Chapter 8

On the syntax-prosody interface in Wushi (Babessi): Tone pattern, dissimilation and spreading

Liliane Hodieb
INALCO, Paris

The interaction between syntax and prosody has been widely studied for Narrow Bantu languages, but quite less for Bantoid languages. Yet, the syntax-prosody interface informs us on significant principles underlying the organisation of language systems. In the light of Match Theory (Selkirk 2011), in this paper we analyze the syntax-prosody interface of Wushi, an under-documented Ring Grassfields Bantu language. We identify and discuss two major processes characterizing it: tone dissimilation and tone spreading.

1 Introduction

One of the essential arguments of the prosodic structure theory is the independence of phonological domain structure from the syntactic structure (Selkirk 2011: 440). In other words, there is no inherent relation between prosodic category types found in phonological representation and the category types of syntactic representation (ibid. 436), which each have distinct properties. However, Match Theory (Selkirk 2011), claims that "in the ideal case, the grammar allows the fundamental syntactic distinctions between clause, phrase, and word to be reflected in, and retrieved from, the phonological representation". This means that prosodic domains mirror syntactic constituents, at least *ideally*, although several cases of mismatches are attested in languages. This non-isomorphism between prosodic domains and syntactic constituents are generally due to the high

Liliane Hodieb. 2024. On the syntax-prosody interface in Wushi (Babessi): Tone pattern, dissimilation and spreading. In James Essegbey, Brent Henderson, Fiona McLaughlin & Michael Diercks (eds.), *Pushing the boundaries: Selected papers from the 51–52 Annual Conference on African Linguistics*, 163–184. Berlin: Language Science Press. DOI: 10.5281/zenodo.14038745

ranking of phonological markedness constraints on prosodic structure which are peculiar to each language.

The study of the syntax-prosody interface has attracted increasing interest among linguists and bantuists in particular. We can mention Cheng & Downing (2007) on Zulu, Hyman & Katamba (2010) on Luganda, Selkirk (2011) on Xitsonga, and Kula & Bickmore (2015) on Bemba, to name a few works on Narrow Bantu languages. On the other hand, as for Bantoid and Grassfields Bantu in particular, studies of the syntax-prosody interface are still rudimentary. This paper provides a description and an analysis of the interaction between syntax and prosody in Wushi, a Ring Grassfields Bantu language of Cameroon. Based on Match Theory, we will focus on simple indicative sentences and modal constructions expressing possibility, which seem to be most significant for this study. We begin by presenting lexical tones in Wushi in §2. Then, in §3 we examine indicative and subjunctive clauses, and more specifically, how prosodic domains interact with syntactic constituents, how well Match constraints are met and which language-specific constraints violate them. The question is also raised as to whether it is not rather the prosodic structure that determines the syntactic constituency in Wushi.

2 The lexical tone system

First of all, it should be noted that we distinguish between (individual) tones and tone patterns. Tone patterns are those associated with morphemes, nominal or verbal, in their underlying representation (UR), and they may vary on the surface.

Underlyingly, Wushi has four tones: two level tones H (high) and L (low) inherited from the proto language – they were reconstructed for Proto-Grassfields (Watters 2003: 236) – and two contour tones that were phonologized in the language, LH (rising) and HL (falling). Surface tones are H, L, M (mid), HL (falling), LH (rising), and HM, plus downstepped $^{!}$H, $^{!}$M and $^{!}$HM. As for tone patterns, I identify two categories: simple tone patterns and complex tone patterns.

2.1 Simple tone patterns

Simple tone patterns consist of only associated tones, i.e., they do not contain any floating tone. They are H, L, HL, and LH. Table 1 provides some examples.

The nouns in the table belong to classes with a zero affix (class 1 for ‘chief’ and ‘man’, class 3 for ‘wood’ and ‘potato’). But even with nouns taking an affix, the tone patterns do not change; this is because tone patterns concern noun roots,

Table 1: Simple tone patterns

Tone	UR	SR	Gloss
H	ŋká	ŋká	wood
L	fùà	fùà	chief
HL	mbə̂	mbə̂	man
LH	ntă	ntă	potato

not including affixes which are underlyingly toneless and copy their tone from the noun root. Assuming toneless noun class affixes differs from the analysis generally adopted for other Grassfields languages which is that tone on noun class affixes is high (Hyman 1980a, Akumbu & Chibaka 2012). Also, the realization of tone patterns may vary on the surface, depending on the tonal environment. The sentences in (1) and (2) illustrate this with the nouns *fùà* 'chief' and *ŋká* 'wood'.

(1) fùá wɔ́ dʒʉ̀:
fùà wɔ́ dʒʉ̀
chief POT come
'Will the chief come?'

(2) mbɔ̀ mɛ̂ ndǒ ŋkâʔ mɛ̂
mbə̂ mɛ̂ ndò ŋká mɛ̂
man DET bite wood DET
'The rat bit the wood.'

In sentence (1), the second mora of the diphthong in *fùà* 'chief' assimilates the high tone of the potential marker and becomes high. This is a general rule in the language, that the closest mora to the potential marker assimilates its tone. The mora is the tone bearing unit in the language. Concerning interrogative sentences, they always end in a L tone. As a matter of fact, when the syntax remains SVO, what indicates the interrogative 'force' of the sentence in the absence of an interrogative word like 'what' is the final L coupled with a lengthening of the vowel bearing it. Declarative sentences on the other hand are not restricted in this regard as they can end in any tone pitch.

In sentence (2), the noun *ŋká* 'wood' is underlyingly H, yet surfaces as HL. We also see that the glottal stop is inserted on the surface to emphasize syllable demarcation. This glottal stop may be responsible for the fall of the tone, but more investigation is needed to confirm it.

In Wushi like in all the other Ring languages, the mid tone is phonetic. In Wushi particularly, it has several sources which are: either a L̰[1] tone raised to M by a preceding H tone in a HL̰ pattern or a L raised to M by a following H̰ in a LH̰ pattern. M in these two contexts – illustrated in Table 2 – could also be interpreted as the fusion of L and H. And, as shown in (3) below, M can also be the realization of the underlying pattern LL̰, but this is much rarer.

