Downstep and Phonological Phrasing in Sandawe*

Yoshihito Dobashi

Sandawe, a Khoisan language spoken in Tanzania, exhibits downstep between words. In this paper, I examine the domain, or the phonological phrase, where downstep takes place and the effects of focus on phonological phrasing within the framework of Alignment Theory (Selkirk 1995, Truckenbrodt 1995, 1999). I propose a constraint *Focus-Right-Edge that prohibits a phonological boundary that immediately follows a focused constituent. I also address the theoretical issues of branching categories and phonological phrasing. I provide a preliminary analysis in terms of a condition on Spell-Out within the framework of Multiple Spell-Out theory (Chomsky 2001a, b, Collins 2001a, b, Uriagereka 1999).

1. Introduction

In this paper, I discuss the downstep between words in Sandawe. I am especially concerned with the domain in which the downstep takes place, which I claim is the phonological phrase. Sandawe is a Khoisan language spoken in Tanzania. It has a tone system, and basically four tones (High, Low, Fall, Rise) are attested (see Elderkin 1989 and Kagaya 1993 for discussion).

In section 2, I introduce the data and briefly discuss a theory of downstep, where linked Low tones, as well as floating Low tones, play a role in inducing downstep. In section 3, I introduce the Generalized Alignment theory (McCarthy and Prince 1993, Selkirk 1995, Truckenbrodt 1995, 1999), and propose a constraint that is related to focus, i.e., *Focus-Right-Edge, which prohibits a phonological boundary that immediately follows a focused constituent. In section 4, I speculate on some theoretical issues concerning phonological phrasing and branching categories. Section 5 is a conclusion.

2. Data

Let us first consider downstep within a word:

(1) ts?áma’sú  'giraffe'  Elderkin (1989: 84 (2.101))

---

*I would like to thank Chris Collins, John Whitman, Draga Zec, and the editors of this volume for invaluable comments.*
Here there is a H-L-H sequence within a word, and the pitch level of the second H is lower than that of the first H, according to Elderkin (1989):

\[(2) \quad \underline{\text{H}} \quad \underline{\text{L}} \quad \underline{\text{H}}\]

\[
\begin{array}{c}
\text{ts’ámà} \\
\downarrow
\end{array}
\quad \begin{array}{c}
sú
\end{array}
\]

The downstep is generally considered to be a result of phonetic interpretation of a phonological representation (Pulleyblank 1986). If H follows L as in (2), H that follows L is phonetically interpreted as having a lower pitch level than H that precedes it. That is, an intervening L tone induces downstep. A downstep is also observed even if we have a H-H sequence:

\[(3) \quad /í’pó \quad \text{‘you didn’t come’} \quad \text{Elderkin (1989: 86 (2.106))}\]

Here, there are two H tones whose sequence is not interrupted by a L tone as in (2), yet downstep is observed.

\[(4) \quad \underline{\text{H}} \quad \underline{\text{H}}\]

\[
\begin{array}{c}
\text{/í’pó}
\end{array}
\]

In this kind of situation, it is generally assumed that a floating L tone, which is not linked to a vowel, exists between the two H tones because if there is such a floating L tone in the representation, it is not necessary to adopt any other assumption that accounts for the downstep between two H tones. Thus, the representation of (3) is in fact (5), which is parallel to (2). And the floating L tone induces downstep (Clements 1980, Pulleyblank 1986):
The second H tone is interpreted phonetically as having a lower pitch level than the first one because of the floating L tone (Clements 1980, Pulleyblank 1986, among others).

Note that downstep does not take place within a word like sómbá ‘fish’, where there are two H tones. It might be the case that the floating L tone in (5) is a boundary tone which is inserted in the prosodic or morphological boundary. (In (5), -po is a second person affix.) And in a word like sómbá ‘fish,’ the two vowels are linked to a single High tone, and there is no room for a floating L tone.

Downstep is also observed between words:¹

(5)    H  L  H
       /í   !   pó

Here, there is a downstep between sómbá and thúmē-sù, while there is no downstep between Sándá and sómbá. Following Elderkin’s (1989) notation, I express the levels of the pitch of phrases by the number that is superscribed on each word. Thus in (6), “1” on Sándá and sómbá means that these words have the highest pitch level, and “2” on thúmē-sù means that it has the second highest pitch.

In (6), there is no downstep between the subject Sándá and the object sómbá ‘fish’. However, if the subject Sándá is focused, then the sentence shows a different downstep pattern.²

¹ The syntactic properties of Sandawe relevant to the present discussion are the followings: First the word order in Sandawe is free although SOV order is the most frequently observed one. Second, pro-drop is allowed. Thus, the subject can be omitted. Third, there are two kinds of suffixes that are related to focus. One is a nominative suffix, which may be attached to the subject. If the subject has this suffix, then it is focused. The other is a pgn-suffix, which agrees with the subject in person, gender and number. It may be attached to non-subjects (e.g., objects, verbs, adverbs, etc.). When a non-subject bears this suffix, then it is focused. See Dalgish (1979), Dempwolff (1916), Kagaya (1990) and Dobashi (2001) for discussions on Sandawe syntax. I underline the focused element in this paper.

² There is an upstep between the adverb and the subject here, which is not discussed in this paper.
There is a downstep between the subject sándá-á and the object sómbá ‘fish,’ and between the object sómbá ‘fish’ and the verb thíímé ‘cooked.’ Let us next consider the following sentence:

(8) \(1\text{?útè-sà}  \quad 2\text{sómbá}  \quad 3\text{thíímé} \)
\(\text{yesterday-3f.sg.} \quad \text{fish} \quad \text{cooked} \)
‘Yesterday she cooked the fish’   Elderkin (1989: 94: (3.3))

Here, ?útè ‘yesterday’ is focused. Then, a downstep pattern similar to that in (7) surfaces. That is, the downstep takes place between all the phrases that follow the focused phrase. However, this is not always the case, as the following sentence shows:

(9) \(1\text{?útè-sà}  \quad 2\text{sándá}  \quad 2\text{sómbá}  \quad 3\text{thíímé} \)
\(\text{yesterday-3f.sg} \quad \text{Sanda} \quad \text{fish} \quad \text{cooked} \)
‘Sanda cooked the fish yesterday’   Elderkin (1989: 96: (3.10a))

Here, the sentence initial phrase ?útè ‘yesterday’ is focused, and downstep occurs between ?útè ‘yesterday’ and the subject sándá which immediately follows it, and between the object sómbá ‘fish’ and the verb thíímé. Downstep does not occur between the subject sándá and the object sómbá ‘fish.’ Thus, downstep applies within some local domain. I call such a domain the phonological phrase (p-phrase) (cf. Truckenbrodt 1999), and assume the following:

(10) The domain for downstep is a phonological phrase.

