicing, we would expect to see an increase in closure duration.

Figure 3: Closure duration (ms)



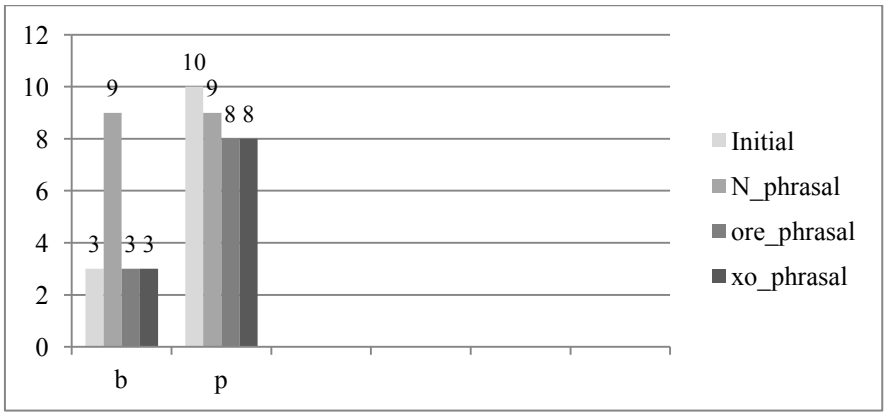
It is also important to consider the percentage of voicing during closure. Table 2 below, shows the percentage of voicing during closure for the bilabial and dental alveolar stops.

Table 2: Voicing during closure %

stop	N_phrasal	ore_phrasal	xo_phrasal
b	80	90	93
p	80	70	65
d	72	90	99
t	76	74	80

The results from Table 2, show that for the voiced stops there appears to be a slight decrease in voicing for the stops in NC position compared to those in intervocalic position. Again, we suspect that the strong voicing in intervocalic position may be due to lenition. Nonetheless, voicing in post nasal position still remains well above 50% for both [b] and [d]. Our data indicate that there is no devoicing in contrast with the results from Coetzee et. al (2010).

Figure 4: VOT Duration (ms)



The measurements for VOT duration shown in Figure 4, indicate that in post nasal position, [b] has a longer VOT duration when compared to word initial and intervocalic. This is a strong indicator that the segment strengthens in this environment. A phonological change is taking place, but it is not devoicing. (The results for [d] and [t] are not included here because of secondary frication that resulted from the contexts that they occurred in, which made it difficult to accurately measure VOT durations[4].)

3.3. Individual speaker results

The results presented so far, are averages across the speakers, any individual differences not reflected. In order to determine if at an individual level, speakers were devoicing we looked at each speaker's production separately. Our definition of what constitutes a voiceless stop is similar to that used by Coetzee et al (2010); that a segment is considered devoiced if it has less than 50% voicing during closure. With this definition in mind, we found that for instances of [b] 4 of the 16 speakers had a devoicing tendency, (i.e. they devoiced some of the tokens in post-nasal position) and 1 speaker devoiced consistently in post-nasal position. For [d] we found that 6 speakers devoiced at least 1 token, but none devoiced consistently[5].

[4] For the productions of most of speakers, there was a period of frication that followed the burst due to the high vowels that [d] occurs after. As a result, we were not able to accurately measure VOT durations.

[5] Coetzee et al (2010) attribute this to a language change, due to the intra-speaker and inter-speaker variation they got in their sample. It would appear then that the language shift is more advanced for the dialect that we are investigating.

These findings are consistent with those of Coetzee et al (2010) who also found that for some speakers /b/ and /p/ were realized as [b] in NC position. The difference between our results and theirs is that they had 25% of the speakers do this, whereas our data indicates voicing to be the norm rather than the exception (with only 1 consistent devoicer, and 4 who devoiced in some instances out of a total of 17 speakers).

4. Nonce words experiment

In order to test for productivity in post nasal position, 20 of the speakers were further recorded producing 50 nonce words, replicating the real words experiment described above. This time, they were shown an action pictured on a card, then given a made up verb for the action. The nonce words were all possible words in the language (as in examples 9 and 10 below).

