

# Prosody and EPP in Swahili

Soo-Hwan Lee  
New York University

## 1 Introduction

The precise motivation for affixation has not been entirely settled. Noyer (1992) and Hankamer & Mikkelsen (2018) argue that the identity of an affix is recognized in syntax-free contexts or in postsyntactic environments. On the other hand, Richards (2010, 2016) proposes a way of identifying affixes by looking into their metrical dependencies initially detected in narrow syntax. Here, I argue alongside Richards (2016) that these suprasegmental features are visible in syntax and that they trigger XP-movements (see also Branan 2018). According to Contiguity Theory (Richards 2016), overt movements triggered by syntactic features such as *[uwh]* and Extended Projection Principle (henceforth EPP) in the Minimalist Program (Chomsky 1993, 1995) are reanalyzed as operations sensitive to the interaction between syntax and prosody. To be more specific, narrow syntax looks at certain phonological information that works in favor of the initial shape of prosody. This suggests that syntactic movement is sensitive to prosodic contiguity prior to spell-out. Richards (2016) discusses some key motivations for movement. They include Probe-Goal Contiguity, Affix Support, and Untethering. Adopting some of the basic assumptions proposed in Match Theory (Selkirk 2009, 2011), Contiguity Theory (henceforth CT) looks into phonological motivations for *wh*-movement and EPP.

In this paper, I argue that Swahili demands additional explanation as to how prosodic requirements are satisfied. In detail, I present an analysis accounting for the *wh*-in-situ phenomenon as well as the presence of EPP in Swahili. With regards to the presence of EPP, I propose that Swahili tense affixes require metrical boundaries on *both* left and right of their peripheries. The metrical boundary on the right is satisfied by the phonological content inside *vP*. The metrical boundary on its left is satisfied by an XP targeting [*Spec,TP*] which gives rise to the desired EPP-effect.

The layout of the paper is as follows: Section 2 discusses a traditional view (i.e., the Minimalist Program) on analyzing movement. In section 3, we look into how CT reanalyzes syntactic movement in relation to phonology. Section 4 deals with the *wh*-in-situ phenomenon in Swahili. Section 5 summarizes the ways in which EPP is satisfied for English and Japanese. In section 6, I provide an analysis as to how EPP is satisfied in Swahili. Finally, section 7 concludes.

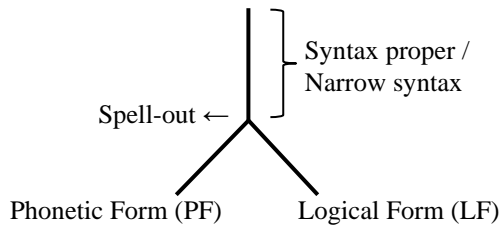
## 2 Syntax without phonology

In this section, I summarize a dominant take on how syntactic movements are analyzed. It has been widely assumed in the Minimalist Program (Chomsky 1993, 1995) that syntax occupies a domain in the grammar which is governed by motivations independent of phonology. The domain in which these syntactic motivations are triggered is referred to as narrow syntax or syntax proper. Assuming the (inverted) Y-model updated in the Minimalist Program, Chomsky notes that the work of syntax is sent off to the phonological domain, namely the Phonetic Form (PF), only after spell-out. Despite the more recent development on derivation by phase or multiple spell-out (see Chomsky 2001, 2008 in particular), one essential tenet remains unchanged: there are separate workspaces for syntax and phonology. Consider the following model of grammar adopted under the generative tradition:

---

\* First and foremost, I thank Sugira Gustave and Peace Fidele for sharing their language with me. I owe many thanks to Norvin Richards, Michael Barrie, Juliet Stanton, and the audience at AMP 2019 for their comments. This work was supported by the Global Research Network program (NRF-2017S1A2A2039972). All remaining errors are my own.

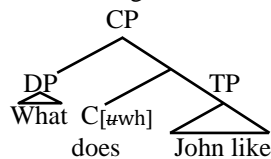
## (1) The (inverted) Y-model



Within syntax proper (or narrow syntax), uninterpretable features are checked. Among the uninterpretable features proposed in the literature,  $[uwh]$  on C (i.e., complementizer) is one that is worth mentioning. In English,  $[uwh]$  is checked by moving a *wh*-word/phrase to the specifier position of CP (i.e., [Spec,CP]). This is done by overt *wh*-movement:

(2) What  $C_{[uwh]}$  does John like  $\leftarrow$ what $\rightarrow$ ? (*wh*-movement)

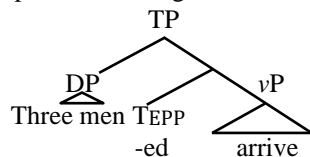
In (2),  $[uwh]$  does not contribute to the semantics and must be eliminated prior to spell-out via *wh*-movement. Hence, it follows that *wh*-movement in English is a part of narrow syntax and its derivational consequence later feeds into PF and LF.<sup>1</sup> Here, C is referred to as the probe which searches down for its goal, namely the *wh*-word *what*. During the syntactic derivation, the probe (i.e., C) and the goal (e.g., *what*) form an Agree relation via a Spec-head configuration (Chomsky 1995, 2000):