I discuss how p-phrasal boundaries are determined in the next section.

---

3 Note that in (9) the object sómbá ‘fish’ does not have the highest pitch level, but it has the pitch level 2 that is one step lower than the highest one in the sentence. I will return to this matter later in this section.
The relevant p-phrases are assigned to (6), (8) and (9) below:

(11) \(( ^1\text{Sándá} )^1 \text{sómbá}^2 \text{thíímé-sù} )\) \(\Rightarrow\) (6)

\(\text{sanda} \quad \text{fish} \quad \text{cook-3f.sg.future} \)

\(\text{‘Sanda will cook the fish’}\)

(12) \(( ^1\text{?útè-sà})^1 \text{sómbá}^2 \text{thíímé}) \) \(\Rightarrow\) (8)

\(\text{yesterday-3f.sg.} \quad \text{fish} \quad \text{cooked} \)

\(\text{‘Yesterday she cooked the fish’}\)

(13) \(( ^1\text{?útè-sà})^1 \text{sándá}^2 \text{thíímé}) \) \(\Rightarrow\) (9)

\(\text{yesterday-3f.sg} \quad \text{Sanda} \quad \text{fish} \quad \text{cooked} \)

\(\text{‘Sanda cooked the fish yesterday’}\)

The object \(\text{sómbá}\) and the verb \(\text{thíímé-sù}\) are grouped into a p-phrase in (11), the whole sentence is grouped into a p-phrase in (12), and the focused adverb \(\text{?útè-sà}\) and the subject \(\text{sándá}\) on the one hand and the object \(\text{sómbá}\) and the verb \(\text{thíímé}\) on the other are grouped into two different p-phrases in (13). Downstep applies between any two adjacent words within each of the p-phrases.

We saw that within a word the downstep is induced by a floating L tone. Similarly, I assume that the downstep between words is induced by a floating L tone, which is inserted between words within a phonological phrase.

Note that in (13) the pitch level of \(\text{sómbá}\) is the same as that of \(\text{Sándá}\). That is, the phonological boundary does not reset the pitch level of the following tone completely. If it did, the pitch level of \(\text{sómbá}\) would be as high as that of the sentence initial high tone. I assume that the pitch level of the p-phrase initial High tone is dependent on the pitch level of the immediately preceding High tone even though there is a p-phrase boundary:

(14) The p-phrase initial High tone is realized at the pitch level that is the same as that of the High tone of the last word in the immediately preceding p-phrase.
3. Phonological Phrasing

3.1 Syntax-Phonology Mapping

In this section, I introduce some basic assumptions about syntax-phonology mapping that are relevant to the present discussion. Basically, I adopt the general assumptions made by Truckenbrodt (1999).

As is generally observed in the literature (Clements 1978, Nespor and Vogel 1986, Selkirk 1986, Inkelas and Zec 1990, among many others), phonological phrasing, or the determination of phonological boundaries within phrases and sentences, is closely related to syntactic phrasing. That is, phonological phrases are determined on the basis of syntactic structures. Following Truckenbrodt (1999), I assume the following:

(15) XP-to-P-phrase Mapping Condition

Mapping constraints relate XPs to phonological phrases, but do not relate XPs to other prosodic entities. (Truckenbrodt 1999: 221)

Here, XP is a lexical category such as NP or VP, but not a functional category such as IP (cf. Selkirk and Shen 1990):

(16) Lexical Category Condition

Constraints relating syntactic and prosodic categories apply to lexical syntactic elements and their projections, but not to functional elements and their projections, or to empty syntactic elements and their projections. (Truckenbrodt 1999: 226)

Furthermore, I assume that phonology cannot refer to syntax directly (see Inkelas 1989, Inkelas and Zec 1995). That is, it refers to prosodic structure (such as phonological phrases created under the condition (15)):

(17) Indirect Reference Hypothesis

Phonological rules refer to only prosodic constituent structure. (Truckenbrodt 1999: 221)

Thus, phonological rules cannot refer to the distinction between syntactic categories, such as NP vs. VP.
3.2 Phonological Phrasing in Sandawe

In this section, I discuss the phonological phrasing in Sandawe. Let us consider the following sentence, where there is no focus:

\[(18) \quad 1Sândá \quad 1sómbá \quad 2thímé-sù \quad (=6)\]

\[
\begin{array}{ll}
\text{sanda} & \text{fish} \\
\text{cook-3f.sg.fut.} & \\
\end{array}
\]

‘Sanda will cook the fish’  
Elderkin (1989: 121 (3.64))

Here, the downstep takes place between the object and the verb, but does not take place between the subject and the object. I assume that (18) has the following syntactic structure:\(^4\)

\[(19) \quad [ \text{IP Sândá} \quad [\text{VP sómbá thíímé-sù }] ]\]

Then, the p-phrase in which downstep takes place here corresponds to the VP. That is, the object and the verb are grouped together, excluding the subject. Downstep takes place within the VP, but does not take place across the VP boundary.

