A modular theory of the relation between syntactic and phonological constituency[1]

Seunghun J. Lee (International Christian University, University of Venda)

Elisabeth Selkirk (University of Massachusetts, Amherst)

Abstract

Patterns of prosodic-structure-sensitive high tone spread in the Bantu language Xitsonga reveal mismatches between syntactic and phonological/prosodic constituency. A modular account of these mismatches is proposed: Match constraints (Selkirk 2011) are re-construed as spell-out constraints relating the output representation of the morphosyntax to the input representation for the phonology. It's argued that in Xitsonga the spell-out constraint $\text{MatchPhrase}_{\text{LEX}}$ relates only phrases with lexical category heads to phonological phrase in the phonological input representation. In the phonology *per se*, a novel class of prosodic structure faithfulness constraints interacts with prosodic structure markedness constraints to produce further constituency mismatches in the output phonological representation.

Keywords

syntax-phonology mismatches, phonological phrase structure, Match constraints, spell-out vs. phonology, Xitsonga, tone spread, tone/edge constraints, prosodic faithfulness constraints, prosodic markedness constraints

[1] This paper was mainly supported by a National Science Foundation grant (#1147083) "Effects of Syntactic Constituency on Phonology and Phonetics of Tone (ECPPT)". Partial support comes from two JSPS KAKENHI Grants 16K02641, 20K00578, an IRC grant IRC202010, as well as the NINJAL collaborative research project 'Cross-linguistic Studies of Japanese Prosody and Grammar'. We thank the two anonymous reviewers and editors of this volume. This work has also benefitted from discussions with the audience members at the ICPP 2020 conference. Of course, remaining errors are of our own.

1. Introduction

1.1 The MSO-PI-PO model

It has been widely assumed that a phonological/prosodic constituent structure at word-level and above forms part of the output representation of the phonological component. In that phonological output representation, call it PO, prosodic constituent structure plays a key role in accounting for the structure-sensitive distribution in a sentence of segmental and tonal phenomena, as well as stress prominence. It's also been recognized that the syntactic constituency of the output representation of the morphosyntax (call it MSO) has a systematic impact on the prosodic constituency of PO at word level and above, though the constituent structures of MSO and PO often show mismatches.[2]

A direct relation between the morphosyntactic constituency of MSO and the prosodic constituency of the output phonological representation PO is assumed in most previous accounts of the relation between syntactic and prosodic constituency, as it is in the more recent Match theory of this interface (Selkirk 2009, 2011).[3] In this paper, though, we argue there are advantages to assuming that the relation between the morphosyntactic constituency of MSO and the phonological constituency of PO is only indirect. We assume instead that the input representation of the phonology (PI) mediates between MSO and PO, as in (1):

[2] For useful reviews of the role of prosodic structure in theories of the syntax-phonology interface, see Elfner (2018), Bennett and Elfner (2019).

[3] See also Bennett et al. (2016), Ito and Mester (2020) and papers in *Phonology*, volume 32, issue 1, a 2015 thematic issue on constituency in sentence phonology edited by Seunghun J. Lee and Elisabeth Selkirk.

(1) MSO (morphosyntactic output representation)

Spell-out |

PI (phonological input representation)

Phonology |

PO (phonological output representation)

The two levels of phonological representation in the model in (1)—a phonological input representation PI and a phonological output representation PO—have played a role in theories of rule-based generative phonology since Chomsky and Halle (1968), and are retained in the optimality-theoretic model of constraint-based phonology (Prince and Smolensky 1993/2004, McCarthy 2001).

Just what, then, is the relation between the phonological input PI and the output MSO of the morphosyntactic derivation? We will assume, consistent with works in the distributed morphology tradition,[4] that the terminal elements, or formatives, of morphosyntactic constituent structure consist of abstract roots and morphosyntactic feature complexes that lack any phonological specification throughout the morphosyntactic derivation. This is consistent with the fact that the phonological properties of these terminal elements/formatives play no role in semantic interpretation or in the morphosyntactic derivation. But these terminal elements of MSO must be paired with an appropriate input representation for the phonology.

The phonological input representation PI is determined with respect to the MSO representation of the grammar in at least two ways. First the abstract terminal elements of MSO must be given phonological expression—i.e., be spelled out—in PI with their distinctive, 'lexical', phonological properties. We can borrow the terms 'vocabulary insertion' or

[4] Halle and Marantz (1993, 1994), Harley and Noyer (1999), Embick and Noyer (2007), Bobalijk (2017).

'exponence' from distributed morphology to refer to this pairing of abstract morphosyntactic formatives with their input phonological form in PI. A second aspect of the spell-out of MSO involves a linearization of the spelled-out formatives and higher constituents, a linearization which would be inherited in the input phonological representation PI. This linearization is determined by grammatical principles that take into account the hierarchical structure of the morphosyntactic output representation. In other words, two essential properties of the phonological input representation—the lexical phonological properties of the formatives of the sentence and a morphosyntax-based linearization of them—indicate that the morphosyntactic output representation interfaces with the input to the phonology, as in (1).

In the context of a model like (1), it is natural to consider the possibility that a third core property of phonological representation—morphosyntax-based phonological/prosodic constituent structure—is also present in the *input* phonological representation PI. This possibility was not entertained in earlier works on the relation between morphosyntactic and prosodic constituent structure in the grammar.[5] Indeed, the interface constraints of Match Theory (in the Selkirk 2011 version) call for the word- phrase-, and clause-level constituents of MSO to correspond to matching prosodic words (ω), phonological phrases (φ), and intonational phrases (ι) structure in PO. In that account, the appearance of constituency mismatches, such as in the Bantu language Xitsonga, where one-word phrases fail to show the phonological properties of a φ, were hypothesized to be the consequence of a language-particular optimality-theoretic constraint ranking, in this case BINARITY(φ) >> MatchPhrase.[6]

[5] See, e.g., the references in footnotes 1 and 2.

[6] As is customary in optimality-theoretic phonology, small capitals are used for the names of phonological constraints, e.g., BINARITY(φ). The names of spell-out constraints will be spelled with word-initial capitals and the absence of space(s) in multi-word names, e.g., MatchPhrase.

In this ranking, an interface constraint relating phrases of MSO to φ in PO is subordinated to a properly phonological markedness constraint on the prosodic φ structure of PO.

Some rethinking of the notion that the syntax-phonology constituency interface directly relates MSO and the output constituent structure of PO has recently emerged, however. The proposal that prosodic constituent structure at word-level and above is instead present in the *input* phonological representation, where it is the spell-out in PI of the morphosyntactic phrase structure of MSO, plays an important role in the Kratzer and Selkirk (2020) analysis of the sentence prosody of sentences of Standard American and British English. It will be reviewed briefly in section 1.2.[7] In what follows, the central goal is to arrive at better insight into the formation of prosodic constituent structure, such as that affecting lexical high tone spread in Xitsonga. We argue this can be done by assuming that the MSO-PI interface is the sole locus of the syntax-phonology constituency interface in the grammar, shown in (1). On this alternative view, MatchPhrase can be understood to be a constraint on the spell-out of MSO as PI.

