

Representing morphological tone in a computational grammar of Hausa

Berthold Crysmann

CNRS, Laboratoire de linguistique formelle (UMR 7110), Paris

ABSTRACT

In this paper¹ I shall discuss the representation of morphological tone in Hausa, as implemented in a computational grammar of the language, referred to as HAG, which has been developed within the framework of Head-driven Phrase Structure Grammar. Based on an in-depth study of segmental and suprasegmental properties manipulated by morphological processes, I shall argue that two fundamental insights from autosegmental phonology need to be seamlessly integrated into typed feature structure grammars of languages with grammatical tone, namely (i) the systematic separation of tonal and metrical information from the string of consonants and vowels, and (ii) the possibility of tonal spreading, i.e. the possibility for a tonal specification to be

Keywords:
Hausa, HPSG,
Autosegmental
Phonology, tone,
implementation

¹ A previous version of this work was presented at the Workshop on “Disentangling the inflectional role of tone”, organised by Jean-Léo Léonard (U Paris Sorbonne & Institut Universitaire de France) and Enrique Palancár (CNRS, SeDyl-CELIA) at the Maison de Recherche, Paris 3 (Sorbonne Nouvelle) in June 2013, as well as at a research seminar on tone and intonation at the Institut für Linguistik, Phonetik, Universität zu Köln. I would therefore like to thank the respective audiences for their helpful and stimulating comments and, in particular, Larry Hyman, Jean-Léo Léonard, Enrique Palancár, Martine Grice, Phuong Ha, and Doris Mücke. A great many thanks also go to the three anonymous reviewers whose comments and suggestions have been invaluable in improving this paper. Finally, I would also like to express my gratitude to Carmela Chateau for her careful proofing and copy-editing.

Work on the development of the grammar is partially supported by the French National Research Agency (ANR) as part of the “Investissements d’Avenir” programme (reference: ANR-10-LABX-0083), project ResHau [re:’zo:] of the Laboratoire d’excellence “Empirical Foundations of Linguistics”.

assigned to an arbitrary number of adjacent tone-bearing units (syllables). To this end, I present a formalisation of tonal melodies in terms of typed list constraints that implement a notion of tonal spreading, allowing for an underspecified description of tonal melodies, independent of the number of tone-bearing units. I shall finally show that this minimal encoding is sufficient, and flexible enough to capture the range of suprasegmental phenomena in Hausa.

1

INTRODUCTION

One of the fundamental motivations for the implementation of grammars in linguistically motivated formalisms is to allow for rigorous testing of the empirical predictions of linguistic theories: given the complex interactions of rich lexica with highly general principles and rules, paired with increasing theoretical coverage of phenomena way past the core into the periphery, manual evaluation of the consequences of a theory is becoming less and less feasible. By running an implemented grammar over a corpus, it is possible not only to detect limitations in coverage, but also to assess the degree of over-generation. This latter aspect is particularly easy to ensure in the context of reversible grammars, i.e. declarative knowledge sources that can be used for both parsing and generation. Despite the widespread use of tone to mark lexical and grammatical distinctions among the languages of the world, none of the grammars implemented to date actually incorporates a treatment of suprasegmentals into the grammar proper. This is due to the fact that, in the sample of languages for which implemented grammars exist, very few are actually tone languages, and the ones that are, like Chinese, only have lexical, not grammatical, tone.

Within theoretical linguistics, by contrast, the study of tone has always enjoyed a more central role, at least over the past 45 years. Investigation into African tone languages (Goldsmith 1976; Leben 1973) has provided the main evidence for the development of multi-tiered approaches to phonology, culminating in the development of autosegmental metrical phonology, subsequently to be adopted for the treatment of supra-laryngeal phenomena as well (Clements 1985). Due to the transformational heritage (Chomsky and Halle 1968), however, some of the descriptive devices that are used to extract generalisations

about tonal phenomena, especially destructive operations or procedural notions such as pre-linking, de-linking and re-spreading, do not lend themselves naturally to direct integration into the kind of formalisms used in the development of linguistically motivated computational grammars, which are generally committed to monotonicity and declarativity.² As established by work on Declarative or One-Level Phonology (Bird and Klein 1994; Scobbie 1993), typed feature structures, as used in HPSG (Pollard and Sag 1987, 1994), provide a convenient and expressive representation for capturing multi-tiered phonological descriptions. However, no sizeable grammar fragment of a tone language has so far been developed within a declarative phonological framework, let alone in the context of implemented computational grammars.

In this paper, I seek to close this gap between grammar engineering and theoretical autosegmental description. In particular, based on the example of Hausa, a language featuring both lexical and grammatical tone, I shall argue that the adoption of an autosegmental approach, i.e. one that separates the representation of tone and vowel length from that of consonantal and vocalic segments, is not only preferable for theoretical reasons, but actually inescapable from a grammar-engineering perspective. Furthermore, I shall show how the intimate connection between suprasegmental phenomena and morphosyntax can be accommodated in TDL (= Type Description Language; Copestake 2002; Krieger 1996), the purely conjunctive and monotonic typed feature formalism underlying HAG. Besides serving the practical purpose of tight integration with morphology and syntax in the implemented grammar, the choice of a lean formalism enables us to explore whether the general formalism is expressive enough to capture the relevant generalisations. More specifically, I shall argue that when suprasegmental operations are aligned with morphological rules, the available mechanisms, though highly restricted by themselves, will nevertheless prove to be sufficient to capture the entire set of surface-true and surface-apparent generalisations on Hausa suprasegmental

² Frameworks differ, of course, as to the extent of uniformity they assume across different linguistic sub-theories: while HPSG defends the hypothesis that all levels of linguistic knowledge (phonological, morphological, syntactic, and semantic) should be expressible in the same formalism, projection-architectures like LFG do not subscribe to this assumption.

phonology. The restrictive nature of our formalisation, which only ever permits one spreading tone per morphological domain, will prove to have interesting theoretical consequences: first, I resolve the debate regarding the direction of tone assignment between right-to-left (Newman 2000) and left-to-right (Leben 1978) in favour of outside-in, with a predominance, in the case of Hausa, of assignment from the right edge. Second, I shall show that all cases with more than one spreading tone in Hausa involve total reduplication, and argue that the proper treatment of this phenomenon independently requires morphological compounding, giving rise to two independent domains, each with a spreading tone of its own.

The paper is organised as follows: in Section 2, I shall give a general overview of Hausa tone, followed by a detailed study of morphological tone in Section 3, capitalising on holistic assignment of tone melodies vs. agglutinative tone. Section 4 provides an overview of HAG, an implemented HPSG of Hausa that crucially integrates suprasegmental information with morphological and syntactosemantic analysis. In Section 5, I shall present the autosegmental representation as implemented in HAG, capitalising on tonal spreading, as well as morphological operations on tone and length, including holistic assignment of tonal melodies and local adjustments of tone and length specifications. A major part of the discussion in this section will be concerned with the integration of prespecified prefixal tone, and its integration with right-to-left spreading, using conjunctive tone list constraints (Section 5.4.2), while Section 5.5 closes with a general discussion regarding the expressive power of the current approach.

2 SUPRASEGMENTAL DISTINCTIONS IN HAUSA

On the suprasegmental level, both tone and vowel length are distinctive. Hausa recognises two level tones, H(igh) and L(ow), as well as a falling contour tone, which is typically analysed as an HL sequence associated with a single (heavy) syllable. Rising contours observable at the surface are the result of interaction between lexical or grammatical tone and intonation (Inkelas and Leben 1990).³ Throughout

³LH sequences associated with a single syllable, e.g. as a result of affixation, undergo obligatory tone simplification (Newman 2000).

this paper, high tone (H) is marked with an acute accent, low tone (L) with a grave, and falling tone (HL) with a circumflex. On the metrical side, Hausa distinguishes between long and short vowels. Long vowels are marked with a macron, whereas vowels unmarked for length are short. Redundant marking of both H and L was chosen to improve readability for scholars from different traditions: in fact, redundant marking of H deviates from common Hausaist practice (see Section 4 for an overview of the range of conventions found in the literature).⁴

Tone in Hausa serves to distinguish both lexical and grammatical meaning: as shown in (1), the lexical meanings associated with the segmental sequence /fari:/ are distinguished by L-H, H-L, and H-H melodies.

- (1) a. *fàrí* (L H) – ‘look (n)’
b. *fáří* (H L) – ‘dry season’
c. *fáří́* (H H) – ‘white/whiteness’ (Wolff 1993, p. 56)

Similarly, grammatical distinctions, such as different TAM (= Tense, Aspect, Mood) categories, are equally distinguished by tonal means, as shown in (2), illustrating H, L, and HL (= fall).

- (2) a. *yá zó* (H H) – he came (relative completive)
b. *yà zó* (L H) – he should come (subjunctive)
c. *yá zó* (H H) – he came (absolute completive)
d. *yâ zó* (HL H) – he might come (potential)

Alongside tone, vowel length is another distinctive suprasegmental property: again, we find minimal pairs, where length serves to distinguish lexical or grammatical meanings, cf. (3) and (4), respectively.

⁴Segmental material, i.e. sequences of consonants and vowels, is represented in standard *Boko* (=Latin script) orthography: hooked letters (*ḃ*, *ḁ*, *ḅ*) represent glottalised stops, bigraph *ts* stands for a glottalised alveolar fricative, whereas *ʸ* denotes a palatalised glottal stop. Other deviations from IPA conventions are *j* (voiced palatal affricate), *c* (voiceless alveo-palatal affricate), and *y* (voiced palatal fricative). The vowel letters *i* and *u* in coda position are actually glides. Geminates are represented by reduplication of the initial letter of a monograph or bi-graph, e.g. *ssh* (=ʃʃ) for geminated *sh* (=ʃ) or *tts* for a geminated glottalised alveolar fricative (=s’).

- (3) a. *fásà* (CVVCVV) – ‘postpone’
b. *fásà* (CVCVV) – ‘smash’ (Newman 2000, p. 400)
- (4) a. *yá zó* – he came (absolute completeive)
b. *yá zó* – he came (relative completeive)

Syllables in Hausa are either light (CV) or heavy (CVC or CVV). Thus, long vowels can only be observed in open syllables. The distribution of tone is also sensitive to syllable structure, ruling out any occurrence of the HL contour tone on light CV syllables.

3 TONE AND INFLECTIONAL MORPHOLOGY IN HAUSA

Owing to its lexical and grammatical functions, Hausa tone is intimately linked to morphological operations, both inflectional and derivational. From a tonal perspective, morphological operations can be classified into two types: *agglutinative* and *holistic*. With agglutinative tone assignment, morphological rules simply add a tone to the base, typically together with some segmental material, whereas with holistic assignment, they may specify a completely new melody for the entire base, thereby overwriting lexical tone specifications. Newman (2000) regards tone assignment as a property of the affixes, and therefore distinguishes between tone-integrating affixes, which holistically affect the tonal make-up of the base, and non-integrating affixes, which leave the tones of the base by-and-large unaffected.⁵

3.1 *Holistic tone assignment: tone-integrating suffixes*

Holistic assignment of tonal patterns by morphological operations can probably best be illustrated by Hausa plural formation: as detailed in Table 1, adapted from Newman (2000, p. 431), the overwhelming majority of Hausa plural formation patterns do indeed feature holistic tone assignment. In fact, among the 15 major plural classes, only

⁵Non-integrating affixes may, of course, locally affect the tones at the juncture, as an instance of internal sandhi. What is crucial to the distinction here, is that agglutinative assignment is an entirely local operation, leaving most of the base’s lexical tone intact, whereas holistic assignment discards the base’s lexical tone specification altogether.

class X and XII preserve the tonal melody of the base. All others assign a melody specific to the plural class to bases of heterogeneous tonal make-up (see the detailed discussion of Class I and II below). As a general convention, I shall represent arbitrarily long sequences of like tones, i.e. the result of tone spreading, using the Kleene plus, which denotes arbitrarily many repetitions of the preceding symbol.

Class	Plural pattern		Example		Gloss
	Segmental	Tonal	Singular	Plural	
I	-oCi	H ⁺	tágà	tágógí	window
II	-ai	L ⁺ -H	dàlìlì	dàlìlái	reason
III	-aCe	H ⁺ -L-H	dámó	dámàmé	land monitor
IV	-(a)Ca	H-L-H	sírdí	sírâdá	saddle
V	-aCu	H-L-H	gúrgù	gúrâgú	cripple
VI	-uCa	H ⁺ -L	hùlá	húlúnà	cap
VII	-aCi	L ⁺ -H	fùré	fùrànní	flower
VIII	-aCCaCi	H-L-H-H	gúntú	gúntàttákí	stub
IX	-u/-i	L ⁺ -H	kùjèrà	kùjèrú	chair
X	-V	-	kwàdó	kwàdí	frog
XI	-āwā	L ⁺ -H / H ⁺	bàdúkù	dùkàwá	leather worker
XII	RED	-	jòjì	jóji-jóji	judge
XIII	-e + RED	L ⁺ -H L ⁺ -H	tsírò	tsiré-tsiré	shoot/sprout
XIV	RED	H ⁺ L ⁺	mākèkè	māká-mākà	expansive
XV	-i + RED	H ⁺ H ⁺	mìnínì	míní-míní	tiny

Table 1:
Hausa plural
formation
patterns

At the segmental level, Hausa employs several different marking devices: suffixation of vowels, with or without reduplication and gemination of the final root consonant (C), as well as total reduplication (RED; classes XII–XV).