In order to account for this, I adopt edge alignment of the following form (relevant here is *Generalized Alignment*: McCarthy and Prince 1993):

\[(20) \quad \text{Where Cat}_1, \text{Cat}_2 \text{ are prosodic, morphological, or syntactic categories and Edge}_1, \text{Edge}_2 \in \{ \text{Right, Left} \}: \]

\[
\begin{align*}
\text{Align} (\text{Cat}_1, \text{Edge}_1; \text{Cat}_2, \text{Edge}_2) & \iff \\
\text{For each Cat}_1 \text{ there is a Cat}_2 \text{ such that Edge}_1 \text{ of Cat}_1 & \text{ and Edge}_2 \text{ of Cat}_2 \text{ coincide.}
\end{align*}
\]

Selkirk (1995) and Truckenbrodt (1995, 1999) propose the constraints on edge alignment of syntactic categories and a phonological phrase (P), as follows:

\[(21) \quad \text{Align-XP, R: Align (XP, R; P, R)}\]

\(^4\) The basic phrase structure I adopt in this section is based on X-bar theory (Chomsky 1986):

(i) \([\text{IP NP}_1 \text{ Infl } [\text{VP NP}_2 \text{ V}]]\), where NP\(_1 = \text{subject}, \text{NP}_2 = \text{object}.)
“For each XP there is a P such that the right edge of XP coincides with the right edge of P.”

(22)  Align-XP, L: Align (XP, L; P, L)

“For each XP there is a P such that the left edge of XP coincides with the left edge of P.”

In order to avoid some unwanted results, the following two constraints are necessary:

(23)  Exhaustivity: Parsing on every prosodic level must be exhaustive.

(24)  *P-Phrase: Avoid p-phrases.

Given these constraints, let us consider the phonological phrasing in Sandawe. I assume the following ranking among the constraints for Sandawe:

(25)  Align-XP, L >> *P-Phrase >> Align-XP, R,

Let us apply this to (18), repeated here in (26):

(26)  \[ \text{IP} \quad ^1\text{Sánda} \quad ^1\text{sóbá} \quad ^2\text{thímé-sù} \quad \]  

\begin{align*}
\text{sanda} & \quad \text{fish} & \quad \text{cook-3f.sg.fut.} \\
\text{‘Sanda will cook the fish’} \\
\text{Align-XP, L} & \quad >> & \quad *P-Phrase & \quad >> & \quad \text{Align-XP, R}
\end{align*}

\begin{align*}
a. & \quad \ast \quad ( & \quad ( & \quad ) & \quad ) & \quad ** & \quad * \\
b. & \quad *( & \quad ( & \quad ) & \quad ( & \quad ) & \quad ** & \quad **! \\
c. & \quad *( & \quad ( & \quad ) & \quad ) & \quad *! & \quad * & \quad ** \\
d. & \quad *( & \quad ) & \quad ( & \quad ) & \quad *! & \quad * & \quad **
\end{align*}

(c) and (d) violate Align-XP, L because the left edges of VP as well as the object NP do not correspond to p-boundaries. And this is a fatal violation for them. Both (a) and (b) satisfy Align-XP,L since in each case, the left edge of the subject NP coincides with the left edge of a p-phrase, and the left edge of the VP, as well as that of the object NP, coincides with the left edge of a p-phrase. (a) violates *P-phrase twice since two phonological phrases are formed, while (b) violates it three times since three
phonological phrases are formed. This additional violation is fatal for (b): the phonological boundary created between the object and the verb is not required by Align-XP, L. Notice that if the ranking were just Align-XP, L >> Align-XP, R without *P-phrase intervening, (b) would be chosen as the optimal candidate since (a) violates Align-XP, R (the right edge of the object NP does not correspond to a p-boundary) while (b) satisfies it. That is, *P-Phrase plays a crucial role in making Align-XP, L effective in the language.

Let us next consider sentence (13), repeated here.

(27) 1?útè-sà 2sándá 2sómbá 3thíímé (= (13))
yesterday-3f.sg Sanda fish cooked
‘Sanda cooked the fish yesterday’ Elderkin (1989: (3.10)

Here, ?ute ‘yesterday’ is focused. Downstep takes place between ?útè-sà and sándá, and between sómbá and thíímé, but not between sándá and sómbá. I assume that (27) has the following syntactic structure:

(28) [IP 1?útè-sà [IP 2sándá [VP 2sómbá 3thíímé]]]
yesterday-3f.sg Sanda fish cooked

Here, the subject sándá occupies the Spec of IP, and ?útè-sà ‘yesterday’ is adjoined to the IP. If the three constraints and their ranking suggested in (25) were applied to this sentence, then the (a)-candidate in the following would be chosen as the optimal candidate, although the (b)-candidate should be the optimal one:

(29) [IP 1?útè-sà [IP 2sándá [VP 2sómbá 3thíímé]]]
yesterday-3f.sg Sanda fish cooked Align-XP, L >> *P-Phrase >> Align-XP, R

a. ( ) ( ) ( ) ( ) *** *
b. <desired>( ) ( ) *! ** **
c. ( ) ( ) *!** * **
d. ( ) ( ) *!** ** *
(a) does not violate Align-XP, L because the left edge of all of the XP’s (the adverb ?útè-sà, the subject NP sándá, and the VP as well as the object NP sómbá) coincide with the phonological boundaries. The desired candidate (b) violates Align-XP, L because the left edge of the subject NP sándá does not coincide with the left edge of a phonological phrase.

In order to account for the effects of focus on phonological phrasing, some constraints relevant to focus should be introduced. First I adopt the following constraint concerning focus (proposed by Pierrehumbert and Beckman 1988 for Japanese. See also Nagahara 1994):\(^5\)

\[(30) \quad \text{Focus-Left-Edge:} \]

A focused constituent is preceded by a p-phrase boundary.

Given this, consider the following sentence:

\[(31) \quad 2\text{swe } 1\text{hi-á? } 1\text{khidzi-`-à? } 2\text{haagàsúkéé j?} \]

\[\text{now } \text{hi-3pl village-SPEC-3pl bring-up j?} \]

Elderkin (1990: 155-156)

‘... and when they brought up villages, the idea of villages, ...’

Here, according to Elderkin (1989: 156) “both hi and khidzi are marked and have tone level 1.” That is, two focused elements are adjacent to each other, and each constituent is preceded by a p-boundary, and both of them have the pitch level 1, indicating that each focused phrase is preceded by a p-boundary. If they were followed by a p-boundary, the prediction would be that there is no downstep between khidzi-`-à? and haagàsúkéé because if so they would have to belong to different p-phrases.

