Phonology and phonetics of tone in Northern Sotho, a Southern Bantu language

Sabine Zerbian

Proceedings of Conference on

Language Documentation & Linguistic Theory 2

Edited by Peter K. Austin, Oliver Bond, Monik Charette, David Nathan & Peter Sells

13-14 November 2009 School of Oriental and African Studies, University of London

Hans Rausing Endangered Languages Project
Department of Linguistics
School of Oriental and African Studies
Thornhaugh Street, Russell Square
London WC1H 0XG
United Kingdom

Department of Linguistics:
Tel: +44-20-7898-4640
Fax: +44-20-7898-4679
linguistics@soas.ac.uk
http://www.soas.ac.uk/academics/departments/linguistics

Hans Rausing Endangered Languages Project:
Tel: +44-20-7898-4578
Fax. +44-20-7898-4349
elap@soas.ac.uk
http://www.hrelp.org

ISBN: 978-0-7286-0392-9

This publication can be cited as:

Sabine Zerbian. 2009. Phonology and phonetics of tone in Northern Sotho, a Southern Bantu language. In Peter K. Austin, Oliver Bond, Monik Charette, David Nathan & Peter Sells (eds) *Proceedings of Conference on Language Documentation and Linguistic Theory 2*. London: SOAS.

or:

Sabine Zerbian. 2009. Phonology and phonetics of tone in Northern Sotho, a Southern Bantu language. In Peter K. Austin, Oliver Bond, Monik Charette, David Nathan & Peter Sells (eds) *Proceedings of Conference on Language Documentation and Linguistic Theory 2*. London: SOAS. www.hrelp.org/eprints/ldlt2_32.pdf

Phonology and phonetics of tone in Northern Sotho, a Southern Bantu language

SABINE ZERBIAN
University of the Witwatersrand, Johannesburg

1. INTRODUCTION

Northern Sotho belongs to the Sotho-Tswana group of Southern Bantu languages (S30 in Guthrie's classification, 1948) alongside Tswana, Southern Sotho and Lozi, and is spoken in the Limpopo Province of South Africa. Like most other Bantu languages Northern Sotho is assumed to have a two-tone systems in which the high tone (H) is the active tone that is specified in the lexicon, and the low tone (L) is a default tone that is inserted late in the derivation of the surface tone pattern. Tonal distinctions carry lexical and/or grammatical meaning. Tone sandhi occur depending on prosodic environment and syntactic constituency.

Tone in Sotho-Tswana is relatively well-described. A range of dictionaries list the tone patterns of words in isolation (cf. Ziervogel & Mokgokong 1975, Kriel & van Wyk 1966 for Northern Sotho; Snyman et al. 1990, Tsaga & Kgonope 1995, and Cole in press for Tswana). Various theses and books describe the tonal grammar of these languages (Mmusi 1992, Cole & Mokaila 1962, Cole in press, Chebanne *et al.* 1995 for Tswana; Khoali 1991 for Southern Sotho; Lombard 1976, Monareng 1992 for Northern Sotho varieties). However, a close comparison of these sources reveals considerable variation in the description of the tone patterns (cf. Schadeberg 1981, Roux 1995, Zerbian & Barnard in press).

The question that then arises is which kind of difference the reported variations reflect: Do they reflect 'real' differences, namely phonological processes and hence separate tonal grammars, or do they reflect 'unreal' differences? 'Unreal' differences easily arise across transcribers, especially if they differ in their background with respect to experience in tone, or in the elicitation method (e.g. introspection versus perception). 'Unreal' differences also arise due to the notational system used, especially if there is no explicit explanation of the notational labels used. 'Unreal' differences furthermore also arise due to theoretical approaches. Examples are differences in the number of tones assumed (cf. Doke 1929 on Zulu), or more recently concerning the role of phonetics in tone. 'Unreal' differences can also arise due to segmental interference.

