

## Issues in the acquisition of the Sesotho tonal system\*

KATHERINE DEMUTH

*Brown University*

(Received 2 January 1991. Revised 15 May 1992)

### ABSTRACT

This paper examines the acquisition of the grammatical tone system of Sesotho, a southern Bantu language where tone sandhi is rich, and where surface and underlying representations are often quite distinct. Results of the longitudinal case study show that rule-assigned tone on subject markers is generally marked appropriately by age two. In contrast, underlying tonal representations on verb roots are learned gradually over time, showing an early Default High tone pattern. The study also finds that, while some tone sandhi rules are in the process of being acquired between 2;6 and 3;0, problems in the mapping between tonal representations and segments persist. The paper raises methodological and theoretical issues not only for the acquisition of tonal systems, but for the acquisition of phonology in general.

### INTRODUCTION

The development of autosegmental phonology (Williams, 1971; Leben, 1973; Goldsmith, 1976) represents one of the most important advances in phonological theory since the generative insights of *The Sound Pattern of English* (SPE) (Chomsky & Halle, 1968). In particular, it has contributed greatly to the understanding of non-linear problems which had previously eluded traditional segmental analysis. Indeed, autosegmental approaches to phonology developed, in part, from attempts to capture the systematic yet apparently complex grammatical tone systems which are characteristic of many African languages. More recent developments in metrical phonology

[\*] Data collection for this work was supported by Fulbright-Hayes and Social Science Research Council (SSRC) grants. Data transcription and analysis has been supported by NSF Grant No. BNS08709938 and by the Social Science Research Council (SSRC). I thank Steven Bird, Tucker Childs, Nick Clements, Julie Croston, Grzegorz Dogil, Elan Dresher, Charles Ferguson, Larry Hyman, Ben Khoali, Thabo Khojane, Chuck Kisseberth, Mark Johnson, Will Leben, Malillo Machobane, Sheila Mmusi, K. P. Mohanan, Jim Morgan, Mpiko Ntseki, Glyne Piggott, Frances W. Pritchett, Mpatletseng Ramaema, as well as Katharine Perera, anonymous reviewers, and several audiences for stimulating discussion, comments and assistance relating to this research. The final interpretations are, of course, my own. Address for correspondence: Katherine Demuth, Department of Cognitive and Linguistic Sciences, Brown University, Box 1978, Providence, RI 02912 USA. Katherine\_Demuth@brown.edu

(Lieberman & Prince, 1977; Hayes, 1982), lexical phonology (Mohan, 1982, 1986; Kiparsky, 1982, 1985; Pulleyblank, 1986) and prosodic phonology (Selkirk, 1984; Nespor & Vogel, 1986) continue to provide powerful tools for researching both phonological and prosodic phenomena.

The field of acquisition has been slow to adopt and integrate new perspectives from theoretical phonology, much the same as it has been slow to adopt and apply theoretical insights in the area of syntax. Much work on the acquisition of phonology has utilized a segmental approach based primarily on insights from structural linguistics (Jakobson, 1941/1968) or SPE (Smith, 1973), though early work influenced by the Firthian School of Prosodic Phonology (e.g. Waterson, 1971), and more recent work by researchers such as Kiparsky & Menn (1977), Spencer (1986), Waterson (1987) and others signal the beginning of a new era in the study of phonological aspects of acquisition. Much of the research conducted on the acquisition of tone took place before the full development of autosegmental phonology and focused on lexical tone languages such as Thai (Tuaycharoen, 1977), Mandarin (Chao, 1973; Clumeck, 1977; Li & Thompson, 1977) and Cantonese (Tse, 1978). Since the late 1970s there has been little work on the acquisition of tone, and no comprehensive study of the acquisition of a complex grammatical tone system like that found in many African languages. More recently several researchers are beginning to address the acquisition of tone in African languages in a more systematic fashion (e.g. Moto, 1988; Chimombo & Mtenje, 1989; Demuth, 1989, 1991; Suzman, 1991), raising questions that have implications for both acquisition and linguistic theory.