Next, the focus causes an extension of the phonological phrase to which it belongs. This kind of phenomenon is observed in other languages. For example, in

\(^5\) Truckenbrodt (1995: 77) proposes a similar constraint for Chichewa (Here, Ø stands for a phonological phrase):

(i) Focus (Chichewa): A focused constituent is followed by a Ø-boundary (Truckenbrodt 1995: 77)

While in Chichewa, a focused constituent is followed by a p-boundary, in Sandawe it is preceded by a p-boundary.
Japanese when some phrase is focused, then it extends the phonological phrase to which it belongs to the end of the sentence (Nagahara 1994):

(32) Naoko-wa nichiyooobi Nagoya de Mari-ni atta
    Naoko-Top Sunday Nagoya-at Mari-with met
    ‘Naoko met with Mari in Nagoya (last) Sunday’

Here, the underlined element is focused, and the p-phrase is extended to the end of the sentence. Nagahara 1994:30 proposes the following constraint to account for this fact:

(33) Focus-To-End:
    No intervening p-boundary between a focused constituent and the end of the sentence. (Nagahara 1994, see also Truckenbrodt 1995)

However, unlike Japanese, the focused phrase in Sandawe does not always extend the phonological phrase to the end of the sentence. In (27), repeated in (34), the phonological phrase is extended up to the VP. That is, it might be the case that Focus-To-End and Align-XP-L compete with each other.

(34) [IP 1?útè-sà  [IP 2sándá  [VP 2sómbá 3thímé ]]
    (yesterday-3f.sg Sanda) ( fish cooked )
    ‘Sanda cooked the fish yesterday’ Elderkin (1989: (3.10))

If Align-XP, L is ranked above Focus-To-End, then we would have the phonological phrasing shown in (35):

(35) (1?útè-sà) (2sándá) (2sómbá 3thímé)

Here, the left edges of all of the XP’s coincide with the left edge of the phonological phrase. However, this is not the desired phrasing since there is a downstep between ?útè-
sà and sándá, and they should be phrased together. If Focus-To-End is ranked above Align-XP, L, then we would have the phonological phrasing:

\[(36) \quad (1\text{ùtè-sà} \quad 2\text{sándá} \quad 2\text{sójmbá} \quad 3\text{thíímé})\]

Here, all the elements that follow the focused element are grouped together. However, this is also not the desired phrasing since there must be a p-boundary between sándá and sójmbá. Then, Sandawe requires a focus constraint that is different from (33).

One way to resolve this problem would be to adopt a category-specific alignment constraint. That is, we could introduce a constraint like Align-VP, L, and rank it above Focus-To-End. Then, Focus-To-End is effective up to a VP boundary, and we would get a desired result. However, it is not appropriate to introduce a category-specific constraint on conceptual grounds. It does not explain why only VP, but not the other lexical categories, should be referred to in the alignment constraint (Inkelas and Zec 1995, among others).

I propose the following constraint:

\[(37) \quad *\text{Focus-Right-Edge}: \text{A focused constituent may not be followed by a p-boundary.}\]

Because some constituent is focused if it is relatively more prominent than the other element, (37), in effect, forces the element that immediately follows the focused element to have a downstepped register of pitch, and the focused phrase gets prominence with respect to the immediately following element.

Note that *Focus-Right-Edge is similar to Focus-To-End in that both of them eliminate potential p-boundaries after the focus. Since I don’t have the data that strongly argue for the existence of Focus-To-End in Sandawe, I don’t use Focus-To-End in my analysis of the Sandawe data.

Notice that (37) can be violated. Consider the sentence in (31), repeated here in (38):
(38) 2swe ¹hi-á? ¹khidží-`à? ²haagásúkéé j?
    now hi-3pl village-SPEC-3pl bring-up j? Elderkin (1990: 155-156)
    ‘... and when they brought up villages, the idea of villages, ...’
    Here, the two focused constituents stand adjacent to each other, and both of them have
the highest pitch level, indicating that *Focus-Right-Edge (37) is ranked below Focus-
Left-Edge:

(39)  Focus-Left-Edge >> *Focus-Right-Edge

Let us now consider the ranking between *Focus-Right-Edge and Align-XP, L.
Consider (34) again, repeated here in (40):

(40)  [IP ¹?útè-sà  [IP ²sándá [VP ²sómbá ³thíímé ]]] (=(34))
    yesterday-3f.sg Sanda fish cooked
    ‘Sanda cooked the fish yesterday’ Elderkin (1989: (3.10))

As is shown below, the desired phonological phrasing violates Align-XP, L but does not
violate *Focus-Right-Edge:

(41)  [IP ¹?útè-sà  [IP ²sándá [VP ²sómbá ³thíímé ]]
    yesterday-3f.sg Sanda fish cooked *Focus-Right-Edge Align-XP, L
a.  *! ( ) ( )

b.  *( ) ( ) ( ) *!

(a) violates Align-XP, L since the left edge of the subject NP sándá does not coincide
with a p-boundary. In contrast (b) does not violate Align-XP, L, but violates *Focus-
Right-Edge since the focused element ?útè-sà is immediately followed by a p-boundary.
Therefore I conclude that the *Focus-Right-Edge is ranked higher than Align-XP, L:

(42)  *Focus-Right-Edge >> Align-XP, L
Given (39) and (42), we have the following ranking:

(43) Focus-Left-Edge >> *Focus-Right-Edge >> Align-XP, L

Since Align-XP, L outranks *P-Phrase, which outranks Align-XP, R, we have the following ranking:

(44) Proposed Ranking
Focus-Left-Edge >> *Focus-Right-Edge >> Align-XP, L >> *P-Phrase >> Align-XP, R

Given this ranking, let us consider every potential phonological phrasing for the sentence (40), repeated here:

(45) [IP 1\?útè-sà [IP 2sándá] [VP 2sómbá 3thíímé ] ] (=(40))