However, only the 'real' differences are of linguistic interest. To these belong phonological processes, stable idiolectal differences within the same speaker or across speakers, tonal change across generations, or geographically different tonal varieties to name a few. They are interesting for phonology, sociophonetics, or dialectology. In better-studied languages such as English, Spanish etc, such differences are often collected by means of experimental studies. These are normally highly controlled for potential interferences from phonetic factors.

The 'unreal' differences are thus reduced as far as possible through:

- acoustic data which are independent of one individual's perception;
- perceptual verification;
- a transparent notational system.

The high standards of experimental work should also be applied to work on lesser-studied languages (cf. Xu 2006) given that technology is readily available. Of course, restrictions of fieldwork always apply (cf. Hildebrandt 2007). If standards are not adhered to there is the danger of artificially increasing the number of unique tonal systems; an approach which would overgeneralize and disguise similarities between languages and varieties.

The current paper thus contributes acoustic data from the study of tone in Sotho-Tswana. This case study illustrates a 'real' difference and an 'unreal' difference in our data, relating them to the role of phonology and phonetics in tone studies. The paper is structured as follows: Section 2 presents the data and summarizes the observations on the realization of a single high tone in Northern Sotho. Section 3 discusses the role of phonology and phonetics with special reference to tone. Section 4 concludes with an outlook.

2. TONE IN THE VERB WORD IN NORTHERN SOTHO

2.1. The grammatical contexts

Bantu languages are agglutinative languages. They are also tone languages, and so lexical tones are assigned to every morpheme. Verb stems only contrast for tone on the stem-initial syllable. Verbal morphemes, such as the obligatory subject concord, the optional tense/mood-morphemes, the object concord or suffixal extensions, are also specified for tone. The verb stem with all its prefixes and suffixes forms one Prosodic Word in Northern Sotho within which tone sandhi apply that can change an underlying tone during the derivation of the surface tone pattern. The orthography of the language is disjunctive and thus gives a wrong impression about prosodic constituency. In the present paper we present the examples according to their prosodic constituency and not according to the Standard Orthography. For the purpose of the present paper we focus on the surface realization of a single high tone in the verb word.

The two verb structures under consideration are given in (1) and (2).[1]

(1) *Re-* *mêm-a* *mokgalabje.*
SC1PL- invite-FV old.man
'We are inviting the old man.'

[1] The abbreviations used in this paper are 1 = first-person, SC = subject concord, OC = object concord, PL = plural, CL = class (+number), FV = final vowel

(2) *Re-* *le-* *lôy-a* *lefsifsing.*
SC1PL- OC.CL5- bewitch-FV in.darkness
'We are bewitching him in the dark.'

The sentence in (1) shows a high-toned verb stem (lexical high tones are underlined) with the final vowel *-a* indicating Present Principal Tense, preceded by a low-toned subject concord. The sentence in (2) shows a low-toned verb stem, preceded by a high-toned object concord and a low-toned subject concord.

2.2. The production study

In a production study we recorded the pronunciation of these sentences by four speakers of Northern Sotho (for details cf. Zerbian & Barnard in press). Each construction in (1) and (2) occurred with three different verb stems (controlled for segmental make-up) and three repetitions. Thus, per speaker we had nine renderings of the above constructions. We tested verb stems of three different syllable length (disyllabic, trisyllabic and quatrosyllabic), which are presented individually below.

The digitized data were analyzed on a PC workstation, using the speech analysis software PRAAT (Boersma & Weenink 2005). The data were first inspected visually (see Xu 2001 for a corresponding procedure). For visual inspection, a PRAAT script (Xu 1999) was run over the data that rendered time-normalized F0 curves for the relevant part of the target sentences. A four-syllable-window (syllable 1=0-20, syllable 2=20-40, syllable 3=40-60, syllable 4=60-80 in the figures below) was chosen for which F0 was calculated, starting with the tone bearing syllable (0-20). The segmentation of the syllables into segments was performed manually by inspecting the speech waveform, while drawing on spectral and auditory information (Ladefoged 1993). Time-normalization for F0 was computed by taking a fixed number of points at equal time proportions. For a clearer graphical picture of F0-syllable alignment, the F0 curves of the four syllables were averaged across the nine repetitions for each type.