9. *bama
10. *pama

The results for the nonce words were very variable. Of the 20 speakers involved, 7 speakers showed alternation in NC position, 9 showed no change, the remaining 4 showed an alternation in some words but not others.[6] As a result of the variability found in the data, we cannot at this point make a conclusion on whether or not there is productivity in NC position.

5. Perception experiment

5.1. Design

In order to confirm if there is neutralization of voiced and voiceless stops in post nasal position, we conducted a perception experiment to test the listeners' ability to discriminate and identify the underlying voicing contrasts. This experiment was run using Praat Experiment MFC, a self run automated experiment which allows a listener to do discrimination and identification experiments.

Three speakers of the 27 were chosen as typical, and their productions of minimal pairs were extracted and played to 10 Setswana listeners in two tasks. In the word identification task, listeners heard extracted tokens of verbs in isolation, from NC phrases with and without the nasal prefix, from "Say_____" phrases and from "To_____ again" phrases. They were then required to identify the infinitive form for the word they heard, (see 11 below) and make a goodness judgment on a scale of 1 to 5 of how sure they were of their answer.

11. e.g. speakers heard: *N_bata* (look for me) or *N_pata* (accompany me)
speakers identify infinitive : *go bata* (to look for) or *go pata* (to accompany)

For the word discrimination task, listeners heard paired tokens. In half of the pairs, the initial consonants were the same, while in the second half the consonants were different. Listeners then had to indicate whether the words they heard were the same or different, and again, how sure they were of their answer.

---

[6] We expected "post-nasal devoicing" to be productive, but most speakers did not neutralize. This could have been due to interpreting the nonce words as foreign language, or English. In urban areas the following words do not change in postnasal position:

foun-ɛla (call) m-foun-ɛla (call me) (phone)
zip-a (zip up) n-zip-a (zip up for me)
*sa*ɪn-ɛ*la* (sign for) n-tsʰaɪn-ɛla (sign for me)
(personal knowledge/observation)

In conversations with the speakers during and after the recording sessions, they indicated that they thought the words were foreign, and did not have to follow Setswana rules, and also since they didn't know the meanings of the words they didn't know how to use them.

5.2. Results

The results from this study show that in initial and intervocalic position, listeners accurately identified and discriminated the words 85% of the time. The more interesting results for this experiment were those involving NC position. For the discrimination task, pairs of words with underlying voicing contrasts were heard as the same (see examples 12 and 13). i.e. Listeners could not tell the difference between the surface realizations of pairs of words such as those in 12 and 13.

12. *m-bata* versus *m-pata*
13. *n-dila* versus *n-tila*

For the identification task, in post nasal position, underlying /b/ was consistently heard as [p], and underlying /d/ as [t][7]. None of the speakers selected the voiced stops as the segment they heard. This identification of the tokens as [p] or [t] was despite the fact that the actual closure was in fact voiced for more that 50% of the duration, as was shown in Table 2.

6. Discussion and Conclusion

We set out to address three questions: whether there is complete neutralization of voiced and voiceless stops in post nasal position, and if so, if it is productive, and finally whether the change is one of devoicing or something else. In order to answer these questions we presented the results of three studies conducted with native speakers of Sengwato.

The results of the perceptual and acoustic study help us provide answers to these questions. As was shown, the data confirms that there is complete neutralization of voiced and voiceless stops in post nasal position. Native speakers were not able to discriminate between voiced and voiceless stops in post nasal position, and heard them as being the same. On the question of whether the neutralization is productive, we cannot conclude either way because the speakers' productions of nonce words showed mixed results.[8] With respect to the more important question of whether in fact this change is one of devoicing, our answer is, no. We have seen from the acoustic data that we analyzed that there is no devoicing that takes place for most of the speakers. Only one speaker was found to devoice consistently. Overall, the voiced stops had 70% or more voicing during closure. While we noted a slight decrease in voicing, it was not to the extent of devoicing (i.e. the duration of the voicing during closure was not less than 50%). We observed that voiceless stops had more voicing in the NC position. We found that the voiced stops occurred with a longer VOT in post nasal position that in the other positions investigated, and a somewhat stronger burst. It was also interesting to see that the closure duration was not lengthened for the voiced stops, in comparison to their duration in intervocalic position. If indeed the voiced stops were devoicing, we would have expected to see lengthened closure duration. But instead, what we do observe is shorter closure duration for the voiceless stop. A shorter closure duration is characteristic of voiced stops. This provides further evidence that the stops do not devoice, but that instead, there is phonetic post nasal voicing, consistent with universal tendencies.