Yesterday-3f.sg Sanda Fish cooked F-L-E *F-R-E Align-XP,L, *P-Ph

a. * ( ) ( ) ( ) ( ) *! ****
b. * ( ) ( ) ( ) *! ***
c. * ( ) ( ) ( ) *! ** ***
d. * ( ) ( ) ( ) *! **
e. * ( ) ( ) ( ) * ***!
f. * ( ) ( ) ( ) * *
g. * ( ) ( ) **!* **
h. * ( ) **!* *

Here, all the candidates satisfy Focus-Left-Edge. Candidates (a)-(d) violate *Focus-Right-Edge since the focused element 1\?útè-sà is immediately followed by a p-boundary. And this is a fatal violation for those candidates. Candidates (c), (d), (g), and (h) violate Align-XP, L because the left edge of the VP and that of the object NP sómbá do not coincide with a p-boundary. Candidates (e)-(h) violate Align-XP, L because the left edge of the subject NP sándá does not coincide with a p-boundary. Notice that candidates (g)
and (h) have three violations of Align-XP, L at this point, and I take the second violation to be a fatal violation for them.

Thus the candidate (45e) and (45f) have survived at this point. (45e) violates *P-Phrase three times (it has three p-phrases) while (45f) violates it twice (it has two p-phrases). Therefore, (45f) is the optimal candidate.

Notice that the pitch level of sòmbá ‘fish’ is not as high as that of the sentence initial element. As I pointed out earlier in (14), the pitch level of the p-phrase initial element is dependent on the pitch level of the preceding element, even though there is a p-boundary.

The following sentence also shows the same downstep pattern:

(46) 1?útè-sà 2sòmbá 2géélétà 3thíímé
       yesterday-3f.sg. fish at the baobab cooked

‘Yesterday she cooked the fish at the baobab’

Suppose that this sentence has the following syntactic structure where the locative expression gélétà ‘at the baobab’ is adjoined to VP, and sòmbá ‘fish’ is adjoined to IP as a result of scrambling over the locative (and the omitted subject pro):

(47) [IP 1?útè-sà [IP 2sòmbá, [IP pro [VP 2géélétà [VP tì 3thíímé ]]]]]

This sentence supports further my analysis under the assumption that the VP-adjoined position is in VP (that is, the left edge of the VP is the higher segment of the VP, and the locative gélétà ‘at the baobab’ is included in the phonological phrase to which the verb thíímé ‘cooked’ also belongs (but cf. Hale and Selkirk 1987)). Then, *Focus-Right-Edge requires that sòmbá belong to the p-phrase to which the focused ?útè-sà belongs, and the left edge of the VP (the higher segment of the VP) corresponds to the left edge of the p-phrase, resulting in the following phonological phrasing, as expected:

(48) (1?útè-sà 2sòmbá) (2géélétà 3thíímé)
The same downstep pattern is also found in the following sentences:

(49)  \(1^\text{gélá-á} \quad 2^?útè \quad 2^\text{sómbá} \quad 3^\text{thíímé}\)

Gele-Nom       yesterday    fish         cooked         Elderkin (1989: 113 (3.43))

‘Gele cooked the fish yesterday’

Here, the sentence initial subject bears a nominative suffix, and is focused. I leave open the exact syntactic structure of this sentence because it is not clear whether the subject undergoes movement over the adverb ?útè ‘yesterday’ or the adverb is adjoined to I’ position or the subject is a major subject. Whatever the syntactic details are, the resulting phonological phrasing is the same. (49) has approximately the following structure (where FP is a projection of a functional category, possibly IP):

(50)  \[
\text{Gele-Nom} \quad \text{yesterday} \quad \text{fish} \quad \text{cooked}
\]

Here, *Focus-Right-Edge requires that ?útè belong to the same p-phrase as gélá-á’s, and Align-XP, L requires that the left edge of the VP coincide with a p-boundary, resulting in the following phonological phrasing:

(51)  \(1^\text{gélá-á} \quad 2^?útè \quad 2^\text{sómbá} \quad 3^\text{thíímé}\)

So far, I have motivated the ranking in (44). In the next section, I apply the proposed ranking to some other downstep patterns.

3.3 Other Downstep Patterns

Now let us consider other downstep patterns under the proposed ranking in (44). If a focused constituent immediately precedes VP, then the focused constituent and the VP should belong to the same p-phrase because *Focus-Right-Edge is ranked higher than Align-XP, L. That is, the alignment of the left edge of VP with the left edge of a p-phrase is blocked by *Focus-Right-Edge. Consider (8), repeated here:
(52) 1?útè-sà            2sómbá  3thíímé    Elderkin (1989: (3.3))
yesterday-3f.sg.  fish  cooked
‘Yesterday she cooked the fish’

\[ [\text{IP}] \ 
\begin{array}{llll}
\text{pro} & \text{[VP NP V]} & F-L-E & *F-R-E \text{ Align-XP,L, } *P-\text{Ph}
\end{array}
\] 

(a) and (b) violate *Focus-Right-Edge because the focused constituent is directly followed by a p-boundary. And this is a fatal violation for them. (c) and (d) violate Align-XP, L because the left edge of the VP as well as that of the object NP does not coincide with a p-boundary. Then, *P-Phrase is the crucial constraint in this case. (c) violates it twice because there are two p-phrases, while (d) violates it once because there is only one p-phrase. Therefore, (d) wins as the optimal candidate.

The same pattern shows up in the following sentence. (I ignore the sentence initial non-focused temporal adverb ?útè. I will discuss this matter at the end of this section.)