The present study develops an autosegmental account of the acquisition of grammatical tone in Sesotho, a southern Bantu language spoken by approximately four million people in the countries of Lesotho and South Africa. After outlining the Prosodic Acquisition Problem and providing a brief introduction to the Sesotho tonal system, the paper focuses on the acquisition of Sesotho tone at underlying and lexical levels of phonology. The paper concludes with a review of the empirical findings and their import for both acquisition and linguistic theory.

#### *The prosodic acquisition problem*

Although the acquisition of prosodic phenomena such as stress, intonation, and tone is in many respects similar to the acquisition of segments (both have discrete units, and both participate in phonological rules), prosodic phenomena differ in that they must be mapped onto segments. Although this problem is relatively trivial in lexical tone languages like Mandarin, where underlying and surface tones are often the same, the problem is more serious in grammatical tone systems like that found in many Bantu languages, where abundant tone sandhi, or permutation of underlying tones, provides a serious challenge for the language learner.

In order to address the Prosodic Acquisition Problem we need to have a model or theory of what prosodic systems, or in this case tonal systems, look like. In addition, that model or theory must be able to account for all the tonal phenomena we find in human language. It is only in this way that we can address the parametric differences found in different tonal systems and discover how children determine which type of system they are learning. Although the study of tonal systems is still developing, we adopt, for the present purposes, a model of lexical phonology (Mohanam, 1982, 1986; Kiparsky, 1982, 1985; Pulleyblank, 1986) as a means for identifying important differences among tonal systems, and as a means for making specific hypotheses about the tonal acquisition process. An adapted model of lexical phonology is presented in Fig. 1.

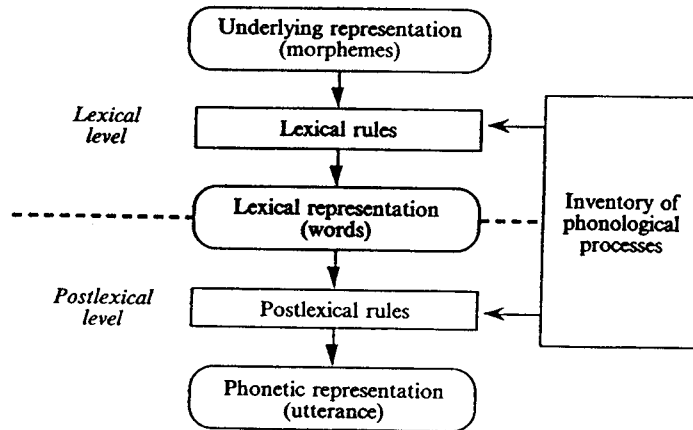


Fig. 1. Model of lexical phonology.

The specific attraction of a model of lexical phonology is that it identifies the different levels at which phonological processes take place. Critically, those processes can be the same at different levels, and can apply more than once (i.e. cyclically) in (at least) the lexical component of the grammar. Thus, tone can be assigned (1) underlyingly, in the lexicon, (2) at the word, or lexical level, where morphemes are combined, or (3) at the postlexical level, where well-formed words enter into the syntax. For English, word stress is assigned at the lexical level, and sentence (phrasal) stress and subsequent pitch contours are assigned at the postlexical level. This contrasts with lexical tone languages such as Mandarin Chinese where pitch is assigned at the underlying level, and is then subject to certain rules of tone sandhi (or tonal rules) at subsequent levels of the phonology. In many Bantu languages tone is also assigned to lexical items underlyingly, but then several tone sandhi rules may apply, some triggered by grammatical phenomena such as different tense/aspect/mood constructions, before postlexical phrasal tone rules apply.

Thus, the model of lexical phonology allows us to identify (at least) three levels of phonology where tone may be assigned or modified in a given language.