2.3. Visual inspection of the results

There are definitely intra-speaker differences but the overall trends depicted in the displays can be described as follows: for context 1 (as depicted in Figure 1), the pitch starts rising on the tone-bearing syllable from a relatively low level (which continues from the preceding low-toned subject marker). The location of the high peak is subject to variation across speakers, occurring somewhere between the second and third syllable, counting from the high-tone-bearing, verbstem initial syllable. The pitch stays high for a while, resulting in a plateau effect.

In context 2 (as depicted in Figure 2), we find the pitch starting out from a relatively low level, just as in context 1. The rise to the maximum peak is considerably delayed across all speakers and only occurs in the third syllable.

Figure 1
Realization of a high tone from verb stems of different lengths

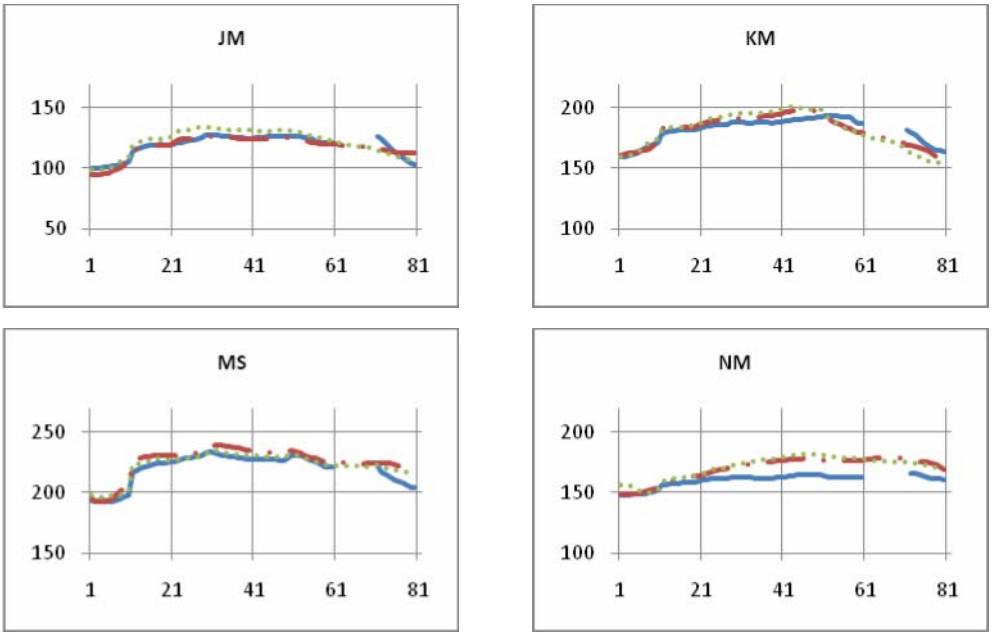
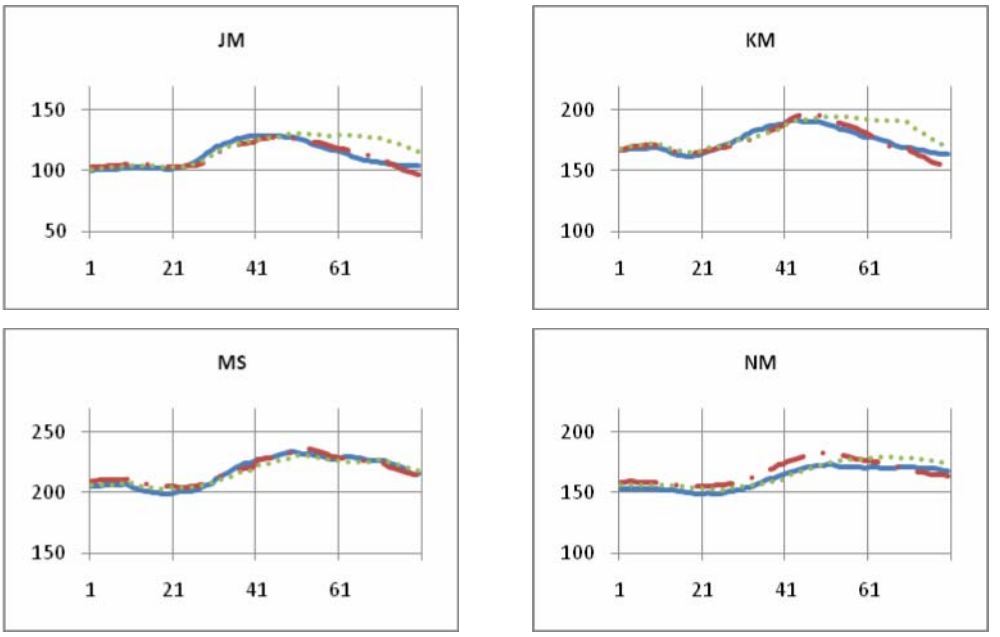