The data in (5) provide examples from the highly regular and productive class I nouns,⁶ which form their plural by affixation of $-oCi$, where C represents reduplication of the last root consonant. Base-final vowels, if any, are replaced by the first vowel of the suffix.

⁶I adopt the classification of Newman (2000). For alternative analyses of the Hausa plural system, see e.g. Jaggar (2001) and Wolff (1993).

- (5) -óCí (H⁺) (Class I) (Newman 2000, p. 432)
- a. gúlà (H-L) – gúlólí (H-H-H)
'drum stick'
 - b. tágà (H-L) – tágógí (H-H-H)
'window'
 - c. gyàlè (L-L) – gyálólí (H-H-H)
'shawl'
 - d. tàmbáyà (L-H-L) – tàmbáyóyí (H-H-H-H)
'question'
 - e. kámfàní (H-L-H) – kámfánóní (H-H-H-H)
'company'
 - f. kwàmìtî (L-L-HL) – kwámítóčí (H-H-H-H)
'committee'

Together with the segmental change, plurals in this highly productive inflectional class are characterised by an all-high tonal melody assigned across the base and the plural affix -óCí. Note further that this tonal assignment is independent of the lexical tone of the base (we find H-L, L-L, L-H-L, and H-L-H). Furthermore, the assignment of an all-high melody appears to be independent of the number of tone-bearing units, there being no difference between trisyllabic and quadrisyllabic words of this plural class. This independence of tonal assignment from the segmental make-up of the word favours an intensional description over an extensional enumeration of tone patterns. One possible description would be assignment of an H tone spreading across the entire domain, as assumed in autosegmental phonology (Leben 1973; Goldsmith 1976).

Another, slightly more complex case of holistic assignment is contributed by class II plurals:

- (6) -ái (L⁺H) (Class II) (Newman 2000, pp. 434–435)
- a. àlhájì (L-H-L) – àlhàzái (L-L-H)
'Hadji'
 - b. dǎlibí (H-L-H) – dǎlibái (L-L-H)
'pupil'
 - c. sánkácè (H-H-L) – sánkàtái (L-L-H)
'reaped corn laid down in a row'

- d. àlmùbàzzàrí (L-L-H-L-H) – àlmùbàzzàrái (L-L-L-L-H)
‘spendthrift’
- e. d́ámì (H-L) – d́ammái (L-H)
‘bundle’

Here, plural formation adds a suffix *-ai*, and assigns a final LH melody, with the L spreading to the left of the word. As before, assignment of the plural tone pattern is independent of the lexical tone of the base (we find H-L, L-H, H-L-H, L-H-L, and L-H-L-H). Similarly, the very same tonal pattern is assigned to bisyllabic, trisyllabic, and pentasyllabic plurals alike, showing even more clearly the independence of melody specification from the syllable count.

Although the need for the incorporation of spreading is illustrated most clearly in the case of inflectional morphology, where the number of tone-bearing units cannot be established a priori, spreading can also be fruitfully put to use to simplify the inventory of lexical tone melodies. Assuming with Newman (2000) that the first tone of any melody automatically spreads to the left, complex patterns such as H-H-L in (6c), and simpler patterns such as H-L in (6e), can be generalised to H⁺-L.

3.2 *Tonal affixation: non-integrating suffixes*

Alongside holistic tone assignment by morphological operations, Hausa also recognises agglutinative tone, i.e., where a tonally specified affix is simply added to a base, leaving the tones of the base fully intact.

An example of a tonally purely agglutinative process is contributed by possessive marking:⁷ as shown in (7), Hausa bound possessives are formed by affixation of a consonantal “linker” that agrees in gender and number with the possessum (feminine singular *-r* vs. *-n* otherwise), plus a pronominal affix, marking person, number and gender of the possessor (e.g. third singular feminine *-t̀à*).

- (7) a. ḱwái (H) – ḱwá-n-t̀à (H-L)
egg(M) egg(M)-L.M-3.S.F
‘(her) egg’

⁷ See Crysmann (2011) for detailed arguments as to why the Hausa linker and possessive markers should be regarded as morphologically attached affixes, rather than postlexical clitics.

- b. rìgá (L-H) – rìgá-r-tà (L-H-L)
gown(F) gown(F)-L.F.SG-3.S.F
'(her) gown'
- c. mótà (H-L) – mótà-r-tà (H-L-L)
car(F) car(F)-L.F.SG-3.S.F
'(her) car'
- d. kâi (HL) – kâ-n-tà (HL-L)
head(M) head(M)-L.M-3.S.F
'(her) head'

The possessive pronominal affix *-tà* comes with its own fixed low tone; the tonal specification of the base, however, remains unaffected. Although the linker, which syllabifies with the coda of the base, triggers shortening of the rhyme, in accordance with Hausa's ban on super-heavy syllables, it leaves the tonal make-up of the rhyme unaffected, as witnessed by the falling tone in (7d).

A slightly more complex case of tonal agglutination can be observed with the specificity or previous reference marker *-r̃/-r̃̀*. Segmentally identical to the linker, this consonantal marker adds a low tone to the coda of the base, turning a final H level tone into an HL contour tone, while leaving low-final bases unchanged.

- (8) a. kwái (H) – kwâ-n (HL)
'the (aforementioned) egg (M)'
- b. rìgá (L-H) – rìgâ-r (L-HL)
'the (aforementioned) gown (F)'
- c. mótà (H-L) – mótà-r (H-L)
'the (aforementioned) car (F)'
- d. kâi (HL) – kâ-n (HL)
'the (aforementioned) head (M)'

However, in contrast to tone-integrating affixes, the tonal effects on the base are highly local in nature.

The last case of non-integrating affixes I shall consider involves a floating tone. Weak verbal nouns in Hausa are formed by affixation of a long H suffix *-wá*. With H-final bases in (9a–c), affixation of *-wá* gives rise to an HL contour tone on the preceding heavy syllable (CVV or CVC), a property that can be traced to the marker's initial low floating tone.

- (9) a. káràntá (H-L-H) – káràntáwá (H-L-HL-H)
‘read’ (GR1)
- b. sáyár (H-H) – sáyâr wá (H-HL-H)
‘sell’ (GR5)
- c. káwô (H-H) – káwôwá (H-HL-H)
‘come’ (GR6)
- d. kámà (H-L) – kámàwá (H-L-H)
‘catch’ (GR1)
- e. gyàrú (L-H) – gyàrúwá (L-H-H)
‘be repaired’ (GR7)

However, with light final syllables (as in grade 7;⁸ cf. (9e)), the floating low tone is suppressed, reflecting the phonotactic constraint of the language that restricts contour tones to heavy syllables (CVC or CVV). Again, tonal adjustments are fully local to the juncture, showing no impact on earlier tones of the base.

Before we move on, I would like to comment briefly on vowel length, the other distinctive suprasegmental property of Hausa: while morphological tone may be assigned holistically or be merely agglutinative, possibly triggering some local adjustments under strict adjacency, all the cases of length alternation we have observed so far are of a strictly local nature. In fact, this appears to be a general property of the language: while tones may be assigned individually or as entire melodies, there is no assignment of rhythmic length patterns that operates across larger domains, let alone “length harmony”, i.e. the spreading of same length specifications. In essence, length is only

⁸Hausa grades can be roughly thought of as inflectional classes, although some grades (namely 4–7) also encode derivational properties, like totality (4), causative or efferential (5), ventive (6) or medio-passive (7), which makes them resemble the *binyanim* of distantly related Semitic languages: each of the grades is associated with characteristic tone patterns, and an equally characteristic alternation of the final vowel, depending on the presence and nature of the direct object, called *frames*: Frame B is used with pronominal affixes, Frame C with locally realised direct objects, and Frame A elsewhere, including object fronting. See Parsons (1960), Newman (2000) and Jaggat (2001) for in-depth descriptions, as well as Abdoulaye (1992) and Crysmann (2005a) for recent synchronic analyses. Note that, for some grades, the exact tonal pattern changes according to frame. Cf. Table 2 for an overview of affixal and tonal patterns.

ever manipulated locally, while tonal manipulations may operate locally or globally, depending on the morphological construction.

3.3

Toneless prefixes

A third type of tonal behaviour triggered by morphological affixation is contributed by prefixation of “toneless” markers, i.e. affixes that are not inherently prespecified for a particular tone: regular and productive formation of pluractionals in Hausa is expressed by prefixation of a CVC reduplicative prefix, where the two consonants are identical to the first root consonant, and the vowel is essentially identical to the first vowel of the root (modulo reduction in closed syllables).

(10) C₁VC₁-

- a. dárɲàcḗ (H-L-H) – dádɲàcḗ (H-H-L-H)
‘press down/oppress’ (GR4 A) (Newman 2000, p. 424)
- b. káràntá (H-L-H) – kákkáràntá (H-H-L-H)
‘read’ (GR1 A/B) (Newman 2000, p. 424)
- c. dàgùrà (L-H-L) – dàddàgùrà (L-L-H-L)
‘gnaw at’ (GR2 A) (Newman 2000, p. 425)
- d. káràntà (H-L-L) – kákkáràntà (H-H-L-L)
‘read’ (GR1 C)
- e. dàgùrí (L-L-H) – dàddàgùrí (L-L-L-H)
‘gnaw at’ (GR2 C)

As illustrated in (10), prefixation of CVC to trisyllabic bases gives rise to quadrisyllabic words. Since melodies for Hausa verbs are maximally tri-tonal, what happens is that the leftmost tone simply spreads to the tonally underspecified prefix, yielding H for H-initial and L for L-initial bases.

Pluractionals of bisyllabic bases display a slightly more intricate pattern: although spreading can still be attested in those paradigms (grades) that are maximally bitonal, such as grades 6 and 7 in (11a–b), or the B and C forms of grade 2 in (11c), grades which witness tri-tonal melodies, such as grades 1 and 4, as well as the A form of grade 2, simply use the tri-tonal melody for pluractionals that we already found with trisyllabic bases in these grades (12). Table 2 provides a synopsis of segmental and tonal patterns for all seven productive grades.

Grade	Frame A ⁹		Frame B		Frame C	
	Seg/Len	Tone	Seg/Len	Tone	Seg/Len	Tone
	$\sigma\sigma / \sigma^+\sigma\sigma$		$\sigma\sigma / \sigma^+\sigma\sigma$		$\sigma\sigma / \sigma^+\sigma\sigma$	
1	-ā	H-L / H ⁺ -L-H	-ā	H-L / H ⁺ -L-H	-a	H-L / H ⁺ -L-L
2	-ā	L-H / L ⁺ -H-L	-ē	L ⁺ -H	-i	L ⁺ -H
3	-a/-i	L-H / L ⁺ -H-L		-		-
4	-ē	H-L / H ⁺ -L-H	-ē	H-L / H ⁺ -L-H	-ē/-e	H-L / H ⁺ -L-L
5	-ř	H ⁺	-shē	H ⁺	-ř	H ⁺
6	-ō	H ⁺	-ō	H ⁺	-ō	H ⁺
7	-u	L ⁺ -H		-		-

Table 2:
Synopsis of
Hausa grades

- (11) a. káwó (H-H) – kákkáwó (H-H-H)
‘bring’ (GR6) (Newman 2000, p. 424)
- b. gyàrú (L-H) – gyàgyàrú (L-L-H)
‘be well repaired’ (GR7) (Newman 2000, p. 424)
- c. jèfè (L-H) – jàjjèfè (L-L-H)
‘throw at’ (GR2 B)
- (12) a. tákà (H-L) – táttàkà (H-L-H)
‘step on’ (GR1 A) (Newman 2000, p. 424)
- b. jèfà (L-H) – jàjjèfà (L-L-L)
‘throw at’ (GR2 A) (Newman 2000, p. 424)

To summarise, the addition of a syllable by pluractional prefixation incurs a switch of tonal pattern to the appropriate melody associated with trisyllabic words for that paradigm cell. If the melody provides for fewer tones than there are tone-bearing units, automatic spreading applies: occasionally, with trisyllabic pluractionals (depending on the tonal pattern of that grade), and always, across all grades, with quadrisyllabic ones.

Having investigated the basic suprasegmental processes associated with morphological operations in Hausa, namely holistic assignment (“tone-integrating affixes”), tonal affixation (“non-integrating affixes”), and spreading of base tones on to inherently toneless prefixes, we are now in a position to explore, in the following sections,

⁹Segmental shape and tone can both be subject to lexical exception in Frame A, in particular in grade 2.

how these processes can be integrated, in an efficient way, into a computational grammar of Hausa, built on a lean typed feature structure formalism.