(53) 2?útè              1sándá-á         2sómbá    3thíímé    (=7)
yesterday  (Sanda-Nom  fish  cooked)
‘Yesterday she cooked the fish’    Elderkin (1989: 96; (3.10))

Here, the subject sándá ‘Sanda’ is focused. I assume the following syntactic structure for this sentence (the adverb omitted):

(54)  
\[ [\text{IP} \ 1\text{sándá-á} \ [\text{VP} \ 2\text{sómbá} \ 3\text{thíímé}]}} ]
\]
Sanda-Nom  fish  cooked

Since the subject sándá is focused, it may not be followed by a p-boundary because of *Focus-Right-Edge. Since this constraint is ranked higher than Align-XP, L, the left edge of the VP may not coincide with the left edge of a p-phrase. Therefore, the following p-phrasing results, as expected:

(55) 1sándá-á 2sómbá 3thíímé
Let us now consider the case where the object, which follows the subject, is focused:

(56) 2útè 1sándá 1sómbá-sà 2thímé Elderkin (1989: 96: (3.10))
yesterday (Sanda) (fish-3f.sg cooked) ‘Sanda cooked the fish yesterday’

I assume that (56) has the following syntactic structure (I ignore the adverb):

(57) [IP 1sándá [VP 1 sómbá 1thímé ]]
Sanda fish cooked F-L-E *F-R-E Align-XP,L, *P-Ph

a. *( ) ) *! * ** *
b. *( ) ) *! * * ** **
c. *( ) ) **
d. *( ) ) *( ) ) *! ***

(a) and (b) violate Focus-Left-Edge since the focused object is not preceded by a p-boundary, and this is a fatal violation. (d) violates *Focus-Right-Edge since the focused object is followed by a p-boundary. Therefore, (c) is the optimal candidate.

Let us now consider the case where the verb is focused:

(58) 2útè 1sándá 1sómbá 1thímé-sà Elderkin (1989: 96: (3.10))
yesterday (Sanda) (fish) (cooked-3f.sg) ‘Sanda cooked the fish yesterday’

The syntactic structure of (58) is as follows (the adverb is omitted):

(59) [IP 1sándá [VP 1sómbá 1thímé-sà ]]
Sanda fish cooked-3f.sg F-L-E *F-R-E Align-XP,L, *P-Ph

a. *( ) ) *! * ** *
b. *( )( ) ) *! * **
c. *( ) ) *! *!* **
d. *( )( ) ) * * ***
(a) and (b) violate Focus-Left-Edge since the focused verb does not follow a p-boundary, and this is a fatal violation for them. All the candidates violate *Focus-Right-Edge since the focused verb is located in the sentence final position. (c), but not (d), violates Align-XP, L since the left edge of the VP does not coincide with a p-boundary. Hence (d) is the optimal candidate.

So far, we have seen the cases that are accounted for by the ranking (44). However, consider the following sentence:

(60)     1sómbá-sà 2sándá 3?útè 3thíímé Elderkin (1989: 95: (3.5))
fish-3f.sg Sanda yesterday cooked
‘The fish, Sanda cooked yesterday’

Here, the non-focused temporal adverb ?útè ‘yesterday’ has a lower pitch level than the preceding phrase. Assuming that the temporal adverbs are adjoined to IP, sómbá-sà ‘fish-3f.sg’ and sándá ‘Sanda’ occupy Spec or adjoined position of some functional category:

(61) [FP 1sómbá-sà [FP 2sándá [IP 3?útè [IPpro Infl [VP t_i 3thíímé ]]]]]
fish-3f.sg Sanda yesterday cooked

Under my analysis presented thus far, it is expected that it would have the phonological phrasing (62a), but not the desired one (62b):6

(62) a.*( 1sómbá-sà 2sándá) (3?útè) (3thíímé) predicted
b.  (1sómbá-sà 2sándá 3?útè) (3thíímé) desired
fish-3f.sg Sanda yesterday cooked

Since *Focus-Right-Edge, unlike Focus-To-End, bans only a p-boundary which immediately follows the focused constituent, it does not exclude the phrasing in (62a) where the boundary does not immediately follow a focused constituent. This might

---

6 I am assuming that the temporal adverb is a lexical category, which is visible to the alignment constraint. If it is a functional category, the desired phrasing is obtained without any additional assumptions.
suggest that something like Focus-To-End is correct in that it deletes not only a p-boundary that immediately follows the focused constituent, but also some other p-boundaries that follow it. However, it might be too hasty to draw such a conclusion since *útè `yesterday’ shows an exceptional behavior when it is not focused. Consider (7)/(53), repeated here as (63):

(63) 2?útè 1sándá-á 2sómbá 3thímé  
  yesterday Sanda-Nom fish cooked  
  ‘Yesterday she cooked the fish’  Elderkin (1989: 96; (3.10))

When it is located in sentence initial position and is not focused, the adverb does not have the highest pitch level. This is different from the non-focused locative expressions and arguments that are located in the sentence initial position:

(64) 1gélétà 1sómbá-sà 2thímé  
  at the baobab fish-3f.sg. cooked  
  ‘the fish, she cooked at the baobab’

(65) 1Sándá 1sómbá 2thímé-sù  
  Sanda fish cook-3f.sg.future  
  ‘Sanda will cook the fish’  Elderkin (1989: (3.64))

They bear the highest pitch level even though they are not focused. I assume that the non-focused temporal adverbial *útè `yesterday’ is exceptional in that it bears some low pitch level when it is not focused. I assume, however, that it has an expected pitch level if it itself is focused or if it is affected by the focus related constraints. That is, if the temporal adverbial is focused, its left edge coincides with the left edge of a p-boundary (Focus-Left-Edge) and it is not followed by a p-boundary (*Focus-Right-Edge). If the non-focused temporal adverbial immediately follows a focused constituent, then it cannot be preceded by a p-boundary, as predicted by *Focus-Right-Edge, and it belongs to the p-phrase to which the preceding focused constituent belongs, as in (49). I leave open some other questions as to why the non-focused temporal adverb has pitch level 3 in (60), why it is exceptional, and so on.
In this section, I have shown that the ranking in (44), i.e., Focus-Left-Edge >> *Focus-Right-Edge >> Align-XP, L >> *P-Phrase, accounts for the various downstep patterns between words in Sandawe.