Crucial to a discussion of tonal systems is a distinction between the language specific 'grammatical' uses of pitch contrasts, and aspects of pitch realization that are more 'universal'. All languages have certain 'universal' intonational properties, for example, a tendency for declination throughout the utterance and a prosodic distinction between different types of speech acts: e.g. declaratives will normally be distinguished intonationally from interrogatives. In the following discussion we maintain a distinction between the discourse 'universals' of intonation, and the language specific uses of tone, focusing the bulk of our discussion on the latter.

From what we know of the acquisition of intonation in English, it appears that some discourse and pragmatic aspects of the system are acquired early. On the other hand, other aspects of postlexical pitch assignment are not fully acquired until around 12 years (Crystal, 1986: 191). Studies of lexical tone languages like Mandarin, Cantonese and Thai uniformly report that distinctions in pitch become recognizably phonemic about the same time as, or before segments, around 1;11-2;2 years (Clumeck, 1980; Crystal, 1986). Rules of tone sandhi, however, are acquired perhaps as late as five years (Li & Thompson, 1977). Preliminary reports on the acquisition of tone in Bantu languages indicate that some aspects of lexical tone and tonal melodies may be acquired by two years (Sesotho - Demuth, 1989, 1991; Chichewa - Chimombo & Mtenje, 1989; Zulu - Suzman, 1991), but we know little about the acquisition of lexical tone rules, postlexical tone rules, or aspects of intonation. Bantu languages, many of which exhibit pitch assignment at all levels of phonology, represent some of the most complex, and therefore some of the most interesting cases for determining how tonal systems are acquired. We turn now to a discussion of the Sesotho tonal system.

#### THE SESOTHO TONAL SYSTEM

There have been several early descriptive studies of the Sesotho tonal system (Letele, 1955; Köhler, 1956; Kunene, 1961, 1972; Tucker, 1969; see also Doke & Mofokeng, 1957). Autosegmental treatments of the Sesotho tonal system focus on the nominal system (Khoali, 1991) and the verbal system (Clements, 1988; Kisseberth, 1989; Khoali, 1991); we restrict our discussion to the latter (cf. also work on closely related Setswana (Mmusi, 1991)).

Sesotho can be described as a grammatical tone language which may be in transition toward evolving a more restricted tonal system, i.e. a system where not every syllable, morpheme, or word has to be encoded for tone in the lexicon. In other words, although it is necessary to posit High (H) tone underlyingly, it is not necessary to posit Low (L) tone underlyingly. This means that verb roots, for instance, can be specified underlyingly as having

either H or Ø tone. Those syllables, or Tone Bearing Units (TBUs) that end up with no tone specification at the surface are generally filled in with a rule of Default Low Insertion. In Sesotho, approximately half of the verb roots fall into the H-tone class; recent verb borrowings (loan words) are also assigned H tone. The major issue addressed in this study is how children determine the underlying tone of verbs.

Sesotho is a pro-drop language with a basic word order of (S)V(O). The verbal complex can be schematically expanded as in (1) and (2) below.<sup>1</sup>

(1) (S) SM-(T/A)-(OM)-V-(ext)-M (O)

(2) Thabo ó-tlá-mo-rék-él-a dijó  
 1T. 1SM-FUT-1OM-buy-BEN-IN 8food  
 'Thabo will buy him/her food'

Although the verb *ho-réka* 'to buy' is an H-toned verb, and surfaces as such in (2), there is not always a one-to-one mapping between surface structure and underlying tonal representations. For instance, if *ho-réka* 'to buy' were used with a 3rd person subject marker, which is also H-toned, the first syllable of the verb would lose its high tone: *bá-reká dijó* 'they are buying food'. Likewise, the Ø-toned verb *ho-batla* 'to want' can surface with an H tone on the first syllable of the verb if it is used with an H-toned 3rd person subject marker: *bá-bátla dijó* 'they want food'. In other words, subject markers may influence the surface realization of tone on the following verb root. The problem for the child is to figure out, given these variable surface tone realizations, what the underlying lexical tone of a verb root may be.