As a first introduction to the workings of the modular, sequential, theory of prosodic structure formation that is implied by the model in (1), we briefly examine the MSO-PI-PO derivation of the output prosodic contrast between Given and new constituents in Standard American and British English, proposed by Kratzer and Selkirk (2020). We then turn to an MSO-PI-PO account of the formation of the prosodic φ structure that guides lexical high tone spread in Xitsonga. It is an account that expands on and revises the Selkirk (2011) treatment. Collectively, Xitsonga and Standard British and American English provide evidence for distinct types of mismatch between the constituent structure of MSO and PO which point to distinct sources of constituency mismatch—in the spell-out module and in the phonology *per se.*

[7] An early version of this proposal is presented in Selkirk (2017). See also Elordieta and Selkirk (2018).

1.2 The MSO-PI-PO model and prosodic structure formation in English

Kratzer and Selkirk (2020) propose a new theory of the impact of information structure on the prosodic constituent structure organization, phrasal stress prominence, and default (epenthetic) tone that appears in a broad variety of sentence types in Standard American and British English, namely all-new sentences, sentences containing [G]-marked, discourse-given, constituents, and sentences containing the [FoC]-marking of alternatives-based focused constituents. The attested phonological contrasts between these morphosyntactically distinct types of sentences must have their source in the *input* phonological representation (PI), given the unique locus of the morphosyntax-phonology interface in the MSO-PI-PO model in (1).

Kratzer and Selkirk propose that a discourse-given, [G]-marked, phrase of MSO is spelled out without any prosodic phrase (φ) structure in the input phonological representation PI, as shown in Figure 2. In that way a [G]-marked phrase contrasts with the presence of the default, matching, φ in PI which spells out any (unmarked) discourse-new phrase of MSO, as shown in Figure 1.[8] The two instances of the example sentence *Sarah mailed the caramels* analyzed here are morphosyntactically identical, except that the direct object *caramels* is [G]-marked in the second case, shown in Figure 2.

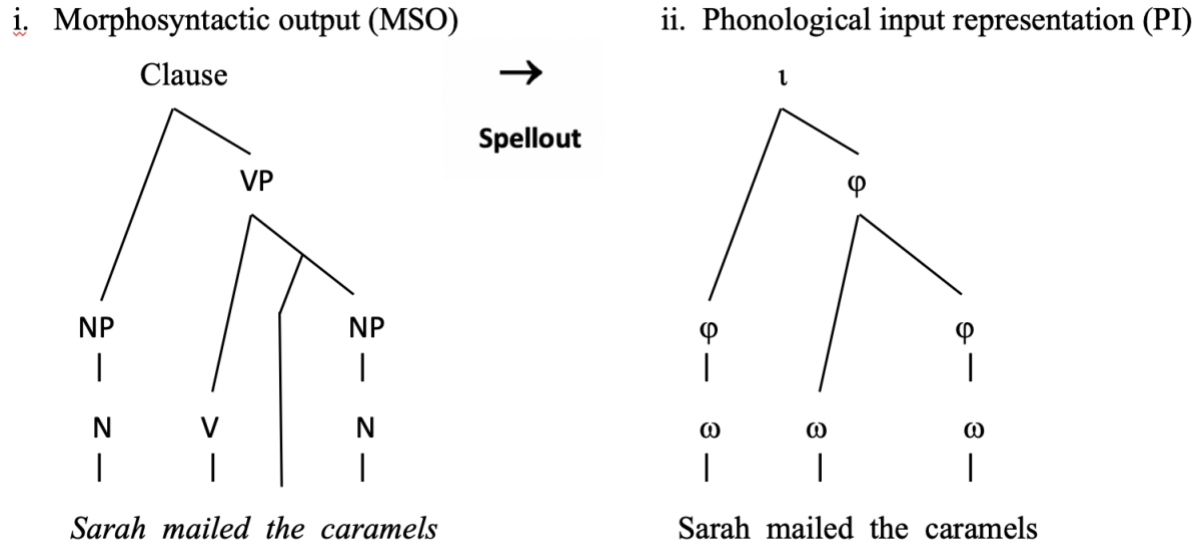
First some basics. Note that the terminal string of MSO, (i) in Figures 1 and 2, is written with italics. This italic orthography should be understood as indicating the non-phonological status of the abstract morphemes that make up the terminal string of the MSO. The phonological input

[8] A phrase of the morphosyntax which lacks any [G] would simply be interpreted as discourse-new (non-given).

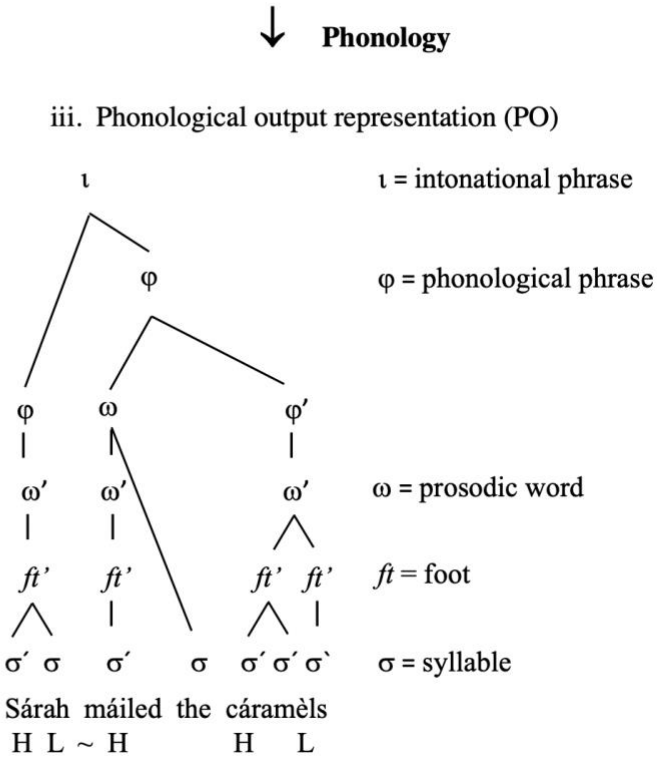
representation PI, (ii), differs from the MSO in containing a phonological spell-out of these morphemes (written in standard orthography here) as well as a phonological spell-out of the morphosyntactic phrase structure that organizes these morphemes into a hierarchical phonological structure made up of the prosodic constituents of types ω, φ and ι. These ω, φ and ι constituents spell out (give phonological expression to) the respective morphosyntactic word-level, phrase-level and clause-level constituents of MSO.[9] The final stage in the MSO-PI-PO derivation of any sentence is the output phonological representation PO seen in (iii).