Another important phenomenon in Wushi and the whole Grassfields is downstep, the lowering of a H triggered by an adjacent L (or L̰) tone. In Wushi downstep is much more frequent in the context of L̰ than with a phonetically realized L. Moreover, downdrift, the gradual declination of H tones in a sequence so that the last H has a lower pitch than the first H, is also recurrent. For example:

(3) ŋə́ tí˩ twɔ́˩ dʒə̀̀
ŋə́ tí twɔ́ dʒə̀̀
3SG NEG yet come
'He has not yet come.'

The first H, the one on the third person singular pronoun, is a "canonical" H. Then the second H is realized on a lower pitch and the third one a little lower again than the second H. We will see in the rest of the discussion that downdrift is in fact conditioned by a fundamental principle of the language system. But before looking at that, we present complex tone patterns in §2.2.

2.2 Complex tone patterns

Complex tone patterns are those that comprise an associated or non-floating tone and a floating tone. The floating tone is always at the final position and so follows the associated tone. The reason for this is historical: Grassfields Bantu languages have lost the final syllable of Proto-Bantu nouns, leading to the change *CVCV > CV. While disappearing, the second syllable left its tone behind, unassociated. This unassociated or floating tone generally manifests itself on the surface as a mid tone, or causes the preceding tone to surface as mid. I identified three complex tone patterns in Wushi: HL̰, LH̰, and LL̰. Their behaviour is portrayed in the following examples.

We can see from Table 2 that the realisation of the floating tone varies. The floating L̰ in *ndó* 'husband' joins the preceding H which raises it to M and both eventually form a contour tone surfacing as HM. So, HL HM. In the LL̰ pattern, the floating L̰ causes the preceding L to manifest as M in *ʃī* 'grave' but does not

[1]The diacritic (circle) under the tone means the tone is floating.

Table 2: Complex tone patterns

Tone pattern	HL̩	LH̩	LL̩
SR	ndó	ntɔ̄ːkə̄	ʃī
UR	ndó L̩	ntɔ̀ː H̩ -kə	ʃì L̩
Gloss	'husband'	'elephant'	'grave'

associate. Thus, the change from LL̩ may be a way to ascertain the presence of a floating tone, here L̩. This is the only instance of such a type we encountered in our data.[2] Concerning the word for 'elephant', the floating H̩ raises the preceding L to M and does not associate. The noun class suffix which, as said earlier, is underlyingly toneless, then copies the mid tone of the noun root. The UR forms of the nouns in Table 2 correspond to their citation forms. Their tonal underlying representation (or tone pattern) was identified following Snider (2018) methodology, based on a major principle which is that tones are more about patterns instead of individual TBUs. In this method, Snider proposes that we begin by analyzing tone patterns in simple CV words (nouns and verbs) having different tones; then we place them in identical environments such as plural form and associative construction where we examine their tonal behaviour. Words that display identical tones in all of these contexts, even if their tones may differ in isolation, have the same tone pattern underlyingly. Then we can proceed with more complex (disyllabic and polysyllabic) words.

3 Prosodic domains vs. syntactic constituents: do they really match?

Two fundamental properties distinguish prosodic from syntactic structure representation: they are the prosodic hierarchy and strict layering (Selkirk 2011: 437).

The prosodic hierarchy is a set of prosodic categories (prosodic domains above the word, influenced by syntactic structure as well as rhythm (Elfner 2018: 3)) which are, from the largest to the smallest:

[2] I have been collecting the data from a 24-year-old native speaker of Wushi since 2018. The civil war going on in the Anglophone region of Cameroon (where Wushi and all Ring languages are spoken) has made the field inaccessible since 2016 and forced thousands of people to leave the region and be dispersed to safer places of the country and also in Nigeria. Therefore, it literally takes a miracle to find native speakers.

- Intonational Phrase (IP, ι)
- Phonological Phrase (PP, φ)
- Prosodic Word (PW, ω)

In Match Theory, prosodic structure is derived from a set of MATCH constraints, "which call for correspondence between syntactic constituents (word, phrase, clause) and prosodic domains (ω, φ, ι)" (Elfner 2018: 5). The mapping constraints can be represented as follows:

(4) MATCH constraints (Ibid.)
MATCH-CLAUSE: syntactic clause ⟶ intonational phrase (ι)
MATCH-PHRASE: syntactic phrase ⟶ phonological phrase (φ)
MATCH-WORD: syntactic word ⟶ prosodic word (ω)

Thus, this theory predicts that every syntactic constituent corresponds to a prosodic domain. More specifically, following the indirect reference theory which underlies Match Theory, syntactic structure is first mapped to the prosodic structure, a separate representation to which phonological processes apply, which produce or determine phonological phrases. However, the high ranking of language-particular phonological markedness constraints on prosodic structure may result in mismatches between syntactic and prosodic structures. Our purpose in this paper is first, to identify the phonological processes that define phonological phrases in Wushi, and second, to analyze the relationship between syntactic constituents and prosodic domains, as well as potential misalignments.

Before tackling those issues, we should briefly discuss Wushi syntax. Grassfields Bantu languages are known to be quite isolating, compared to Narrow Bantu languages which are rather agglutinating. This is particularly evident in the verb morphology. Indeed, in Bantu languages the verb consists of a root to which prefixes and suffixes are attached, which represent tense, aspect, mood, negation, evidentiality, indexes of agreement with the subject and object, etc. Many of these morphemes have a V (vocalic) shape and for some others, CV. Nurse (2008: 40) represents the structure of the Bantu verb as follows:

(5) Pre-SM + SM + NEG2 + TA + OM + root + extension + FV + post-FV

In Grassfields Bantu languages, the situation is quite different. Almost all verbal morphemes are independent morphemes with a CV shape. So that the structure of the verb phrase or the sentence exhibits the syntax as in Narrow Bantu, the main difference being that constituents agglutinate on the verb root in one

case, whereas they are separate elements in the other. The following sentences taken from Aghem (Hyman 1979: 71), Babanki (Akumbu & Chibaka 2012: 144) and Wushi (personal notes) illustrate the isolating tendency of (Ring) Grassfields languages. Note that for Aghem, the author does not provide the UR of sentences. The numbers indicate the noun class.