(3) A Spec-head configuration between *what* and C

Another instance of movement is triggered by what is referred to as EPP. According to Chomsky, EPP on T (i.e., tense) requires TP to have a specifier (i.e., [Spec,TP]). In English, EPP can be satisfied either by inserting an expletive in [Spec,TP] (see Deal 2009 for a different take on expletives) or by moving an XP up to [Spec,TP]. Consider the following examples:

(4) a. There  $T_{EPP}$  arrived three men. (*there*-insertion)  
b. Three men  $T_{EPP}$  arrived  $\leftarrow$ three men $\rightarrow$ . (XP-movement)

In (4a), EPP is satisfied via *there*-insertion. This suggests that the probe (i.e., T) and the goal (i.e., *three men*) are capable of undergoing Agree without relying on movement. In the case of (4b), however, EPP is satisfied by placing the probe and the goal in a Spec-head configuration via overt movement similar to how the probe and the goal satisfy Agree via *wh*-movement in (2):

(5) A Spec-head configuration between *three men* and T

<sup>1</sup> Logical Form (LF) is where semantic interpretations are evaluated. Here, we do not discuss the role of LF in detail as the primary concern of this paper is devoted to the interaction between phonology and syntax. For an extensive discussion on LF and covert (*wh*-)movement, refer to Huang (1982).

It is quite significant to note that English T (e.g., *-ed*) is realized on the verb via Affix-hopping (Chomsky 1957), Morphological Merger (Halle & Marantz 1993) or Lowering (Embick & Noyer 2001) in more recent terms. The discussion of realizing T on V in English has a long history, but I argue alongside Halle & Marantz (1993) and Embick & Noyer (2001) that the process takes place post-syntactically at PF. In section 6, the ordering of multiple operations (e.g., movement and Affix Support) is scrutinized and it is crucial to keep in mind that Morphological Merger (or Lowering) comes after all of the syntactic operations are done. If we assume otherwise, certain operations will be prevented from taking effect which will lead to an undesirable result. So as to scrutinize the derivation in narrow syntax, the tree structures given in this paper including (5) reflect the derivation prior to spell-out and not the actual linearization process at PF.

By now, it is not difficult to understand that syntactic features are capable of triggering movement operations such as the ones we saw in (2) and (4b). While an enormous progress has been made in the literature using [*uwh*] and EPP, there is no principled way of actually predicting the presence or the absence of movement under the minimalist framework. For instance, there is no well-thought out generalization as to *why* English requires overt *wh*-movement whereas Korean does not. Also, there is no concise way of telling *why* English requires EPP whereas Welsh does not. Without concrete evidence to support the precise motivation for overt movement, Chomsky's assumptions remain somewhat stipulative. Perhaps, this is a natural consequence of the traditional view which has imposed a strong divide between syntax and phonology. Here, I would like to emphasize that the role of phonology should not be entirely ignored in narrow syntax. In the next section, we look into a theory that takes a departure from the conventional wisdom. In particular, it demonstrates that syntax and phonology put together disentangle some of the problematic knots left unnoticed in the previous literature.

### 3 Contiguity Theory

Richards (2010, 2016) attempts to bring certain phonological components into the workspace of syntax. In doing so, Richards dispenses with assumptions solely grounded upon syntax. His framework (CT) establishes a theoretical stance that calls for a revision of the (inverted) Y-model which has been widely assumed in the minimalist circle. In Richards' view, overt movements triggered by [*uwh*] and EPP are reanalyzed as operations sensitive to phonological factors. In detail, they are suprasegmental features which condition movement in narrow syntax. This suggests that the driving cause for syntactic movement is relevant to prosodic requirements which must be satisfied prior to spell-out. Hence, syntax refers to specific aspects of phonology even before the derivation reaches PF.

Clearly, CT announces a separation from the traditional account discussed in section 2. Note, however, that CT is not the only framework that highlights the connection between phonology and syntax. In fact, Richards discusses motivations for movement which are compatible with Match Theory (Selkirk 2009, 2011). In particular, CT adopts the view that syntactic constituents map onto phonological constituents:

- (6) a.  $X^0$  (syntactic head)  $\rightarrow \omega$  (phonological word)  
 b.  $XP$  (syntactic phrase)  $\rightarrow \varphi$  (phonological phrase)

Furthermore, CT introduces additional prosodic requirements that syntax must satisfy. Among those requirements, Probe-Goal Contiguity and Affix Support are mainly utilized in this paper:

(7) Probe-Goal Contiguity

Given a probe  $\alpha$  and a goal  $\beta$ ,  $\alpha$  and  $\beta$  must be dominated by a single  $\varphi$ , within which  $\beta$  is Contiguity-prominent.

(8) Affix Support

If a head is an affix, there must be a metrical boundary in the direction in which it attaches.