4 HAG – A COMPUTATIONAL GRAMMAR OF HAUSA

The treatment of tone discussed in this paper is part of an emerging implemented computational grammar of Hausa, developed in the framework of Head-driven Phrase Structure Grammar. The underlying typed feature logic is a purely conjunctive variant of TDL (= Type Description Language; Krieger 1996), as currently implemented in several processing systems, such as the Linguistic Knowledge Builder (=LKB; Copestake 2002), the PET parser (Callmeier 2000), and the ACE parser and generator (the latter being described in Crysmann and Packard 2012). Owing to its declarative nature, the grammar is fully reversible, i.e. it can be used for both parsing and generation. Furthermore, the symbolic grammar resource is complemented by stochastic models for parse selection and realisation ranking, developed on the basis of the Redwoods treebanking technology (Oepen *et al.* 2002). An overview of the grammar and the major constructions it covers can be found in Crysmann (2012a). The grammar is freely available at <http://svn.emmtee.net/trunk/11f/hag/> under an open-source licence. An online demonstrator of the grammar is hosted at <http://hag.delph-in.net/logon>.

Although the lexicon is still rather small, the grammar already covers a wide range of core constructions of the language. With respect to morphology, the grammar implements inflectional morphology in both the nominal and the verbal domain, including the infamously rich set of plural formation patterns. On the segmental side, morphological rules cover all morphophonological processes attested in the language, including affixation, gemination, as well as partial and full reduplication.

With respect to morphosyntax, the grammar boasts a systematic treatment of direct object marking, the so-called Hausa frames (Parsons 1960). The inflectional approach to Hausa verb frames developed in Crysmann (2005a) has been generalised to nominal categories, including gerunds, pronominal adjectives and possessives (Crysmann 2011), as well as prepositional nouns (Crysmann 2012b).

On the purely syntactic side, the grammar covers local complementation and modification, as well as non-local processes, such as wh-extraction, focus fronting, and relativisation, with both gap and resumptive strategies (Crysmann 2012c, 2015).

Owing to the central status of tone and length for marking morphological and morphosyntactic properties, the grammar has been developed from the ground up to support suprasegmental representations. Particular care has been given to the fact that suprasegmental information is represented to different degrees in textual input: while standard Latin orthography (*Boko*) does not represent tone or length at all, length, but not tone, is marked in the Arabic script (*Ajami*). Scholarly as well as educational work on Hausa, by contrast, tends to fully mark tone and length, although the marking regimes differ: while long vowels are typically marked by macrons, leaving short vowels unmarked, as in the grammars by Newman (2000), Jaggar (2001), and Caron (1991), the Hausa language course by Cowan and Schuh (1976), or the Hausa–French dictionary by Caron and Amfani (1997), there are also clearly alternative marking schemes: Jungrathmayr and Möhlig (1976) use geminated vowel letters to mark length, and Newman and Ma Newman (1977) mark brevity (with a comma below the vowel). In *Ajami*, long vowels are signalled by a combination of letters (*ya*, *wau*, *alif*) and diacritics, distinguishing 5 vowel qualities, whereas short vowels are solely marked by diacritics, distinguishing 3 qualities. As for tone, the most wide-spread convention used in the Hausaist literature is to mark low and falling tone (with grave and circumflex accents), leaving high tone unmarked.

Given this diversity regarding the amount of suprasegmental information being marked, as well as the way it is signalled, a highly flexible approach is called for, if we want to be able to maximally exploit suprasegmental information, if present in the input, while at the same time ensure robustness towards partially marked or unmarked input. Moreover, given the locality of morphological processes regarding segmental material and length specifications as opposed to the potentially non-local assignment of melodies, an autosegmental separation of these pieces of information is inevitable. To this end, the grammar employs a token rewrite system (Adolphs *et al.* 2008) to convert a diacritically marked input string into a featural representation, separating the segmental level from the levels of tone and length spec-

ifications. The grammar can be configured at runtime as to which inferences should be drawn from overt suprasegmental specifications in the input: consistent full marking, where suprasegmentally unmarked segments are interpreted as the complement of the tone and length markings found, giving maximal disambiguation, and partial marking, where only overt marking is taken into consideration, permitting sporadic marking of tone or length by the user. While the former is best suited to the processing of edited texts, the latter is intended for interactive input to the grammar, where only critical tones may be marked and strict adherence to a consistent marking regime would appear cumbersome.

Grammar-internally, tone and length are systematically represented at the lexical, morphological and morphosyntactic levels. Thus, on the basis of the interaction of lexical with local and non-local grammatical constraints, suprasegmental information missing from the input can be recovered to a great extent by symbolic means. The residual ambiguity is addressed by means of discriminative parse selection models extracted from a treebank. Regenerating from disambiguated parses provides full reconstruction of tone and length specifications, obeying the full set of constraints imposed by the grammar, both locally and globally. Based on the tight integration of suprasegmental information, the grammar not only reaches a high level of linguistic adequacy, but will ultimately be suitable in a number of application scenarios for which this information is crucial, including text-to-speech synthesis (TTS), and computer-assisted language learning (CALL).

5 TOWARDS AN EFFICIENT AND FLEXIBLE REPRESENTATION OF TONE

5.1 *Tonal tiers in typed feature structures*

The data structures I shall adopt for the representation of suprasegmental information are lists: one list for tone sequences, and another list for vowel length information. Lists already have some intrinsic properties that make them suitable as a representation of tiers: first, in contrast to sets, they permit multiple occurrences of like elements, and, even more importantly, they are ordered, capturing the temporal organisation of the tier. Second, they constitute a much simpler

data structure than trees, which again seems to be a desirable property: while there is some evidence suggesting that the distribution of tones may depend on a hierarchical structure, be it morphological or prosodic, there seems to be very little evidence in Hausa as to a hierarchical structure of tonal sequences themselves.¹⁰

In (typed) feature structure formalisms, lists can be recursively implemented using FIRST/REST or HEAD/TAIL notation: the first element of the list is represented as the value of the HD feature, while the list remainder, itself a list, is represented as the value of TL. As illustrated in Figure 1, the second element will be found under TL.HD, the third element under TL.TL.HD etc.¹¹ Pure feature structure encoding of lists in HEAD/TAIL notation only directly exposes one end of the data structure, essentially corresponding to a stack or LIFO (= Last In First Out) in terms of data structure: if we recursively build up these lists element by element, we can easily access the last member that has been added to the resulting list, but direct access to the first element ever added will not be straightforward.

$$\langle \text{high, low, low, ...} \rangle \equiv \left[\begin{array}{l} \text{HD } \textit{high} \\ \text{TL } \left[\begin{array}{l} \text{HD } \textit{low} \\ \text{TL } \left[\begin{array}{l} \text{HD } \textit{low} \\ \text{TL } [] \end{array} \right] \end{array} \right] \end{array} \right]$$

Figure 1:
HD/TL encoding of lists

Throughout this paper, I shall use the list constructor ([]) as an abbreviatory device to partition a list into an initial sequence of elements and a list remainder, a notation familiar from Prolog (Clocksin and Mellish 1981, pp. 52–53), that is also regularly used in HPSG. As for the representation of morphological rules, I assume feature structure descriptions where the entire term represents the properties of the derived structure (the mother in a unary rule), whereas the properties of the morphological daughter, i.e. the base, are embedded under a

¹⁰ Of course, as pointed out by one of the reviewers, trees can be encoded using lists of lists. At the level of tonal tiers, however, appropriateness conditions on list types will ensure that lists cannot be nested. See below in this section on typed lists.

¹¹ I am using period as a path separator, to avoid confusion with the list constructor ([]).

feature DTR (cf. Riehemann 1998; Koenig 1999). Information shared between the mother and the daughter are captured by means of reentrancies (= *token identity* or *structure sharing*), expressed using boxed coreference tags, e.g. $\boxed{\tau}$ in Figure 2.¹²

Figure 2:
Basic list
operations

$$\left[\begin{array}{l} \text{TONE } < \textit{low} \mid \boxed{\tau} > \\ \text{DTR } \left[\text{TONE } \boxed{\tau} \right] \end{array} \right] \equiv \left[\begin{array}{l} \text{TONE } \left[\begin{array}{l} \text{HD } \textit{low} \\ \text{TL } \boxed{\tau} \end{array} \right] \\ \text{DTR } \left[\text{TONE } \boxed{\tau} \right] \end{array} \right]$$

(a) Add tone (push)

$$\left[\begin{array}{l} \text{TONE } \boxed{\tau} \\ \text{DTR } \left[\text{TONE } < \textit{high} \mid \boxed{\tau} > \right] \end{array} \right] \equiv \left[\begin{array}{l} \text{TONE } \boxed{\tau} \\ \text{DTR } \left[\text{TONE } \left[\begin{array}{l} \text{HD } \textit{high} \\ \text{TL } \boxed{\tau} \end{array} \right] \right] \end{array} \right]$$

(b) Remove tone (pop)

$$\left[\begin{array}{l} \text{TONE } < \textit{low} \mid \boxed{\tau} > \\ \text{DTR } \left[\text{TONE } < \textit{high} \mid \boxed{\tau} > \right] \end{array} \right] \equiv \left[\begin{array}{l} \text{TONE } \left[\begin{array}{l} \text{HD } \textit{low} \\ \text{TL } \boxed{\tau} \end{array} \right] \\ \text{DTR } \left[\text{TONE } \left[\begin{array}{l} \text{HD } \textit{high} \\ \text{TL } \boxed{\tau} \end{array} \right] \right] \end{array} \right]$$

(c) Change tone (pop & push)

Given these rather straightforward assumptions, we are already in a position to capture the kind of operations characteristic of

¹²Lexical rules, as employed here, are description-level rules, as opposed to meta-level rules: i.e. they essentially function like unary branching syntactic rules, with the added functionality of attaching orthographic changes using a variant of string unification (Calder 1989). Reentrancies between mother and daughter are fully specified in the rule descriptions or the rule types they inherit from. We exclusively rely on the type system to minimise redundancy, in contrast to Meurers (2001), who proposes compilation of a special lexical rule format. While the LKB supports default unification (Lascarides *et al.* 1996; Lascarides and Copestake 1999), neither PET nor ACE do: therefore, the development of HAG is fully monotonic, i.e. devoid of defaults.

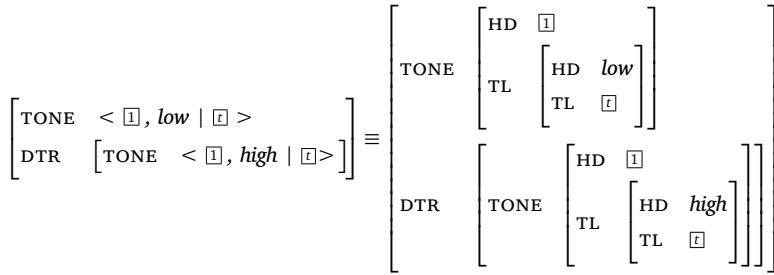
tone non-integrating affixation.¹³ In essence, there are three basic tonal operations that can be captured by means of morphological rule schemata operating on the tones of the base: addition of a tone (push an element on to the stack), deletion of a tone (pop an element off the stack), and modification, as a combination of push and pop. Figures 2 and 3 illustrate the equivalence (\equiv) of the simplified list notation with the underlying feature structure representation.

One can add tones to the beginning of the list (Figure 2a), by way of structure sharing the TONE list of the base with the TL of the resulting sign plus a specification of the HD element; one can remove initial elements from the list, by having the resulting sign's TONE list structure shared with the TL of the base only, excluding the HD element of the base (Figure 2b); or one can replace elements (Figure 2c), by combining the two operations.¹⁴ As illustrated in Figure 3a below, while access is not random, it is by no means limited to the first element: any element can be manipulated, as long as it is found at a definite distance from the beginning of the list. Likewise, one can easily define

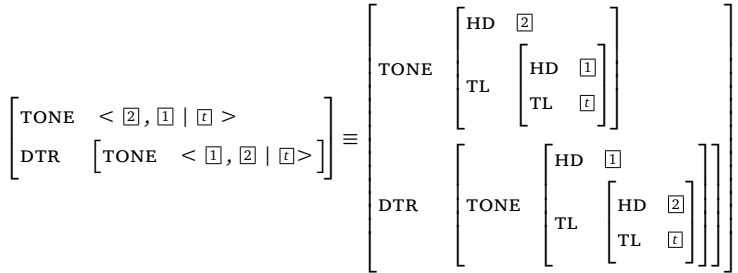
¹³In contrast to One-Level Phonology (Bird and Klein 1994), I shall assume a weak version of Phonological Compositionality, permitting "feature-changing" operations as the result of the application of morphological rules: while, strictly speaking, any feature value can only ever be made more specific in a monotonic feature logic like the one assumed here, the effect that a constraint imposed on the morphological base may not hold true for a form derived from that base can indeed be captured by assigning different values to the representations of the morphological mother and the daughter, i.e. by not fully equating their phonological representations.

¹⁴The way tonal modification is presented here, i.e. independent of syllable count or segmental changes is intentional: in the spirit of autosegmental phonology, alterations on one tier may, but need not, be paralleled by according alterations on a different tier. Basic tonal operations, as presented here, are building blocks, which may occur in conjunction with segmental and metrical changes in concrete morphological rules. In Hausa, addition of tones, as represented in the grammar, is usually accompanied by adding a tone-bearing unit (see e.g. suffixation of possessives in Figure 10). Changing a tone need not be, as witnessed by the previous reference marker in Figure 11. Adding or deleting a tone without manipulating the metrical structure would entail shifting of the remainder of the tonal specification. However, Hausa does not seem to provide any clear evidence for this.

