

A contrastive analysis of the sound structure of Sotho-Tswana for second-language acquisition¹

A B S T R A C T The paper addresses second language teaching of phonetic, phonological and prosodic features in the Sotho-Tswana languages (Southern Bantu) from a linguistic perspective. It motivates the inclusion of phonetic, phonological and prosodic background knowledge in second language teaching, and singles out potential areas of learners' difficulties based on a comparative analysis, assuming a Germanic language as the learner's first language. It adopts the idea that learner problems can be overcome by a mixture of linguistic background explanation and practical, contrast-directed exercises.

1. Introduction

During the middle of the last century, mastery of the sound system of a second language (L2²) was considered first priority in L2 teaching and learning. Two factors induced a shift in attention in L2 teaching and learning: With a change in focus in linguistics towards syntax during the second half of the century, acquisition of phonology received less attention. Also, a general acceptance that learners past the age of puberty are unlikely to achieve native-like pronunciation (Critical Period Hypothesis) independent of the effort put into it, led to a decrease in scholarly attention (Saville-Troike, 2006:142).

¹ The article was written while being in the Department of Linguistics at the University of the Witwatersrand, Johannesburg. My own experiences of studying Northern Sotho and Tswana with different teachers in different settings went into this paper. I am grateful to A. Joubert, W. Pretorius and Rev. D. Mascher for their contribution in forming the views expressed in this paper.

² This article is primarily concerned with the acquisition of a Sotho-Tswana language against a fully acquired first language which is of Germanic origin. It thus comprises both adults and children as learners. The term second language is used throughout the paper in this sense. Other terms might be common in this context as well, such as third language or foreign language.

Pronunciation nevertheless remains to be an important aspect in L2. This is especially true for L2 languages which are mostly learned for interpersonal competence (as opposed to academic competence). Sotho-Tswana languages, though being official languages of South Africa and the national language of Botswana, are mainly used for interpersonal communication. In interpersonal competence, the oral mode is dominant and with it listening and speaking skills (Saville-Troike, 2006:137). Thus, in order to decode and encode the sounds of a language active knowledge about the sound system is necessary. Fries (1945:3) goes so far as to say that “in learning a new language ... the chief problem is not at first that of learning vocabulary items. It is, first, the mastery of the sound system. [...] It is, second, the mastery of the features of arrangement that constitute the structure of the language”. He even states that, for English, “one can achieve mere fluency in a foreign language too soon. [...] Such students with fluency in vocabulary but with no basic control of either the sound system or the structure, are almost without exception hopeless so far as ever achieving a satisfactory control of English is concerned” (1945:3). Though this generalisation might be too strong to prove true, it nevertheless makes the point of how important mastery of pronunciation is. After all, L2 learning is about communicating and being understood.

Learners often only want to acquire a level of proficiency in a language that allows them to communicate effectively in the language. Native-like pronunciation is very rarely the goal and hardly ever considered necessary. The article acknowledges this pragmatic approach to language learning. However, there might be students who have the will and the talent to learn more features of the language in question. Teachers should be equipped to teach these features and it would be desirable if there was some language learning material available that would address these needs as well. There is, for example no material available for any of South Africa’s African languages that teaches tone.

The present paper addresses the acquisition of phonetic, phonological and prosodic features by L2 learners of any of the languages of the Sotho-Tswana group. The group of Sotho-Tswana languages comprises Northern Sotho, Southern Sotho and Tswana, three closely related and mutually intelligible Southern Bantu languages spoken in South Africa, Botswana and Lesotho. They are considered similar enough to be treated together in the current article. The article concentrates on learners with a Germanic L1 because being non-cognate to the Bantu languages considerable difficulties are predicted, especially in what concerns the tone system. L2 acquisition of learners with a Bantu mother tongue will have qualitatively different problems. Sotho-Tswana is taught to speakers with a Germanic L1 at institutions of tertiary education throughout South Africa, at the University of Pretoria, University of South Africa, University of Johannesburg, at North-West University Potchefstroom, at the University of Lesotho in Roma, at the University of Botswana in Gaborone. Furthermore, Northern Sotho is taught at Humboldt-University Berlin (Germany), and according to website information also Michigan State University³ offers Southern Sotho, and Yale University⁴ offers Tswana and Southern Sotho on demand.

³ <http://africa.msu.edu/afrilang.php>, accessed 10/06/2009

⁴ <http://www.yale.edu/macmillan/african/languages/index.html>, accessed 10/06/2009

The phonological aspects in (1) have been singled out as areas of potential problems for a learner with a Germanic L1 background that will be discussed in the present article.

(1) Potential areas of problems in Sotho-Tswana L2 for Germanic L1 speakers

- a) Grapheme-to-phoneme correspondence
 - (i) 7-vowel system
- b) Sound production
 - (i) Place/manner of articulation
 - (ii) Airstream mechanisms
 - (iii) Voice/ aspiration
- c) Syllable/rhythmic structure
 - (i) NC clusters
 - (ii) Syllable-timed language
- d) Phonological processes
 - (i) Vowel harmony
 - (ii) Postnasal devoicing
- c) Prosody
 - (i) Tone
 - (ii) Lengthening
 - (iii) Question prosody

The areas in (1) have been singled out by a contrastive analysis of English versus Sotho-Tswana. Contrastive analysis is an approach to L2 acquisition that involves predicting learner problems based on a comparison of L1 and L2 in order to determine similarities and differences. The assumption is that there is positive transfer of structures that are present in both L1 and L2. These structures should not pose a problem for the learner. However, the learner will encounter problems if a linguistic structure in L1 is inappropriate in L2 (interference) or if it does not exist in the L1. The linguistic features in (1) do not exist in the Germanic languages, or are used differently (interference), so they are predicted to cause problems to the learner, according to the approach of Contrastive Analysis.

A problem with the approach of Contrastive Analysis is that the predictions it makes are not always validated by evidence from actual learner errors. The Error Analysis proves to be a more useful approach in this respect as it includes an internal focus on the learners' creative ability to construct language. However, I am not aware of any study that has been conducted whose results would allow error analysis for L2 learners of Sotho-Tswana. Thus, the current paper will start out from a Contrastive Approach, preparing the stage for empirically driven error-analyses in the future. The following sections will address each of the potential problematic areas in (1) by laying out the linguistic background. The paper provides the necessary linguistic background that makes language teachers sensitive to the problems at hand and comfortable in addressing them in class.

2. Grapheme-to-phoneme correspondence

The orthography of the Sotho-Tswana languages came with the missionaries. Being of European origin, the missionaries used the Latin alphabet to transcribe the indigenous languages.

In Sotho-Tswana languages the spelling matches the pronunciation relatively well in contrast to English, where an orthography is used which has not been updated regularly throughout the centuries to reflect changes in pronunciation. The phoneme-grapheme correspondence in Tswana is shown in (2).

(2) Grapheme-to-phoneme correspondence in Tswana (Cole, 1955:21ff.)