[Note to editor: ideally Figure 1 would be placed here,]

Figure 1. *Sarah mailed the caramels* (Kratzer and Selkirk 2020, sec. 6)



[9] Note that in the case of an all-new sentence like that in Figure 1, all the MSO phrases that are headed by lexical category items (N, V, A) have the status of φ in the (intermediary) phonological input representation (PI), not unlike what we see in Xitsonga (cf. discussion in section 3).



The spelling out of morphosyntactically unmarked discourse-new phrases as phonological phrases (φ) in PI would be the consequence of the syntax-phonology constituency interface constraint MatchPhrase which is proposed in Selkirk (2011). In the MSO-PI-PO model in (1), however, a constituency interface constraint like MatchPhrase would hold only of the relation between MSO and PI. The more recent statement of MatchPhrase given in Kratzer and Selkirk (2020) is (2):

(2) MatchPhrase

For every instance of a Phrase in MSO there is exactly one instance of a phonological phrase (φ) that spells it out phonologically in PI.

The phonology *per se* is in general responsible for whether or not the phonological properties of PO are inherited intact from PI, or instead undergo phonological constraint-induced

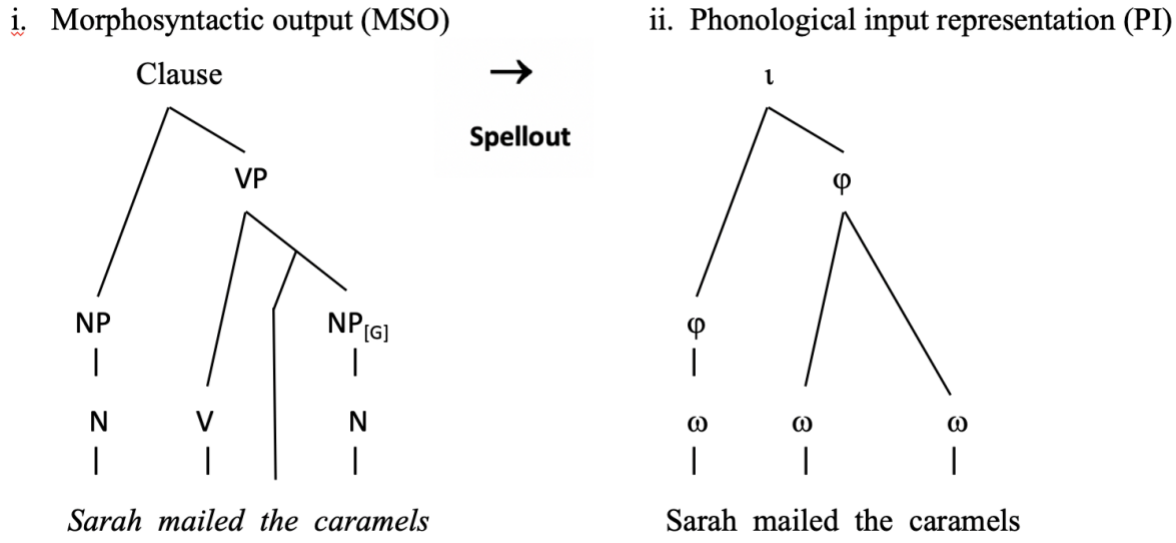
changes which produce differences between PI and PO. We assume here an optimality-theoretic, constraint-based, theory of the phonology, in which a language-particular ranking of phonological markedness and faithfulness constraints selects the optimal candidate for PO from the set of possible candidates that are consistent with the input PI (Prince and Smolensky 1993/2004, McCarthy 2001). In the case of the all-new sentence in Figure 1, differences between PI and PO involve the introduction in PO of phonologically predictable properties that were totally absent from the input phonological representation PI—the presence of syllable and foot constituents, the presence of prosodic head prominence ('main stress') within the foot (*ft)*, the prosodic word (ω) and the phonological phrase (φ) constituents, as well as the obligatory presence of an H accent tone associated to the head-prominent syllable of a φ and the presence of a L edge tone at the right edge of a φ. In the PI representations of all-new, pragmatically neutral sentences of Standard American and British English head prominence and tone are absent. In this variety of English, the presence and distribution of head prominence and tones in PO is predictable, entirely determined by phonological markedness constraints of the phonology *per se.* For example, the H accent tones are plausibly epenthesized in order to satisfy a high ranked markedness constraint on PO that calls for a syllable bearing φ-level prominence/stress to bear tone, while φ-level head prominence itself is predictably determined with respect to the output prosodic φ structure. (For details and discussion of this analysis, see Kratzer and Selkirk 2020, sec. 6)

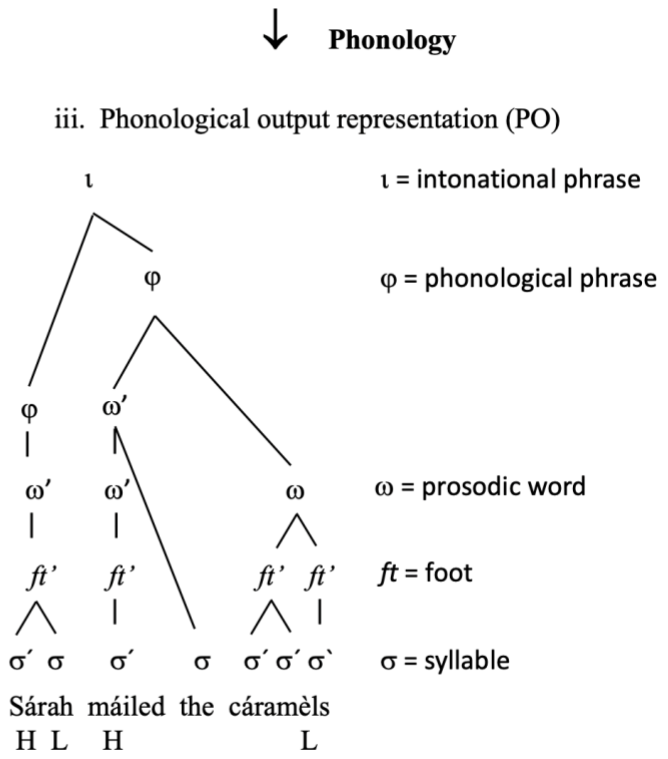
Consider next the MSO-PI-PO derivation of a sentence with the same MSO constituent structure but where the object phrase is [G]-marked (interpreted as discourse-given by the semantics/pragmatics), as shown in Figure 2. The spelling out of a [G]-marked morphosyntactic phrase as the *absence* of any corresponding phonological phrase in PI can be understood as an instance of prosodic morphology, given that it pairs the morphosyntactic feature [G]—a

formative—with a prosodic constituent structure configuration (see McCarthy and Prince 1996, inter alia).

[Note to editor: ideally Figure 2 would be placed here.]