(6) bvʉ́ ʻtɨ́ mâa zɨ́ bɛ́-kɔ́ à lím-ghɔ́
bvʉ́ ʻtɨ́ mâa zɨ́ bɛ́-kɔ́ à lím-ghɔ́
dog SM P1/FOC eat fufu and yams
ʻThe dogs *did* eat fufu and yams.ʼ

(7) nyámsə́ pfɨ̌ʼə́ ágə́ŋ
nyàm.sə́ sə́ pfɛ́ʼ ə à.gə́ŋ
animal.10 10.SM chew PROG 06.grass
ʻCattle eat/are eating grass.ʼ

(8) vī ŋɔ̄ nɔ̌ nâ tʃá
vì ŋ.ɔ́ nɔ̌ nâ tʃá
wife 1.POSS.1SG PROG cook fish
ʻMy wife is cooking fish.ʼ

As we can see, tense-aspect markers and subject/object-verb agreement[3] markers are not affixes attached to the verb root but they are rather independent morphemes. Another notable point is the absence of subject/object-verb agreement in Wushi, which makes the order or position of the constituents of the sentence the only indicator of syntactic relationships in the language. The scarcity of concord marking is a general feature of Wushi, not only in the verbal system but also in the nominal system. With noun phrases, concord is evident, at least in a shape similar to what obtains in Narrow Bantu languages, only in possessive and demonstrative phrases. In the other types of modification (especially adjective and genitive) something quite unusual happens: the noun class suffixes – among the nine noun classes six are suffixes – display an exceptional mobility, in most cases, leaving the head noun to appear on the modifier or at the right end of the phrase, after the determiner which is invariable, i.e., it does not strictly speaking agree with the noun it modifies. In some instances of genitive phrase, the two nouns even exchange their suffix. For lack of space, and because this is not the focus of our discussion here, we will simply mention two examples in order to give a glimpse of the phenomenon. The first example is an adjectival phrase and the second is a genitive phrase.

[3]I use "agreement" and "concord" interchangeably.

(9) ntsə́ fə́-mə́ mə̂
ntsə́-mə́ fə́ mə̂
blood-6a black DET
'the black blood'

(10) núà wùfúá mə̂ -fə̀
núá-fə̀ wù-fúá mə̂
bird-19 2-chief DET
'the bird of the chiefs'

In (9), the suffix of the head noun *ntsə́-mə́* 'blood' (class 6a and sometimes class 19) attaches to the adjective *fə́* 'black' which is followed by the determiner. In the genitive phrase, the head noun suffix moves right to the end of the phrase, after the determiner. These two examples could be interpreted as two instances of the different "stops" made by the head noun suffix. All things being equal, however, it is quite patent that the preferred direction of noun class suffixes in noun phrases is rightward and probably this very movement actually indicates concord.

3.1 Indicative sentences

Concerning the actual syntactic structure of sentences in the indicative mood, they are made up of the subject (S), followed by the negative marker (NEG), aspect marker (A) – unlike the other Ring languages some of which have up to seven tense markers, tense is not morphologically marked in Wushi: it is an aspect-prominent language, probably the only one of this type in the group, at least from the data on Ring languages available so far (the aspectual system of Wushi is described in Hodieb 2021) – the verb root (V) that may itself be followed by the object(s) (O). Schematically:

(11) S + NEG + A + V + O

The majority of sentences we will be using for the argumentation have either the first person (*mə̀*) or the third person (*ŋə́*) singular pronoun as subject. The latter bears a L tone in the object position, whereas the former does not change from subject to object position.

3.1.1 Sentences involving one aspectual-modal marker

To begin with, the perfective marker is the only non-segmental aspect marker: it consists of a LH tone ascribed to the verb root, as seen in the following examples.

Table 3: Sentences with perfective

UR	SR	Gloss
ŋə́ jɔ̀ʔ	ŋə́ jɔ̌ʔ	'He sings.'
ŋə́ tʃɔ̀ʔ	ŋə́ tʃɔ̌ʔ	'He speaks.'
ŋə́ tàʔ ŋgɔ̀ːkə	ŋə́ tǎʔ ŋgɔ̀ːkə̀	'He wants the banana.'
ŋə́ zə́ ŋgɔ̀ːkə	ŋə́ zə̌ ŋgɔ̀ːkə̀	'He eats the banana.'
ŋə́ jɛ́ ŋgɔ̀ːkə	ŋə́ jɛ̌ ŋgɔ̀ːkə̀	'He sees the banana.'
ŋə́ bé	ŋə́ bě	'He sleeps.'
H H/L (L)	H LH (L)	

The first three verb roots are underlyingly L, while the others are underlyingly H. Nevertheless, they all surface with LH in the perfective. This shows us that the perfective marker is very steady. Also, it does not fluctuate, whatever the tonal or morphosyntax environment.

Let us now consider the negative forms of these sentences.

Table 4: Sentences with perfective and negation

UR	SR	Gloss
ŋə́ tí jɔ̀ʔ	ŋə́ tī jɔ̀ʔ	'He does not sing.'
ŋə́ tí tʃɔ̀ʔ	ŋə́ tī tʃɔ̀ʔ	'He does not speak.'
ŋə́ tí tàʔ ŋgɔ̀ːkə	ŋə́ tī tàʔ ŋgɔ̀ːkə̀	'He does not want the banana.'
ŋə́ tí zə́ ŋgɔ̀ːkə	ŋə́ tī zə̄↓ ŋgɔ̀ːkə̀	'He does not eat the banana.'
ŋə́ tí jɛ́ ŋgɔ̀ːkə	ŋə́ tī jɛ̄́ ŋgɔ̀ːkə̀	'He does not see the banana.'
ŋə́ tí bé	ŋə́ tī bē↓	'He does not sleep.'
H H L/H (L)	H M L/M↓ (L)	

In Table 4, we go from H H H to H M M↓ on one hand and from H H L to H M L on the other hand – we are omitting the object here. The subject pronoun being always the same, *ŋə́* with H, what differentiates the two patterns is actually the tone of the verb. First, the H of the negative marker has to distinguish from the H of the subject by becoming M, and second, the verb which was initially H just like the negative marker, in its turn distinguishes from the preceding negative marker now bearing M, by becoming M↓. So, we see that the tonal interaction

from the subject to the verb is one where each element adjusts itself according to what precedes. More precisely, the tone of the negative marker turns from H to M because of the H of the preceding subject pronoun from which it must differ. Indeed, in the presence of a L tone pronoun instead, the negative marker keeps its H tone and does not change to M as seen in (12):

(12) wà tí nɔ̌ jʉ̀ fá
wà tí nɔ̌ jʉ́ fá-kə
1PL NEG PROG do thing-7
'We are doing nothing.'