Although Sesotho makes use of several basic grammatical tonal melodies, dependent on the tense/aspect/mood of the construction, we restrict the present discussion to the examination of the affirmative present and futures *-tla-* and *-ea-*, focusing on tonal phenomena that apply at underlying and lexical levels of the phonology. After a brief discussion of lexical tone assignment to verb roots, we discuss rule-assigned tone on subject markers. We then illustrate the rules of High Tone Doubling (HTD) and Obligatory Contour Principle (OCP) Effects. An outline of the processes to be examined is listed in Table 1.

[1] Glosses are as follows: BEN benefactive, CONJ conjunction, DEM demonstrative pronoun, FUT future tense, ext verbal extensions, IN indicative, LOC locative, M mood, O lexical object, OM object marker, PASS passive, PERF perfective aspect, PN independent pronoun, POSS possessive, PREP preposition, PRES present tense, S lexical subject, SM subject marker, T/A tense/aspect, V verb root, ' high tone, + mid tone (low tone unmarked). Numbers indicate the noun class to which different nouns belong (e.g. *motho* 'person' (and other singular human nouns) = noun class no. 1, *batho* 'people' (and other plural human nouns) = noun class no. 2, *dijo* 'food' = noun class no. 8, etc.). First and second person singular/plural SMs and OMs are therefore marked as 1s/p and 2s/p respectively. A modified version of Lesotho orthography has been used.

TABLE I. *Subset of Sesotho tonal processes*

A. Underlying level (lexically assigned tone)
Verb roots
B. Lexical level (rule assigned tone)
(i) Subject markers
(ii) High tone doubling (HTD)
(a) Verb roots
(b) Subject markers
(iii) Obligatory contour principle (OCP) effects and tier conflation
(a) High tone delinking from subject markers
(b) High tone delinking on verb roots

*Underlying level*

*Lexical tone assignment to verb roots.* As mentioned above, tone on Sesotho verb roots is assigned at the underlying level. Tone is then predictably associated with the first syllable of the verb root at the lexical level. Pitch contours are subsequently realized by application of the different tone rules, depending on the tonal melody of the construction (i.e. its tense/aspect/mood). Although some Bantu languages make a distinction between H- and L-toned verb roots underlyingly, in Sesotho it is not necessary to specify H versus L, but only H versus  $\emptyset$ . Syllables (TBUs) left unspecified for tone at the surface are filled in with a late postlexical rule of Default Low Insertion. This is illustrated in Fig. 2, where perpendicular lines (|) = initial

H-toned roots	$\emptyset$ -toned roots	
ho-bona	ho-batla	
H		Underlying representation
ho-bona	ho-batla	
		Lexical tone association
H		
ho-bona	ho-batla	
/		High tone doubling (HTD)
H		
ho-bona	ho-batla	
/		Default low insertion
L H	L L L	
ho-bona	ho-batla	
* \		Phrasal final lowering
L H L%	L L L	
ho-bóna+	ho-batla	
'to see'	'to want'	

Fig. 2. Lexical tone assignment to verb roots.

ACQUISITION OF THE SESOTHO TONAL SYSTEM

tone associations, slant lines (/) = tones that arise through spreading or late association, and (%) = a phrasal boundary.

The examples in the derivation in Fig. 2 are relatively transparent as to their underlying tonal specification. However, as will be seen in the following section, most verbs undergo tone sandhi, or the permutation of tone, resulting in multiple surface tone patterns for a given verb root. Children learning grammatical tone languages must therefore abstract away from these surface forms in order to posit the correct underlying tone of a particular verb root. We might expect this type of lexical acquisition to be a difficult process, taking place gradually over a long period of time. In contrast, we might expect the acquisition of tone in rule-governed domains to be acquired more easily, and perhaps earlier. We turn now to a discussion of these tonal rules.

*Lexical tone rules*

*Subject markers.* The tone of subject markers (SMs) is determined by the person and by the tense/aspect/mood of the construction. In the present affirmative, 1st and 2nd person SMs take  $\emptyset$  tone, and 3rd person is marked for H. This is shown in Fig. 3.