(a) Consonants

<i>Plosives</i>	<i>Fricatives</i>	<i>Nasals</i>	<i>Affricates</i>
<p> → /pʰ/	<f> → /f/, [ɸ]	<m> → /m/	<ts> → /tsʰ/
<t> → /tʰ/	<s> → /s/	<n> → /n/	<tš> → /tʃʰ/
<k> → /kʰ/	<š> → /ʃ/	<ny> → /ɲ/	<tsh> → /tsʰ/
<ph> → /pʰ/	<g> → /x/	<ng> → /ŋ/	<tšh> → /tʃʰ/
<th> → /tʰ/	<h> → /h/	<i>Glides</i>	<kg> → /kxʰ/
<kh> → /kʰ/	<i>Liquids</i>	<w> → /w/	<j> → /dʒ/
 → /b/, [β]	<r> → /r/	<y> → /j/	
<tl> → /tʰ/	<d> → /ɾ, d/		
<tʰh> → /tʰ/	<l> → /l/		

(b) Vowels (Setswana Terminology and Orthography No 4: 5)

<i> → /i/	<u> → /u/
<e> → /e/	<o> → /o/
<ê> → /ɛ/	<ô> → /ɔ/
<a> → /a/	

From (2), it emerges that there is a close relationship between graphemes and phonemes in the consonant system which facilitates acquisition once the correct pronunciation of the consonant and vowel phonemes has been learned (cf. section 3 below).

However, with the vowels a specific problem occurs in second-language acquisition. The current official Setswana terminology and orthography No 4 (1988:5) advises that a

“circumflex sign (ˆ) on the vowels ê and ô should be used regularly in scientific works such as grammars and dictionaries, which as books of reference, should reflect the pronunciation accurately. In non-scientific works, however, such as readers, novels, etc. the circumflex should be used only where there is likely to be confusion with different words having identical spelling [...]”.

However, current second-language teaching material hardly indicates vowel quality (but cf. Cole & Mokaila, 1962 and Mascher, in prep.). One of the reasons I often hear for not differentiating between vowel qualities in teaching material is that they are also not reflected in orthography.

The data in (3) show minimal pairs to underline the fact that the graphemes <e> and <o> conflate two phonemes each if used without the circumflex.

- (3) Minimal pairs in Tswana (Setswana Terminology and Orthography No. 4: 6; tones added from Kgasa & Tsonope, 1995)

phóló	'ox'	vs.	phóló	'health'
lema	'plough'	vs.	léma	'spoil/shape horns'
noká	'river'	vs.	nóka	'hip'
tshéla	'cross'	vs.	tshêla	'pour in'

If second-language teaching is based solely on “non-scientific” orthography, the orthographic conflation poses an immense problem and can easily become a source of frustration for the L2 learner both in production and perception. The fact that these obvious phonological differences are not encoded in an orthography which even reflects certain consonant allophones (see the /d/-alternation in section 5.1) makes the learner wonder if there is a way to predict which mid-vowel occurs, perhaps based on tone or phonological context. The learner is wasting his/her energy unnecessarily. The difference in vowel quality is phonemic, thus not predictable and must simply be learned as part of the vocabulary item.

Cole and Mokaila (1962) consequently use the suggested diacritics in their Tswana teaching material to indicate the difference consistently. Mascher (in prep.) goes a step further and even uses completely separate symbols in order to make the contrast between the two sounds very clear, even to the eye.

The teacher should point out this ambiguity in non-scientific orthography right from the beginning in order to avoid confusion on the side of the learner. I would even suggest adopting the suggested orthographic convention for the open mid-vowels even if this means a partial rewriting of existing teaching material. This allows the learner to pronounce a given word correctly before taking on irreversible pronunciation errors which will become fossilised. Fortunately, dictionaries follow the advice of the Department of Education and indicate vowel quality in their entries (cf. Pukuntšú ya Polelopedi ya Sekolo 2008 for the most recent example). Thus, with some effort on the side of the teacher the learner can easily be provided with this information. It will be only at a very late stage in the acquisition process, namely once the learner is very familiar with the vocabulary, that diacritics can be dispensed of.

Next to visual cues, the accurate pronunciation of especially these contrasting mid-vowels must receive attention in the classroom as well as their perception in contrast to other vowels in the language. Awareness of differences in vowel quality in production obviously helps to develop listening skills so that an active teaching of vowel quality helps to identify the sounds in listening correctly. Language laboratories which practise speaking and listening skills in foreign languages are well established across language schools all over the world. Similar training is needed for the sounds of Sotho-Tswana. More varied and interesting exercises than the pronunciation drill in Cole and Mokaila (1962:3) need to be devised for this purpose.

To give a practical example, African Voices (2004) has published a very useful CD, *Everyday Phrases in Tswana*. It is accompanied by a booklet with the orthographic rendering of the phrases together with an English translation. In the booklet and on the CD the semantically important differences in vowel qualities are explained. In rendering the phrases in ‘non-scientific’ orthography only, however, the authors impose on the casual learner the burden

to develop the listening skills to distinguish these two mid-vowel qualities in rapid speech first, and then to remember them for each and every word without a visual cue (unless s/he transcribes them for him-/herself). A lot of time is wasted which could have been more beneficially used for the acquisition of several phrases of Tswana instead.

3. Sound production

The Sotho-Tswana languages have sounds that are unknown to speakers of Germanic languages like English. A specific combination of manner and place of articulation might be unknown, as with lateral fricatives and bilabial fricatives, or the airstream mechanism might be new, as in ejectives. These features or their combination do not occur in English. They thus pose a problem for the learner (cf. Mojapelo & Jordaan, 2004). Therefore, it is suggested to give a detailed explanation of the production of these sounds in the classroom, accompanied by specialised pronunciation exercises which are repeatedly administered.

A helpful guide for explaining the production of unknown sounds is Catford (1988). The following explanations and exercises which are listed are taken from his 'Practical Introduction to Phonetics' (Catford, 1988). The three problematic aspects of sound production will be discussed in detail in the following sections.

3.1 Place/manner of articulation

With respect to place and manner of articulation, the following sounds might pose challenges to the L2 learner of the Sotho-Tswana languages given they do not occur in some of the Germanic languages: lateral fricatives, the bilabial fricative [β], the trilled [r], and the velar fricative [x]. Example words are provided in (4) from Northern Sotho.

(4) Northern Sotho (Ziervogel & Mokgokong, 1975)

[ɬ]	hlóǰō – 'head'
[β]	gō bába – 'be bitter'
[r]	gō r̄era – 'to preach'
[x]	gō gǒga – 'to pull'; also in the affricate [kx] kgómó – 'cow'

Laterals other than [l] are a rare feature in the languages of the world, also on the African continent (Maddieson, 2008). The pronunciation of the lateral fricatives poses a problem for the L2 learner with a Germanic language background. Often they are substituted by [ʃ] (Mjiyako, p.c.). The pronunciation of the lateral fricative [ɬ] can be derived from the lateral approximant [l]. One starts by producing a prolonged [l], and while doing this one expands the tongue – striving to push the sides of the tongue out sideways, so that they squeeze between the molar teeth. With a little trial and error one can narrow the lateral channels of [l] so much that a lateral fricative type hiss begins to be heard along with the sound of voice. This is a voiced lateral fricative [ɬ]. By manipulating voicing (see section 3.2) one arrives at the target sound (Catford, 1988).