With the distal marker which is *kə̀* bearing a L tone underlyingly, the verbs in the affirmative in (4) behave as in Table 5.

Table 5: Sentences with distal

UR	SR	Gloss
ŋə́ kə̀ jɔ̀ʔ	ŋə́ kə̀ jɔ̌ʔ	'He sang.'
ŋə́ kə̀ tʃɔ̀ʔ	ŋə́ kə̀ tʃɔ̌ʔ	'He spoke.'
ŋə́ kə̀ tàʔ ŋgɔ̀ːkə	ŋə́ kə̀ tǎʔ ŋgɔ̀ːkə	'He wanted the banana.'
ŋə́ kə̀ zə́ ŋgɔ̀ːkə	ŋə́ kə̀ zə́↓ ŋgɔ̀ːkə	'He ate the banana.'
ŋə́ kə̀ jɛ́ ŋgɔ̀ːkə	ŋə́ kə̀ jɛ́↓ ŋgɔ̀ːkə	'He saw the banana.'
ŋə́ kə̀ bé	ŋə́ kə̀ bé↓	'He slept.'
H L L/H (L)	H L LH/H↓ (L)	

When a L verb follows the L distal marker it changes to LH, whereas a H verb is produced on a lower pitch than the initial subject pronoun H. This could suggest that L verbs differ from H verbs in the way they dissimilate: while the former insert a final H to become contour LH tone at sentence final position, H tone verbs either lower their pitch to H↓ (when preceded by L), or become M↓ (when preceded by M).

3.1.2 Sentences involving two aspectual-modal markers

In the following sentences, the distal marker is accompanied by the negative marker which it follows, so, two TAM markers are involved. Note that the negative marker *tí* is considered to occupy the modal position of the sentence as part of TAM, right after the subject as shown in Table 6.

Table 6: Sentences with distal and negation

	UR	SR	Gloss
a.	ŋə́ tí kə̀ jɔ̀ʔ	ŋə́ tī kə̄ jɔ̀ʔ	'He did not sing.'
b.	ŋə́ tí kə̀ tʃɔ̀ʔ	ŋə́ tī kə̄ tʃɔ̀ʔ	'He did not speak.'
c.	ŋə́ tí kə̀ tàʔ ŋgɔ̀ːkə	ŋə́ tī kə̄ tàʔ ŋgɔ̀ːkə̀	'He did not want the banana.'
d.	ŋə́ tí kə̀ zə́ ŋgɔ̀ːkə	ŋə́ tī kə̄ zə́↓ ŋgɔ̀ːkə̀	'He did not eat the banana.'
e.	ŋə́ tí kə̀ jɛ́ ŋgɔ̀ːkə	ŋə́ tī kə̄ jɛ́↓ ŋgɔ̀ːkə̀	'He did not see the banana.'
f.	ŋə́ tí kə̀ bé	ŋə́ tī kə̄ bé↓	'He did not sleep.'
	H H L L/H (L)	H M M L/M↓ (L)	

Interestingly, the negative marker and the distal marker which are respectively H and L in the UR both appear on the surface as M. Thus, in this case, the two morphemes do not have to differentiate. Immediately after the M distal marker, L verbs remain L, whereas H verbs lower their pitch, just like in the preceding instances in Table 5. Our assumption, which we will discuss in detail in §4, is that tone dissimilation demarcates prosodic domains and within a prosodic domain, each element carries the same tone, through tone spreading. In other words, whenever tone dissimilates, it signals the frontier of a prosodic domain or rather the beginning or edge of a prosodic domain.

Tone dissimilation between prosodic domains is further exemplified in the sentences in Table 7.

Table 7: Sentences with distal and potential

	UR	SR	Gloss
a.	ŋə́ wɔ́ʔ kə̀ dʒʉ̀	ŋə́ wɔ̀ʔ kə̀ dʒʉ́	'He will be coming.'
b.	ŋə́ wɔ́ʔ kə̀ zə́	ŋə́ wɔ̀ʔ kə̀ zə́	'He will be eating.'
	H H L L/H	H L L H/L	

Because *wɔ́ʔ* (potential marker) and *kə̀* (distal marker) belong to the same prosodic group, they have to bear the same tone. First, they have to dissimilate with the preceding group which is the subject pronoun group having a H tone. This dissimilation is achieved though the change of *wɔ́ʔ*, the first element of the second group (the aspectual group) from H to L. Then, since *kə̀* already bears a

similar tone, no tone spreading is needed. Next, the verb group also dissimilates with the preceding which is L, and turns from L to H in sentence (a) in Table 7, whereas in sentence (b) in Table 7, dissimilation does not occur as the verb group already carries a different tone.

There are a few exceptions to this rule as shown in Table 8.

Table 8: Exceptions

	UR	SR	Gloss
a.	ŋə́ tí kə̀ fè ŋgɔ̀ːkə	ŋə́ tī kə̀ fě ŋgɔ̀ːkə	'He did not sell the banana.'
b.	ŋə́ tí kə̀ tʃěʔ ŋgɔ̀ːkə	ŋə́ tī kə̀ tʃěʔ ŋgɔ̀ːkə	'He did not cut the banana.'
c.	ŋwɛ́ wɔ́ʔ kə̀ tʃɔ̀ʔ	ŋə́ wɔ̀ʔ kə̀ tʃɔ̀ʔ	'They will be talking.'