While other requirements are also introduced and utilized in CT, I argue that (7) and (8) are all that is necessary in accounting for the *wh*-in-situ phenomenon as well as the EPP-effect displayed in Swahili (see section 4 and 6). In addition to these requirements, Richards (2019) provides a typological classification of

the prosodic activity of XPs by measuring the prosody (F0) of the following languages:<sup>2</sup>

(9) Left active languages based on pitch peaks (Richards 2019:10)

Korean	1.24
English	1.19
Norwegian	1.19

(10) Right active languages based on pitch peaks (Richards 2019:11)

Zulu	1.07
Basque	1.06
Italian	1.06
Bulgarian	1.04
Icelandic	1.03
Portuguese	1
French	0.95
Russian	0.87

The figures in (9) and (10) indicate the average ratio of the first and the second words' pitch peaks (in Hertz) measured within a noun phrase (e.g., a DP). In this experiment, Richards uses adjectives and nouns to extract these measures. In doing so, he verifies that languages with the average ratio of 1.19 or more are classified as left active whereas those with the average ratio of 1.07 or less are classified as right active. Thus, we see a distinction between the two families of languages.

In addition to pitch peaks, stressed vowel lengths are considered as another criterion in determining whether a given language bears a right or left active edge. Once again, Richards uses DPs consisting of an adjective and a noun (e.g., *Canadian saber*) to confirm that German and English XPs have a prosodically left active edge whereas Russian, Portuguese, and Italian XPs have a right active edge. The average vowel length ratio between the first and the second words' primary stresses are shown below (see Richards 2019 pp.19–23 for further discussion on the setup and the result of the experiment):

(11) Left active languages based on vowel lengths (Richards 2019:21)

German	1.02
English	1.01

(12) Right active languages based on vowel lengths (Richards 2019:21)

Russian	0.87
Portuguese	0.80
Italian	0.72

Here, the threshold for distinguishing right active languages from left active languages seems to be somewhere between the figures of 0.87 (e.g., Russian) and 1.01 (e.g., English). Although more work is to be done on determining prosodic activities of XPs, Richards provides convincing measures and empirical data guiding us to a generalization that connects prosody to syntax. In the next section, I utilize the general concepts and mechanisms laid out in CT to clarify the prosodic motivation for the *wh*-in-situ phenomenon shown in Swahili. Instead of coping with purely syntactic features, I show that the optionality of *wh*-movement is in fact phonologically motivated.

#### 4 *Wh*-in-situ in Swahili

Swahili is a Bantu language with a canonical SVO word order. In terms of forming interrogative sentences, Swahili does not require *wh*-movement. From a minimalist perspective, this is due to the absence of a strong [wh] feature in narrow syntax. Cross-linguistically, Chicheŵa, French, Korean, and

<sup>2</sup> The list is by no means exhaustive.

Japanese display the same pattern. Consider the following examples in Swahili which show the optionality of *wh*-movement:

- (13) The optionality of *wh*-movement in Swahili
- a. huyu mtu mrefu a-na-pend-a nini  
 1.this 1.man 1.tall 3SG-PRES-like-FV what  
 ‘What does this tall man like?’
  - b. nini huyu mtu mrefu a-na-pend-a  
 what 1.this 1.man 1.tall 3SG-PRES-like-FV  
 ‘What does this tall man like?’

Here, I reiterate that there is no principled way of explaining why *wh*-movement is optional (i.e., not obligatory) in languages such as Swahili. In order to tackle this issue, CT attempts to establish a generalization by making use of both syntactic as well as prosodic factors. According to Richards (2010, 2016), there are four ways of categorizing languages based on their syntactic headedness (see Kayne 1994 for a different view) and prosodic boundaries (or prosodic activity discussed in section 3):

(14) The language typology of syntactic headedness and prosodic boundaries

	<b>C to right of TP</b>	<b>C to left of TP</b>
<b>Prosodic boundaries on right of XPs</b>	Basque (Movement required)	Chicheŵa, French (Movement not required)
<b>Prosodic boundaries on the left of XPs</b>	Korean, Japanese (Movement not required)	English, Danish (Movement required)

Among the four options in (14), languages categorized as ‘C to the right of TP + Prosodic boundaries on the right of XPs’ or ‘C to the left of TP + Prosodic boundaries on the left of XPs’ require *wh*-movement. Here, the presence of a prosodic boundary between a non-adjacent probe (i.e., C) and goal (*wh*-XP) disrupts Contiguity (e.g., Probe-Goal Contiguity) which forces the goal to move to a position that satisfies Contiguity, namely [Spec,CP]. On the other hand, languages with ‘C to the right of TP + Prosodic boundaries on the left of XPs’ or ‘C to the left of TP + Prosodic boundaries on the right of XPs’ do not require *wh*-movement. This is mainly because no prosodic boundary intervenes between the probe and the goal. Hence, Contiguity is satisfied without the need for overt movement.