Figure 2: *Sarah mailed the caramels*$_{[G]}$ (Kratzer and Selkirk 2020, sec. 7)





Kratzer and Selkirk propose that, in the grammar of this variety of English, it is a feature-specific MSO-PI spell-out constraint—call it DephraseGiven—which requires a [G]-marked phrase of MSO to be realized with an absence of φ status in PI.

(3) [G]=No-φ (DephraseGiven) (Kratzer and Selkirk 2020)

A [G]-marked constituent in MSO corresponds to a prosodic constituent in PI which is not a φ and contains no φ.

A language particular constraint-ranking DephraseGiven >> MatchPhrase within the spell-out module could ensure that there is no φ in PI that corresponds to a [G]-marked phrase of MSO.

As for the PO of Fig 2-(iii), the terminal string in PO that corresponds to the [G]-marked object NP in MSO retains the φ-less status that it has in PI. In the phonology *per se*, the absence

in PI of φ structure for [G] would be maintained in the output PO by the high rank of an anti-insertion faithfulness constraint with respect to any phonological markedness constraint which might call for the presence of φ. (The assumption that prosodic phrase structure is present in both input and output representations requires an extension to prosodic constituency of the McCarthy and Prince (1995, 1999) theory of faithfulness, or correspondence, between distinct linguistic representations like PI and PO.)

What about the position of phrasal head prominence (stress) on the verb *mailed* and the associated presence of the H accent tone within the φ corresponding to the VP as a whole in PO (iii) in Figure 2? Kratzer and Selkirk point out that a simple phonological account is possible, rather than an account based on the givenness-driven metrical stress shift originally proposed by Ladd (1980) and others later. Cross-linguistic phonological research has shown that in same-level-sister configurations within prosodic words and feet, head prominence ('main stress') within the constituent can just as well fall on the leftmost or the rightmost sister. So in Standard American and British English, it should come as no surprise that in the phonology *per se*, head prominence would come to fall on the leftmost ω in the same-level-sister configuration (ω´ ω)φ in PO that corresponds to the VP in Figure 2.[10]

Summing up, we have seen that the two modules of the MSO-PI-PO model of the relation between the output morphosyntactic representation of a sentence and its phonological output representation—spell-out and phonology—together allow for simple, principled, solutions to classic issues relating to the prosody of all-new sentences vs. sentences with discourse-given

[10] Leftmost-ω also happens to be the locus of head prominence within noun compounds consisting of sister words -- (ω´ω)ω-- like *bike rack*, *back pack*, etc.

material in Standard American and British English. Constraints of the interface spell-out module relate morphosyntactic properties of the output morphosyntactic representation to phonological properties of the input representation for the phonology *per se*. As for the constraints which have direct impact on the output phonological representation PO, these are purely phonological markedness or faithfulness constraints of the phonology *per se*, relating just the PI and PO levels of phonological representation (Kratzer and Selkirk 2020, sec. 6-7).

The notions that a spell-out module defines the interface between the MSO and PI of a sentence and that the phonology *per se* defines the relation between PI and PO are by no means new here. What is new is the notion that morphosyntactically-determined prosodic constituent structure at word-level and above appears in PI, the input representation for the phonology. The upcoming sections of this paper provide additional support for an MSO-PI-PO model of the relation between the morphosyntactic constituency of MSO and the prosodic constituency of PO. They focus on the empirical and theoretical benefits of a two-stage derivation of the prosodic φ constituency in PO that is found in the Bantu language Xitsonga. Mismatches between the phrasal organization of PO on the one hand and the phrasal organization of MSO on the other will be argued to have their source in either one or both of the two modules of the MSO-PI-PO model—the system of spell-out constraints determining φ structure in PI, and the phonological constraint system that determines φ structure in PO.

1.3 The MSO-PI-PO model and prosodic structure formation in Xitsonga: a preview

Tone-related evidence of the distinction between the PI and PO representations of the sentence in Xitsonga is shown in examples (4) and (5). In the PI representations (4a) and (5a), lexical H tone

in the verb *lava* ‘want’ is located on the verbal root morpheme and is indicated by an orthographic acute accent. This acute accent is also used to indicate the extent of H tone spread in the output PO representations. The underlining of an individual vowel in both PI and PO is a convenient informal device used to show the location of the lexical tone in PI that lies behind the change in tonal representation observed in PO.[11] It is the mere existence of phenomena like the spread of lexical H tone from its lexical position of origin to an adjacent sequence of syllables which lack lexical tone that motivates a distinction between input and output phonological representation in generative phonologies.

(4) a. PI: ni-l<u>á</u>v-a nguluve
1sg-want-FV pig
b. PO: ni-l<u>á</u>v-á ngúlú:ve ‘I want a pig’

(5) a. PI: hi-l<u>á</u>v-a hlambeto yi-ntsó<u>ngó</u>
1pl-want-FV cooking.pot cl9-small
b. PO: hi-l<u>á</u>v-á hlambeto yi-↓ntsó:ng<u>ó</u> ‘We want a small cooking pot’

As we see in (4) and (5), there are contrasts in the extent of H tone spread observed in the PO representations. In Xitsonga, the phenomenon of rightward lexical H tone spread provides essential evidence for the presence and distribution of prosodic φ structure in PO. Section 2 of this paper will show in detail that in a broad array of sentential contexts, between-word H tone spreads rightward only into single-word phrases; it is blocked at the left edge of binary phrases. This difference in H tone spread in PO was analyzed by Selkirk (2011) as the effect of two distinct

[11] The colon ‘:’ indicates the penultimate lengthening found in PO in the penult syllable of an intonational phrase. ‘↓’ is a symbol indicating the presence of downstep in the phonetic interpretation of a H tone span that is preceded by another H tone span in the PO of the sentence.

types of phonological markedness constraint on PO representation. One of them, BINARITY(φ), requires a φ in PO to consist of two ω[12]. The other, CRISPEDGE-L(φ;H), blocks instances of H tone spans through the left edge of a φ of PO.[13]

Figures 3 and 4 illustrate the proposed MSO-PI-PO derivation in Xitsonga of the φ structure of PO that governs the distribution of H tone spread in (4) and (5). Consider first the derivation in Figure 3 of the output prosodic structure of the sentence *Nilava nguluve* 'I want a pig' in (4).

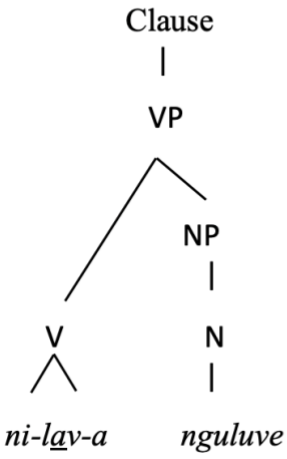
Place Figure 3 here :

Figure 3: *Nilava nguluve.* 'I want a pig.'