With the verbs in (a, b) from Table 8, the distal marker does not change from L to M after the M negative marker. Consequently, the verbs themselves display unexpected tonal patterns: *fè* 'sell' becomes LH – instead of remaining L, and, recalling that each tone changes according to the preceding tone, this change is due to the preceding distal marker which also keeps its L and does not change to M as expected. However, the second verb, *tʃěʔ* 'cut' keeps its underlying LH tone. The reason why *kə̀* does not change to M in these two contexts is still to be clarified. Likewise, the tone verb group consisting of the verb *tʃɔ̀ʔ* in sentence (c) from Table 8 is expected to become H, dissimilating with the preceding L aspectual group consisting of *wɔ̀ʔ* and *kə̀*. Probably the tone of the subject pronoun HM has a certain influence on the pattern of the sentence, but again, we do not have enough arguments for now to confirm or reject this hypothesis.

Another exceptional case concerns the progressive (or imperfective) marker *nə̌*. Like the perfective which is always LH – recall that the perfective is marked only tonally and not segmentally – the tone on this morpheme is also unchanging, whatever the tonal, segmental or syntactic environment. This clearly appears in Table 9 and Table 10.

Interestingly, the unchanging form of the progressive is also attested in Babanki (Akumbu, Hyman & Kießling 2020), a Central Ring language. There should be a historical explanation for this phenomenon, probably dating back to Proto-Ring. Besides that, the tone on the verb following *nə̌* deserves our attention. Indeed, it is never LH like the progressive marker, but always carries a different tone. This is another evidence of the dissimilation process at work in the construction of indicative sentences.

Table 9: Tone dissimilation with progressive

	UR	SR	Gloss
a.	mə̀ nɔ̌ nâ	mə̀ nɔ̌ nâ	'I am cooking.'
b.	ŋɡə́ nɔ̌ dʒʉ̀	ŋɡə́ nɔ̌ dʒʉ̂	'He is coming.'
c.	wə̀ nɔ̌ tàʔ ŋɡə́	wə̀ nɔ̌ tàʔ ŋɡə́	'You want it.'

Table 10: Tone dissimilation with progressive and distal

	UR	SR	Gloss
a.	ŋɡə́ kə̀ nɔ̌ dʒʉ̀	ŋɡə́ kə̀ nɔ̌ dʒʉ́	'He was coming.'
b.	ŋɡə́ kə̀ nɔ̌ zə́	ŋɡə́ kə̀ nɔ̌ zə́	'He was eating.'
c.	wà tí nɔ̌ jʉ̀ fákə	wà tí nɔ̌ jʉ̂ fá	'We are doing nothing.'

3.1.3 Sentences involving three aspectual-modal markers

So far, we have been looking at prosodic groups consisting of one or two morphemes. It is worth mentioning the case where more than two morphemes are found in a group. As a matter of fact, the maximum number is three, and we will see that this case stands out from the findings we have hitherto been discussing, namely, the occurrence of tone dissimilation between prosodic groups and tone spreading inside a group.

When there are three elements in a group, particularly in the aspectual group, they never assimilate, contrary to what was observed when they are two in the group.

Table 11: Prosodic groups with three morphemes

	UR	SR	Gloss
a.	ŋɡə́ tí wɔ́ʔ kə̀ dʒʉ̀	ŋɡə́ tī wɔ̀ʔ kə́ dʒʉ̄	'He will not be coming.'
b.	ŋɡə́ tí wɔ́ʔ kə̀ zə́	ŋɡə́ tī wɔ̀ʔ kə̀ zə́	'He will not be eating.'
c.	ŋɡə́ tí kə̀ nɔ̌ dʒʉ̀	ŋɡə́ tī kə̀ nɔ̌ dʒʉ́	'He was not coming.'
d.	ŋɡə́ tí kə̀ nɔ̌ zə́	ŋɡə́ tī kə̀ nɔ̌ zə́	'He was not eating.'

The aspectual prosodic group is made up of *tí wɔ́ʔ* and *kə̀* (respectively, the negative marker, potential marker and distal marker) in the first two sentences,

and in sentences (c) and (d) from Table 11 it is made up of *tí* and *kə̀*. In sentence (a) from Table 11, the distal marker (*kə̀*) is H, but L in (b); the only difference between these two sentences is the verb, which is underlying L Table 11a and H in Table 11b. We could straightaway assume the process of anticipatory dissimilation to be responsible for this discrepancy: the distal marker turns from L to H dissimilating in anticipation from the following L verb which subsequently becomes M. In Table 11b, since the verb is already H, i.e., underlyingly, and thus different from the L distal marker, nothing changes. In sentences (c) and (d) from Table 11, each element in the aspectual group surfaces with its underlying tone, except the negative marker which in all four cases – in fact in all cases in general – lowers from H to M because of the preceding H of the subject pronoun. In other words, even though assimilation does not always take place within the aspectual group when three morphemes are involved, two dissimilatory changes seem to be inescapable: 1) between the subject pronoun and the negative marker, and 2) between the last aspectual marker (i.e., the last element of the aspectual group) and the verb. This is confirmed when we compare for instance the sentence in Table 11a with the one in (12), where the subject pronoun is L (*wà* 'we') and for this very reason, i.e., because they already differ, the underlying H of the negative marker *tí* following it does not change, contrary to Table 11a where this same negative has to dissimilate from the preceding H subject pronoun which belongs to a different prosodic group, hence its realization as M.

Before closing this section on indicative sentences, one last context we could examine and that corroborates our assumption is the potential aspect, marked by *wɔ̀(ʔ)*. The final glottal stop is optional, and one of its functions in Wushi is syllable demarcation.