4. Multiple Spell-Out and Phonological Phrasing

In this section, I speculate on the status of branching categories.\(^7\)

(66) \[1\text{Sándá} \quad 2\text{sómbá} \quad 2\text{thíímé-sù} \]
\[
\text{Sanda} \quad \text{fish} \quad \text{cook-3f.sg.future} \\
(\quad)(\quad)
\]
\[\text{‘She will cook Sanda’s fish’} \quad \text{Elderkin (1989: 123: (3.69))}\]

Here, the object \textit{Sánďá sómbá} ‘Sanda’s fish’ is branching, and it corresponds to a phonological phrase. The data in the previous section do not include branching NP’s. As I will show below, the present approach that is based on an alignment constraint fails to distinguish branching and non-branaching categories. I argue that a version of multiple Spell-Out gives an account for this distinction.

4.1 Branching Categories and Alignment Constraints

Let us consider the following examples, where there is no focus:

(67) a. \[1\text{Sándá} \quad 2\text{sómbá} \quad 2\text{thíímé-sù} \]
\[
\text{Sanda} \quad \text{fish} \quad \text{cook-3f.sg.future} \\
(\quad)(\quad)
\]
\[\text{‘She will cook Sanda’s fish’} \quad \text{Elderkin (1989: 123: (3.69))}\]

b. \[1\text{Sándá} \quad 1\text{sómbá} \quad 2\text{thíímé-sù} \]
\[
\text{Sanda} \quad \text{fish} \quad \text{cook-3f.sg.fut.} \\
(\quad)(\quad)
\]
\[\text{‘Sanda will cook the fish’} \quad \text{Elderkin (1989: 121 (3.64))}\]

In (67a), we have a branching NP \textit{Sánďá sómbá} ‘Sanda’s fish’ and there is a downstep within the NP, but there is no downstep between the NP and the verb \textit{thíímé-sù}. That is,
the branching object is not p-phrased with the verb. In contrast, if the object is not branching as in (67b), it is p-phrased with the verb.

Let us first consider the structure of the branching NP in (66a). Suppose that it has the following DP structure (Abney 1987):

(68) \[
\begin{array}{c}
\text{DP} \\
\text{NP} \\
\text{Sándá} \\
\text{D} \\
\text{NP} \\
\text{sómbá}
\end{array}
\]

Here the possessor is in the Spec of the DP. Assuming that this entire DP is generated as a complement of the verb, the sentence has the following structure:

(69) \[
\begin{array}{c}
\text{IP} \\
\text{pro} \\
\text{Infl} \\
\text{VP} \\
\text{DP} \\
\text{NP} \\
\text{Sándá} \\
\text{NP} \\
\text{sómbá} \\
\text{thímé-sù}
\end{array}
\]

Given the ranking Align-XP,L >> *P-Phrase, phonological phrasing of the sentence would be as follows:

(70) \[
\begin{array}{cccc}
\text{IP} & \text{pro} & \text{Infl} & \text{VP} [\text{DP [NP Sándá] [NP sómbá]}] \text{ thímé-sù}] \\
\text{Align-XP,L} & \text{*P-Ph}
\end{array}
\]

Here (b) is the desired candidate, where the entire DP is p-phrased independently of the verb, but it is not chosen as the optimal candidate here. It violates Align-XP, L because the left edge of the NP sómbá ‘fish’ does not coincide with a p-boundary. Rather, (c) is

---

7 See Bickmore (1990), Inkelas and Zec (1990), among others, for discussion on branching categories and phonological phrasing.
chosen as the optimal candidate since it does not violate Align-XP, L, and it has less violations of *P-Phrase than (d) which does not violate Align-XP, L either.

The approach that is based on alignment constraints does not provide a straightforward account for this problem. In languages like Sandawe where the word order is O-V and Align-XP, L >> *P-Phrase >> Align-XP, R, when there is no focus in a sentence, there is no way to create a phonological boundary between O and V since the left edge of the VP always coincides with the left edge of the object NP. It could be the case that the branching object moves out of the VP so that the verb is not grouped together with the branching object, but it does not seem reasonable that kind of movement operation applies only if the object is branching.

In order to solve this problem, I introduce some recent approaches to syntax-phonology mapping in the next subsection.

4.2. Multiple Spell-Out and Phonological Phrasing

Within the framework of the minimalist program (Chomsky 1995, 2000, 2001a, b), the operation Spell-Out plays a role of relating syntax to phonology. Lexical items that have phonological, semantic, and syntactic features are extracted out of Lexicon, and undergo some syntactic operations such as a binary operation Merge which assembles lexical items into a larger syntactic object, and at some point of derivation Spell-Out strips away the phonological features of the lexical items, and sends them to the phonological component (PF).

Collins (2001a, b) (cf. Chomsky 2000, 2001a, b, Uriagereka 1999) argues that Spell-Out must apply in a cyclic manner (Multiple Spell-Out). More specifically, Spell-Out applies every time all the properties of a lexical item are satisfied:

(71) Spell-Out occurs every time a lexical item is saturated.

Suppose that Spell-Out defines a phonological phrase (Collins 2001a, b, Uriagereka 1999). Then, phonological phrasing is determined in a cyclic manner as syntactic derivation goes on. Since the phonological phrasing is defined in terms of
lexical category (see (16)), let us assume that Spell-Out of a lexical category, but not a functional category, defines phonological phrasing. In order to formulate this, it is necessary to define the notion Spell-Out of X:

(72) Spell-Out of X spells-out X and lexical items with which X has syntactic relations.

Thus in (73), where X’s properties are satisfied by Y (X’s complement) and Z (X’s specifier), Spell-Out of X spells-out X, Y, and Z, resulting in a string of Z-X-Y:

(73)

If X is a functional category, then Spell-Out of X does not define a phonological phrase, but if X is a lexical category, then Spell-Out of X defines a phonological phrase. Let us call a phonetic string PS that is spelled-out by Spell-Out of X PS(X), and a phonetic string that was spelled-out before Spell-Out of X PS’. I suggest that Spell-Out of X, X a lexical category, defines a phonological phrase P in the following manner:

(74) a. P includes both PS(X) and PS’.
     b. P includes PS(X), and does not include PS’.