[12] There are well-known prosodic binarity requirements at the level of the foot (see, e.g., Hayes (1995)), the prosodic word (ω) (see, e.g., Ito and Mester (1992/2003)), and the phonological phrase (φ) (see e.g., Ghini (1993)). So we can with confidence entertain the hypothesis that in Xitsonga a single-word (non-binary) phrase of MSO does not count as a φ in PO, as shown in (4), and that, by contrast, a two-word phrase of MSO must count as a φ in PO, as shown in (5).

[13] CRISPEDGE-L(φ;H) and other φ-structure-sensitive constraints on H tone spread that are operative in Xitsonga are taken up in Section 2 below.

i. Morphosyntactic output (MSO)



→ **Spellout**

ii. Phonological input representation (PI)

↓ **Phonology**

iii. Phonological output representation (PO)

With the loss in PO of the non-binary φ_2 of PI that spells out the MSO direct object NP as a φ, there is no obstacle to the spreading of lexical H tone from its PI location in the root of the verb into the object noun.[14]

The agent of the constituency mismatch in Figure 3 between the PI and PO of the φ-structure corresponding to the single-word object phrase in MSO is the phonological constraint BINARITY(φ), which holds of the output phonological representation (PO):

(6) BINARITY(φ)

A phonological phrase (φ) must be binary.

[A prosodic constituent is binary if it immediately dominates exactly two prosodic constituents.]

In the MSO-PI-PO model assumed in this paper, a prosodic markedness constraint on PO like BINARITY(φ) may be responsible for the absence ('deletion') of a φ in PO only if it outranks any input-output faithfulness constraint that would call for the retention in PO of a φ that is present in PI. The faithfulness constraint MAX(φ) in (7) would play that role (Elordieta and Selkirk 2018, Kratzer and Selkirk 2020). Such a faithfulness constraint on the correspondence between PI and PO is not proposed in McCarthy and Prince (1995, 1999), but falls into the broad class of anti-deletion MAX constraints proposed there.

[14] As for the failure of H tone to spread into the final syllable of the φ, this will be discussed in Section 2, along with other aspects of structure-sensitive restrictions on H tone spread.

(7) MAX(φ) [= 'No deletion of φ'] (Revised formulation[15])

Any constituent (node) φ_n of PI must also be present in PO.

The optimality-theoretic tableau in (8) illustrates the consequences of the constraint ranking BINARITY(φ) >> MAX(φ) in the case of the single-word object sentence in (4). φ2 in PI corresponds to the single-noun object NP. For the sake of readability, the labelled bracketings that appear in such tableaux will include not ω-brackets surrounding prosodic words.

(8) Verb + unary object: $[_{VP}\ [ni\ [\underline{la}va]_V]_V\ [_{NP}\ [nguluve]_N\ ]_{NP}]_{VP}$ (= (4))

PI	(φ1 ni lá̲va (φ2 nguluve)φ2)φ1	BINARITY (φ)	MAX (φ)
PO	a. (φ1 nilává (φ2 nguluve)φ2)φ1	*φ2!	
	☞ b. (φ1 nilává ngúlúve)φ1		*φ2

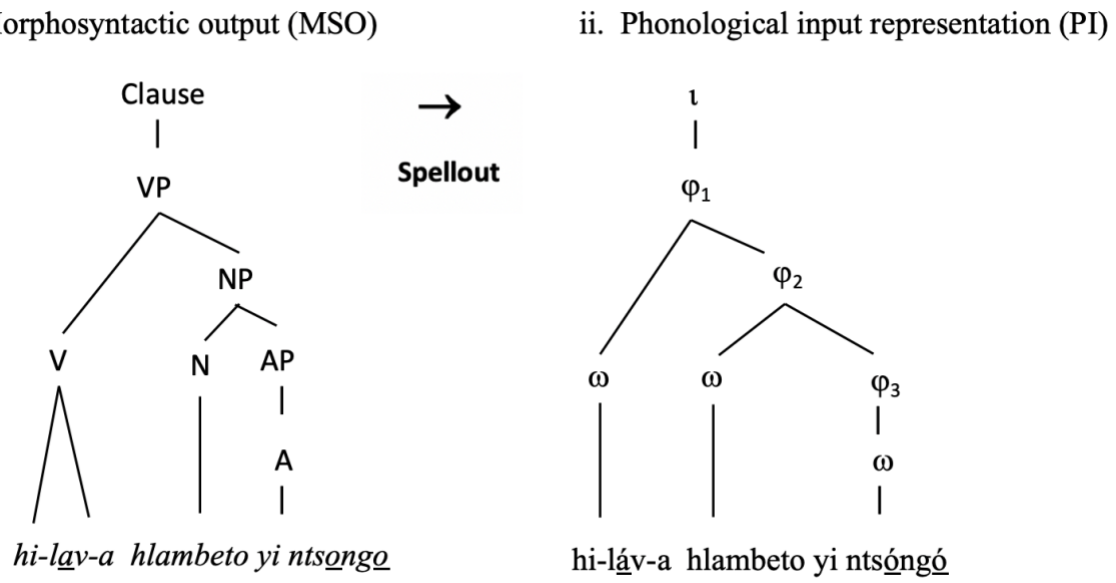
The nonoptimal candidate (8a) contains a nonbinary φ2 in PO and therefore incurs a violation of BINARITY(φ). The PO candidate (8b) has lost that φ2 (in violation of lower-ranked MAX(φ)) and as a consequence satisfies BINARITY(φ). Given the higher ranking of BINARITY(φ) in the grammar of Xitsonga, (8b) is the optimal candidate. The MSO-PI-PO derivation depicted in Figure 3 and the constraint tableau in (8) provide a first illustration of the role for the constraint-

[15] Thanks go to a reviewer of this paper who pointed out that MAX(φ) must refer to the disappearance of a prosodic node φ, and not its segmental content, as was proposed in the version of MAX(φ) exploited in Kratzer and Selkirk (2020), for example.

ranking of the phonology *per se* of Xitsonga in creating constituency mismatches between the prosodic structure of PO and morphosyntactic phrase structure of MSO.

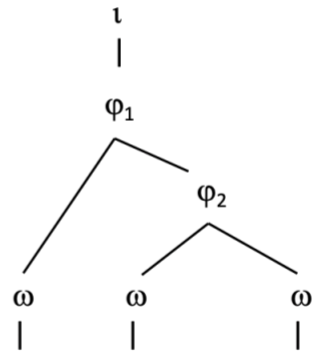
Consider next the derivation in Figure 4 of the output prosodic φ structure and H tone spread in the Xitsonga sentence *Hilava nguluve hlambeto yintsongo* 'We want a small cooking pot' in (5), where the direct object NP consists of a head noun and an adjectival modifier phrase. MatchPhrase spells out the nested phrase structure of the direct object in MSO with the matching nested φ structure in PI, as seen in Fig 4-(ii).