The trilled [r] is produced by allowing the tip of the tongue to flap against the ridge behind the upper teeth. It can be trained by raising the tip of the tongue slightly and letting it rest loosely against the back of the ridge behind the upper teeth. On powerful exhalation and, again, by trial and error adjustments of the precise position and tension of the tongue, and of the exhalation

pressure, the student might be able to start the tongue-tip regularly, periodically flapping against the alveolar ridge. This will result in a voiceless trill [r̥]. By producing a prolonged [r̥ r̥ r̥ r̥] and manipulating voicing in the middle of it, one finally reaches at the target sound. “Incidentally, in experimenting with trills it is often best to start with the voiceless ones, since it is easier to set organs trilling with the more powerful voiceless air-stream than with the voiced one” (Catford, 1988:69).

The bilabial fricative [β], which still occurs in varieties of the Sotho-Tswana languages, is formed with a narrowing at the lips through which the air escapes, producing a friction noise. It can be trained by starting out with a bilabial stop [p], building up pulmonic pressure behind the closed lips. Allowing the lips to separate very slightly, so that a turbulent air-stream escapes through this narrow channel, will generate a voiceless bilabial fricative [ϕ]. If voice is added, it will become a voiced bilabial fricative.

The velar fricative [x] is reached by starting out with the tongue in a position for a [k]. While taking care not to shift the tongue forward or backward, a very small channel between the tongue and the soft palate is opened up. If the airstream coming from the lungs is forcefully directed through this channel, the sound of a voiceless dorso-velar fricative [x] is heard. The affricate [kx] is produced by a quick sequence of the velar plosive [k] and the velar fricative [x]. It is followed by aspiration (Cole, 1955:33, Ziervogel *et al.*, 1969:7).

3.2 Voice

Aspiration is a feature of voice (e.g. Ashby & Maidment, 2005). Aspiration refers to a weak friction sound heard on the release of a plosive sound. Aspiration is associated with voiceless plosives. Aspiration can be illustrated by holding a sheet of paper in front of the mouth when producing a [p] and [p^h] respectively. Whereas the paper does not move when pronouncing the unaspirated [p], it does move with [p^h].

Southern Bantu languages use both aspirated and unaspirated consonants phonemically, i.e. in order to distinguish meaning⁵. Minimal pairs are given in (5).

(5) Northern Sotho (Ziervogel & Mokgokong, 1975)

go khuta – ‘to hide’	vs.	go kuta (or kúta) – ‘to have a haircut’
phelá! – ‘Live!’	vs.	péla – ‘rock-rabbit’
thaba – ‘mountain’	vs.	taba – ‘a matter’
tshéla – ‘cross over’	vs.	tsela – ‘road’

In orthography, aspiration is consistently encoded by <h> (except for the affricate <kg> which is always aspirated). Thus, it is easy for an L2-learner to see in the orthography how the sounds differ. What may cause difficulties, however, is if the learner cannot actively distinguish between these two sounds in production or perception. One reason for this might be that,

⁵ At least in the Sotho-Tswana languages, aspirated and unaspirated stops do not contrast directly. The aspirated consonant is a pulmonic consonant, whereas the unaspirated is reported to be an ejective (see next section). However, when the distinction between pulmonic and ejective consonants vanishes, the distinction between aspirated and unaspirated is kept intact (cf. Wissing, 2005, Tlale, 2005).

though aspiration is used with voiceless consonants in English, it is not phonemic, i.e. that it does not differentiate between words. Also, aspiration occurs only in certain positions: in English, voiceless plosives are aspirated in all positions except after [s]. Thus, the plosives in *spin*, *stop*, and *school* are produced without aspiration in English.

Aspiration seems to be indeed a major problem for L2 learners (Mjijako, p.c.) Because of the negative inference from the mother tongue, the L2 learner has to be made aware of the difference and attention needs to be paid to the correct articulation of the respective sounds. As a first step, the learner has to be made aware of aspiration as such. Catford (1988:59f) suggests the following exercise to acquire the difference between aspirated and unaspirated stops through own experience rather than through abstract, theoretical explanation:

- (i) Breathe in then begin a prolonged voiceless exhalation – a pulmonic pressure initiated air-stream flowing up through a wide-open glottis. Momentarily superimpose a [p] on the airstream, thus [h h p H h h]. [...] Repeat the experiment, but this time 'switch on' voice immediately after release of the stop, thus: [h h p H ə ə ə]. This should produce an aspirated [p]. Do the same with [t] and [k].
- (ii) Breathe in as before, but now begin a prolonged whispered exhalation. [...] Momentarily superimpose a [p] on this whispered airstream. Now repeat the experiment, but this time 'switch on' voice just as you release the stop. If you have indeed kept the glottis in the whisper configuration up to that point, voice should start immediately – with no delay. In other words, the result should be an unaspirated [p]. Do the same with [t] and [k].

Once the learner has become conscious of the difference, carefully designed production and perception exercises are necessary to provide practice, both in production and perception. The correct production of aspiration in consonants is a feature that needs to be paid attention to continuously throughout tuition.

3.3 Airstream mechanisms

In Germanic languages, all sounds are produced on exhalation, i.e. by means of the pulmonic egressive airstream. However, sounds can also be produced using initiation other than the air from the lungs, as is the case for many consonants of the Sotho-Tswana languages.

The Sotho-Tswana languages distinguish between aspirated pulmonic consonants and ejectives. Ejectives are produced with an egressive glottal airstream mechanism which means that the glottis acts as the initiator of the sound. The glottis is firmly closed while the closure is formed in the oral cavity. When the larynx rises slightly, a small quantity of air is trapped between the closed glottis and the oral closure. If the oral closure is released, the entrapped high-pressure air bursts forward in a short, sharp explosion. Ejectives are thus not surprisingly acoustically more prominent than pulmonic stops.

Germanic languages only have pulmonic stops. Thus, the production of ejectives has to be practised by the L2 learner if s/he wants to produce the sounds correctly.⁶ In order to practise

⁶ Interestingly, velar ejectives seem to pose less of a problem for L2 learners with an Afrikaans background (Mjijako, p.c.). Afrikaans voiceless stops have a short voice onset time, i.e. they are not aspirated. It remains to be investigated in how far the characteristics of Afrikaans stops enhance the production of ejectives.

their production ejectives are most easily discernible as such at the velar place of articulation. Catford (1988) suggests the following steps: first, the controlled closure of the glottis needs to be practised, then the larynx needs to be lowered and raised at will (both steps are also important for the production of implosives), and lastly, these two steps need to be combined in order to produce one speech sound – an ejective.

In order to practice closing the glottis at will, Catford (1988: 24) suggests to start out "with making a series of quiet coughs; that is, a series of glottal stops [ʔ]. Now the closure for [ʔ] is formed but not released, holding the closure for some time. One way of testing whether the glottis is really closed is by opening the mouth and making sure that the soft palate is raised (as for an oral consonant) and that the tongue does not touch the roof of the mouth at any point. This is to ensure that there is a clear passageway from the larynx through the mouth. While the mouth is held open and unobstructed in any way and the glottis is closed, flick a finger against the neck, either on the side of the larynx or just above it. Done properly this renders a note whose pitch can be changed by lip rounding or tongue movement. If the glottis is open, the flicking of the finger against the larynx will result in a dull sound. Raising and lowering the larynx at will can be trained by uttering the highest note one can reach followed in a rapid change by the lowest note. This action needs to be repeated several times with a finger lightly touching the 'Adam's apple' in order to feel the larynx moving up and down. The next steps are to silently utter the notes, and then to imagine uttering the notes.