Figure 2
Realization of a high tone from an object concord



A three-factor ANOVA confirmed a fundamental difference between the two contexts which is highly significant for all speakers in what concerns the peak alignment in the first two syllables (cf. Zerbian & Barnard in press).

The results illustrate that underlying tone and surface tone do not always coincide. We observe two tone sandhi: first, high pitch (corresponding to a high tone) is not only realized on the syllable the high tone is associated with underlyingly but also on following syllables, as shown in Figure 1. This phenomenon is referred to in the phonological literature as High Tone Spread (HTS). In the phonetics literature, a parallel behavior is referred to as peak delay (cf. Myers 1999). Second, as shown in Figure 2, an underlying high tone might not correspond to high pitch on the syllable it is associated with but only on a later syllable. We will refer to this behavior as peak shift. This is made explicit in the transcription in (3).

(3) (a) */re-mêma mokgalabje/* → *[re-mémá mókgalabje]*
(b) */re-le-lôya lefsifsing/* → *[re-le-lóyá lefsifsing]*

2.4. Peak delay and peak shift in the literature on Sotho-Tswana

High Tone Spread has been reported for all Sotho-Tswana varieties. However, the right boundary of the spread (i.e. spread up to one or two syllables following the tone-bearing syllable or even unbound) has been reported to vary across different varieties and even within a variety (cf. Zerbian & Barnard in press for an overview). In our study, the right boundary also varies across our speakers, but no statistically significant results could be obtained that are constant across all speakers. Instead, we found that the location of the pitch peak depends on various interacting factors, such as the presence of a syllable onset, the voicing of an onset consonant, and the number of syllables in the verb stem, as additional syllables lead to a less steep slope in declination.

The peak shift (or tone shift) that we can observe with the object concord has not been reported for Northern Sotho previously. However, it is reported for the Southern neighbour Southern Sotho and the Eastern neighbour Tsonga (Khoali 1991, Kisseberth 1994). The difference in peak alignment that is depicted in Figure 2 is statistically significant and constant across the four speakers.

The question poses itself how to evaluate these generalizations as to their status in the tonology of this language. Peak shift seems to involve a categorical shift of the tone (i.e. pitch peak) from the underlying syllable to the following syllable. It occurs clearly visible and constantly across all four speakers. The shift cannot be related to other factors, such as syllable make-up or length of the verb stem. This is good reason to assume that this shift is a 'real' difference (as stated in section 1) and is phonological in nature, i.e. that there is a morpho-phonological rule in this language which targets a high tone originating from an object concord. Tone shift with the object concord thus seems to emerge as an interesting parameter of variation in the Southern Bantu languages.