Figure 4: *Hilava nguluve hlambeto yintsongo.* 'We want a small cooking pot'



↓ **Phonology**

iii. Phonological output representation (PO)



hi-láv-á hlambeto yi ntsóngó

'We want a small cooking pot'

In the phonology, however, the constituent φ_2 corresponding to the object phrase is binary, so it is maintained in PO. It prevents the rightward spread of H tone from the verb into that object φ. Note also that φ_3, which corresponds to the single-word modifier phrase within the object phrase, is nonbinary; its presence in PO (9a) is ruled out by BINARITY(φ). The optimal candidate (9b) lacks the φ of the modifier phrase, and so violates MAX(φ).

(9) Verb + binary object: $[_{VP}\ [hi\ [lava]_V]_V\ [_{NP}\ [hlambeto]_N\ [_{AP}\ [yi\ ntsongo]_A]_{AP}]_{NP}]_{VP}$ (= (8ii))

PI	(φ1 hi lág v a (φ2 hlambeto (φ3 yi ntsó:ngó)φ3)φ2)φ1	BINARITY (φ)	MAX (φ)
PO	a. (φ1 hilává (φ2 hlambeto (φ3 yintsó:ngó)φ3)φ2)φ1	*φ3!	
	☞ b. (φ1 hilává (φ2 hlambeto yintsó:ngó)φ2)φ1		*φ3
	c. (φ1 (φ4 hilává hlámbéto)φ4 yintsó:ngó)φ1		*φ2,*φ3

Tableaux (8) and (9) show that by assuming that the S-P interface constraint MatchPhrase spells out a phrase of MSO as a φ in the input PI representation of the phonology *per se*, we can account for phonological-constraint-driven mismatches between MSO and PO constituency within a restrictive model of grammar in which the syntax-phonology interface in all its aspects involves just relations between MSO and PI. We need only exploit a natural extension of the theory of input-output faithfulness that would allow reference to prosodic structure constituency in PI (as with MAX(φ)) in order to embrace a fully phonological-constraint-driven account of the MSO-mismatching absence in PO of a prosodic constituent φ whose presence violates the prosodic markedness constraint BINARITY(φ).

In what follows in section 2 we will see that the creation of MSO-PO constituency mismatches in the phonology *per se* as a result of the ranking BINARITY(φ) >> MAX(φ) is systematic throughout the sentence types of Xitsonga. In section 3, on the other hand, we consider data that argue that the spell-out module is also responsible for producing systematic MSP-PO constituency mismatches in Xitsonga. We will see that only MSO phrases that are headed by the lexical category items N, A, V may have a reflex as φ in phonological representation. In the MSO-PI-PO model these facts require a version of MatchPhrase that is restricted in terms of the morphosyntactic features of the phrase (which are projected from its head). Given the [N]/[A]/[V]-sensitive spell-out of phrases in Xitsonga and the [G]-marking-specific 'dephrasing' in PI that is produced by spell-out in Standard American and British English it is worth entertaining the general hypothesis that the spell-out module itself is a possible source of mismatch between the constituency of MSO and the constituency of PO. It is in spell-out that a relation between the morphosyntactic features of MSO and the phonological properties of PO is established.

2. A general account of φ-sensitive H tone spread in Xitsonga in the phonology per se

2.1 The phonology of H tone spread

Supplied with the analysis above of the distribution of φ-structure in the PO representations of Xitsonga sentences as governed by the phonological constraint ranking BINARITY(φ) >> MAX(φ), we can turn to the analysis of rightward H tone spread in the language and its dependence on the φ structure of the sentence in PO. This section reviews core factual generalizations about the distribution of H tone spread in Xitsonga, which are drawn for the most part from the extremely valuable account in Kisseberth (1994), to a lesser degree from Cole-Beuchat (1959), Beuchat (1962) and Herbert (1992), as well as from results of fieldwork on Xitsonga that we have carried out ourselves[16].

The analysis presented below is based on the Selkirk (2011) optimality-theoretic constraint-based analysis of structure-sensitivity in the distribution of H tone spread in Xitsonga sentences, which draws heavily on the insights of Kisseberth (1994). Our subsequent fieldwork on Xitsonga largely corroborates the 2011 analysis. Section 2.2 includes data that had not been previously been available on between-word H tone spread in a variety of syntactic contexts, and

[16] Fieldwork data reported in this paper was collected between 2013 and 2015 in Limpopo, South Africa. Our main consultant is a female speaker in her 30's from Mhinga, South Africa, which is a municipality located just outside of the Kruger National Park that separates South Africa from Mozambique. The Mhinga variety displays tonal patterns that are comparable to those observed in the Mozambican variety of Xitsonga investigated in Kisseberth (1994); hence we believe that data from our fieldwork likely reflects the Xitsonga system reported in earlier studies. All data obtained from these fieldwork sessions are archived in the Xitsonga page on the archive hosted by the Research Institute for Languages and Cultures of Asia and Africa that is affiliated with the Tokyo University of Foreign Studies (https://ecppt.aa-ken.jp/xitsonga.html). Earlier results of the fieldwork were presented in Lee and Selkirk (2014).

shows the basic soundness of our multi-faceted account: an account of φ-formation in the phonology along with an account of the sensitivity of H tone spread to the φ-structure of PO.

The rightward spreading of H tone in Xitsonga is arguably forced by a H-tone-spread-inducing phonological markedness constraint, dubbed here HTS-RT, which calls for unrestricted rightward H tone spread. The formulation in (10) makes no reference to φ-structure.

(10) HTS- RT: "A H tone spreads [expands its span] to the right."

Assign a violation to any toneless tone-bearing unit (tbu) that is preceded by the right edge of a H tone span.

* tbu$^{H\backslash}$ … tbu

This is a formulation of free high tone spread in the spirit of Cassimjee and Kisseberth (1998). The term 'tone span' denotes the (multiple) association(s) of a single high tone to a (sequence of) tone-bearing unit(s).[17] The supercripting of $^{H\backslash}$ to the right of a tbu, as shown in (11) indicates the right edge of a H tone span on that preceding tbu. In Xitsonga, the left limit of a H tone span, written $/^{H}$, coincides with a lexical H tone. When expositional clarity is sought, we will use the tone span notation given here, in addition to the common use of orthographic acute accents to indicate the association of a H tone to a tbu:

(11)

	PI	PO
Acute accent 'spelling' of H tone:	<u>vá</u> lulamisa	<u>vá</u>lúlámísa
Tone span representation of H tone:	<u>va</u>$^{/H\backslash}$ lulamisa	<u>va</u>$^{/H}$lulami$^{H\backslash}$sa

[17] Cassimjee and Kisseberth (1998) use the term 'tonal domain' rather than the more theoretically neutral term 'tone span' that we exploit here.

Of course any difference between the span of a lexical H tone in PI and the broadening of the H tone span to additional tone-bearing units in PO would involve a violation of some input-output faithfulness constraint(s) (McCarthy and Prince 1995, Myers 1997). In the case of spreading, a toneless tbu of PI would become associated with H tone in PO, and a H tone in PI would become associated with a new tbu in PO. Following Myers (1997), we will assume that the input-output faithfulness constraint DEP-IO (A) rules against any tone-tbu association in PO which is not present in PI:

(12) DEP-IO (A): An association in the output must have a correspondent in the input.