Figure 3:
Some complex
list operations



(a) Change second tone



(b) Metathesis

metathesis of a pair of tones (Figure 3b), provided their position is known, which I assume to be the case.¹⁵

There is one important inherent limitation to lists represented as simple feature structure terms: in lean formalisms without any functional or relational constraints,¹⁶ such as the one we are using here, lists are essentially stacks, i.e. access of any element is straightfor-

¹⁵In case the relevant elements are found in a finite number of positions, situations can of course be enumerated using multiple rules.

¹⁶Current alternative implementations of typed feature formalisms include TRALE (Penn 2004), which actually supports relational constraints and disjunction. However, relational constraints provide an additional recursive structure besides the main recursion on the rule backbone. Thus, while lean formalisms will force hidden costs out into the open, such costs can be hidden by relational constraints. However, as we shall see, no purely phonologically motivated recursion steps are needed, other than what is already offered by type expansion. For the purposes of this paper, which aims at assessing the minimal computational power needed to address autosegmental phonology in Hausa, the choice of a purely conjunctive typed feature structure unification formalism without relational constraints must appear preferable, for methodological reasons.

ward from the beginning of the list, but quite hard to determine, in a general fashion, from the end of the list. Thus, in the context of these formalisms, it is of central concern whether tones are represented from left to right or rather from right to left. Given that Hausa is predominantly suffixal, and that spreading also proceeds from right to left, I shall assume for now¹⁷ that the most appropriate encoding of the tonal tier in this language will be from right to left, essentially assuming that tones on the TONE list are represented in inverse order of surface tone, i.e. list-initial tones correspond to rightmost surface tones, whereas list-final tones correspond to leftmost surface tones. Given an encoding in this order, tonal suffixation, including local modifications, as observed for non-integrating affixes, can be straightforwardly captured, using the basic operations illustrated in Figure 2. Sample analyses will be provided in Section 5.3.2.

Having established so far that a right-to-left encoding of tone is most suitable for the treatment of Hausa, facilitating the description of local tonal modifications observed for non-integrating suffixes, we shall now move on to tone-integrating suffixation, including spreading.

An interesting property of typed feature formalisms is that they inherently provide for parameterised list constraints, enabling us to impose some constraint over an arbitrary number of elements. As we shall see shortly, this property is key to our implementation of spreading.

To start with, consider the type hierarchy of basic list types, as given in Figure 4. The supertype *list* is defined to have exactly two immediate subtypes: either an empty list (*e-list*), or a non-empty list (*ne-list*). While the former has no appropriate features, the latter intro-

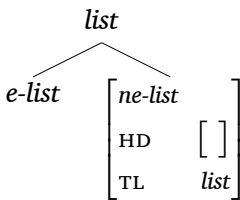


Figure 4:
Basic list type declarations

¹⁷ See, however, Sections 5.2 and 5.4.2 for detailed discussion and a generalised treatment.

duces the features HD and TL. According to the logic of typed feature structures (Carpenter 1992), whenever a feature such as HD is required to be present by some constraint, the type of the feature structure must be at least *ne-list*, and all other constraints associated with this type are enforced. Conversely, once a list has been specialised to *e-list*, it will be incompatible with either HD or TL features, since *e-list* and *ne-list* do not have a common lower bound.

List types and appropriateness, however, will show their full potential once we associate additional properties with a type. More concretely, I shall build on typed list constraints, a powerful, yet efficient way to impose constraints on the members of lists of arbitrary size.¹⁸

As stated in (13) and (14), one can provide type definitions for tone lists consisting of an arbitrary number of H, or an arbitrary number of L, yielding the type hierarchy in Figure 5. In essence, these list types will provide us with a concise and efficient formalisation of tone spreading.

- (13) a. $h^*list := list$.
 b. $h^*e-list := h^*list \wedge e-list$.
 c. $h^*ne-list := h^*list \wedge ne-list \wedge \begin{bmatrix} HD & high \\ TL & h^*list \end{bmatrix}$

- (14) a. $l^*list := list$.
 b. $l^*e-list := l^*list \wedge e-list$.
 c. $l^*ne-list := l^*list \wedge ne-list \wedge \begin{bmatrix} HD & low \\ TL & l^*list \end{bmatrix}$

As defined in (13), the type h^*list can be expanded either into $h^*e-list$,¹⁹ a subtype of the empty list type, or else into the non-empty list type, $h^*ne-list$, which restricts the HD element to be *high*. The remainder of the $h^*ne-list$ is in turn restricted to be of type h^*list , prop-

¹⁸To the best of my knowledge, list types were first explored in a systematic way by Flickinger (2000). See also Crysmann (2005b) on extended applications of this technique, e.g., for the implementation of type identity.

¹⁹Throughout this paper, the asterisk on tone specifications (*) is the Kleene star familiar from regular languages, denoting an arbitrary number of repetitions of tone symbols of this type, including zero. This notation should not be confounded with that used for pitch accents in intonation.

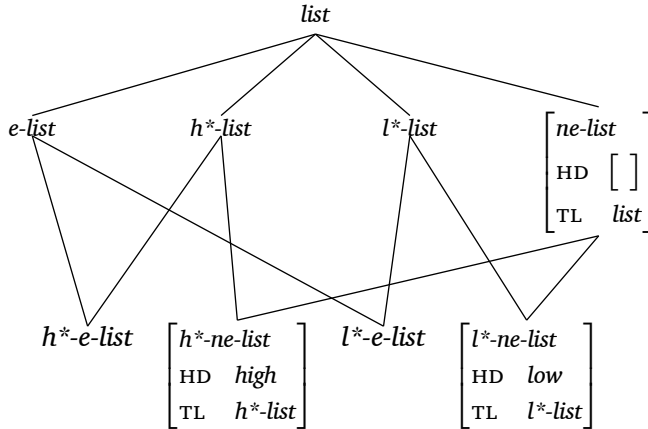


Figure 5: Hierarchy of tone list types

agating the tonal requirement further down the list: if the remainder is empty, nothing much happens, but if it is non-empty, it must be a subtype of both *ne-list* and *h*-list*, resolving it to *h*-ne-list*, fixing the value of the HD element (to *high*), and setting the value of the new list remainder to *h*-list*. Thus, tone list types provide a concise and effective way to push properties across lists of arbitrary length, in essence stating properties independently of the number of list members. More importantly, these constraints are latent, such that expansion is delayed until tonal specifications are accessed by other constraints, with full expansion being reserved to synchronisation with the metrical tier.

Based on this minimal inventory of tone list types, we are now in a position to formalise Hausa tone assignment, including automatic spreading, within the context of typed feature structure grammar. To give a more concrete example, let us consider the case of grade 6 verbs, featuring holistic assignment of a single spreading H tone for both basic verbs and pluractionals. The tonal constraint associated with grade 6 verbs is that of an *h*-list*: depending on the number of tone-bearing units (e.g. two for a basic verb like *kāwō*, and three for the pluractional derived from it), unification of the general *h*-list* constraint with tone lists of appropriate length will yield specialisation of *h*-list* to exactly as many high tones as required.

In order to understand the exact workings of list type constraints, let us briefly work through the example in Figure 6a: unification of *h*-list* with the outer level of the two-element list enforces specialisation of *h*-list* to *h*-ne-list*, entailing specialisation of the top-level HD

Figure 6:
Automatic
expansion
of tone list
constraints

$$h^*-list \wedge \begin{bmatrix} ne-list \\ HD \quad [] \\ TL \quad \begin{bmatrix} ne-list \\ HD \quad [] \\ TL \quad e-list \end{bmatrix} \end{bmatrix} \equiv \begin{bmatrix} h^*-ne-list \\ HD \quad high \\ TL \quad \begin{bmatrix} h^*-ne-list \\ HD \quad high \\ TL \quad h^*-e-list \end{bmatrix} \end{bmatrix}$$

(a) Bisyllabic *káwó*

$$h^*-list \wedge \begin{bmatrix} ne-list \\ HD \quad [] \\ TL \quad \begin{bmatrix} ne-list \\ HD \quad [] \\ TL \quad \begin{bmatrix} ne-list \\ HD \quad [] \\ TL \quad e-list \end{bmatrix} \end{bmatrix} \end{bmatrix} \equiv \begin{bmatrix} h^*-ne-list \\ HD \quad high \\ TL \quad \begin{bmatrix} h^*-ne-list \\ HD \quad high \\ TL \quad \begin{bmatrix} h^*-ne-list \\ HD \quad high \\ TL \quad h^*-e-list \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

(b) Trisyllabic *kákkáwó*

value to *high* and of TL to *h*-list*. This specialisation of the TL value will in turn trigger unification with the value *ne-list* which yields the most general subtype of *h*-list* and *ne-list*, namely *h*-ne-list* of Figure 5. The constraint associated with this latter type will again be applied to the feature structure under TL, specialising TL.HD to *high*, and propagating the tone list type on to TL.TL. Unification with the empty list under TL.TL finally resolves to *h*-e-list*.

Before moving on, let us briefly discuss how the suprasegmental representations are synchronised. As established at the end of Section 3.2, length specifications are only ever modified locally and, more crucially, do not require underspecification, in contrast to tonal spreading, for example. This makes the syllable or length tier an ideal timing tier. By contrast, holistic assignment and tonal spreading favour an underspecified description. During morphological construction, we therefore state constraints using underspecified descriptions, treating representations of tone separate from length. At the level of the maximal morphological word, however, we use the length of the LEN list, our timing tier, to determine the exact length of the TONE list, effectively expanding list constraints to exactly the number of tones re-

quired.²⁰ The delayed synchronisation of tone lists with the timing tier is not only clearly in the spirit of an autosegmental approach, but it also provides a highly efficient implementation, avoiding repeated recursion over autosegmental tiers. Furthermore, given the monotonic nature of HPSG’s underlying feature structure formalism, a single, delayed evaluation is fully sufficient.

5.2 *Direction of tone assignment*

We have seen so far that the tonal effects of morphological processes in Hausa assign a privileged status to the right edge: tone non-integrating affixes simply add tones on the right, potentially supplanting or modifying a base-final tone, whereas melodies assigned by tone-integrating affixes are typically assigned one by one, starting with the suffix, with the leftmost tone spreading leftwards on to any preceding syllables. In fact, as suggested by Newman (1986, 2000) and Jaggar (2001), the standard direction of association in Hausa is from right to left. While this view may contradict the conventions proposed by Clements and Ford (1979), who suggest a universal left-to-right spreading convention, right-to-left association and spreading nevertheless appear to be particularly well motivated by the empirical patterns of the language.

To start with, tone patterns assigned to lexical bases are probably somewhat inconclusive, as argued by Schuh (1989, p. 257). He discusses abstract L H sequences of quadrisyllabic lexical nouns and observes that all three logically possible patterns are attested in the Hausa lexicon: L-H-H-H (*bùlágúřó* ‘trip’), L-L-H-H (*gwàlámńyá* ‘speaking unintelligibly’), and L-L-L-H (*ànnàshùwá* ‘feeling happy’).

The situation, however, becomes much clearer once we consider grammatically assigned tone patterns: out of the 15 plural patterns

²⁰ Technically, synchronisation of TONE and LEN lists is performed by unification: in order to keep type hierarchies of tone and length values distinct, we actually maintain a “shadow” list (–LEN) as part of our LEN definition, that will have the same length as the main list, yet does not constrain the type of the list elements, as shown below.

- i. *len-list* := *list* \wedge [–LEN *list*]
- ii. *len-e-list* := *len-list* \wedge *e-list*
- iii. *len-ne-list* := *len-list* \wedge *ne-list* \wedge [–LEN < [] | [] > , TL.–LEN []]

identified by Newman (2000) (cf. Table 1), 7 have an initial tonal plateau and final alternation of H and L (classes II, III, VI, VII, IX, XI, XIII), whereas only 1 class observes the opposite pattern (class VIII), showing a fixed H-L-H-H melody. However, since the plurals in this class are all quadrisyllabic, no really strong case can be made for spreading of a final H. The remaining plural tone patterns are inconclusive, since they are either monotonal (all plateau), or show no evidence of spreading (no plateau). Furthermore, the right-to-left perspective on assignment neatly aligns with the fact that any unambiguous examples of tone-integrating affixes are suffixal.²¹ Considering the tonal patterns within the verbal grade system (cf. Table 2), we find this pattern confirmed: 8 out of 10 minimally bi-tonal patterns show alternating tones on the right, while having a tonal plateau, a potential indicator of spreading, on the left. Furthermore, the two somewhat exceptional H-L-L patterns do feature spreading on the left with quadrisyllabic verbs. Moreover, the point of transition between H and L is always at a fixed distance when counting from the right, yet would be variable when proceeding from the left.