Table 12: Prosodic groups with three morphemes

	UR	SR	Gloss
a.	ŋə́ tí wɔ́ʔ jɔ̀ʔ	ŋə́ tī wɔ̄ʔ jɔ̀ʔ	'He will not sing.'
b.	ŋə́ tí wɔ́ʔ tʃɔ̀ʔ	ŋə́ tī wɔ̄ʔ tʃɔ̀ʔ	'He will not speak.'
c.	ŋə́ tí wɔ́ʔ tàʔ ŋgɔ̀ːkə	ŋə́ tī wɔ̄ʔ tàʔ ŋgɔ̀ːkə̀	'He will not want the banana.'
d.	ŋə́ tí wɔ́ʔ zə́ ŋgɔ̀ːkə	ŋə́ tī wɔ̄ʔ zə́↓ ŋgɔ̀ːkə̀	'He will not eat the banana.'
e.	ŋə́ tí wɔ́ʔ jɛ́ ŋgɔ̀ːkə	ŋə́ tī wɔ̄ʔ jɛ́↓ ŋgɔ̀ːkə̀	'He will not see the banana.'
f.	ŋə́ tí wɔ́ʔ bé	ŋə́ tī wɔ̄ʔ bé↓	'He will not sleep.'
	H H H L/H (L)	H M M L/H↓ (L)	

Tonal dissimilation and tone spreading manifestly organize the prosodic structure of sentences at least in the indicative mood: tone dissimilation demarcates what we refer to here as prosodic groups, and tone spreading occurs within a group. Hence the M tone shared between *tī* and *wɔ̄ʔ*, respectively the negative marker and the potential marker, constituting what we will call the *aspectual group*. The M tone produced on the two successive morphemes could be seen as an instance of tone spreading applying within the group and not across. This aspectual group is preceded by the *subject group* which is H in the above examples, and is followed by the *verb group* which is L or H$^{↓}$ depending on the underlying tone of the verb. The verb group in its turn may be followed by the object which we consider extrametrical; in other words, it is invisible to the tone processes occurring in the rest of the sentence. What justifies this choice is the fact that the tone of the object does not dissimilate with the preceding tone, that of the verb. This is seen in the above examples with the L object *ngɔ̀ːkə̀* 'banana' and is equally observed when the object is H.

Table 13: Extrametricality of the object

	UR	SR	Gloss
a.	ŋə́ tí wɔ́ʔ tàʔ tʃá	ŋə́ tī wɔ̄ʔ tàʔ tʃá	'He will not want the fish.'
b.	ŋə́ tí wɔ́ʔ zə́ tʃá	ŋə́ tī wɔ̄ʔ zə́$^{↓}$ tʃá$^{↓}$	'He will not eat the fish.'
c.	ŋə́ tí wɔ́ʔ jɛ́ tʃá	ŋə́ tī wɔ̄ʔ jɛ́$^{↓}$ tʃá$^{↓}$	'He will not see the fish.'
	H H H L/H H	H M M L/H$^{↓}$ H/H$^{↓}$	

The tone on the object *tʃá* 'fish' remains H except when preceded by a downstepped H; this is the result of downdrift, whereby the pitch level of a H in a sequence of Hs is dependent on a preceding H tone. This shows us that the object is out of the scope of dissimilation. Put differently, tone dissimilation concerns only the subject pronoun group, the aspectual group, and the verb group, as far as indicative sentences are concerned. Note that these are prosodic groups, i.e. they make up the prosodic constituency of indicative sentences, and do not necessarily match with syntactic constituents, as assumed in the theory of prosodic structure as well as theories of the syntax-prosody interface. We will come back to this discussion later.

3.2 Modal constructions: possibility

In this section we will focus on modality, specifically the expression of possibility. Grassfields Bantu languages are well-known for their highly analytic morphology and this is particularly evident in modal constructions. Indeed, while in Narrow Bantu languages modality is mostly expressed through affixes, this is achieved in most Ring Grassfields languages through independent particles generally preceding the verb. Modal constructions expressing possibility in Wushi are particularly interesting for our present study as they show a tonal pattern very close to what we have described for simple indicative sentences although their syntax is quite different. The modal marker *làkə̀* comes first, and then the rest of the sentence displays the canonical order, i.e. SVO as schematized in (13).

(13) làkə̀ + S + V + O

It should be noted that *làkə̀* is not limited to possibility but is used for other types of modality including necessity, permission, and certainty. I explain in my dissertation its initial position as signalling its scope, which goes from the next element, i.e., the subject, to the end of the sentence. In other words, the entire sentence is its scope, meaning that the modality expressed affects the meaning of the clause as a whole. Thus, the representation in (13) could otherwise be as follows:

(14) làkə̀ {S V O}

The brackets here mean the whole clause, i.e., SVO, is the scope of the modality *làkə̀*, rather than the verb only.

Let us now consider some examples.

Table 14: Tone dissimilation and modality

	UR	SR	Gloss
a.	làkə̀ wə́ gɛ̀	làkə̀ wə́ gɛ̀	'You can go/You are free to go.'
b.	làkə̀ ŋə́ tí dʒʉ̀	làkə̀ ŋə́ tì dʒʉ̂	'He cannot come.'
c.	làkə̀ mə̀ dʒʉ̀ mbó:kə́	làkə̀ mə́ dʒʉ̀ mbó:kə́	'Can I come and play?'
d.	làkə̀ ŋə́ tí tʃɔ̀ʔ	làkə̀ ŋə́ tì tʃɔ̂ʔ	'He cannot speak.'
c.	làkə̀ mə̀ tʃà?sə́ ŋwà?nə̀ jə	làkə̀ mə́ tʃà?sə́ ŋwà?nə́ jə̂	'Can I borrow your book?'
	L L/H (H) L	L H L H L	

Here again, the dissimilation process taking place is striking. We go from patterns like L H H L (sentences b and d in Table 14) and L L L H (sentence c in

Table 14) in the UR with consecutive identical tones, to a L H L H pattern on the surface, obviously forbidding successive identical tones – we count *làkə̀* as a single L since this is a unique marker,[4] that of modality. The (H) corresponds to the negative marker, present only in sentences (b) and (c) in Table 14. We assume that the L H L H pattern actually underlies modal constructions, so that every one of them aligns with it. Sentence (c) in Table 14 is another significant illustration. It reveals that the tone pattern on the noun *ŋwàʔnə̀* 'book' changes from LL to LH, so as to fit in the pattern of the whole sentence, although the two L in the L L pattern belong to the same morpheme. In addition, evidence also comes from the first person singular pronoun *mə̀* which is underlyingly L and always changes to H when following *làkə̀*. Since *làkə̀* carries a L tone, the L tone on *mə̀* must dissimilate. This is not the case for the third person singular pronoun *ŋə́* which is H in the UR, and given that its tone is already different from that of the preceding modal marker, nothing changes. In other words, the pattern L H L H is a fixed pattern for modal constructions with *làkə̀*, unlike indicative sentences where the nature of each tone constituting the tonal pattern depends on the preceding tone, and therefore in this case the pattern is built up as the sentence unfolds. Nevertheless, there is an invariant tonal process governing the construction of these patterns, be they fixed or "context-dependent", and that is dissimilation. How could the Match Theory account for these phenomena? The next section is dedicated to this question.