In combining the two actions of closing the glottis and raising the larynx, the pronunciation of an ejective can be achieved. One starts with closing the glottis and keeping it closed throughout. With the closed glottis, the back of the tongue raises to the velum to form the velar closure. The larynx will be raised slightly. While maintaining the glottal closure, the oral closure is suddenly released, resulting in a short explosion. At last, the glottis will be opened again."

Again, once the student can produce ejectives in isolation, s/he needs carefully designed practice sessions on both production and perception together with ongoing attention to this linguistic feature throughout the course. Recent work has doubted the occurrence of ejectives in Sotho-Tswana (Tlale, 2005, Wissing, 2005), at least in some dialects or varieties. However, until further research has settled that matter, I think it is fair to teach the traditionally reported sounds.

4. Syllable structure

Bantu languages differ from Germanic languages in their rhythmic structure. Whereas Germanic languages are so-called stress-timed languages, Bantu languages are syllable-timed languages. Going into the difference is beyond the scope of this article and teaching these differences in rhythmic aspects is probably beyond even an advanced language course (cf. discussion in Gut *et al.*, 2007). For now, suffice it to say that in stress-timed languages, stressed syllables occur at roughly equidistant intervals whereas in syllable-timed languages, syllables occur in roughly equal intervals. The distinction between stress-timed and syllable-timed languages coincides with differences in syllable structure which is directly relevant even to the beginning learner.

Syllable-timed languages show less complex phonotactics, that is permissible consonant clusters in onsets or codas of syllables. Consequently, the syllable structure of syllable-timed languages such as Bantu languages is less complex. Normally, we find syllables of the general structure CV in syllable-timed languages, where C stands for a consonant and V stands for a

vowel. Thus, a single consonant is followed by a vowel. Contrary to languages like English which allow complex consonant clusters at the beginning of syllables (e.g. 'strange' [stɹeɪŋdʒ]), Bantu languages thus show simpler syllable structure and it is not predicted to cause a problem for L1 English speakers who acquire a Bantu language. Despite their relatively simple syllable structure, however, Bantu languages show some features with respect to syllable structure that might cause difficulties nevertheless. These are treated in the following.

4.1 Affricates

The languages of the Sotho-Tswana family show a relatively large amount of affricates, i.e. of plosives being followed by homorganic fricatives. Examples are given in (6)

(6) Northern Sotho (Ziervogel & Mokgokong, 1975)

In analyses of Northern Sotho they have been described as single segments (Lombard, 1985).

< tsh >	→ [tʃ ^h]	go tshéla – 'to cross over'
< tʃh >	→ [tʃ ^h]	go tʃhába – 'to flee'
< kg >	→ [kx ^h]	kgômó – 'cow'
< ts >	→ [ts ^ʰ]	tsele – 'road'
< tʃ >	→ [tʃ ^ʰ]	go tʃéa – 'to take'
< pʃh >	→ [pʃ ^h]	go pʃhatla – 'to break into pieces'
< psh >	→ [ps ^h]	go ípshína – 'to be happy'
< pš >	→ [pʃ ^ʰ]	go bópša – 'to be moulded'
< bj >	→ [βʒ]	bjalá – 'beer'

However, though there is some coarticulation and thus overlap in articulating gestures, phonetically the affricates can be considered a simple sequence of stop + fricative (Ladefoged & Maddieson, 1996:354).

For these affricates, the L2 learner needs to acquire the skill of producing a plosive followed by a fricative at the designated places of articulation in rapid sequence.

4.2 Syllabic sonorants

The Sotho-Tswana languages also systematically allow sonorants to occur as syllable nuclei. Although this also occurs in English due to vowel reduction processes (cf. 'little' – [lɪ.t̩]), the phenomenon is much more frequent in Sotho-Tswana. Examples are given in (7).

(7) Northern Sotho (Ziervogel & Mokgokong, 1975)

mollô	[mɔ̩.lɔ̩]	–	'fire'
mmê	[m̩.mé]	–	'mother'
nna	[n̩.ná]	–	'I'
o a nnyaka	[n̩.na.ka]	–	'he looks for me'
e nngwe	[n̩.nwí]	–	'another'

⁷ Although these last four sound sequences do not meet the definition of an affricate they are often listed as such in grammar books and are also subsumed here under this label for simplicity's sake.

Syllabic sonorants carry their own tone (see section 6.1) which might be distinct from the tone of the following syllable. Careful pronunciation of the syllabic sonorant together with its tone is thus important for adequate pronunciation. Moreover, syllabicity of sonorants can also be semantically relevant in that it differentiates between meanings, as in (8).

(8) ba.ná – ‘children’ vs ba.ǀná – ‘men’

The L2 learner has to be made aware of the fact that double sonorants and sonorants preceding consonant clusters in orthography are always syllabic and should be pronounced accordingly.

5. Phonological processes

5.1 /l/d-alternation

Whereas in English and German the consonants /l/ and /d/ are separate phonemes, in the Sotho-Tswana languages there is no phonemic contrast between them. Their phonetic realisation is contextually determined: [l] appears before non-high vowels, and [d] appears before high vowels. Furthermore, neither consonant may come at the end of a word or before another consonant (Odden, 2005:50). Examples are given in (9) from Tswana.

(9) Tswana

go lema	– ‘to plough’
selépe	– ‘axe’
go kwála	– ‘to write’
lekwálô	– ‘letter’
loná	‘you (pl.)’
but: mokwádi	– ‘writer’
go dumêla	– ‘to agree’

The phonological alternation is of importance to the L2 learner because some morphological processes systematically trigger the alternation. One example is the formation of nouns from verbs (deverbatives) that denotes the person who is carrying out an action. This is a very productive process and is triggered by the suffix *-i*. Whenever a verbal root ends in an *-l-* it can be foreseen that this will change into *-d-*, as in (10a). Another context arises with the past tense suffix *-ile* or the causative infix *-is-* which also causes the alternation systematically and productively. Examples are given in (10).

(10) Northern Sotho

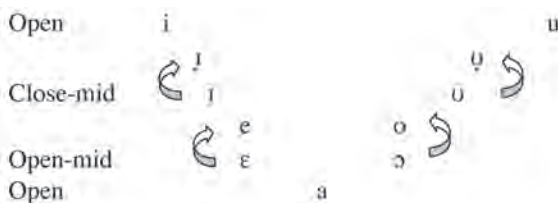
(a)	ngwála	→ mo-ngwád-i	‘to write’
	rwála	→ mo-rwád-i	‘to carry’
(b)	go bóla	→ bód-ilê	‘to be rotten’
	go êla	→ êd-ilê	‘to flow’
	go fêla	→ fêd-ítšê	‘to finish’

The l/d-alternation is reflected in the orthography (except for Southern Sotho which writes <l> in all environments). Beyond the orthography, however, the L2 learner may have difficulties recognising the root when they are not made aware of this systematic alternation of the two phonemes. It is thus advisable to draw the learner’s attention to the alternation whenever one comes across it. The learner will soon learn to recognise it him-/herself.

5.2 Vowel harmony

As presented in section 2, Sotho-Tswana languages have seven vowels of four different heights. However, in certain phonological contexts the mid-vowels show allophonic variations that are higher in quality. They do not merge with existing vowels but create new phonetic vowel heights. Thus, there are eleven surface vowels in Sotho-Tswana surfacing at six phonetic heights (see Khabanyane, 1991; Tlale, 2005 for phonetic studies). The eleven surface vowels are shown in (11).