Another piece of evidence in favour of an alignment between suffixation and tone assignment comes from multiple suffixation of tone-integrating affixes: as discussed in Newman (1986), regular Hausa past participles are formed by affixation of a tone-integrating reduplicative *-aCC* suffix, where C reduplicates the last consonant of the base, and an agreement marker (*-ē, -iyā, -ū*). In the singular, the base is characterised by an LH tone pattern, whereas the masculine and feminine agreement markers *-ē* and *-iyā* carry high tone. In the plural, however, we observe a final H on the agreement marker *-ū*, preceded by an all-L tone sequence.

²¹ There are two not fully productive derivational patterns featuring both prefixation and holistic tone assignment, namely agentive nouns and a subclass of ethnonyms. However, since both patterns are circumfixal, i.e. they simultaneously involve suffixation, they do not provide conclusive counter-evidence to the otherwise systematic observation that tone-integrating affixes are suffixal. Moreover, as witnessed by the contrast between masculine and feminine agentive nouns, choice of tonal pattern is most likely associated with the suffixal part of the circumfix, rather than with the segmentally and tonally constant prefix. I will discuss these in full in Section 5.4.2, including a generalised formal analysis of tone spreading and prespecification.

<i>masculine</i>	<i>feminine</i>	<i>plural</i>	<i>gloss</i>
dàfáffé	dàfáffíyá	dàfáffú	'cooked'
gàgàrré	gàgàrríyá	gàgàrrú	'rebellious'
yàgàlálélé	yàgàlálíyá	yàgàlálú	'torn into pieces'

Table 3:
Hausa past participles

Newman (1986) explains these patterns by means of the interaction of two tone-integrating suffixes: a tone-integrating LH participle marker *-aCC*, and a tone-integrating LH plural marker *-ū*, which is independently attested in noun class IX, with identical segmental and suprasegmental properties. While non-integrating singular agreement affixes *-ē* and *-iyā* leave the LH participial tone pattern intact, suffixation of the tone-integrating plural marker *-ū* replaces the characteristic participial tone pattern with that of the plural marker, again applied from right to left.

Leben (1978) has suggested unifying the case of Hausa's apparent preference for right-to-left assignment and spreading with the left-to-right convention (Clements and Ford 1979) standardly assumed within autosegmental phonology at the time. He reanalyses the Hausa participle facts, using a combination of lexical/morphological prespecification and automatic left-to-right spreading of an initial floating H: taking the example of Hausa past participle formation in Table 3 above, the tonal pattern of a form, such as *gàgàrré*, is analysed as having a pre-linked tone on the past participle marker *-aCC*, preceded by a floating H that associates (and spreads) from left to right. Similarly, the pre-linked L will spread to the end of the word. While this may work particularly well for the case at hand, given the identical final tones, any explicit account of Hausa lexical tone assignment and tone-integrating affixes will still have to establish the exact location of pre-linked tones, which are found in this case at a fixed distance from the edge only when counting from the right (cf., again the examples in Table 3). Thus, even if there is a credible analysis of spreading from the left, pre-linking still needs to proceed relative to the right edge. Finally, any approach that draws on pre-linking requires a non-monotonic logic, since association of unassociated tones has to check first whether or not a tone has already been assigned, which is not necessary when using tone list constraints, as proposed here, which are assigned in a fully monotonic fashion.

There is, however, additional evidence for connecting spreading to a standard right-to-left association in Hausa: while we do find unambiguous examples of toneless *prefixes* in the language (see Section 3.3), there is no such unambiguous evidence for toneless *suffixes*. In contrast to pluractional prefixes, which display an alternation of tone depending on that of the following tone, cases of toneless suffixes are mostly non-existent, or inconclusive, like the past participle agreement markers discussed above, where the tonal specification is invariant (always H in this case).

Another piece of evidence cited by Leben (1978) in favour of left-to-right association comes from vowel epenthesis:²² the Hausa lexicon has a rather small number of words that are consonant-final. Among these, a subset has an alternate form, the use of which becomes obligatory, for phonotactic reasons, in combination with the linker *-n/-r* (see Section 3.2 above).

(15) HL-final bases

a. fâm – fámì
‘pound’

b. àlhàmîs – àlhàmíshì
‘Thursday’

(Leben 1978, p. 207)

(16) H-final bases

a. bàbúr – bàbúrí
‘motorbike’

b. àlján – àljání
‘imp’

(Leben 1978, p. 207)

(17) L-final bases

a. mālám – mālámí
‘teacher’

b. fénsir – fénsirí
‘pencil’

(Leben 1978, p. 207)

In order to provide a unified account of the final tone observed in (15)–(17), Leben (1978) suggested that the epenthetic vowel itself is

²²Thanks to one of the anonymous reviewers for pointing this out as potential evidence favouring left-to-right association.

toneless, which immediately accounts for the distribution of the falling contour tone over the last two tone-bearing units in (15), and using left-to-right spreading, for the tonal identity of the epenthetic vowel to the preceding syllable in (16). As for L-final bases in (17), he suggested spreading of the final L, paired with subsequent application of Low Tone Raising (Leben 1971), a hypothesised productive rule of Hausa that raises a low on heavy (CVC or CVV) final syllables when preceded by a low tone.

However, there are several reasons to question the validity of this analysis: first, epenthesis hardly enjoys the status of a productive rule of the language. According to Newman (2000, p. 307), the majority of consonant-final bases (over 250) in Hausa do not give rise to epenthesis (e.g. *māshîn* ‘motor cycle’), but rather use syntactic means to encode, e.g. possessive and previous reference marking. Second, the segmental make-up of the epenthetic form is not fully predictable from that of the consonant-final form, whereas the short form can be predicted from the long form: variation includes the quality of the epenthetic vowel, which is mostly *i*, but sometimes *ū* (18), the quality of the final consonant (19), and the length of the penult, which is mostly long, but sometimes short (20).

- (18) *hàrâm* – *hàrâmù*
‘unlawful’ (Newman 2000, p. 307)
- (19) *ràsî* – *ràsîdî*
‘receipt’ (Newman 2000, p. 307)
- (20) *mùtûm* – *mùtúmì*
‘man’ (Newman 2000, p. 307)

Newman (2000) therefore reanalyses the final vowel as a latent one, essentially proposing clipping rather than epenthesis, which enables him to account in a straightforward way for the limited productivity, as well as their segmental, metrical, and tonal properties: all of these can be derived on the basis of general, undisputed, fully regular phonotactic properties of the language, invoking general restrictions on coda segments (consonants), syllable weight (vowel shortening in closed syllables), and tone (simplification of LH).

Finally, and most importantly, the status of Low Tone Raising as a (synchronic) phonological rule of Hausa in itself is not unproblematic:

as discussed in Newman and Jaggar (1989), systematic and sporadic exceptions to this rule can be observed throughout the grammar of Hausa: they cite seven phenomena in total (see Schuh 1989, for arguments discarding lengthening in questions as intonational in nature) where this rule is indisputably violated at the surface, including regular plural formation of augmentative adjectives, as illustrated in (26), some imperatives, ideophonic adjectives, adverbs, and action nouns, and recent loans from English. Furthermore, they argued that several putative applications of this rule necessitate unorthodox assumptions regarding word boundaries. Schuh (1989), in a reply to Newman and Jaggar, argued that the generalisation expressed by Low Tone Raising can be saved, once the conditions are suitably refined: he suggested in particular that Low Tone Raising may only apply to sequences of singly associated tones, not to spreading of a single L.²³ However, this specific refinement will not help in the case at hand, since Leben's analysis crucially depends on the combination of spreading and Low Tone Raising in order to derive the surface patterns in (17), in particular, since he explicitly argued in his 1978 paper in favour of multiple association over copying.

To summarise, lexical and morphological assignment of tone in Hausa strongly militates for a right-to-left regime. With spreading, both directions remain as an option, with right-to-left spreading keeping an edge over its competitor, both in terms of a better match between association and spreading, and the empirical asymmetry regarding the privileged existence of toneless prefixes vs. suffixes. The idea of language-specific directions of association and spreading may run counter to universalist ideas about uniformity: faced with the empirical evidence in Hausa, however, it should appear as equally odd to enforce a universalist left-to-right view of spreading, while still maintaining the opposite picture for lexical and morphological tone assignment. I shall therefore conclude that the analysis advanced by Newman (1986, 2000) still remains valid.

Within the context of the current formal approach couched in terms of tone list constraints, prevalence of a single, albeit language-

²³Despite this qualification, Schuh (1989) equally rejected the status of Low Tone Raising as a phonological *rule* of Hausa, picturing it rather as a lexical constraint of the language.

specific regime is actually a welcome result: given that a stack-like typed list encoding of the tonal tier like the one proposed here confines augmentative, modifying, and subtractive operations to one end of the list, and spreading to the other, we shall give preference to any analysis that treats assignment and association in a symmetrical way. Moreover, a representation encoding both lexical and morphological holistic assignment in the same direction (from right to left) not only facilitates the implementation of spreading by means of list types, it is also beneficial to the treatment of agglutinative tone in a suffixing language such as Hausa, essentially exposing the rightmost tone(s) as the top of the stack, directly available for modification and addition.

5.3 *Suffixes and tone*

Now that we have established a preferential direction for tone assignment and spreading, I shall show how the phenomena we have considered so far can be represented in the context of a Head-driven Phrase Structure Grammar (HPSG) of Hausa, crucially building on the aforementioned typed tone lists.

5.3.1 *Tone-integrating affixes*

The first type of morphological rules I shall discuss pertains to tone-integrating suffixes, i.e. holistic assignment of melodies. Throughout this section I shall leave out the description of morphosyntactic and purely morphological properties, focusing on segmental and suprasegmental changes instead.²⁴ As depicted in Figure 7, the suprasegmental effects of regular $-\bar{o}C\bar{i}$ suffixation are captured by means of constraints on the TONE and LEN lists, which are both encoded from right-to-left: regarding vowel length, the rule ignores the last length specification of the base and adds two specifications for long vowels. Length specifications for syllables other than the last are shared by the morphological mother ($\bar{[l]}$). The segmental changes induced by the rule are

²⁴ For ease of exposition, I shall employ a feature structure encoding of the segmental representation, rather than the string substitution patterns that are used in the implemented grammar, which are conceived as a variant of string unification (Calder 1989). See Copestake (2002) for details of the orthographic machinery used in DELPH-IN (<http://www.delph-in.net>) grammars and processing platforms.

represented schematically on the SEG list: the final vowel of the base is suppressed and the last root consonant is reduplicated (\overline{c}), with the vowels *o* and *i* interspersed.

Figure 7:
Morphological rule for
noun class I plural formation

$$\left[\begin{array}{l} \text{SEG} \quad \overline{s} \oplus \langle \overline{c}, o, \overline{c}, i \rangle \\ \text{SUPRA} \quad \left[\begin{array}{l} \text{TONE} \quad h^* \text{-list} \\ \text{LEN} \quad \langle \text{long}, \text{long} \mid \overline{L} \rangle \end{array} \right] \\ \text{DTR} \quad \left[\begin{array}{l} \text{SEG} \quad \overline{s} \oplus \langle \overline{c} C, V \rangle \\ \text{SUPRA} \quad \left[\begin{array}{l} \text{TONE} \quad \text{list} \\ \text{LEN} \quad \langle [] \mid \overline{L} \rangle \end{array} \right] \end{array} \right] \end{array} \right]$$

The assignment of an all-H melody to the morphological complex, however, directly makes use of the typed tone lists introduced in Section 5.1 above. Note moreover that, with tone-integrating suffixes, which indiscriminately ignore the tonal specification of the base, the tonal description of the base is highly underspecified.

Class II suffixation is a variation on the same theme (Figure 8): the base-final vowel is suppressed and replaced with *-ai*. Accordingly, the length specification of the final syllable is constrained to be short: again, this is modelled by suppressing the length specification of the final syllable of the base, together with the addition of a *short* element to that of the morphological complex.

Figure 8:
Morphological rule for noun class II
plural formation

$$\left[\begin{array}{l} \text{SEG} \quad \overline{s} \oplus \langle a, i \rangle \\ \text{SUPRA} \quad \left[\begin{array}{l} \text{TONE} \quad \langle \text{high} \mid l^* \text{-list} \rangle \\ \text{LEN} \quad \langle \text{short} \mid \overline{L} \rangle \end{array} \right] \\ \text{DTR} \quad \left[\begin{array}{l} \text{SEG} \quad \overline{s} \oplus \langle V \rangle \\ \text{SUPRA} \quad \left[\begin{array}{l} \text{TONE} \quad \text{list} \\ \text{LEN} \quad \langle [] \mid \overline{L} \rangle \end{array} \right] \end{array} \right] \end{array} \right]$$

The $L^+ \text{-H}$ melody is again assigned independently of the tonal make-up of the base. Instead of employing a tone-list constraint for the

entire list, we specify non-spreading tones at the beginning of the list (here: *high*) followed by a tone-list constraint on the remainder (*l*-list*).

5.3.2 Non-integrating suffixes

The second major type of morphological rules pertains to tone non-integrating affixes. A trivial case is that of consonantal suffixes, which do not add any tonal specification at all, as witnessed by the linker in Figure 9: while a final *-r* is added to the list of segments, and the length specification of the final syllable is adjusted to *short*,²⁵ the tonal specification of the base is merely passed on in its entirety ($\bar{[]}$).