---

[7] Speakers selected the underlying voiceless stop as the form they heard, even though they indicated that the words were the same. We had predicted that speakers may prefer one meaning or the other, but all listeners consistently chose the voiceless infinitive even though it is clear there is neutralization. This could be due to:

a) Devoicing, in which case we would be wrong ☹
b) orthography, i.e. both N-bala and N_pala are spelt mpala, so given the two options bala ~pala they picked the one that matched the orthography.
c) An oversight on our part: we should have included a 3rd option that would allow the listeners to indicate that could it be either one of the underlying voicing contrasts.

[8] In other studies still to be undertaken we plan to tell the speakers that the words are from an old form of Setswana and provide a definition for the words to see if they would apply Setswana phonological rules in this case.

We therefore conclude that in NC position, stops tend towards fortition rather than devoicing. Post nasal fortition in Setswana has been well attested in traditional literature (Cole 1955). In our study, the tendency towards fortition is reflected by an increase in VOT duration, a small increase in burst and closure duration for the voiced stops.

In a similar study, Coetzee et. al (2010) concluded that the dialect of Setswana that they studied, showed evidence of post nasal devoicing. Even though they studied Setswana as spoken in Potchefstroom, South Africa (not clear which dialect) they note that they do not expect their results to be similar to the dialect we study here because of the geographical distance and political boundaries between Shoshong and Potchefstroom. It will be interesting to see if in studies of other dialects that we intend to carry out, whether speakers will produce results more similar to Coetzee et al. (2010) for the Potchefstroom area, or our results for Sengwato. Even though Coetzee et al. (2010) noted post nasal devoicing, they observed post nasal devoicing is unstable and possibly undergoing a change. Our data seems to suggest that in Sengwato the change is more advanced than the dialect they studied.[9]

References

Boersma, Paul & Weenink, David (2012). Praat: doing phonetics by computer [Computer program]. Version 5.3.17, retrieved 12 June 2012 from http://www.praat.org/

Coetzee, Andries W. & Pretorius, Rigardt. (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*. Capetown: Longman.

Gouskova, Maria. Zsiga, Elizabeth & Boyer, One Tlale. (2011) Grounded Constraints and the Consonants of Setswana. *Lingua, 121, 2120-2152*.

Hyman, Larry. M. 2001. On the limits of phonetic determinism in phonology: *NC revisited. In Elizabeth Hume and Keith Johnson, eds., *The Role of Speech Perception in Phonology*. New York: Academic Press, 141-185.

Pater, Joe. (1999) Austronesian nasal substitution and other NC effects. In R Kager, H. van der Hulst & W. Zonneveld (Eds.). *The Prosody-Morphology interface (pp310-343). Cambridge University Press*.

Sole, Maria-Josep, Hyman, Larry M. & Monaka, Kemmonye C. (2010) More on post-nasal devoicing: The case of Shekgalagari. *Journal of Phonetics, 38, 604-615*.

Zsiga, Elizabeth, Gouskova, Maria, & Tlale, One. (2006) On the Status of voiced stops in Tswana: Against *ND. In C. Davis, A. R. Deal & Y Zabbal (Eds), NELS 36 *Proceedings of the 36 annual meeting of the North East Linguistics Society* (pp. 721-734) Amherst: GLSA.

www.ethnologue.com

[9] Voiceless stops appear to be voiced post nasally, having more voicing that in other positions. Statistical analysis still needs to be done to determine where there are significant differences.