4 Discussion

Prosodic domains in Bantu languages – also referred to as tone groups in other studies, for example in Chen (1987), and Hyman & Katamba (2010) – are characterized by a variety of tonal processes including tone spreading at the phrasal level for example in Copperbelt Bemba M40 (Kula & Bickmore 2015), L deletion and H plateauing in Luganda JE15 (Hyman & Katamba 2010), penultimate lengthening in Xitsonga S53 (Selkirk 2011), to give but a few examples. What is common to these processes is that they occur within a prosodic domain and only within that domain, and therefore characterize it as a prosodic domain. Some of these processes occur specifically at the left or right edge of a constituent. For instance, Truckenbrodt (1999: 237) explains that in Kimatuumbi, shortening applies on long stem vowels, except in the prosodic word immediately preceding the right edge of a p-phrase. Thus, shortening applies on *mpúunga* 'rice' in the

[4] Although not a unique morpheme. Indeed, *làkə̀* is made up of two morphemes: *là* the hypothetic marker, and *kə̀* the distal marker.

phrase *mpunga wá baándu* 'rice of people' as the word for 'rice' in this context is the prosodic word immediately preceding the right edge of the phonological phrase (φ).

(15) (mpunga wá baándu) φ 'rice of people'

Such phonological processes are equally attested in Wushi. When analyzing indicative sentences, we saw that tone spreads within a group; this concerns particularly what we referred to as the aspectual group, when consisting of two elements: generally the first element is the negative marker *tí*, followed by either the distal marker *kə̀* and/or the potential marker *wɔ́ʔ*, but we can also have, in the absence of the negative marker, the potential marker followed by the distal marker. We observed that tone spreading in this group is quite regular when two morphemes are involved, but less consistent when all the three morphemes appear. However, in all cases, one process necessarily obtains, and this is tone dissimilation between prosodic groups or domains. In fact, the Obligatory Contour Principle (OCP) underpins the construction of (indicative) sentences, so that identical consecutive tones are avoided on the surface.

To represent these requirements and constraints in the syntax-prosody framework, we consider three major prosodic domains (four including the object group) which are phonological phrases, as far as indicative sentences without modality are concerned.

(16) a. (ŋə́) φ (tī kə̄) φ (zə́ˡ) φ (ŋgɔ̀ːkə̀) φ 'He did not eat the banana.'
b. (ŋə́) φ (wɔ̀ʔ kə̀) φ (dʒʉ́) φ 'He will be coming.'
c. (ŋə́) φ (tī wɔ̄ʔ) φ (jɔ̀ʔ) φ 'He will not sing.'

The peculiarity of the organization of prosodic constituency in Wushi is that each phonological phrase is determined by the preceding one instead of by what happens within, in other words the indicator of the frontier of a phonological phrase is not within the phrase in question but outside: it is tone dissimilation. More specifically, as the sentence unfolds, tone dissimilates whenever we are crossing a phrase boundary, and until the frontier is crossed, every element receives the same tone (hence tone spreading postulated earlier). Accordingly, tone dissimilation indicates that we have accessed the next phonological phrase. Thus, as the utterance is produced, tone is assigned to each prosodic group based on this fundamental principle. For example, the phonological phrase (*tī kə̄*) exhibits a M because of the H of the preceding group. *tí*, which is the first element to dissimilate to *tī* thereby signalling the left edge of a new phrase, changes to M because of the preceding H. Then, this M spreads until the next occurrence of

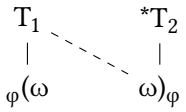
dissimilation, indicating the next boundary. It is in this sense we are claiming that each phonological phrase is dependent on what precedes. This is why when the subject pronoun is L, the negative marker does not need to change to M, and thus, it remains H as in the sentence in (12).

One crucial question would be, given the rigid tone on the progressive marker which is always LH, where do we put the next boundary? We could answer that since *nǒ* is never subject to change, it should be considered as an exceptional morpheme, and therefore it would be inappropriate to base our hypothesis on it. The most regular pattern prevailing is rather one where the negative marker and the aspectual markers share the same tone, hence their grouping together into one prosodic group. Consequently, although in sentence (12) there is no clear indication of a phonological phrase consisting of *tí* and *nǒ*, the general and most consistent observation allows us to group them together in the prosodic constituency.

Another central question pertains to the correspondence between prosodic structure and syntactic structure and Match Theory. If we consider both structures in the sentences in (16), we will have the following:

(17) a. $_{S}$[(ŋá)$_{\varphi}$] $_{VP}$[(tī kā)$_{\varphi}$ (zá$^{\downarrow}$)$_{\varphi}$] $_{O}$[(ŋgɔ̀ːkà)$_{\varphi}$] ‘He did not eat the banana.’
b. $_{S}$[(ŋá)$_{\varphi}$] $_{VP}$[(wɔ̀ʔ kà)$_{\varphi}$ (dʒú)$_{\varphi}$] ‘He will be coming.’
c. $_{S}$[(ŋá)$_{\varphi}$] $_{VP}$[(tī wɔ̄ʔ)$_{\varphi}$ (jɔ̀ʔ)$_{\varphi}$] ‘He will not sing.’

Following Dixon (2009), we consider the syntactic verb phrase to consist of the verb and the grammatical morphemes associated to it, such as tense, aspect, mood, and modality. Here, misalignment between the syntactic verb phrase and prosodic structure requires specifying constraints on the prosodic constituency that will repair the violation of MATCH(Phrase, φ). Two constraints will be posited here: ALIGN-L(φ, PrW) followed by SPREAD(φ). T in the representation below stands for tone.