H-toned SMs o-batla ...	$\emptyset$ -toned SMs ke-batla ...	Underlying representation
o-batla ...   H	ke-batla ...	Lexical tone association
o-batla ...  / H	ke-batla ...	High tone doubling (HTD)
o-batla ...  /   H L	ke-batla ...       L L L	Default low insertion
ó-bátla ... 'S/he wants X'	ke-batla ... 'I want X'	

Fig. 3. Rule-assigned tone on subject markers.

It is possible that the acquisition of tone on subject markers would parallel that found for verb roots. On the other hand, subject markers represent a closed class, and the application of tone is ruled governed. Furthermore, the inherent tone of subject markers is generally identical to that realized on the surface. Therefore, we might predict that children would acquire the correct

tone for subject markers earlier and more easily than that on verb roots. But what about the acquisition of tone sandhi rules? We turn now to a discussion of High Tone Doubling.

*High tone doubling (HTD) on verb roots.* Sesotho has a rule of High Tone Doubling (HTD), where an H tone associated with the first syllable of an H-toned verb root will double, or spread to the next syllable. This process is illustrated in Fig. 4.

ke-rekela ... H	Underlying representation
ke-rekela ...   H	Lexical tone association
ke-rekela ...  / H	High tone doubling (HTD)
ke-rekela ...    / L H L	Default low insertion
ke-rékéla ... 'I'm buying X for Y'	

Fig. 4. High tone doubling (HTD) on verb roots.

The HTD rule (i.e. spreading of an H tone only to the adjacent syllable) is distinct from the rule of Iterative High Tone Spread (i.e. spreading of an H tone to the end of the word) – a rule that applies in the tonal melody that includes the perfective (e.g. *ke rékélé...* 'I bought X'). One of the acquisition problems would be to determine that these are two rules that apply in different contexts. We might expect that children would initially collapse these two rules into one, being unaware of the different tense/aspect/mood domains to which they apply.

*High tone doubling (HTD) from subject markers.* Inherent in the theory of lexical phonology is the possibility for cyclic application of rules: after applying to a particular domain, a certain phonological rule can apply again to a different or larger domain. We have just discussed the application of HTD to the domain of the verb root. Once the subject marker and tense/aspect marker have been added to the verb root, the rule of HTD has another chance to apply. In this case the H tone on the subject marker spreads to only the following syllable, either the tense/aspect marker (T/A), if one is present, or onto the verb root itself (V). This is shown in Fig. 5.

ACQUISITION OF THE SESOTHO TONAL SYSTEM

HTD to verb	HTD to tense/aspect	
o-lemma ...	o-a-lemma	Underlying representation
o-lemma ...	o-a-lemma	
l	l	Tone assignment
H	H	
o-lemma ...	o-a-lemma	
l/	l/	High tone doubling (HTD)
H	H	
o-lemma ...	o-a-lemma	
l/ l	l/ l l	Default low insertion
H L	H L L	
ó-léma ...	ó-á-lemma	
'S/he's ploughing X'	'S/he's ploughing'	

Fig. 5. High tone doubling (HTD) from subject markers.

The acquisition of HTD from subject markers is of particular interest as it holds the key to understanding children's development of underlying tonal representations on verbs. Specifically, spreading of the H tone from the subject marker onto the verb root should ONLY take place if the verb root is toneless (i.e. has  $\emptyset$  tone); if the verb root is L, spreading should not apply (i.e. the first syllable of the verb will already bear a tone). We turn now to a consideration of another class of tonal rules, those that are generally characterized as Obligatory Contour (OCP) effects.