Since Swahili is head-initial with no obligatory *wh*-movement, we expect there to be a prosodic activity on the right of XPs so that Contiguity is satisfied without any explicit application of syntactic operations. Hence, I predict that Swahili patterns with Chicheŵa and French. So as to verify this assumption, Swahili prosody is examined below:

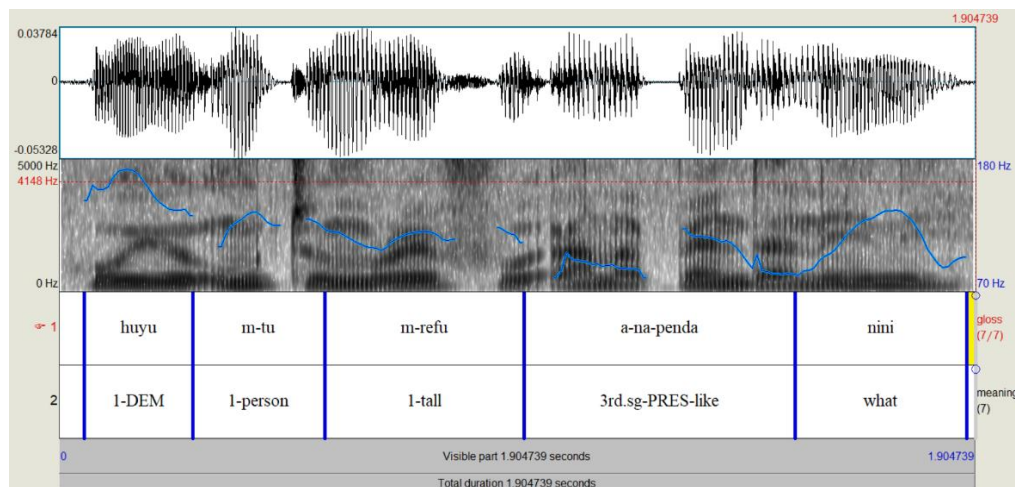


Figure 1: Prosody of (13a)

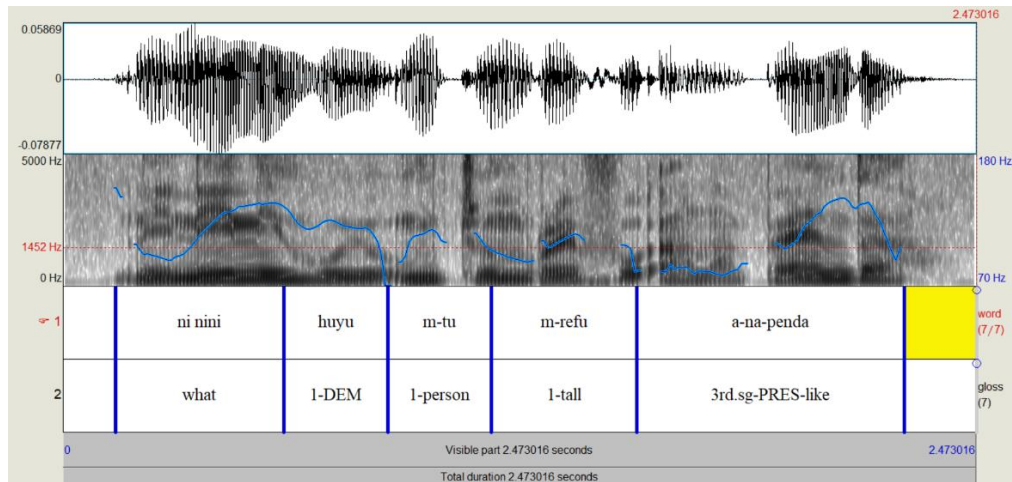


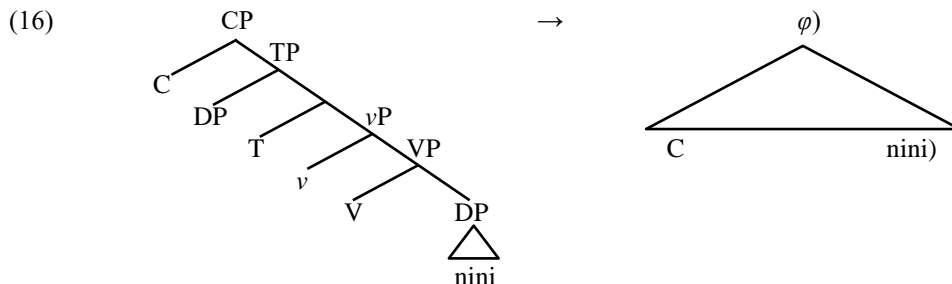
Figure 2: Prosody of (13b)

In both Figure 1 and Figure 2, *nini* has a prosodic boundary on its right edge.<sup>3</sup> As for the noun phrase *huyu mtu mrefu*, I only considered the pitch peaks for the lexical categories, namely the noun (*mtu*) and the adjective (*mrefu*), following Richards’ (2019) diagnostics introduced in section 3. Here, I disregard the demonstrative (*huyu*), since demonstratives in general are classified as a functional category which serves a deictic function that triggers a unique prosodic contour of its own. The average ratio of the pitch peaks for the noun (the first word) and the adjective (the second word) in Figure 1 and Figure 2 is 1.068. Based on the list given in (10), Swahili ranks safely between Zulu and Basque. Thus, I add Swahili to the list of right active languages.

(15) Right active languages based on pitch peaks

Zulu	1.07
Swahili	1.068
Basque	1.06

The prosodic measures for Swahili given in (15) coupled with its syntactic headedness (C to the left of TP) make *wh*-in-situ possible.<sup>4</sup> This adds weight to the claim that syntactic well-formedness is conditioned by prosodic requirements. To elaborate, C and the *wh*-word in Figure 1 are dominated by a single  $\varphi$  within which Probe-Goal Contiguity is satisfied and Affix Support is vacuously satisfied (i.e., C is empty).