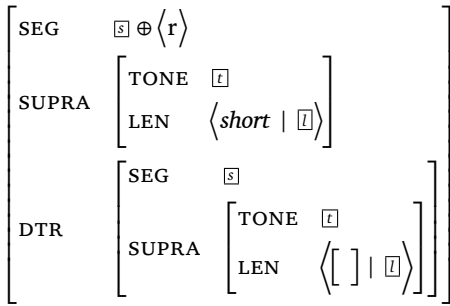


Figure 9:
Morphological rule for
“genitive” linker *-r*

A slightly more interesting example is contributed by the possessive pronominal affix, as shown in Figure 10: here, the tone of the suffix, and its length specification are added to the respective suprasegmental lists, with the TONE ($\bar{[]}$) and LEN value ($\bar{[]}$) of the base being identified only with the list remainder.

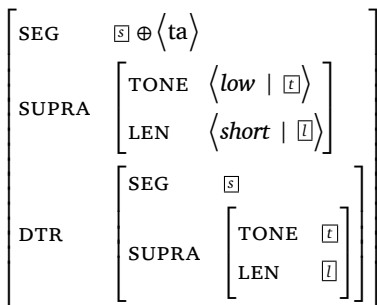


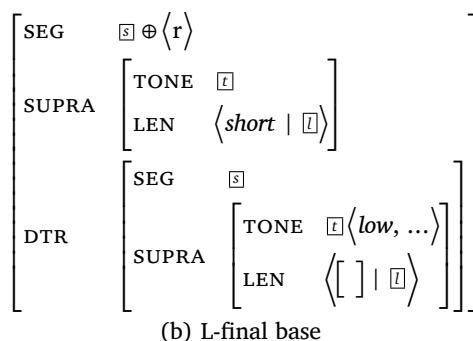
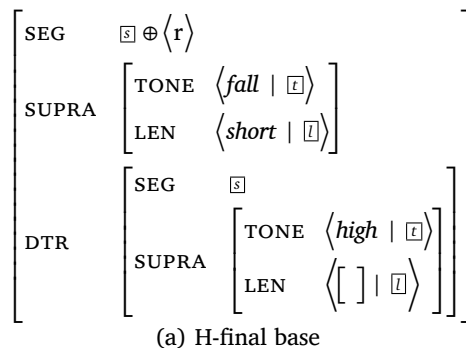
Figure 10:
Morphological rule for possessive
pronominal suffix *-tà*

The final example of a non-integrating tone pertains to the previous reference marker *-ɔ̃/-ɔ̃̀*: in terms of segmental information and

²⁵ Recall that there are no long vowels in Hausa closed syllables

length specifications, the rule in Figure 11 is identical to that for the linker in Figure 9, for obvious reasons. Tonally, however, this rule is clearly distinct: while for L-final bases the tonal specification of the base is carried along unaltered, affixation of the previous reference marker changes a final H to a fall.

Figure 11:
Morphological rules for previous
reference marker \rightarrow



Having presented morphological rules for both tone-integrating and tone non-integrating affixation, we are in a position to illustrate how these rules interact. As an example, let us consider the possessive of a class II plural noun, like *tambayōyîn* ‘the aforementioned questions’. By way of illustration, we shall embed the tone-integrating plural rule inside the non-integrating rule for the previous reference marker, as shown in Figure 12.

Starting with the outer rule of previous reference marking, we see that it constrains the final tone of the morphological complex to be *fall* (under SUPRA.TONE.HD), just in case the base is H-final (under DTR.SUPRA.TONE.HD). Similarly, it constrains the final length under

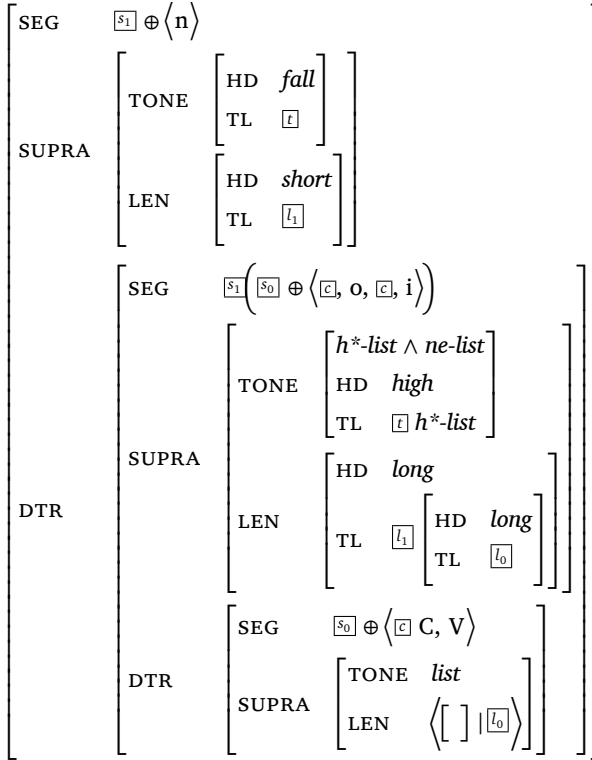


Figure 12:
Interaction between tone-integrating
and non-integrating affixation

SUPRA.LEN.HD to *short*, a specification that replaces the final length specification of the base (cf. Figure 11 for the rule in isolation). The inner rule of class I plural formation constrains its mother to be all-H: cf. the constraint *h*-list* under DTR.SUPRA.TONE. Given that the outer rule references the first element (and, for that matter, the list remainder) on this underspecified *h*-list*, the list type is coerced into the type *ne-list*. Given our hierarchy of list types in Figure 5, the most general subtype that is both of type *h*-list* and *ne-list* is *h*-ne-list*. Thus, the conjunction of two types will automatically resolve to the most general subtype of the two, if defined, or otherwise yield a unification failure. Resolution to the subtype, however, automatically enforces any further constraints associated with this type, such as the constraint that the HD value be *high* and the TL value a list of type *h*-list*, enforcing the constraint on DTR.SUPRA.TONE.TL, in our example. Since the constraint on the expanded type *h*-ne-list* is compatible with that of the

outer rule, rule application succeeds and the underspecified spreading list constraint is pushed one element down. Let us suppose that we had tried to apply the outer rule to a base specified as *l*-list*: in this case, a constraint restricting the first element of this list to *high* will equally trigger the expansion of *l*-list* to *l*-ne-list*. However, the value of HD imposed by the outer rule (*high*) will fail to unify with that of the type constraint (*low*), blocking application of the rule in this case, as desired. Thus, underspecified list constraints expand, as required, whenever any of their members are accessed. As a result, constraints on tonal identity are virtually present, without our having to keep track of the number of instances they may apply to, ultimately providing us with a very general and efficient approach.

5.4 *Prefixes and tone*

Having discussed the two major modes of operation for suffixation, we shall now turn to the more restricted cases of prefixation. First, I shall discuss how the present approach to spreading naturally extends to toneless prefixes, and then address the phenomenon of tonally pre-specified prefixes that I have glossed over in the discussion so far. Finally, I shall generalise the present approach in such a way as to permit morphological operations on tone with both prefixation and suffixation, and show how this integrates with our approach to spreading.

5.4.1 *Toneless prefixes*

I have argued in the previous section that the overwhelmingly suffixal nature of Hausa, both segmentally and suprasegmentally, favours a representation of tone and length that facilitates access on the right, and I have therefore suggested encoding both TONE and LEN lists from right to left. While this is certainly a reasonable decision, we still need to provide a solution for the few cases of prefixation that nevertheless exist in the language.

One such instance of prefixation was observed in Section 3.3: pluractionals in Hausa are formed by prefixation of a segmentally underspecified reduplicative syllable (CVC-). From a tonal perspective, we observed that pluractionals constitute another case of holistic assignment of melodies, including spreading. While prefixation of segmental material poses no problem for the formalism in use (the implementation of string unification provides both prefix and suffix con-

structs), this is not the case for suprasegmental information, such as vowel length, which is represented using a feature structure encoding of lists. As stated in Section 5.1, in lean typed feature formalisms without relational constraints, arbitrary manipulations are easy at the beginning of the list, yet harder at the end of lists of indeterminate or arbitrary length. If, however, we only need to add elements to the end of a list, unification provides a solution: *difference lists* (familiar from Prolog, for example, Clocksin and Mellish 1981, Chapter 3.8) extend the functionality of ordinary lists by maintaining a pointer to the (open) end of the list, represented here as the feature LAST. As illustrated in Figure 13, concatenation of two lists then proceeds by unifying the LAST feature of the first difference list with the LIST feature of the second list. Since the LAST feature of the first difference list is token identical with the list remainder of that difference list's LIST value, the second list will just wind up at the end of the first. In order to facilitate further list concatenation, the LAST feature of the newly formed difference list will be identical to the LAST feature of the second list.

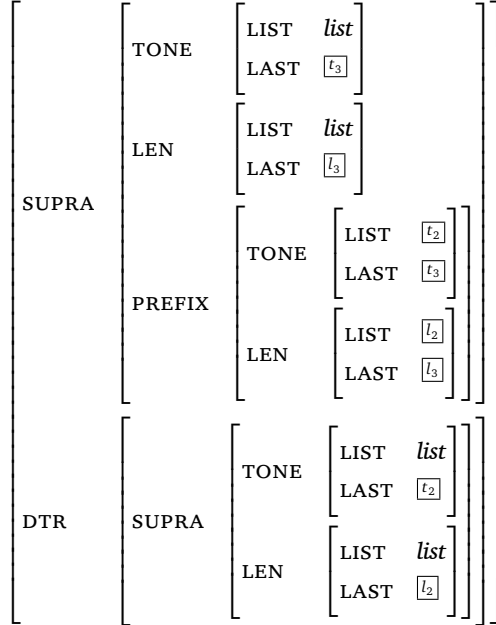
$$\begin{aligned} & \left[\begin{array}{l} \text{LIST} \langle \text{long, long} \mid \boxed{1} \rangle \\ \text{LAST} \boxed{1} \end{array} \right] + \left[\begin{array}{l} \text{LIST} \boxed{1} \langle \text{short} \mid \boxed{2} \rangle \\ \text{LAST} \boxed{2} \end{array} \right] \\ &= \left[\begin{array}{l} \text{LIST} \langle \text{long, long, short} \mid \boxed{2} \rangle \\ \text{LAST} \boxed{2} \end{array} \right] \end{aligned}$$

Figure 13:
List concatenation using
difference lists

How is this applied now to the task at hand? In order to integrate length prefixation by means of difference lists, the first thing we have to do is to provide a list representation for the suprasegmental prefix information, and a principle that ensures concatenation using unification. To this end, I shall introduce the feature PREFIX which takes as its value a suprasegmental structure consisting of TONE and LEN difference lists. In a strictly analogous fashion, I shall generalise the TONE and LEN lists under SUPRA to be difference lists. As depicted in Figure 14, concatenation of the prefixal tone and length lists can be effected by a principle that identifies the end of the suprasegmental lists of the base with the beginning of the prefixal lists. Since we still want to be able to modify, add, or delete tone and length specifications at the beginning of the suprasegmental list, we do not constrain

the beginning of the lists under SUPRA to be identical to the beginning of the lists under DTR, but rather leave this to the individual morphological rules.

Figure 14:
Concatenation of DTR and PREFIX
difference lists



Let us now consider the morphological rule for pluractionals, shown in Figure 15: apart from the prefixation of segmental material (on SEG), the rule specifies a prefixal length specification for a *short* vowel under SUPRA.PREFIX.LEN, as well as an underspecified prefixal tone.²⁶

According to the principle in Figure 14, the lists of prefix tone and syllable specifications under SUPRA.PREFIX will be appended to the lists under DTR.SUPRA, and the LAST value of the resulting lists on SUPRA will be set to the LAST values on the PREFIX lists. Since pluractional prefixation does not involve any changes at the right end of the base and, more importantly, no suprasegmental ones, we equate the beginning of the lists on SUPRA with those on BASE.

²⁶This is mainly for symmetry: since tone assignment is holistic, using open list constraints, and TONE lists are synchronised with LEN, we might just as well have specified an empty difference list here.

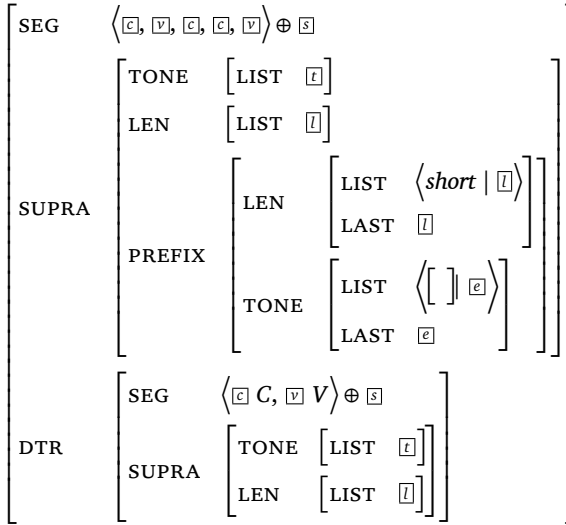


Figure 15:
Pluractional rule for trisyllabic bases

Holistic assignment of tone melodies to verbs depends essentially on the inflection class (grade), the particular paradigm cell (frame in Hausaist terminology) and the syllable count. Since prefixal syllables will wind up on the SUPRA.LEN list of pluractionals, we can select the appropriate melody, just as we would do with ordinary bisyllabic or trisyllabic verbs. Furthermore, given bijection between tone and length specifications at the word level, tone-list constraints will be expanded to exactly the right number of syllables, triggering spreading whenever the number of length (= syllable) specifications surpasses the number of tones.²⁷

5.4.2 Tone-specified prefixes

The second type of prefixation we are going to investigate concerns tone-specified prefixes, as witnessed by two derivational processes of the language: agentive nouns and ethnonyms.²⁸

²⁷ In a purely technical sense, the syllable count can of course never surpass the number of tones, owing to the fact that we consistently use tone-list constraints, which can denote lists of arbitrary length. But, for the purposes of clarity, I shall continue to use the current wording.