I assume that (74) is a parameter: [+/- PS’ included]. Given this assumption, let us consider the following hypothetical derivation, where L, M, N = lexical categories, F = functional category:

(75) a. F L b. M F L c. M F L

In (75a), a functional category F and a lexical category L are merged. Notice that before this Merge operation applies, Spell-Out does not apply to each of F and L, perhaps
because Spell-Out is an operation that feeds a result of syntactic computation to interface levels. I formulate this intuitive idea in the following way, as a formal condition on Spell-Out:

\[(76) \text{ Spell-Out may not apply to a lexical item which has not entered into any syntactic operation yet.}\]

In (75b), a lexical category M is merged. Notice that M, which has not undergone any syntactic operation before it is merged, has not been spelled-out when it is merged because of (76). Suppose that F is saturated after M is merged. Then, Spell-Out of F occurs, yielding a phonetic string M-F-L. Note that since this is Spell-Out of a functional category, no phonological phrase is defined. In (75c), a lexical category N is merged with a syntactic object created in (75b). Suppose that N is saturated as a result of this Merge operation. Then Spell-Out of N occurs. Since N is a lexical category, Spell-Out of N defines a phonological phrase. Suppose that a language has a value [+ PS’ included] = (74a). Then, Spell-Out of N results in a phonological phrase that includes both N and M-F-L = PS’. Suppose that a language has a value [-PS’ included] = (74b), then Spell-Out of N results in a p-phrase that includes N, but not M-F-L.

Let us now consider the contrast in (67), repeated here:

\[(77) \begin{align*}
\text{a.} & \quad 1\text{Sándá}^1 2\text{sómbá}^2 2\text{thíímé-sù} \\
& \quad \text{Sanda} \quad \text{fish} \quad \text{cook-3f.sg.future} \\
& \quad (\quad) (\quad) \\
& \quad \text{‘She will cook Sanda’s fish’} \quad \text{Elderkin (1989: 123: (3.69))} \\
\text{b.} & \quad 1\text{Sándá} \quad 1\text{sómbá} \quad 2\text{thíímé-sù} \\
& \quad \text{Sanda} \quad \text{fish} \quad \text{cook-3f.sg.fut.} \\
& \quad (\quad) (\quad) \\
& \quad \text{‘Sanda will cook the fish.’} \quad \text{Elderkin (1989: 121 (3.64))}
\end{align*}\]

The branching object is not p-phrased with the verb in (77a), while the non-branching object is p-phrased together with the verb in (77b). Let us first consider the derivation of (77a):
In (78a), D and sómbá ‘fish’ are merged. In (78b), Sándá is merged with the syntactic object formed in (78a), resulting in DP Sándá sómbá, ‘Sanda’s fish.’ At this point, D is saturated, and Spell-Out of D occurs. Since D has syntactic relations with Sándá and sómbá, it results in a phonetic string Sándá-D-sómbá, where D is phonetically empty. However, since this is Spell-Out of a functional category, it does not define a phonological phrase. In (78c), the verb is merged with the DP formed in (78b). And as a result of this merger, the V is saturated. Then Spell-Out of V occurs. It spells-out DP-V. Since this is Spell-Out of a lexical category, it defines a phonological phrase. I assume that Sandawe is [- PS’ included] (=74b)); that is, a phonological phrase is formed so that it includes V itself but excludes a phonetic string that has already been spelled-out, i.e., the DP. This yields the desired result.

Let us next consider (77b). I just show the derivation of the VP, which is relevant to the present discussion:

(79) sómbá thíímé-sù

---

8 An example of [+PS’ included] might be Japanese (Selkirk and Tateishi 1990):
(i) a. [IP Aoyama-ga [VP [Yamaguchi-no aniyome-o] yonda]]
Aoyama-Nom Yamaguchi-Gen sister-in-law-Acc called
( ) ( ) --- p-phrasing
‘Mr. Aoyama called Mr. Yamaguchi’

b. [IP Aoyama-no Yamaguchi-ga [VP aniyome-o yonda]]
Aoyama-Gen Yamaguchi-Nom sister-in-law-Acc called
( ) ( ) --- p-phrasing
‘Mr Yamaguchi from Aoyama called his sister-in-law.’

(i.a) has a branching object while (i.b) has a non-branching object. The phonological phrases indicated above are the domain for downstep. Irrespective of whether the object is branching or not, the object is phrased together with the verb. That is, the Spell-Out of V results in a p-phrase that includes not only V but also PS’, i.e., the object.
Here, the object sómbá ‘fish’ and the verb thímé-sù ‘cook’ are merged. Recall that as I formulated in (76), Spell-Out does not apply before this merge operation takes place. Suppose that as a result of merger, the verb is saturated. Then, Spell-Out of V applies, spelling out a phonetic string sómbá-thímé-sù. Since V is a lexical category, Spell-Out of V defines a phonological phrase. Assuming that Sandawe is [- PS’ included], the resulting p-phrase includes V and the object sómbá, both of which are spelled-out by Spell-Out of V (see (72)), but does not include a phonetic string that has been spelled-out. However, there is no phonetic string that was spelled-out prior to Spell-Out of V. Therefore, the p-phrase which has V and the object sómbá is obtained, as expected.

Notice that the assumption in (76) plays an important role in distinguishing branching and non-branching objects. Since the non-branching object does not undergo Spell-Out before it is merged with V because of (76), (72) forces it to be spelled-out together with V, and the object and the V are phrased together. In contrast, the branching object is spelled-out before it is merged with V, and it is not phrased together with V because of [-PS’ included] = (74b).

5. Conclusion

In this paper, I gave an account for downstep in Sandawe. I proposed a constraint *Focus-Right-Edge, which prohibits a phonological boundary right after the focused constituent. I showed that the ranking (44), repeated here, accounts for a wide range of the data.

(80) Proposed Ranking:
Focus-Left-Edge >> *Focus-Right-Edge >> Align-XP, L >> *P-Phrase >> Align-XP, R

I speculated on the status of branching categories, and I argued that a version of multiple spell-out approach provides an account for the difference between branching and non-branching categories with respect to phonological phrasing.
References