In the constraint ranking of Xitsonga phonology, it must be that the markedness constraint HTS-RT outranks the faithfulness constraint DEP-IO (A): HTS-RT >> DEP-IO (A).

Given this constraint ranking allowing for free rightward H tone spread, any restrictions on H tone spread are hypothesized to be the consequence of other, higher-ranked, phonological markedness constraints. In Xitsonga, rightward H tone spread occurs systematically, *except* where it is blocked by prosodic structure-sensitive phonological markedness constraints which regulate the relation between tone and properties of prosodic structure. These markedness constraints are the familiar constraints NONFINALITY (Kisseberth 1994, Cassimjee and Kisseberth 1998, Yip 2002, Hyde 2011), CRISPEDGELEFT (Ito and Mester 1999, Selkirk 2011) and the OCP (Leben 1973, Yip 1988, Myers 1997 and many others). Our hypothesis is that the influence of the prosodic structure of the sentence on its tonal representation is embodied in structure-sensitive markedness constraints like these and their ranking within the constraint system of the phonology.

For example, the failure of H tone to spread rightward all the way to the right edge of the phrase in Xitsonga seen in Figure 3 is analyzed by Kisseberth (1994) as a consequence of a

NONFINALITY constraint specific to the right edge of a phrase (see also Cassimjee and Kisseberth 1998). We call the constraint NONFINALITY(φ,H) and formulate it as in (13):

(13) NONFINALITY(φ,H)

The right edge of a H tone span may not coincide with the final syllable of a phonological phrase (φ).

* …..^{H}\ )φ

The ranking NONFINALITY(φ,H) >> HTS-RT prohibits the right edge of a H tone span from coinciding with the final syllable of a phrase, but it would allow for H to spread onto the final syllable of a word that is not the last one in a φ, as in the pre-object verb in Figures 3 and 4.

We must consider a possible alternative constraint-based account of long-distance spread of H tone to the penultimate syllable of a φ, namely a markedness constraint that explicitly calls for the penultimate syllable of a φ to be associated with a H tone, due perhaps to the prosodic prominence of that syllable (Archangeli and Pulleyblank 1994, Volk 2011, Hyman 2009). With such a constraint the lexical H tone would in effect be 'pulled' to the penult syllable of φ. While this account of the penult-positioning effect within a φ is a possible one, it can't account for all instances of H tone spread observed in the Xitsonga. These are cases where the attested H tone span does not in fact reach as far as the phrasal penult. Instead, the generalization in Xitsonga is that H tone spreads unboundedly, *as far right as it can*, up to the point where it is blocked by yet further constraints.

An additional H-spread-limiting constraint is at work in (5) and Figure 4 above, where the lexical tone of the subject marker *vá* spreads rightward to the final syllable of the verb, but is blocked from spreading into the two-word object φ that follows the verb. Why is it that HTS-RT

can spread from the verb into an object consisting of a single noun, as in (4) and Figure 3, but not spread if the object contains more than one word, as in Figure 4. Selkirk (2011) proposes that the φ-status of the two-word object has an impact on rightward H-tone spread from the verb through the effect of a structure-sensitive phonological markedness constraint on the tone-prosodic structure relation, namely CRISPEDGELEFT(φ,H), formulated in (14).

(14) CRISPEDGELEFT(φ,H)

> A H tone span must not include tone-bearing units that precede and follow the left edge of a phonological phrase φ:
>
> * /H … (φ … H\

The CRISPEDGE-Left(φ,H) constraint (14) belongs to the CRISPEDGE family of constraints posited by Ito and Mester (1999). It rules out a representation in which the left edge of a phonological phrase φ falls within a H tone span.[18] In the sentence in Figure 4, CRISPEDGELEFT(φ,H) would not allow the lexical H tone of the subject marker *vá-* to spread beyond the verb itself into the following binary φ: (φ vátísá (φ nguluve yi↓ntsó:ngó)φ)φ . In that example, H tone has simply spread as far as it could without violating any phonological markedness constraint on the tone-prosodic structure relation.

The remaining phonological markedness constraint that blocks the rightward spread of H tone in Xitsonga is the OCP (cf. Myers 1997), which we will show has the formulation in (15):

[18] In further examples to be examined in section 2.2, we will see additional types of cases where the CRISPEDGELEFT(φ,H) constraint blocks H tone spreading through the left edge of φ. In these the preceding word in which the spreading H tone originates is located at the right edge of an intonational phrase ι or at the right edge of a φ which does not also coincide with the right edge of a ι. The L-edge-of-φ formulation of the CRISPEDGE constraint in (18) captures exactly what is crucial to the edge-based blocking of H tone spread in Xitsonga.

(15) OCP(ω, H)

Within a prosodic word ω, two H tone spans may not be syllable-wise adjacent.[19]
* (ω … σ$^{H\ /H}$σ …)

The examples in (16) below consist of a verb with lexical H tone and a following object noun that contains a lexical H tone as well. The object nouns of the sentences in (16a) carry lexical H tone on their final syllable.[20] The lexical H tone of the subject marker or the verb root spreads rightward into that object noun, but is blocked from extending to its penultimate syllable, which immediately precedes the lexical H of the noun on its final syllable. Instead the H of the verb reaches only as far as the pre-penult syllable. The blocking of H tone spread where a syllable-adjacent sequence of distinct H tones is avoided looks like a familiar OCP effect, like that observed with H tone spread in many other Bantu languages (see Myers 1997 inter alia).[21]

(16) H tone spread blocked by OCP(ω,H)

a. Verb with lexical H, noun with lexical H on final syllable

PI: vá$^{/H\}$ xava ma tandzá$^{/H\}$ PO: vá$^{/H}$xává má$^{H\}$ta:$^{\downarrow}$ndzá$^{/H\}$ 'they buy eggs'
3pl buy cl6-egg

PI: ni vó$^{/H\}$na va lalá$^{/H\}$ PO: nivó$^{/H}$ná vá$^{H\}$ la:$^{\downarrow}$lá$^{/H\}$ 'I see enemies'
1sg see cl2-enemy

b. Verb with lexical H, noun with lexical H on initial (and final) syllable

---

[19] Lee (2014) examines H tone spread into longer nouns with a final H tone, which may possibly be nominal compounds. In these cases H tone does not spread into a prosodic word when a H tone is present further to the right, which may suggest that adjacency for OCP is simply at the prosodic word level, not requiring syllable adjacency as well. But it may also suggest a more complex prosodic structure for these cases which could be relevant to blocking.

[20] The presence in output phonological representation of a lexical H on a φ-final syllable is a violation of NONFINALITY(φ,H). Clearly, faithfulness constraints that call for (i) the presence of an input lexical tone in the output representation and (ii) the retention of the lexical association between the tone and its input tone-bearing unit (Myers 1997) must outrank NONFINALITY(φ,H) in Xitsonga.