²⁸ I have deferred discussion of these data to this point, essentially for expository purposes: as we shall see shortly, an integrated account of prespecified prefixal tone with spreading calls for a revision of our treatment of spreading that would otherwise have been difficult to grasp.

Agentive nouns in Hausa are derived by prefixation of a high short syllable *ma*, and a gender/number marking suffix, which is *-ē* for masculine singular, *-ìyā* for feminine singular, and *-ā* for plural. The prefixal *ma-* can also be found in other deverbal nominalisations, like instrumentals (*mabùdī* ‘key/opener’) and locatives (*makērā* ‘smithy’), the latter of which, however, differ from agentives not only in the choice of the suffix, but also in terms of their tonal specification, carrying an all-H melody. Amongst the agentive nominalisations, the masculine singular and plural forms deserve particular attention: as described e.g. by Newman (2000) and Jaggar (2001), these forms are characterised by an LH melody, with the initial L spreading to the left, up to, but excluding the prefixal high *ma-*. Note that, again, we are dealing with tone-integrating affixes here, indiscriminately replacing the lexical tone specification with a holistic melody.

- (21) a. *húkùntá* ‘to judge’ – *máhùkùncí* ‘judge (M)’ –
máhùkùntá ‘judges’ (Newman 2000, p. 52)
- b. *tsòràtá* ‘be afraid’ – *mátsòràcí* ‘a coward (M)’ –
mátsòràtá ‘cowards’ (Newman 2000, p. 52)

While it is difficult to find examples of agentive nouns derived from quadrisyllabic bases, a spreading analysis of the LLH pattern observed with trisyllabic bases can instead be motivated by the tonal properties of agentive nouns with bisyllabic or monosyllabic bases.

- (22) a. *kérà* ‘to forge’ – *mákèrí* ‘smith (M)’ –
mákèrá ‘smiths’ (Jaggar 2001, p. 108/p. 13)
- b. *só* ‘to want’ – *másòyí* ‘lover (M)’ –
másòyá ‘lovers’ (Jaggar 2001, p. 108)
- c. *shá* ‘to drink’ – *máshàyí* ‘drinker (M)’ –
máshàyá ‘drinkers’ (Jaggar 2001, p. 108)

As can be witnessed in (22), agentive nouns of monosyllabic and bisyllabic bases essentially assign an LH-pattern. Under a spreading account, this pattern directly generalises to trisyllabic bases, as witnessed by (21), yielding L⁺-H.

The second piece of data illustrating spreading up to a prespecified prefixal tone is contributed by certain ethnonyms. As shown

in (23), singular ethnonyms of this type²⁹ are marked by a short low prefix *bà-* and a tone-integrating suffix *-ē* with an HL melody for masculine, and *-iyā* with HLH for feminine.

- (23) a. Fàránsà ‘France’ –
Bàfàránshè (M), Bàfàránshiyá (F), Fàránsáwá (PL)
‘French’ (Newman 2000, p. 171)
- b. Jámùs ‘Germany’ –
Bàjámúshè (M), Bàjámúshiyá (F), Jámúsáwá (PL)
‘German’ (Newman 2000, p. 171)
- (24) Dámágàráám ‘Damagaram’ –
Bàdámágàrè (M), Bàdámágàriyá (F), Dámágàráwá (PL)
‘person from Damagaram’ (Newman 2000, p. 171)

With quadrisyllabic instances of this derivational pattern, we observe an LHHL surface melody: although the medial H-H sequence in (23) may already be indicative of spreading, the pentasyllabic example in (24) shows more convincingly that we are again facing spreading to the left, up to the prespecified prefixal tone.

The phenomenon of spreading up to some prespecified tone presents a challenge to the implementation of spreading proposed so far: since constraints in the formalism are inviolable,³⁰ the *h*-list* constraint accounting for the sequence of H will inevitably impose a *high* tone specification for the prefix, leading to a unification failure. Crysmann (2009) suggested circumventing this issue by assigning non-spreading tones in these few cases. Even if not particularly elegant, a solution along these lines is nevertheless feasible, since the complexity of the tone lists is quite limited: in fact, pentasyllabic cases are quite rare – the example in (24) appears to be the only one cited by Newman (2000) – and the existing patterns could be broken down into three sub-patterns, according to syllable count. Similar observations regarding the complexity of tone patterns can be made for

²⁹ Besides the pattern discussed here, which is the more common, though still not fully productive one, there is also a non-integrating tone pattern, i.e. one where base tone is preserved. See Newman (2000, p. 142) for details.

³⁰ The LKB does provide default constraints, which could be put to use here. However, none of the efficient processing platforms, such as PET (Callmeier 2000) or Packard’s ACE (Crysmann and Packard 2012), support this.

agentive nouns. Furthermore, the formation of ethnonyms is limited in productivity, with new forms favouring an analytical construction instead:

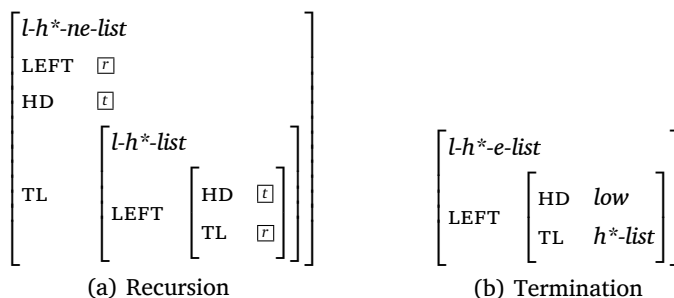
- (25) a. Pàlàsǎínù ‘Palestine’ –
 ǎán/mùtùmìn Pàlàsǎínù (M.SG), Pálàsǎínáwǎ (PL)
 ‘Palestinian’ (Newman 2000, p. 174)
- b. Bósníyà ‘Bosnia’ –
 ǎán/mùtùmìn Bósníyà (M.SG), Bósníyáwǎ (PL)
 ‘Bosnian’ (Newman 2000, p. 174)

While these observations regarding productivity and complexity are certainly valid, they can hardly obscure the fact that we would miss a generalisation here. Even if it can be shown, with respect to Hausa, that this will not lead to any serious problems regarding weak generative capacity, the current solution will certainly not scale up to other tone languages that feature both spreading and prespecified prefixal and suffixal tone.

I therefore propose a generalised approach to spreading that caters for the possibility of prespecified tone. To this end, I shall integrate into the feature structure of tone lists an additional stack of tones to be assigned from the left inwards: the feature LEFT.

As depicted in Figure 16, an H tone spreading list with a leftmost L (*l-h*-list*) does not directly specify the quality of the tone on its list elements (HD), but will rather assign the elements specified on LEFT one-by-one backwards from the end of the (primary) tone list.

Figure 16:
 H tone spreading
 with prespecified
 initial L



The prespecified tones themselves are associated with the end of the tone list: i.e. the empty list type *l-h*-e-list* specifies on the LEFT list,

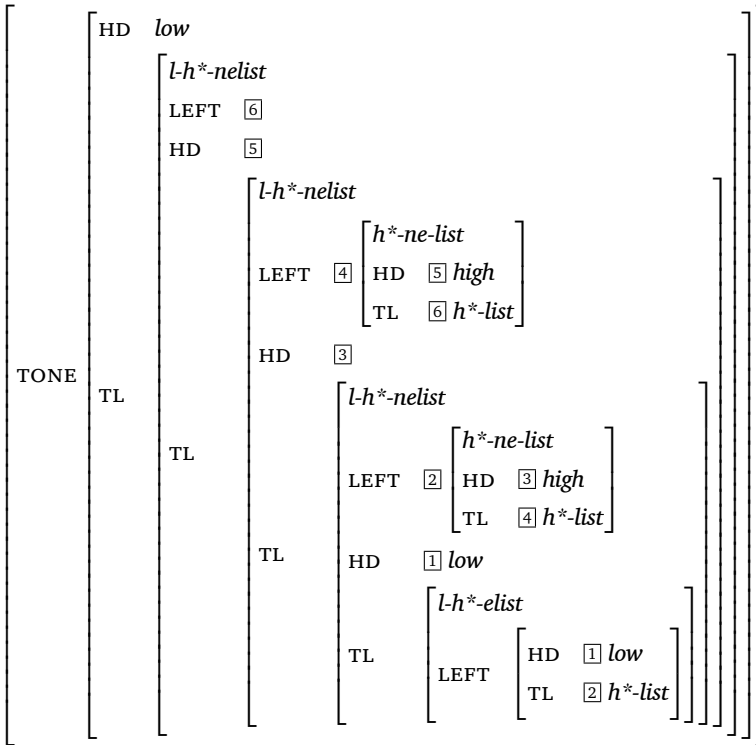


Figure 17:
Sample analysis
of L-H⁺-L
assignment to
quadrissyllabic
words

from left to right, the prefixal tones, followed by a list constraint pertaining to the spreading tone. As dictated by the constraint on the non-empty *l-h*-ne-list*, these tones are assigned one-by-one to HD, starting from the end of the tone list. Once the prefixal non-spreading tones on the LEFT list are exhausted, this constraint will insert instances of the spreading tone.

Let us briefly consider a sample analysis of a quadrissyllabic ethnonym with an L-H⁺-L melody, as shown in Figure 17: the morphological rule of ethnonym formation will assign a final L tone to the beginning of the tone list (HD), together with an *l-h*-list* constraint on the remaining tones (TL). Owing to synchronisation with the LEN list, the tone list will be constrained to be exactly 4 elements long, thus automatically triggering recursive expansion of the *l-h*-list* constraint, specialising intermediate types to *l-h*-ne-list*, and the end of the list to *l-h*-e-list*, thereby instantiating the LEFT list with $\langle low \mid h^*-list \rangle$. Based on the constraints associated with the type *l-h*-ne-list*, these

tones are inserted backwards, one-by-one on to the TONE list, triggering expansion of *h*-list* to *h*-ne-list* as necessary.

The generalised approach to spreading I have just proposed essentially synthesises the proposals by Newman (2000) and Leben (1978): while primary organisation is indeed right-to-left, including recursive propagation of tone-list constraints, actual instantiation with concrete surface tone constraints proceeds left-to-right. Thus, rather than stipulating a universal direction of assignment, the formal means offered by a lean unification formalism inherently constrain assignment to an outside-in regime.

5.5 *The expressive power of tone-list constraints*

We have seen so far how a purely unification-based approach to tone assignment and spreading by means of tone-list constraints can provide for an efficient implementation of morphological tone within the context of a computational grammar of Hausa. In this section, we shall investigate now on a more abstract level what functionalities the current approach provides, in general, and try to assess to what degree this approach may scale up to the computational treatment of morphological tone in general.

Owing to the pure feature structure encoding of lists, non-agglutinative operations on tone and length, i.e. deletion or modification, are generally easier to specify at the beginning of the list, rather than at its end, thereby favouring right-to-left encoding for predominantly suffixal, and left-to-right encoding for predominantly prefixal languages. Furthermore, we have established that prefixation of syllable length specifications in a suffixing language can be integrated with an overall right-to-left organisation, by means of difference lists. In the absence of spreading, the difference list approach could be straightforwardly applied to tone as well. If, however, as we have seen in Hausa, such prefixal tones are prepended to spreading tones, it is vital to feed these prefixal tone specifications as floating tones of a generalised tone spreading constraint.

Having now established that suprasegmental representations can be augmented at either end, one aspect in need of further elaboration is the question of “feature-changing” operations: while alignment with a preference regarding morphological composition will ensure

that the majority of such operations will be supported by the chosen direction of suprasegmental encoding, the question still remains as to what degree such operations will be possible at the disfavoured end. Essentially there are two answers to this question: strictly regarding the treatment of prefixation formulated above, where contributions on the PREFIX difference lists are *immediately* concatenated with the suprasegmental lists of the base, subsequent feature-changing operations targeting the prefix will be inexpressible. However, a simple change regarding the principle governing list composition will greatly extend the expressive power of the current approach: instead of immediately concatenating prefixal tones, we can delay concatenation up to the next prefixal morphological rule. Once we do this, any immediately subsequent prefixation rule will have a privileged data structure to operate on, namely the suprasegmental specification of the previously attached prefix, which can then be modified or deleted, if appropriate.