(11) Surface vowel system of Sesotho



The vowel harmonic processes are given in (12) following work by Khabanyane (1991) (but see also Tlale, 2005).

- (12) (a) /ɪ ʊ/ → [ɪ̥ ʊ̥] / __ /i u/
 (b) /ɛ ɔ/ → [ɛ̥ ɔ̥] / __ /i u/, /ɪ ʉ ts/

The rule in (12a) states that the close mid-vowels are slightly raised in quality when preceding high vowels. Rule (12b) states that similarly the open mid-vowels are slightly raised when they precede front vowels, other mid-vowels or the nasals /ɪ ʉ/ and the affricate /ts/. The harmonic process is thus regressive, i.e. following sounds influence the pronunciation of preceding sounds. The influence of the vowel triggering the harmony goes back leftward through prior syllables until it meets the vowel [a] (cf. Cole, 1955, Tlale, 2005). Examples of vowel rising are given in (13).

(13) Tswana (Tlale 2005, section 2.4)

/bita/ + ilê	[bɪ̥tɪ̥]	'have chocked'
/bolóka/ + -i	[babʊ̥lɔ̥ki]	'saviors'
/rɛka/ + i	[mɔ̥rɛki]	'buyer'
/ts ^{hw} ɛni/	[ts ^{hw} ɛ̥ni]	'baboon'
/pɛu/	[pɛ̥u]	'seed'
/ts'ɛbɛ/ + -ng	[tsebɛ̥ŋ]	'in the ear'
/t'iro/ + -ng	[tiro̥ŋ]	'at work'

In second-language acquisition it is important, however, to keep in mind that the additional vowel heights induced by vowel harmony are not phonemic in the Sotho-Tswana languages, i.e. that they do not differentiate between meaning. Thus, for the L2 learner there is no danger of confounding lexical meanings of words if s/he does not master this aspect perfectly. However, the variation in vowel qualities might lead to confusion in perception and it is thus most of all the perception that needs to be worked on. Making the student aware of vowel rising in certain phonological contexts (e.g. most prominently in the locatives) keeps potential confusion to a minimum.

5.3 Postnasal devoicing

The Sotho-Tswana languages show so-called sound hardening in the context of the first person singular object marker, the reflexive marker and the class 9/10 class prefixes (see e.g. Dickens, 1984). In these contexts the following sound changes in (14) occur, illustrate by the imperative plus first person object marker.

(14) Tswana

(a)	b > p	/N-bóts-e/	→	Mpótsé! – ‘Ask me!’
(b)	r > th	/N-rékêl-ê/	→	Nthékélé! – ‘Buy for me!’
(c)	s > tsh	/N-sény-e/	→	Ntshényé! – ‘Destroy me!’
(d)	f > ph	/N-f-é/	→	Mphé! – ‘Give me!’
(e)	V > k+V	/N-ádim-e/	→	Nkádímé! – ‘Lend me!’
(f)	l > t	/N-latêl-e/	→	Ntatélé! – ‘Follow me!’

The occurring sound changes are counterintuitive from a phonological point of view, as in the context of nasals one would rather expect voicing. Articulatorily, airflow leakage through the nasal cavity should promote, not inhibit, voicing post-nasally. However, the postnasal devoicing in examples as in (14) has been attested experimentally for at least some Tswana speakers (cf. Coetzee *et al.*, 2007). Further research has to provide a satisfying explanation and analysis for the sound changes depicted in (14).

Postnasal devoicing is reflected in the orthography. Consequently, it does not pose a problem for the pronunciation when reading a text, for example. From own experience I know that these changes posit a considerable problem in production and perception, especially when these morphologically complex forms are parsed while speaking or listening. A unified explanation as well as specific exercises might help the adult L2 learner to master this problem through linguistic insight.

5.4 Palatalisation

Southern Bantu languages are known for the process of palatalisation. Traditionally, palatalisation refers to a secondary articulation in which the front of the tongue is raised to the hard palate. However, it can also be used for describing a process in which the primary articulation is changed in such a way that it becomes more palatal. This is the case in the Southern Bantu languages. Palatalisation in these languages occurs most of all in the passive (15a) and in diminutive formation (15b).

(15) Tswana

(a)	/-rôb-w-a/	→	[-rôjwa]	‘be broken’
	/-bóp-w-a/		[-bótšwa]	‘be moulded’
	/-tlhôph-w-a/		[-tlhôtšhwa]	‘be chosen’
	/-lef-w-a/		[-lešwa]	‘be paid’
(b)	/kubú-ana/		[kujwána]	‘small hippo’
	/legápu-ana/		[legátšwána]	‘small water melon’
	/tshêphê-ana/		[tshêtšhwána]	‘small springbok’
	/sefôfu-ana/		[sefôšwana]	‘small blind person’

Palatalisation exists in all Sotho-Tswana languages but the sounds resulting from palatalisation differ quite considerably across the varieties (cf. Kotzé & Zerbian, 2008). A detailed discussion is thus beyond the scope of the present paper.

With respect to second-language learning, however, as palatalisation is reflected in the orthography it does not pose a problem for the pronunciation when reading a text, for example. However, as with postnasal devoicing the L2 learner might face difficulties recognising nominal and verbal stems once they have been palatalised. A basic knowledge of the places of articulation and the induced sound changes is thus helpful for the L2 learner in order to be able to derive underlying or resulting forms him-/herself.

6. Prosody

The term prosody is used to refer to pitch changes and durational changes across the utterance. Though discussed here last, some approaches to L2 acquisition start with making the learner familiar with the sentence melody of the target language before single phonemes are focused on (cf. Tomatis, 1991). More commonly, however, pronunciation training mainly focuses on segmental aspects. Aspects of word and sentence prosody, even in languages like English, are taught and addressed less frequently or considerably later in the acquisition process (cf. Gut & Trouvain, 2007).

For the languages of the Sotho languages, as for Bantu languages in general, it needs to be stated though that word and sentence prosody is manipulated in order to encode lexical, grammatical, syntactic and/or semantic/pragmatic meaning. As such, the teaching of tone and prosody is of the utmost importance for the correct and appropriate acquisition of the vocabulary and grammar of the target language. It cannot be left until a specific level of proficiency has been reached.

The importance of tone for the lexicon and grammar in the Sotho-Tswana languages makes it even more surprising that tone is not taught in second-language teaching of these languages. In an experiment Roux (2001) found that tone bears relatively little weight for the discrimination of minimal pairs. This is due to the fact that there are only a few tonal minimal pairs in these languages and context will always help to disambiguate. However, this does not mean that tone should not be taught. Every mother-tongue speaker of the language can confirm that inappropriate word and sentence melody contributes to a foreign accent and negatively influences intelligibility. In speech technology, it has been found that tonal modelling and the correct prosody with respect to duration increase the judged naturalness of synthetic speech (E. Barnard, p.c.).