The status of High Tone Spread (or peak delay) as a 'real' difference is much more debatable, especially with respect to its right boundary. We see inter- and

intraspeaker variation that does not amount to any consistent statistical significance. The peak delay can be characterized by gradiency. Moreover, the peak contour depends on other factors such as syllable structure and length of the utterance. If we want to capture this difference in the speech signal in a phonological rule (as has been done in the rule of High Tone Spread in the phonological literature) do we consider it to be of the same 'quality' as the high tone shift mentioned above? The following section discusses this question as an interaction of phonology and phonetics.

3. PHONOLOGY-PHONETICS INTERFACE

The tone contours shown above differ in the gradiency of their pattern: Whereas the tone shift from the object concord is a categorical shift in the peak from one syllable to the next, the location of the pitch peak and extension of the pitch plateau in High Tone Spread is gradient, i.e. subject to variation across speakers, within a speaker, and also dependent on the number of syllables of the verb stem and the voicing of the word-initial consonant. Gradiency is the critical property that distinguishes phonetic from phonological patterns. Phonetic patterns are gradient, whereas phonological patterns are categorical.

Furthermore, phonological but not phonetic patterns are sensitive to morphological categories (Myers 2000: 262). Phonetic patterns, on the other hand, are sensitive to the surrounding phonetic context. And indeed, we find that High Tone Spread is reduced if two high tones are adjacent and are realized with a low tone target intervening (Zerbian & Barnard, under review).

The current paper follows the view that both phonology and phonetics are part of the universal grammar and that a linguistic theory cannot do away with one or the other. Why is it important to differentiate between phonological and phonetic patterns in tone? It has repercussions for our assumed architecture of language and for phonological theory, i.e. what is part of the psychological, representational component 'phonology' and what is part of the implementation module 'phonetics'? A theoretical account of tone should avoid redundancy between the two components and should strive for the most economical representation, along the lines of Occam's Razor.

Myers (2000: 259ff) states that the only way to differentiate between a phonological and phonetic pattern is to look for the best-fit in a complete model, i.e. comparing a phonetics-only with a phonology + phonetics model. Such a complete model is outside the scope of the present paper. However, differentiating into phonetic and phonological patterns in tone helps explaining inter- and intraspeaker variability that we find in the one pattern as compared to the homogenous data in the other pattern. Instead of interpreting gradient patterns as different idiolects or even dialects which imply entirely separate tonal grammars, it is more economical to simply account for intraspeaker variability in the component of phonetic implementation.

The examples in (4) summarize the difference by giving the underlying form of the sentences in (1) and (2) above, as taken from the lexicon, then giving the phonological form after phonological rules have applied, and eventually giving the actual output once rules of phonetic implementation have applied.

(4) (*a*) */re-mêma mokgalabje/* → */re-mḗma .../* → *[re-mḗmá ...]*
(b) */re-lẹ-lôya lefsifsing/* → */re-lẹ-lṓya .../* → *[re-le-lṓyá ...]*

4. OUTLOOK

As Myers (2000) observes with respect to methodology in tone studies:

> Such a comparison [i.e. comparing a phonetics-only with a phonology + phonetics model, SZ] would not have been possible on the basis of tonal transcription data alone, since such data would not have allowed us to see a gradient pattern if it had been there. (Myers 2000: 252)

Thus, controlled stimuli, acoustic data, and data from several speakers are needed to come to terms with phonetic and phonological patterns in a language. For the adequate description of the phonology of a language it is important to write phonological rules only for those processes that are indeed not derivable by phonetic implementation in order not to overload the phonological component with redundant information. It needs to be stressed that the status of High Tone Spread in a language's grammar might well be language-specific (cf. Bickmore 2009).

Other cases in Bantu tone that would warrant a closer look as to their phonological versus phonetic status, both within a single language and across languages are

- downstep versus declination and reset
- tone deletion or fusion in sequences of two high tones
- depressor consonants and tone.

These remain topics for future investigation.

REFERENCES

Bickmore, Lee. 2009. Is All Tone Doubling Simply Peak Delay? Paper presented at ACAL 40, University of Illinois.