<sup>3</sup> The minimum and the maximum pitch for the *wh*-word, *nini*, in Figure 1 is 83.46 Hz and 137.40 Hz. Similarly, the pitch for the *wh*-word in Figure 2 ranges from 91.21 Hz to 139.48 Hz. The appearance of the sentence initial morpheme *ni* in Figure 2 indicates the start of a fronted *wh*-word.

<sup>4</sup> This does not mean that *wh*-movement is entirely banned in Swahili. As long as the probe (e.g., C) and the goal (e.g., *wh*-word) are linearly adjacent to one another, Contiguity is satisfied. Thus, the optional movement of *nini* shown in (13b) is possible. The same argument can be made for Chicheŵa and French optional *wh*-movement.

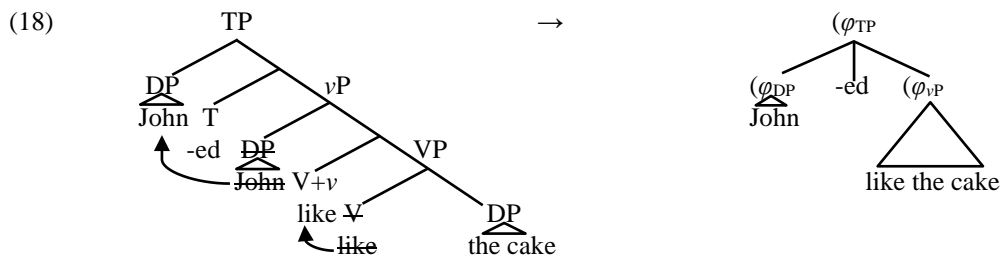
However, it will soon be pointed out that further elaboration is inevitable when accounting for EPP. Simply put, overt movement is necessary when satisfying prosodic requirements between T (i.e., probe) and the subject (i.e., goal) which is not the case for C and the *wh*-word. To settle this discrepancy in Swahili, the following section focuses on how the EPP-effect is motivated in CT using previously analyzed languages.

## 5 Satisfying the EPP

**5.1 EPP in English and Japanese** Richards (2016) provides an extensive cross-linguistic study on the prosodic motivation for EPP. Since English and Japanese EPP have been analyzed under this account, let us briefly go over the discussion on these languages. First, consider the following example in English:

(17) John liked the cake.

By examining DPs consisting of an adjective and a noun (e.g., *an open doorway*), Richards (2019) verifies that English has a prosodic left active edge (see section 3). Also, recall that there are two essential motivations for movement in CT: Probe-Goal Contiguity and Affix Support. The syntactic and prosodic trees schematized below illustrate how these two requirements are met in narrow syntax:



In (18), the movement of the subject *John* satisfies both Probe-Goal Contiguity and Affix Support.<sup>5</sup> Hence, applying a single syntactic operation satisfies both prosodic requirements established in CT which leads to a desirable outcome.

Similar to English, Japanese has a suffixal tense morpheme (e.g., *-da*). However, the two languages differ in terms of their syntactic headedness. Due to this divergence, the movement of the subject in Japanese is not so obvious. Consider the Japanese example given below:

(19) Taroo-ga koohii-o non-da  
 Taroo-NOM coffee-ACC drink-PST  
 ‘Taroo drank coffee.’

According to (14), languages categorized as ‘C to the right of TP + Prosodic boundaries on the left of XPs’ do not require overt movement in syntax proper. Hence, we have to rely on an additional operation which triggers the desired EPP-effect in Japanese. In CT, the operation is referred to as Untethering:<sup>6</sup>

<sup>5</sup> When Probe-Goal Contiguity and Affix Support are both satisfied, Multitasking results:

(i) Multitasking (Richards 2016:342)

At every step in a derivation, if two operations A and B are possible, and the conditions satisfied by A are a superset of those satisfied by B, the grammar prefers A.

<sup>6</sup> Additionally, Prosodic Untethering is introduced in CT. For our current analysis, Prosodic Untethering and Untethering give rise to the same effect.

(i) Prosodic Untethering (Richards 2016:339)

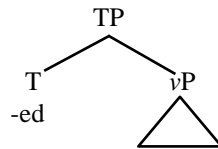
Given two sisters in the prosodic tree, X and Y, delete all existing ordering statements that make reference to either X or Y.

## (20) Untethering

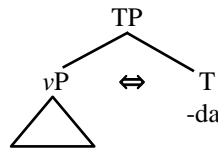
Given two sisters, X and Y, delete all ordering statements that refer to either X or Y.