It is actually incorrect to say that tone carries little meaning in these languages. The tone in the verb paradigm as exemplified in Chebanne *et al.* (1997) is a thorough example that this is not true. Among the reasons why tone is not taught is that the tone system is complex. Every prosodic system that differs from the learners' L1 system is difficult to acquire. This is true for the Chinese tone system, but also for English word stress if learnt by speakers of a non-Germanic language. By comparison, the fact that the English vocabulary is so rich due to its language history, has not led to a simplification of its vocabulary in teaching. A further objection given against teaching tone is that even mother-tongue speakers cannot explain the tone system of their language. However, no untrained mother-tongue speaker of whatever

language can consciously access and list the rules that govern the various linguistic aspects of his/her language. That does not mean that we cannot teach rules or generalisations that have been established in linguistic studies. A last objection given against teaching tone is that the student numbers for African languages are low already, so those who do study African languages should not be scared away with the complexities of a tone system.

Assuming a goal to teach and learn a language in all its complexity, the tone systems of the respective languages would need to be re-studied. Subsequently, teaching materials would need to be expanded or developed, and an appropriate methodology would need to be worked out. Given the salience of tone especially in the verb word, the teaching of tone and morpho-syntax will influence each other. I am aware of two initiatives to develop pedagogically oriented teaching materials which include tone, one for Tswana which is very advanced (Mascher, in prep., based on but superseding Cole & Mokaila, 1962) and one for Zulu (Buell & Sosibo, in prep.). Further ideas for the methodology in the teaching of tone can also be obtained from the teaching of Chinese. Also, the awareness of the importance of tone needs to be raised so that dictionaries include tone for each lemma (the latest Zulu dictionary includes tone (*Isichazamazwi sesiZulu*, 2006), whereas the latest Northern Sotho and Tswana dictionaries omit it (*Pukuntšu ya Polelopedi ya Sekolo, Thanodi ya Setswana*, 2008)). Otherwise learners lack the necessary supportive resources to put the knowledge to use. It is without doubt that the correct tonal intonation is important for a thorough appreciation and mastery of the language.

Three aspects will be discussed and illustrated in more detail in the following section: tone at word and sentence level, durational changes, and pitch and durational changes for the indication of sentence types.

6.1 Word and sentence tone

Most of the world's languages are tone languages. The Southern Bantu languages are classified as having simple tone systems, consisting of a high and a low tone. This classification is based on prevalent analyses the past 40 years that found that Bantu languages display a two-tone system, in which high tones are specified in an underlying manner and low tones are inserted by default. Earlier descriptions by scholars and missionaries assumed (far) more tones (see e.g. Doke 1992 who posited 9 tonal levels for Zulu, and Endemann (1911) who stated three tones for Northern Sotho). As these previous studies indicate, the surface tonal system is far from being trivial. However, for the language learner the simple tonal inventory is good news in contrast to the more complicated tonal inventories of Mandarin Chinese or Vietnamese, for example.

To my knowledge, second-language teaching of Bantu languages, both at school and university level, disregards tone (except for some few lexical minimal pairs). However, acquisition of tone and prosody will remain impossible if the linguistic system behind tone or prosody in general, is not taught, practised and learned. Only once the basis of the tone system has been laid, the acoustic cues of prosody in the target language, such as a high or a low tone can and must be trained in the language laboratory by means of carefully designed pronunciation exercise. An explanation of the tone system of the Sotho-Tswana languages is beyond the scope of the present paper (but see Khoali, 1991 for Southern Sotho, Lombard, 1976 for Northern Sotho, and Cole & Mokaila, 1962 and Chebanne *et al.*, 1997 for Tswana).

The purpose of the present section is to outline and give examples of the three types of tonal mistakes that are predicted for learners of Sotho-Tswana as a second language: mistakes with tonal minimal pairs that lead to unintended lexical meanings, mistakes with grammatical tone which lead to incorrect grammatical constructions, and mistakes with tone patterns that merely lead to the learner sounding foreign.

The examples in (16) show how tone is used for the distinction of lexical meaning in Northern Sotho. Obviously, if the tone pattern is produced incorrectly on a lexical item, miscommunication might result due to an unintended lexical meaning. As Roux (2001) pointed out, context might disambiguate in many cases but the learner should not rely on that from the start.

(16) Lexical minimal pairs in Northern Sotho (Ziervogel & Mokgokong, 1975)

lapá - 'court-yard'	lapa - 'become tired'
bōna - '(to) see	bōná - 'they'
thápō - 'kernel'	thapō - 'rope'
hlába - 'illegitimate child'	hlaba - '(to) stab'

Tone of lexical items is not predictable and thus must be learned together with the vocabulary item. In parallel to the rules regarding diacritics in respect of vowel quality, Chebanne *et al.* (2003:i) advise that

“tone should be marked in technical texts such as dictionaries, grammar books and other texts dealing with language pedagogy”.

Unfortunately, not all dictionaries include tonal information for their entries due to a variety of reasons. The dictionaries that include tonal information are Ziervogel and Mokgokong (1961, 1975) and Kriel and van Wyk (1966) for Northern Sotho, as well as Snyman *et al.* (1990) and Kgasa and Tsonope (1995) for Tswana. Unfortunately Ziervogel and Mokgokong (1961, 1975) and Snyman *et al.* (1990) seem to be out of print, Kriel and Van Wyk (1966) is Tswana-Afrikaans only. Learners might thus have to start their own tone-marked dictionary.

There is some teaching material available for Tswana that indicates tone. Sandilands (1953), though not outlining tone patterns of Tswana, indicates the surface prosody in some of his exercises. Cole and Mokaila (1962) include tone-marking on each syllable of their example sentences by using acute and grave accents. Also Mascher (in prep.) marks every syllable for tone, using a tone marking system that is even clearer to the learner’s eye, namely adding a diacritic for high tones on top of the vowel concerned and a diacritic for low tones underneath the vowel in question.

The real challenge of the Sotho-Tswana tone system is not the tonal inventory as such, but the tonal morphology and tonal changes. Tonal distinctions are made between different grammatical constructions and/or different tenses in Sotho-Tswana. Examples from Tswana are given in (17).

(17) Grammatical minimal pairs (Tswana, according to Cole & Mokaila 1962)

(a)	re rúta - 'we teach'	ré rúta - 'while we teach'
(b)	ke morúti - 'I am a teacher'	ké morúti 'He/she/it is a teacher.'

Producing the examples in (17) with an incorrect tone pattern will evoke an unintended grammatical construction and may lead to miscommunication. Again, often the context will disambiguate any confusion that may arise from using an inappropriate tone pattern. However,

teachers will not be lenient if the learners' confusion with the two grammatical constructions result in other grammatical errors, such as the choice of the incorrect subject concord in (18a) or the use of an inappropriate negation for the copula expressions in (18b).

- (18) (a) ó rúta ...- 's/he teaches...' á rúta - 'while s/he is teaching...'
 (b) ga ké morúti - 'I am not a teacher' ga sé morúti - 'S/he is not a teacher.'

A set of tonal rules applies to the underlying lexical and grammatical tones and leads to the eventual tonal pattern with which an utterance is pronounced. Learners' mistakes at this level will probably mainly result in the learner sounding foreign, and only in exceptional cases lead to misunderstanding. Tone deviances from the underlying lexical and grammatical tones occur mainly due to three – sometimes conflicting – tendencies in the grammar: the tendency of a high tone to spread onto the following syllable (19a), the tendency of a final syllable not to be targeted by this high tone spread (19b) and by the avoidance of more than two adjacent high tones (19c).