Boersma, Paul & David Weenink. 2005. Praat: doing phonetics by computer. Computer software. www.fon.hum.uva.nl/praat/.

Chebanne Andy M., Creissels, Denis & H.W. Nkhwa. 1997. *Tonal Morphology of the Setswana Verb*. Munich: LINCOM Europe.

Cole, Desmond T. & Dingaan M. Mokaila 1962. *A Course in Tswana*. Washington: Georgetown University.
Cole, Desmond T. in press. Sechuana & English Illustrated Dictionary.
Doke, Clement M. 1929. Text-book of Zulu Grammar. London: Longmans, Green & Co.
Guthrie, Malcolm. 1948. *The classification of the Bantu languages*. London: International African Institute. Reprinted 1967.
Hildebrandt, Kristine A. 2007. Phonology and Fieldwork in Nepal: Problems and Potentials. In Peter Austin, Oliver Bond & David Nathan (eds.) *Proceedings of the Conference on Language Documentation and Linguistic Theory*, 33-44. London: SOAS.
Kgasa, Morulaganyi L.A. & Joseph Tsonope (1995). *Thanodi ya Setswana*. Gaborone: Longman Botswana.
Khoali, Benjamin T. 1991. A Sesotho Tonal Grammar. Unpublished PhD thesis, University of Illinois at Urbana-Champaign.
Kisseberth, Charles W. 1994. On Domains. In Jennifer Cole & Charles W. Kisseberth (eds.), *Perspectives in Phonology*, 133-166. Stanford: CSLI.
Kriel, Theunis J. & Egidius B. van Wyk. 1989. *Pukuntšu – Woordeboek Noord-Sotho – Afrikaans, Afrikaans – Noord-Sotho*. 4th edition. Van Schaik: Pretoria.
Ladefoged, Peter. 1993. *A Course in Phonetics*. Boston: Thomson Wadsworth.
Lombard, Daan P. 1976. Aspekte van Toon in Noord-Sotho. Unpublished PhD thesis, University of South Africa.
Mmusi, Sheila O. 1992. Obligatory Contour Principle effects and violations: The Case of Setswana Verbal Tone. Unpublished PhD thesis, University of Illinois at Urbana-Champaign.
Monareng, William M. 1992. A Domain-based Approach to Northern Sotho Tonology: A Setswapo Dialect. Unpublished PhD thesis, University of Illinois at Urbana-Champaign.
Myers, Scott. 1999. Tone Association and F0 Timing in Chichewa. *Studies in African Linguistics* 28(2): 215-239.
Myers, Scott. 2000 Boundary disputes: The Distinction between Phonetic and Phonological Sound Patterns. In N. Burton-Roberts, P. Carr and G. Docherty (eds.) *Phonological knowledge – Conceptual and Empirical Knowledge,* 245-272. Oxford: Oxford University Press.
Roux, Justus C. 1995. Prosodic data and phonological analyses in Zulu and Xhosa. *South African Journal of African Languages* 15(1): 19-28.
Schadeberg, Thilo C. 1981. Tone in South African Bantu Dictionaries. *Journal of African Languages and Linguistics* 3: 175-180.
Snyman, J.W., Shole, J.S. & Le Roux, J.C. 1990. *Dikišinare ya Setswana – English – Afrikaans*. Pretoria: Via Afrika.
Xu, Yi. 2006. Principles of Tone Research, In *Proceedings of International Symposium on Tonal Aspects of Language*, 3-13. LaRochelle, France.

Zerbian, Sabine & Etienne Barnard. in press. The realization of a single high tone in Northern Sotho. To appear in *Southern African Linguistics and Applied Language Studies*.

Ziervogel, Dirk & Mokgokong, Pothinus C. (1975). *Groot Noord-Sotho Woordeboek/ Comprehensive Northern Sotho Dictionary/ Pukuntsu Ye Kgolo Ya Sesotho Sa Leboa*. Pretoria: Van Schaik.