Untethering applies to T in languages such as Korean and Japanese which deletes all ordering statements. The untethered T in (19) enables the subject *Taroo* to raise to [Spec,TP] similar to how *John* in (17) raises to [Spec,TP]. The difference between Japanese and English boils down to whether Untethering is applied or not. Taking into consideration the analyses put forward by Christophe et al. (1997, 2003) and Nespor et al. (2008), Richards (2016) provides explicit evidence in suggesting that such is the case. Although this paper does not explore this topic in full length, it is crucial to note that Affix Support can be satisfied via different derivational processes:

## (21) a. English Affix Support



## b. Japanese Affix Support (via Untethering)



To briefly summarize, English displays EPP in order to satisfy both Probe-Goal Contiguity and Affix Support (see footnote 5 on Multitasking) whereas Japanese displays EPP only to satisfy Affix Support. Probe-Goal Contiguity need not be satisfied at this point for Japanese, since Contiguity is already satisfied when the probe and the goal are adjacent to one another immediately after Untethering.<sup>7</sup>

**5.2 EPP in Swahili** Interestingly, the lack of obligatory *wh*-movement in Swahili (see section 4) does not suggest that all XP-movements are ruled out. In fact, Carstens (2005, 2011) demonstrates that Bantu languages in general display EPP. Swahili subject raising is one instance. The prosody of (22b) is shown in Figure 3.<sup>8</sup>

- (22) a. Juma a-ta-nunu-a maembe  
 Juma 3SG-FUT-buy-FV 6.mango  
 ‘Juma will buy mangoes.’  
 b. Juma a-li-end-a shule-ni  
 Juma 3SG-PST-go-FV 9.school-LOC  
 ‘Juma went to school.’

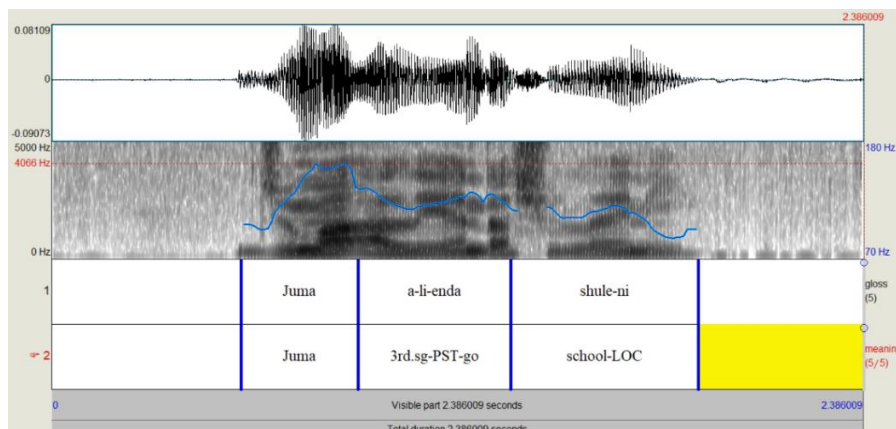


Figure 3: Prosody of (22b)

<sup>7</sup> Richards (2016:122) notes that French T must be adjacent to its goal when the goal precedes it. This is because a prosodic active edge intervenes between the two. In theory, Japanese T and its goal must be adjacent to one another at some point during the derivation before spell-out if Probe-Goal Contiguity is not achieved.

<sup>8</sup> The minimum and the maximum pitch for the subject DP, *Juma*, in Figure 3 is 101.72 Hz and 147.40 Hz.

In Figure 3, the subject DP, *Juma*, has a right active edge. Similar to how C and the *wh*-word satisfy Probe-Goal Contiguity, T and *Juma* are prosodically contiguous without having to move the subject *Juma* to [Spec,TP]. For this very reason, Probe-Goal Contiguity cannot explain the prosodic motivation for Swahili EPP. Alternatively, we have to rely on Affix Support. Similar to English Ts and untethered Japanese Ts, I argue that Ts in Swahili trigger movement of XPs up to [Spec,TP]. Here, XPs are not limited to just agent-bearing DPs. In fact, locatives are also subject to EPP:<sup>9</sup>

- (23) a. nyumba-ni pa-na watu wengi  
 9.house-LOC 16-have 2.person 2.many  
 ‘At the house are many people.’ (Carstens 1997)
- b. mji-ni ku-me-kuf-a watu wengi  
 3.town-LOC 17-PRES-die-FV 2.person 2.many  
 ‘Many people have died in the town.’ (Ashton 1947)

A crucial difference between English and Swahili is that the former realizes its tense morpheme as a suffix, whereas the latter does not. To put it in other words, English Ts necessarily follow the verb, whereas Swahili Ts do not. The following examples illustrate the difference between the two tense morphemes:

- (24) a. John like-**d** Mary.  
 b. Juma a-**li**-pend-a Halima  
 Juma 3SG-PST-like-FV Halima  
 ‘Juma liked Halima.’

Despite the different morphological status of English and Swahili Ts, they share the commonality of being independent from the verb in narrow syntax. In addition to Lowering discussed in section 2, it is known in the literature that English Vs do not raise to T as opposed to French Vs (Pollock 1989). The distribution of the verbs *speak* and *parle* in relation to the temporal adverbs *often* and *souvent* illustrate this point:

- (25) a. John (often) speaks (\*often) Italian.  
 b. Jean (\*souvent) parle (souvent) l’italien  
 Jean (\*often) 3SG.speak (often) Italian  
 ‘Jean (often) speaks Italian.’