- (19) (a) Tone assimilation: Mo-sádfi ó-bóná... 'The woman sees...'
 (b) Extratonicity: Ò-bóná mo-sádfi. 'S/he sees a woman.'
 (c) Meeussen's Rule: mo-sádfi wámé 'my woman'

Once the teaching of tone is implemented into the curriculum of second-language teaching, further research will be necessary concerning the pedagogy of the teaching (cf. Nguyen & Macken, 2008 for Vietnamese).

6.2 Lengthening

The most important durational feature of the Southern Bantu languages is the length of the penultimate syllable of a clause. Vowel length is not contrastive in the Sotho-Tswana languages except for some few lexical items. A vowel (or sonorant) is predictably lengthened when it appears in the penultimate syllable of a sentence. It is then nearly twice as long in duration as in any other position (Tlale, 2005). This prominent length has often been equalled with stress (Doke, 1992). However, apart from duration this syllable does not show any other phonetic correlates of stress, such as raised amplitude or pitch.

Penultimate lengthening is a feature that is relatively easy to acquire (as compared, for example to tone) and which has huge benefits to the learner for making his/her speech sound more natural (Pretorius, p.c.). Work in speech technology has shown that the judged naturalness of an utterance increases significantly once penultimate lengthening has been implemented (Barnard, p.c.). However, there are at least two aspects that the student must be aware of. The first aspect relates to the fact that nasals can be syllabic in these languages (cf. 4.2). If a syllabic nasal occurs in the penultimate position, this nasal is lengthened, as shown in (20).

(20) Penultimate lengthening

mollo	[mó:..lɔ]	'fire'
mmê	[m̩:..mɛ]	'mother'
nna	[n̩:..ná]	'I'
e nngwe	[é:ŋ:..ŋwí]	'another'

Secondly, the lengthening rule always applies to the penultimate syllable counted from the clause boundary. It is thus not restricted to certain word classes (unlike stress in English), like nouns and verbs, but can occur on grammatical morphemes as well. Experience has shown that monosyllabic verbs often cause a problem, as the shift of lengthening is most prominent here (Pretorius, p.c.). Compare:

- | | | | | |
|------|-----|------|------------------------|-------------------------------|
| (21) | (a) | (i) | Ke ya tórópó.ng. | 'I go to town.' |
| | | (ii) | Ke a: ya. | |
| | (b) | (i) | Ke tlo já na:ma. | 'I will eat meat.' |
| | | (ii) | Ke tlo: já. | |
| | (c) | (i) | Ke sá nwá teye yá: mé. | 'I am still drinking my tea.' |
| | | (ii) | Ke sa: nwa. | |

Special attention should thus be paid to the acquisition of the rule of penultimate lengthening. However, it is relatively easy to get a feeling for this, and it is a very rewarding achievement. The mastery of this aspect of intonation is also of importance because the lack of penultimate lengthening actually carries semantic meaning in these languages as will be discussed in the next section.

6.3 Question prosody

Two different kinds of questions are commonly distinguished across the languages of the world: so-called *wh*-questions (22a +b) and yes/no-questions/ polar questions (22c).

(22) Question types

- | | | |
|-----|---------------------|------------------------|
| (a) | O nyaka eńg? | 'What do you want?' |
| | SC2 want what | |
| (b) | O fihl-ilé neńg? | 'When did you arrive?' |
| | SC2 arrive-PST when | |
| (c) | Na o na lé mphó? | 'Do you have a gift?' |
| | QP SC2 have gift | |

Wh-questions always contain a question word like *who*, *what*, *when*, *where*, *how* (which in English mostly start with *wh*, hence the name). With *wh*-questions, the speaker signals to the hearer that s/he needs a specific piece of information. Polar questions, on the other hand, put a whole proposition out for acceptance or rejection. These two types of questions differ in the markers they use. Whereas *wh*-questions always contain a question word, polar questions are either unmarked, or they bear the question marker *na/ a/ afa* in Northern Sotho (cf. Zerbian, 2006a).

These two question types are also differentiated intonationally in Southern Bantu languages. *Wh*-questions show normal declarative intonation (i.e. normal register, downdrift across the utterance, penultimate lengthening). This is comparable to the situation in English and German where *wh*-questions are also unmarked intonationally. Yes/no-questions, on the other hand, display specific question prosody. It is important to note that this question prosody is different from the rising intonation found in English. In a study on polar question intonation in Xhosa, Jones *et al.* (2001a) found polar question intonation comprises three characteristics acoustically: a short duration of the normally lengthened penultimate (cf. section 6.2), pitch on the penultimate vowel and the overall raised pitch of questions.

The high pitch marking by means of a raised register is common cross-linguistically for the indication of questions (Ohala, 1983), though not necessarily in the African languages (cf. Riialand, 2009). Perceptually, Jones *et al.* (2001b) showed that the pitch on the first syllable already provides a strong cue to interrogativity. However, the duration and pitch on the penultimate syllable can either reinforce or override perceptual judgments made earlier in the utterance.

For the L2 learner, two aspects are thus of importance: First, in yes/no-questions, the penultimate syllable of the utterance has to be pronounced short. Secondly, the learner should try to raise the overall pitch register when posing a yes/no-question. The greeting includes a yes/no-question (*O kae?*) and thus gives an excellent daily practice of the implementation of these features. My own experience has shown that misunderstanding will arise if these suprasegmental features are not implemented. As an additional device, question particles should be used to mark polar questions as such.

6.4 Focus intonation

In languages like English pitch accents are placed on words that reveal new information. Intonation thus serves to distinguish semantic differences in information status, as is shown in (23) (where caps mark the constituent that bears the pitch accent).

- (23) (a) Who ate chocolate?
PETER ate chocolate.
(b) What did Peter eat?
Peter ate CHOCOLATE.

It was long believed in literature that tone languages would not make a similar distinction by use of prosody, because changes in pitch were already used to differentiate lexical and grammatical meaning. However, work on Chinese (Xu, 1999) or the Bantu language Chichewa (Kanerva, 1990) has shown that focus can nevertheless be indicated prosodically. In Northern Sotho, however, it has been shown that prosody is not used to indicate that a constituent represents new information in discourse (Zerbian, 2006a). Instead, (morpho-)syntactic means are exploited. An example from everyday speech is given in (24).

- (24) Radio advertisement from a bank institute
(a) How can we help YOU?
(b) Re ka go thuša bjang, wena?

It will probably not lead to misunderstandings if the L2 learner applies the intonational focus grammar of English to the Sotho-Tswana languages. The language learner has to be aware, however, that it will not be possible to distinguish a sentence by intonation alone. S/he has to learn that morpho-syntactic means must be employed more intensely than in English. For a more detailed treatment of the use of morpho-syntactic means for focus marking in Northern Sotho see Zerbian (2006b).

7. Conclusion

The article presented a contrastive analysis of the Sotho-Tswana languages with Germanic languages like English, Afrikaans or German with the goal to predict possible learner errors in the L2 acquisition of any of the Sotho-Tswana languages. It provided linguistic background knowledge to the areas of differences in order to facilitate understanding of potential learners'

problems. The article argued specifically for the inclusion of two phonologically relevant linguistic features into the curriculum: the indication of the phonemic differences in vowel quality with the mid-vowels and the inclusion of tone.