To summarise, typed list constraints provide a powerful mechanism to incorporate analyses inspired by autosegmental phonology into HPSG-style computational grammars, building solely on unification. The constraints and representations I have proposed in this paper permit automatic spreading, as well as addition, modification and subtraction of tonal material at both the right and left end of the tonal representation. Crucially, this formalisation differs from traditional autosegmental approaches in two respects: first, descriptions are surface-oriented. Thus, instead of pre-linking, delinking, and re-spreading, I use direct specification of tones and melodies, plus local modification by morphological rules. Second, the association convention used here is a simple bijective relation between tones and tone-bearing units. As a consequence, there is no double association, but contour tones are represented as such. Generalisation over tone classes, e.g., *fall* and *low*, can instead be achieved by means of a tonal type hierarchy.

There is, however, one remaining limitation: owing to the monotonic nature of the constraint formalism, there can only ever be one spreading tone within a morphological domain. In Hausa, plurals of augmentative adjectives provide a case where more than one indisputable instance of spreading exists within a morphological complex.

- (26) a. *táŋkwálélè* (SG.M) – *táŋkwálá-táŋkwàlà* (PL)
‘large and round’ (Newman 2000, p. 25)
- b. *bùllúkí* (SG.M) – *bùllúká-bùllúkà* (PL)
‘huge’ (Newman 2000, p. 25)
- c. *tsálélè* (SG.M) – *tsálá-tsàlà* (PL)
‘tall and slender’ (Newman 2000, p. 76)

As illustrated in (26), these plurals are formed using total reduplication, assigning an H⁺ tonal pattern to the base on the left and an L⁺ pattern to the reduplicant on the right.

However, a purely phonological approach to reduplication is already computationally expensive, independently of the question of tone assignment: in contrast to partial reduplication, there is no principled upper bound on the number of cross-serial dependencies between segments. Thus, from a purely segmental perspective, it is advisable to approach such formations in terms of a compound structure involving two like bases. If, however, total reduplication is best viewed as involving compounding of two minimal morphological words, the instance of multiple spreading within the larger morphological complex can be broken down into two independent spreading processes within the two minimal morphological words that contribute to the compound structure. It remains to be seen whether there are languages (other than Hausa) that feature unmistakable instances of multiple spreading, i.e. extended sequences of like tones that are demonstrably independent of syllable count, outside the domain of total reduplication, or, more generally, outside morphologically compound structures.

6

CONCLUSION

In this paper, I have argued for a highly efficient encoding of suprasegmental information for the treatment of tone and length in a computational grammar of Hausa using typed list constraints. Investigating a range of morphological processes in the language, I have shown that manipulations of length and segmental material are of a highly local nature, as opposed to tonal operations that require non-local specification of melodies, including automatic spreading, in addition to agglutinative processes. I have argued that typed list constraints provide an efficient way of underspecifying spreading, and I have shown

how “feature-changing” processes, such as modification and deletion, as well as prespecified tones, can be integrated with monotonic list constraints, thus covering the full range of phenomena pertaining to lexical and grammatical tone in Hausa.

REFERENCES

- Mahamane L. ABDOULAYE (1992), *Aspects of Hausa Morphosyntax in Role and Reference Grammar*, Ph.D. thesis, SUNY Buffalo, NY.
- Peter ADOLPHS, Stephan OEPEN, Ulrich CALLMEIER, Berthold CRYSMANN, Dan FLICKINGER, and Bernd KIEFER (2008), Some Fine Points of Hybrid Natural Language Parsing, in Nicoletta CALZOLARI, Khalid CHOUKRI, Bente MAEGAARD, Joseph MARIANI, Jan ODIJK, Stelios PIPERIDIS, and Daniel TAPIAS, editors, *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)*, European Language Resources Association (ELRA), Marrakech, Morocco, ISBN 2-9517408-4-0, <http://www.lrec-conf.org/proceedings/lrec2008/>.
- Steven BIRD and Ewan KLEIN (1994), Phonological Analysis in Typed Feature Systems, *Computational Linguistics*, 20:455–491.
- Jonathan CALDER (1989), Paradigmatic Morphology, in *Fourth Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, pp. 58–65, Association for Computational Linguistics, Manchester.
- Ulrich CALLMEIER (2000), PET – A Platform for Experimentation with Efficient HPSG Processing Techniques, *Journal of Natural Language Engineering*, 6(1):99–108.
- Bernard CARON (1991), *Le haoussa de l'Ader*, Dietrich Reimer, Berlin.
- Bernard CARON and A. H. AMFANI (1997), *Dictionnaire français-haoussa: suivi d'un index haoussa-français*, IFRA-Ibadan, Ibadan.
- Bob CARPENTER (1992), *The Logic of Typed Feature Structures with Applications to Unification-based Grammars, Logic Programming and Constraint Resolution*, volume 32 of *Cambridge Tracts in Theoretical Computer Science*, Cambridge University Press, New York.
- Noam CHOMSKY and Morris HALLE (1968), *The Sound Pattern of English*, Harper and Row.
- George Nicholas CLEMENTS (1985), The Geometry of Phonological Features, *Phonology Yearbook*, 2:223–252.
- George Nicholas CLEMENTS and Kevin C. FORD (1979), Kikuyu Tone Shift and its Synchronic Consequences, *Linguistic Inquiry*, 10:95–108.

- William F. CLOCKSIN and Christopher S. MELLISH (1981), *Programming in Prolog*, Springer, Heidelberg, 5th edition.
- Ann COPESTAKE (2002), *Implementing Typed Feature Structure Grammars*, CSLI Publications, Stanford.
- J. Ronayne COWAN and Russell SCHUH (1976), *Spoken Hausa*, Spoken Language Services, Ithaca.
- Berthold CRYSMANN (2005a), An Inflectional Approach to Hausa Final Vowel Shortening, in Geert BOOIJ and Jaap VAN MARLE, editors, *Yearbook of Morphology 2004*, pp. 73–112, Kluwer.
- Berthold CRYSMANN (2005b), Syncretism in German: a Unified Approach to Underspecification, Indeterminacy, and likeness of Case, in Stefan MÜLLER, editor, *Proceedings of the 12th Int'l Conference on Head-driven Phrase Structure Grammar (HPSG), Aug 22–24, Lisbon*, pp. 91–107, CSLI publications, Stanford.
- Berthold CRYSMANN (2009), Autosegmental Representations in an HPSG for Hausa, in *Proceedings of the ACL-IJCNLP workshop on Grammar Engineering Across Frameworks (GEAF 2009)*, pp. 28–36, ACL.
- Berthold CRYSMANN (2011), A Unified Account of Hausa Genitive Constructions, in Philippe DE GROOTE, Markus EGG, and Laura KALLMEYER, editors, *Formal Grammar. 14th International Conference, FG 2009, Bordeaux, France, July 25-26, 2009, Revised Selected Papers*, volume 5591 of *Lecture Notes in Artificial Intelligence*, pp. 102–117, Springer.
- Berthold CRYSMANN (2012a), HaG — an Implemented Grammar of Hausa, in Michael R. MARLO, Nikki B. ADAMS, Christopher R. GREEN, Michelle MORRISON, and Tristan M. PURVIS, editors, *Selected Proceedings of the 42nd Annual Conference on African Linguistics (ACAL 42)*, pp. 321–337, Cascadilla Press, Somerville, MA.
- Berthold CRYSMANN (2012b), On the Categorical Status of Hausa Genitive Prepositions, in Bruce CONNELL and Nicholas ROLLE, editors, *Selected Proceedings of the 41st Annual Conference on African Linguistics (ACAL 41)*, pp. 29–39, Cascadilla Press, Somerville, MA.
- Berthold CRYSMANN (2012c), Resumption and Island-hood in Hausa, in Philippe DE GROOTE and Mark-Jan NEDERHOF, editors, *Formal Grammar. 15th and 16th International Conference on Formal Grammar, FG 2010 Copenhagen, Denmark, August 2010, FG 2011 Ljubljana, Slovenia, August 2011*, volume 7395 of *Lecture Notes in Computer Science*, pp. 50–65, Springer.
- Berthold CRYSMANN (2015), Resumption and Extraction in an Implemented HPSG of Hausa, in *Proceedings of the ACL-IJNLP workshop on Grammar Engineering Across Frameworks (GEAF-2015), Beijing, China, ACL*.
- Berthold CRYSMANN and Woodley PACKARD (2012), Towards Efficient HPSG Generation for German, a Non-configurational Language, in *Proceedings of the*

24th International Conference on Computational Linguistics (COLING 2012), pp. 695–710, Mumbai, India.

Daniel P. FLICKINGER (2000), On Building a More Efficient Grammar by Exploiting Types, *Natural Language Engineering*, 6(1):15–28.

John A. GOLDSMITH (1976), *Autosegmental Phonology*, Ph.D. thesis, MIT.

Sharon INKELAS and William R. LEBEN (1990), Where Phonology and Phonetics Intersect: The Case of Hausa Intonation, in Mary E. BECKMAN and John KINGSTON, editors, *Between the Grammar and the Physics of Speech*, Papers in Laboratory Phonology, pp. 17–34, Cambridge University Press, New York.

Philip JAGGAR (2001), *Hausa*, John Benjamins, Amsterdam.

Herrmann JUNGRAITHMAYR and Wilhelm J. G. MÖHLIG (1976), *Einführung in die Hausa-Sprache: Kursus für Kolleg und Studium*, Reimer, Berlin.

Jean-Pierre KOENIG (1999), *Lexical Relations*, CSLI publications, Stanford.

Hans-Ulrich KRIEGER (1996), *TDL — A Type Description Language for Constraint-Based Grammars*, volume 2 of *Saarbrücken Dissertations in Computational Linguistics and Language Technology*, DFKI GmbH, Saarbrücken.

Alex LASCARIDES, Ted BRISCOE, Nicholas ASHER, and Ann COPESTAKE (1996), Order Independent and Persistent Typed Default Unification, *Linguistics and Philosophy*, 19(1):1–90.

Alex LASCARIDES and Ann COPESTAKE (1999), Default Representation in Constraint-Based Frameworks, *Computational Linguistics*, 25:55–105.

William R. LEBEN (1971), The Morphophonemics of Tone in Hausa, in C.-W. KIM and Herbert STAHLKE, editors, *Papers in African Linguistics*, pp. 201–218, Linguistic Research, Edmonton.

William R. LEBEN (1973), *Suprasegmental Phonology*, Ph.D. thesis, MIT.

William R. LEBEN (1978), The Representation of Tone, in Victoria FROMKIN, editor, *Tone: A Linguistic Survey*, pp. 177–219, Academic Press, New York.

Walt Detmar MEURERS (2001), On Expressing Lexical Generalizations in HPSG, *Nordic Journal of Linguistics*, 24(2):161–217.

Paul NEWMAN (1986), Tone and Affixation in Hausa, *Studies in African Linguistics*, 17(3):249–267.

Paul NEWMAN (2000), *The Hausa Language. An Encyclopedic Reference Grammar*, Yale University Press, New Haven, CT.

Paul NEWMAN and Philip JAGGAR (1989), Low Tone Raising in Hausa: A Critical Assessment, *Studies in African Linguistics*, 28(3):227–252.

Paul NEWMAN and Roxana MA NEWMAN (1977), *Modern Hausa–English Dictionary*, University Press, Ibadan and Zaria, Nigeria.

- Stephan OEPEN, E. CALLAHAN, Daniel FLICKINGER, Christopher MANNING, and Kristina TOUTANOVA (2002), LinGO Redwoods: A Rich and Dynamic Treebank for HPSG, in *Beyond PARSEVAL. Workshop at the Third International Conference on Language Resources and Evaluation, LREC 2002*, Las Palmas, Spain.
- Fred W. PARSONS (1960), The Verbal System in Hausa, *Afrika und Übersee*, 44:1–36.
- Gerald PENN (2004), Balancing Clarity and Efficiency in Typed Feature Logic Through Delaying, in *Proceedings of the 42nd Meeting of the Association for Computational Linguistics (ACL'04), Main Volume*, pp. 239–246, Barcelona, Spain, <http://www.aclweb.org/anthology/P04-1031>.
- Carl POLLARD and Ivan SAG (1987), *Information-Based Syntax and Semantics*, volume 1, CSLI, Stanford.
- Carl POLLARD and Ivan SAG (1994), *Head-Driven Phrase Structure Grammar*, CSLI and University of Chicago Press, Stanford.
- Susanne RIEHEMANN (1998), Type-Based Derivational Morphology, *Journal of Comparative Germanic Linguistics*, 2:49–77.
- Russell SCHUH (1989), The Reality of ‘Hausa Low Tone Raising’: A Response to Newman & Jaggar, *Studies in African Linguistics*, 20:253–262.
- James M. SCOBIE (1993), Constraint Violation and Conflict from the Perspective of Declarative Phonology, *Canadian Journal of Linguistics*, 38:155–167.
- Ekkehard WOLFF (1993), *Referenzgrammatik des Hausa*, LIT, Münster.

This work is licensed under the Creative Commons Attribution 3.0 Unported License.
<http://creativecommons.org/licenses/by/3.0/>