Similar to English Vs, Swahili Vs do not undergo movement to T. Henderson (2003, 2013) reports that Swahili undergoes T-to-C movement in the absence of V-to-T movement within relative clauses:

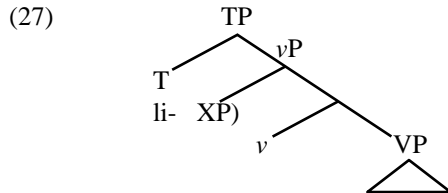
- (26) T-to-C without V-to-T movement  
 mchana **a-li** ye **soma** kitabu  
 1.girl 3SG-PST she read 7.book  
 ‘the girl that read the book’ (Henderson 2003, 2013)

At this point, the prosodic status of Swahili T (e.g., *li*) in syntax proper becomes crucial to our analysis. First, I accept Richards’ analysis that Affix Support in a head-initial language triggers the desired EPP-effect only when the T requires a metrical boundary on its left. A less studied approach to handling this issue is to say that Swahili Ts require metrical boundaries on *both* edges. Richards abstracts away from mentioning affixes requiring multiple metrical boundaries. Nonetheless, I propose that Swahili tense morphemes are such instances which properly account for the EPP facts as well as the phonological linearization of the verbal morphology (e.g., *alipenda* ‘liked’, *anapenda* ‘likes’, *atapenda* ‘will like’). In the next section, we discuss in detail why positing multiple metrical boundaries in narrow syntax best captures the EPP-effect in Swahili.

<sup>9</sup> Lee & Lee (2019) claim that not all locatives in Swahili are DPs. In particular, those that are structurally equivalent to or smaller than the size of a DP undergo locative inversion which satisfies EPP.

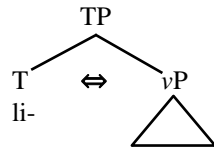
## 6 Affix Support in Swahili

**6.1 A single metrical boundary on T?** On the surface, Swahili tense morphemes look like prefixes. However, the metrical dependency of a prefix alone cannot solve the puzzle introduced in the previous section. Apparently, we are faced with the problem that Affix Support can be satisfied without overt XP-movement. Consider the following tree diagram which eventually leads to a crash in derivation due to the absence of XP in [Spec,TP]:



To circumvent this fallout, an additional mechanism which allows an XP to target [Spec,TP] is necessary. Note that applying Untethering does not give us the desired derivation either. Even if the ordering of T and vP are switched around, we would not expect anything to precede T. This is mainly because T would require a metrical boundary only on its right if it were a prefix. In this respect, Swahili tense morphemes are different from those of Japanese:

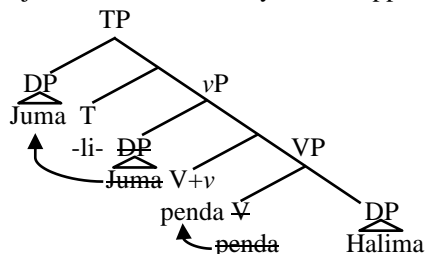
(28) No movement triggered via Untethering



If we maintain the view that Ts in Swahili bear a single metrical boundary, no prosodic motivation for EPP results. With a bit of leeway from Richards' (2016) original analysis, I argue that Ts may require more than one metrical boundary. In the case of Swahili, I propose that there is one metrical boundary for the left periphery (e.g., XP-*li*) and another for the right (e.g., *li*-YP) in narrow syntax. The derivational consequences of having multiple metrical boundaries on T along with the syntactic identity of XP and YP are provided in the following subsection.

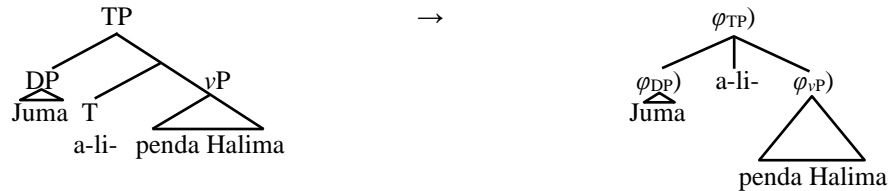
**6.2 Multiple metrical boundaries on T** From a typological standpoint, Swahili and French are both syntactically head-initial and prosodically right active. One major difference between the two, however, is that T in French is a suffix whereas T in Swahili is not. This makes French EPP relatively easy to account for, since Affix Support is satisfied by moving a phonological constituent up to [Spec,TP]. Now, suppose that Swahili T does require a metrical boundary on its left. This essentially allows us to handle Swahili EPP in the same manner as we would handle French EPP. As opposed to French T, however, Swahili T is not a suffix and would thus require an *additional* metrical boundary on its left edge. As a result, it would require metrical boundaries on both left and right. Consider the following tree structure based on (24b):

(29) The subject DP raises to satisfy Affix Support



In (29), two metrical boundaries are present on T and Affix Support is satisfied on both sides: the subject DP (i.e., *Juma*) provides the left metrical boundary and *vP* (i.e., *penda Halima*) provides the right.<sup>10</sup> Subsequently, the probe T and the goal *Juma* form a Spec-head configuration which desirably gives rise to Agree (see section 2). Further note that the prosodic counterpart of the syntactic structure satisfies Contiguity:

(30) Agree and Contiguity satisfied



In (30), the subject agreement marker *a-* is realized only after the subject DP *Juma* raises to [Spec,TP] which is well in line with how agreement works in Bantu languages in general (Carstens 2005, 2011). All in all, a well-formed derivation results using the requirements introduced in Richards' phonology-sensitive design of grammar.