REFERENCES

- African Voices. 2004. *Everyday phrases in Tswana*. Audio-CD.
- Ashby, M. & Maidment, J. 2005. *Introducing phonetic science*. Cambridge: Cambridge University Press.
- Buell, L. & Sosibo, Z. in prep. Fundamentals of Zulu tone. Unpublished manuscript.
- Catford, J.C. 1988. *A practical introduction to phonetics*. Oxford: Oxford University Press.
- Chebanne, A.M., Creissels, D. & Nkhwa, H.W. 1997. *Tonal morphology of the Setswana verb*. Munich: Lincom Europe.
- Chebanne, A., Mokitimi, M., Matlosa, L. Nakin, R., Wakumelo Nkolola, Mokgoatšana, & Machobane, M. 2003. *The standard unified orthography for Sotho/Tswana languages*. Cape Town: CASAS.
- Coetzee, A.W., Lin, S. & Pretorius, R. 2007. Post-nasal devoicing in Tswana. In J. Trouvain & W.J. Barry (eds) *ICPhS, XVI*: 861-864.
- Cole, D.T. 1955. *An Introduction to Tswana grammar*. Cape Town: Longman.
- Cole, D.T. & Mokaila, D.M. 1962. *A course in Tswana*. Washington: Georgetown University Press.
- Dickens, P. 1984. The history of so-called strengthening in Tswana. In: *Journal of African Languages and Linguistics*, 6: 97-125.
- Doke, C.M. 1927/1992. *Textbook of Zulu grammar*. Cape Town: Maskew Miller Longman.
- Endemann, K. 1911. *Wörterbuch der Sotho-Sprache*. Hamburg: L. Friedrichsen & Co.
- Fries, C. 1954. *Teaching and learning English as a foreign language*. Ann Arbor, MI: University of Michigan Press.
- Gut, U. & Trouvain, J. (eds.) 2007. *Non-native prosody*. Berlin: Mouton.
- Gut, U., Trouvain, J. & Barry, W.J. 2007. Bridging research on phonetic descriptions with knowledge from teaching practice – The case of prosody in non-native speech. In: Gut, U. & Trouvain, J. (eds) *Non-Native Prosody*. Berlin: Mouton: 3-21.
- Jones, J., Louw, J. & Roux, J. 2001a. Queclaratives in Xhosa: An acoustic analysis. *SAJL supplement*, 36: 3-17.
- Jones, J., Louw, J. & Roux, J. 2001b. Perceptual experiments on queclaratives in Xhosa. *SAJL supplement*, 36: 19-31.
- Kanerva, J.M. 1990. *Focus and phrasing in Chichewa phonology*. New York: Garland Publishing.
- Kgasa, M.L.A. & Tsonope, J. 1995. *Thanodi ya Setswana*. Gaborone: Long man Botswana.
- Khabanyane, K.E. 1991. The five phonemic vowel heights of Southern Sotho: an acoustic and phonological analysis. *Working Papers of the Cornell Phonetics Laboratory*, 5: 1-35.
- Khoali, B.T. 1991. A Sesotho tonal grammar. Unpublished PhD thesis, University of Illinois at Urbana-Champaign.
- Kotzé, A.E. & Zerbian, S. 2008. On the trigger for palatalization in the Sotho languages. *Journal of African Languages and Linguistics*, 28(1): 1-28.
- Kriel, T.J. & van Wyk, E.B. 1966. *Pukuntšu – Woordeboek (Noord-Sotho – Afrikaans, Afrikaans – Noord-Sotho)*. Pretoria: Van Schaik.

- Ladefoged, P. & Maddieson, I. 1994. *The sounds of the world's languages*. Malden: Blackwell.
- Lombard, D.P. 1976. Aspekte van toon in Noord-Sotho. Unpublished PhD thesis, University of South Africa.
- Lombard, D.P. 1985. *An introduction to the grammar of Northern Sotho*. Pretoria: Van Schaik.
- Maddieson, I. 2008. Lateral consonants. In: Haspelmath, M. & Dryer, M.S. & Gil, D. & Comrie, B. (eds) *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library, chapter 8. <http://wals.info/feature/8>. Accessed on 04/12/2008.
- Mascher, D. in prep. A five-year course in Tswana second language. Manuscript, Vryburg, South Africa.
- Mbatha, M.O. 2006. *Isichazamazwi sesiZulu*. Pietermaritzburg: New Dawn Publishers.
- Mojapelo, M.L. & Jordaan, K. 2004. Multilingualising South Africa: Teaching basic communication skills in an African language. Paper presented at 34th CALL, Leiden, Netherlands.
- Nguyen, H. & Macken, M.A. 2008. Factors affecting the production of Vietnamese tones – a study of American learners. *Studies in Second-Language Acquisition*, 30: 49-77.
- Odden, D. 2005. *Introducing phonology*. Cambridge: Cambridge University Press.
- Ohala, J.J. 1983. Cross-language use of pitch: An ethnological view. *Phonetica*, 40: 1-18.
- Pukuntšu ya Polelopedi ya Sekolo – *Sesotho sa Leboa le Seisimane* (Oxford Bilingual School Dictionary – Northern Sotho and English) 2008. Cape Town: Oxford University Press.
- Rialland, A. 2009. African 'lax' question prosody: its realization and its geographical distribution. *Lingua* 119(6): 928-949.
- Roux, J. C. 2001. Xhosa: A tone or a pitch-accent language? *SAJL Supplement* 36: 33-50.
- Sandilands, A. 1953. *Introduction to Tswana*. Tigerkloof, South Africa: The London Missionary Society.
- Saville-Troike, M. 2006. *Introducing second-language acquisition*. Cambridge: Cambridge University Press.
- Setswana Terminology and Orthography No. 4*. 1988. Department of Education and Training and Department of Education: Bophuthatswana.
- Snyman, J.W., Shole, J.S. & Le Roux, J.C. 1990. *Dikišinare ya Setswana – English – Afrikaans*. Pretoria: Via Afrika
- Thanodi ya Setswana* 2008. Nutrend Publishers: Pietermaritzburg.
- Tlale, O. 2005. The phonetics and phonology of Sengwato, a dialect of Setswana. Unpublished PhD thesis, Georgetown University.
- Tomatis, A. 1991. *Nous Sommes Toutes Nés Polyglottes*. Paris.
- Wissing, D.P. 2005. Aspiration of English voiceless stop consonants in Southern Sotho: a case study. *SAJAL*, 3: 189-205.
- Xu, Y. 1999. Effects of tone and focus on the formation and alignment of contours. *Journal of Phonetics*, 27: 55-105.
- Zerbian, S. 2006a. Questions in Northern Sotho. *Linguistische Berichte* 208: 385-405.
- Zerbian, S. 2006b. Expression of information structure in Northern Sotho. *ZASPiL* 45. Berlin: ZAS.
- Ziervogel, D. & Mokgokong, P.C. 1961. *Klein Noord-Sotho-Woordeboek*. Pretoria: Van Schaik.
- Ziervogel, D., Lombard, D.P. & Mokgokong, P.C. 1969. *A handbook of the Northern Sotho language*. Pretoria: Van Schaik.
- Ziervogel, D. & Mokgokong, P.C. 1975. *Pukuntšu ye kgolo ya Sesotho sa Leboa – Groot Noord-Sotho-Woordeboek – Comprehensive Northern Sotho Dictionary*. Pretoria: Van Schaik.