(18)


ALIGN-L(φ, PrW): align a tone with the left-most prosodic word of a phonological phrase.
SPREAD(φ): spread the tone to the entire phonological phrase.

This constraint stipulates that the tone of the prosodic word on the left edge determines the tone of the phonological phrase and is followed by spreading

which is restrained to the elements of the phonological phrase, so that two different tones (here T_1 and T_2) cannot co-occur within the same phonological phrase. A different tone will therefore have to be associated to a different (following) phonological phrase.

Concerning modal constructions, since the L H L H pattern is fixed for all the constructions, we are claiming here that it actually constitutes the intonational phrase. Thus:

(19) a. ι {làkə́ wə́ gɛ̀} ι
b. ι {làkə́ ŋə́ tí dʒʉ̀} ι
c. ι {làkə́ ŋə́ tì tʃɔ̂ʔ} ι

Indeed, contrary to the "context-dependent" application of dissimilation in indicative sentences without modality, which consequently may yield several different tonal patterns on the surface, the rigid L H L H pattern associated to modal constructions indicates that the whole clause is the target. And the corresponding prosodic domain is the intonational phrase. Therefore, MATCH(Clause, ι) constraint is met.

Before concluding this paper, a crucial question already raised by Elfner (2012: 11) is whether prosodic structure plays a role in determining the surface structure "traditionally syntactic domains such as word order…" We could even go further and ask ourselves if syntactic constituency is not, at least in certain languages, and in particular contexts, the reflection of prosodic structure. In other words, what if it was rather the syntactic structure that matches the prosodic structure, which therefore, would be the one deriving the syntactic structure? A necessary implication of such an analysis would be that prosody pre-exists syntax, as argued by Martin (2013).

5 Conclusion

The aim of this paper was to analyze the syntax-prosody interface in Wushi, an under-documented Ring Grassfields Bantu language. Our focus was on indicative sentences with and without modality, and we established the analysis in the framework of Match Theory (Selkirk 2011). It followed that tone dissimilation governs the prosodic structure of sentences, and, more generally, OCP (Obligatory Contour Principle). This forbids consecutive phonological phrases to display the same tone. On the other hand, tone spreading obtains within each phonological phrase, except in a few special cases. We found out that MATCH(Phrase, φ) is overriden by ALIGN-L(φ, PrW) and SPREAD(φ), while MATCH(Clause, ι) is

met especially in modal constructions expressing possibility. More descriptions of the prosody-syntax interface in Grassfields languages, as well as the study of intonation would help us have a wider understanding of the interaction between tone and syntax not only in this language group, but in African languages at large.

Abbreviations

POT	potential	P1	today past
DET	determiner	POSS.1SG	1st person singular possessive
SM	subject marker		
NEG	negation	SR	surface representation
TA	tense aspect		
OM	object marker	UR	underlying representation
FV	final vowel		

References

Akumbu, Pius W. & Evelyn Fogwe Chibaka. 2012. *A pedagogic grammar of Babanki: A Grassfields language of Northwest Cameroon*. Rüdiger Köppe Verlag.

Akumbu, Pius W., Larry M. Hyman & Roland Kießling. 2020. The segmental and tonal structure of verb inflection in Babanki. *UC Berkeley PhonLab Annual Report* 2(1). 1–33.

Chen, Matthew Y. 1987. The syntax of Xiamen tone sandhi. *Phonology* 4. 109–149. DOI: 10.1017/S0952675700000798.

Cheng, Lisa & Laura Downing. 2007. The prosody and syntax of Zulu relative clauses. *SOAS Working Papers in Linguistics* 15. 51–63.

Dixon, R. M. W. 2009. *Basic linguistic theory: Methodology*, vol. 1. Oxford, New York: Oxford University Press.

Elfner, Emily. 2012. *Syntax-prosody interaction in Irish*. University of Massachusetts, Amherst. (Doctoral dissertation).

Elfner, Emily. 2018. The syntax-prosody interface: Current theoretical approaches and outstanding questions. *Linguistics Vanguard* 4(1). 1–14. DOI: 10.1515/lingvan-2016-0081.

Hodieb, Liliane. 2021. On the aspectual system of Wushi (Babessi), a Ring Grassfields Bantu language of Cameroon. *Language in Africa* 2(2). 43–65. DOI: 10.37892/2686-8946-2021-2-2-43-65.

Hyman, Larry M. 1979. *Aghem grammatical structure* (Southern California Occasional Papers in linguistics 7). Los Angeles: Department of Linguistics, University of Southern California.

Hyman, Larry M. 1980. *Noun classes in the Grassfields Bantu borderland* (Southern California Occasional Papers in Linguistics 8). Los Angeles: Department of Linguistics, University of Southern California.

Hyman, Larry M. & Francis X. Katamba. 2010. Tone, syntax, and prosodic domains in Luganda. *UC Berkeley PhonLab Annual Report* 6. 27–56. DOI: 10.5070/P74fh5b2q7.

Kula, Nancy & Lee Bickmore. 2015. Phrasal phonology in Copperbelt Bemba. *Phonology* 32(1). 147–176. DOI: 10.1017/S095267571500007X.

Martin, Philippe. 2013. La structure prosodique opère-t-elle avant ou après la syntaxe ? *TIPA: Travaux interdisciplinaires sur la parole et le langage* 29. DOI: 10.4000/tipa.943.

Nurse, Derek. 2008. *Tense and aspect in Bantu*. Oxford: Oxford University Press.

Selkirk, Elisabeth. 2011. The syntax-phonology interface. In John Goldsmith, Jason Riggle & Alan C. L. Yu (eds.), *The handbook of phonological theory*, 2nd edn., 435–484. Oxford: Blackwell.

Snider, Keith L. 2018. *Tone analysis for field linguists*. Dallas: SIL International Publications.

Truckenbrodt, Hubert. 1999. On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry* 30(2). 219–255. DOI: 10 . 1162 / 002438999554048.

Watters, John R. 2003. Grassfields Bantu. In Gérard Philippson & Derek Nurse (eds.), *The Bantu languages*, 225–256. London: Routledge.