## 7 Conclusion

In this paper, I verified that syntactic phrases in Swahili bear a prosodic right active boundary. Hence, Swahili *wh*-in-situ is accounted for under CT; C (the probe) and the *wh*-word (the goal) satisfy Probe-Goal Contiguity without the presence of *wh*-movement in narrow syntax. In relation to overt movement, I also discussed a syntax-prosody driven motivation for Swahili EPP in part by observing EPP in English, Japanese, and French. Due to the troublesome issue of analyzing Swahili EPP with the operations currently available in CT, I presented a new way of handling the phenomenon. Specifically, I argued that Affix Support applies to affixes bearing one or more metrical boundaries. This implies that affixes in need of multiple metrical boundaries validate syntactic movement. From a cross-linguistic point of view, other Bantu languages such as Chicheŵa and Zulu are likely to display the same phonological motivation for movement. Close examination on metrical dependencies and EPP may shed further light on the status of Ts in Bantu and possibly other languages. While much is left to be explored, I hope the analysis put forward in this paper has demonstrated a way of accounting for the phonological motivations necessary for conditioning syntactic movement in Swahili.

## References

- Ashton, Ethel O. 1947. *Swahili grammar*. 2nd edition. London: Longman.
- Branan, Kenyon. 2018. Relation preservation. Doctoral dissertation, MIT.
- Carstens, Vicki. 1997. Empty nouns in Bantu locatives. *The Linguistics Review* 14:361–410.
- Carstens, Vicki. 2005. Agree and EPP in Bantu. *Natural Language and Linguistic Theory* 23:219–279.
- Carstens, Vicki. 2011. Hyperactivity and hyperagreement in Bantu. *Lingua* 121:721–741.
- Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.
- Chomsky, Noam. 1993. A minimalist program for linguistic theory. In Ken Hale & Samuel Jay Keyser (eds.), *The View from Building 20*, 1–52. Cambridge, MA: MIT Press.
- Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2000. Minimalist inquiries. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by Step: Essay on Minimalism in Honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A Life in Language*, 1–52. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos Otero & Maria-Luisa Zubizarreta (eds.), *Foundational Issues in Linguistic Theory*, 133–166. Cambridge, MA: MIT Press.

<sup>10</sup> Contiguity between T and the subject is not at risk, since they are immediately adjacent to one another. This is parallel to what we observed for Japanese in section 5.1.

- Christophe Anne, Maria Teresa Guasti, Marina Nespors, Emmanuel Dupoux & Brit van Ooyen. 1997. Reflections on phonological bootstrapping: Its role in lexical and syntactic acquisition. *Language and Cognitive Processes* 12:585–612.
- Christophe Anne, Marina Nespors, Maria Teresa Guasti & Brit van Ooyen. 2003. Prosodic structure and syntactic acquisition: The case of the head-direction parameter. *Developmental Science* 6:211–220.
- Deal, Amy Rose. 2009. The origin and content of expletives: evidence from “selection”. *Syntax* 12:285–323.
- Embick, David & Rolf Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 32:555–595.
- Halle, Morris & Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. In Kenneth Hale & Samuel Jay Keyser (eds.), *The View from Building 20*, 111–176. Cambridge, MA: MIT Press.
- Hankamer, Jorge & Line Mikkelsen. 2018. Structure, architecture, and blocking. *Linguistic Inquiry* 49:61–84.
- Henderson, Brent. 2003. PF evidence for phases and Distributed Morphology. In Keir Moulton & Matthew Wolf (eds.), *Proceedings of the 34th annual meeting of the North East Linguistic Society*, Amherst, GLSA, 255–265.
- Henderson, Brent. 2013. Agreement and person in anti-agreement. *Natural Language and Linguistic Theory* 37:461–522.
- Huang, James Cheng-Teh. 1982. Logical relations in Chinese and the theory of grammar. Doctoral dissertation, MIT.
- Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.
- Lee, Soo-Hwan & Doo-Won Lee. 2019. Nominal mismatches in Swahili locatives. *Proceedings of the 93rd annual meeting of the Linguistic Society of America* 4:1–11.
- Nespors Marina, Mohinish Shukla, Ruben van de Vijver, Cinzia Avesani, Hanna Schraudolf & Caterina Donati. 2008. Different phrasal prominence relations in VO and OV languages. *Lingue e Linguaggio* 7:1–29.
- Noyer, Rolf. 1992. Features, positions, and affixes in autonomous morphological structure. Doctoral dissertation, MIT.
- Pollock, Jean-Yves. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20:365–424.
- Richards, Norvin. 2010. *Uttering trees*. Cambridge, MA: MIT Press.
- Richards, Norvin. 2016. *Contiguity theory*. Cambridge, MA: MIT Press.
- Richards, Norvin. 2019. Detecting Contiguity-Prominence. Ms., MIT.
- Selkirk, Elisabeth. 2009. On clause and intonational phrase in Japanese: The syntactic grounding of prosodic constituent structure. *Gengo Kenkyu* 136:35–73.
- Selkirk, Elisabeth. 2011. The syntax-phonology interface. In John Goldsmith, Jason Riggle & Alan C. L. Yu (eds.), *The handbook of phonological theory*, 2nd edition, 435–484. Oxford: Blackwell.