*The obligatory contour principle (OCP)*

In the phonology of many languages, identical phonological entities such as vowels, tones, and prominently stressed syllables, are prohibited from occurring next to each other. Languages deal with this problem in different ways, in some cases 'fusing' two like elements to yield only one, in other cases modifying one such that two like elements are no longer adjacent. With regard to tone, the restriction is often one that prohibits two H tones from being adjacent on the tonal tier. Sesotho employs two rules of High Tone Delinking (HT delinking) to avoid having an HH representation on the tonal tier. HT delinking from subject markers reverses the effect of the HTD rule, i.e. the subject marker retains its H tone, but the tense/aspect marker loses the H it gained through doubling. The tense/aspect marker is left unspecified for tone, and the sequence surfaces as HLH. This is illustrated in Fig. 6.<sup>2</sup>

[2] Note that this allows HTD to apply and then undoes its effect with the use of a delinking rule. This is, in effect, a 'repair' strategy, a solution some phonologists would prefer to avoid by simply not letting the rule apply in the first place (i.e. adopting a blocking rule

## CHILD LANGUAGE

ba-a-bona	
H	Underlying representation
ba-a-bona	
l l	Lexical tone association (verb root)
H H	Tone assignment (subject marker)
ba-a-bona	
l/ l/	High tone doubling (verb root)
H H	High tone doubling (subject marker)
ba-a-bona	
l* l/	High tone delinking (OCP effect)
H H	
ba-a-bona	
l l l/	Default low insertion
H L H	
ba-a-bona	
l l l* \	Phrasal lowering
H L H L%	
bá-a-bóna+	
'They see/understand'	

Fig. 6. High tone delinking from subject markers (OCP effect).

HT delinking on the verb root applies in a somewhat different domain: when an H-toned verb root is adjacent to an H-toned subject marker, it is the underlying tone on the VERB ROOT that gets delinked, again producing an HLH surface pattern. This is shown in Fig. 7.

In illustrating the workings of the Sesotho tonal system we have shown that there is not necessarily a one-to-one correspondence between underlying tone and its surface realization. This raises many questions about how such complex grammatical tone systems are learned. For example, how do children determine underlying tonal representations? Does the learning of underlying representations depend on learning some of the tone sandhi rules of the language? Is this process only mastered by the age of five, as suggested by Li & Thompson (1977) in the case of Mandarin, or are aspects of it acquired earlier? Are there any special 'learning strategies' that children employ in learning such a system? The following case study examines each of these questions.

instead). However, as will be seen below, Sesotho still needs a delinking rule to account for other OCP effects. Furthermore, the acquisition evidence indicates that the child may be applying HTD before OCP effects are acquired. Thus, it would appear that the acquisition scenario is indeed one of 'apply and repair', rather than 'block'.

ACQUISITION OF THE SESOTHO TONAL SYSTEM

ba-bona ...	
H	Underlying representation
ba-bona ...	
l l	Lexical tone association (verb root)
H H	Tone assignment (subject marker)
ba-bona ...	
l l/	High tone doubling (verb)
H H	
ba-bona ...	
l #/	High tone delinking (OCP effect)
H H	
ba-bona ...	
l l l	Default low insertion
H L H	
bá-boná ...	
'They see X'	

Fig. 7. High tone delinking on verb roots (OCP effect).

THE CASE STUDY

*The subject*

The data for this study come from a monolingual Sesotho-speaking boy named Hlobohang, the only child of a rural mountain family in Lesotho. A five-year-old male cousin, mother and grandmother were living in the household during the 12-month course of the study. A profile of Hlobohang's and other Sesotho-speaking children's morphological and syntactic development are reviewed in Demuth (1992*b*).

*The procedure*

Audiorecordings were conducted during spontaneous, naturalistic interactions between Hlobohang and his older cousin, mother, grandmother, and peers. Hlobohang was taped for three to four hours at five-week intervals over a period of 12 months. The data consulted for this study include sessions when Hlobohang was aged 2;1, 2;6 and 3;0. The data were drawn from the utterances that included a full verb phrase. The total number of utterances consulted for each session was 243, 496 and 582, respectively. Though the actual number of utterances per session increased over time, the number of utterances employing only simple present and future actually decreased, other tenses and H-toned object markers being used in an increasing number of Hlobohang's utterances.

Audiorecordings were made with a Superscope directional microphone and a Superscope/Marantz cassette recorder. The child's utterances, which