[21] Note that the second H tone in a (possibly nonadjacent) sequence of two H tone spans is downstepped with respect to the first, as indicated by the down arrow that precedes the syllable bearing the second H in the sequence.

PI: vá$^{/H\backslash}$ kuma vú$^{/H}$ swá$^{H\backslash}$ PO: vá$^{/H}$kúmá$^{H\backslash}$ ↓vú$^{/H}$:swá$^{H\backslash}$ ‘they get porridge’
3pl get cl14-porridge

PI: ni vó$^{/H\backslash}$na nyó$^{/H}$ká$^{H\backslash}$ PO: ni vó$^{/H}$ná$^{H\backslash}$ ↓nyó$^{/H}$:ká$^{H\backslash}$ ‘I see a snake’
1sg see snake

Turning to (16b), where a lexical H tone appears in word-initial position in the object noun, the H tone from the subject marker or the verb root spreads rightward onto the verb-final syllable, despite its adjacency to the following initial H-tone syllable of the noun. This fact provides evidence that the OCP in Xitsonga only rules out H-tone sequences that appear within the same prosodic ω in PO.

In cases of H tone spread like those in (16a) and (16b), as well as in (5)/Figure 4, H tone spread is clearly not driven by the need to satisfy a constraint that the phrasal penult syllable bear H tone. Instead, the pattern is consistent with the idea embodied in HTS- RT (10) that in Xitsonga H tone spreads rightward freely in PO (except if blocked by some higher ranked markedness constraint(s)). More specifically, when structure-sensitive markedness constraints like these outrank HTS-RT in the phonology, as in (17), restrictions in the distribution of H tone spans in PO like those seen in Xitsonga will appear.

(17) NONFINALITY(φ,H), CRISPEDGELEFT(φ,H), OCP(ω,H) >> HTS-RT >> DEP-IO(A).

Summing up, the phonological constraint ranking in (17) displays the set of prosodic structure-sensitive markedness constraints on the tone-φ relation in PO that provide an account of the limits on the free H tone spread in PO that is called for by the lower-ranked markedness

constraint HTS-RT. As for the phonological constraint ranking BINARITY(φ) >> MAX(φ) that figures in tableaux (8) and (9), it is repeated here as (18).

(18) BINARITY(φ) >> MAX(φ)

The ranking in (18) is responsible for producing a phonologically driven, mismatching, fine-tuning in PO of the prosodic φ structure that spells out the morphosyntactic phrase structure of MSO in PI. The combination of the constraint rankings in (18) and (17) constitute the phonological analysis of H tone spread and its structure-sensitivity in Xitsonga. Together with the MatchPhrase theory of the spell-out relation between the phrase structure of MSO and φ structure in PI, they provide an account of the 'indirect' relation illustrated in Figures 3 and 4 between the phrase structure of MSO, on the one hand, and the φ structure and patterns of H tone spread in PO, on the other.

2.2 Generalizing to other syntactic contexts

The analysis presented here of the phonology of φ-structure and φ-sensitive H tone spread in Xitsonga holds quite generally, not just in contexts where H spreads from a head verb into a verbal complement. Due to language-particular accidents in the tonal properties of nouns and 'modifier' expressions in Xitsonga, however, we are not able to observe in PO whether H tone spreads from a head noun to a modifier phrase. First, there is no H tone spread within the noun root itself. The association of H tone to tone-bearing unit is contrastive in nouns in Xitsonga; lexical H tone span may appear on any tbu of a noun in the input representation PI, including the final syllable.[22]

[22] It is not uncommon cross-linguistically for nouns to show a greater array of lexical contrasts than verbs (see, e.g., Smith 2001). Smith argues that PI-PO faithfulness constraints operate differently in nouns and

Moreover, it is not possible to even pose the question whether a noun-final lexical H tone can spread into a following modifier. This is because there are no adjectives, numerals or other modifiers in Xitsonga that, by virtue of their own lexical tone properties in PI, are able to host a H tone span extending rightward in PO from a final lexical H tone in the preceding head noun. No adjectives or numerals are toneless and none have lexical H tone far enough to the right for the OCP to not block H tone spread into that word.[23]

Importantly, there are instances of H tone spread in PO in Xitsonga where the tone span extends between words that are in different morphosyntactic phrases of MSO. This allows us to expand the empirical scope of our investigation. For example, in the two schematic representations of sentences containing the double object construction in (19a-b), a noun that corresponds to the first object phrase has a lexical H tone on its final tbu in the input phonological representation PI.

(19)

a. verb [*noun*]$_{NP1}$ [*noun*]$_{NP2}$

PI: ni hlawulela hosí$^{/H\}$ hlambeto
1s select chief cooking pot

PO: ni hlawulela hosí$^{/H}$ hlámbé:$^{H\}$to
'I select for the chief a cooking pot'

b. *verb* [*noun*]$_{NP1}$ [*noun* [*mod*]$_{MP}$]$_{NP2}$

PI: ni hlawulela hosí$^{/H\}$ hlambeto yin'we
1s select chief cooking pot one

PO: ni hlawulela hosí$^{/H\}$ hlambeto yi:n'we.
'I select for the chief one cooking pot.'

In (19a) that final H tone of the first object spreads into the second object phrase, but in (19b) it doesn't. In anticipation of the discussion of the full syntactic and prosodic phrase structure of the double object construction coming up in section 3.1, the prosodic φ structure representations of PI and PO in (19a-b) have been left out. But the simple display of the noun-final locus of lexical H

verbs. In Xitsonga nouns, a lexical H tone span may fall on one or more tone-bearing units in a noun, regardless of position, while lexical H tone in verb roots in Xitsonga is restricted to root-initial position.

[23] Representative adjectives in Xitsonga are as follows: *-ntshwá* 'new', *-nkúlu* 'big', *-néné* 'good', *-ntsóngo* 'small' etc. Numerals *-mbirhí* 'two' and *-nhárhu* 'three' are adjectival modifiers; numbers above 'four' are nouns and require the associative construction.

tone in the first object that is reflected in the PI, along with the familiar BINARITY(φ)-based variation in H tone spread into the second object in the PO in (19a) vs. (19b), provides evidence for the existence of between-phrase H tone spread in Xitsonga. In (19b), where a noun and a modifier phrase make up the second object, the φ status of that two-word object in PO and the ranking CRISPEDGELEFT(φ,H) >> HTS-RT mean that no between-phrase H tone spread is possible. But in (19a), between-phrase H tone spread is possible, because the second object is *not* a φ in PO and so CRISPEDGELEFT(φ,H) does not come into play. The prediction of the phonological analysis motivated so far is that, in Xitsonga, between-phrase H tone spread is possible between words in any morphosyntactic context as long as the word on the right does not lie in φ-initial position in PO.[24]

A more dramatic instance of 'between phrase' H tone spread appears in sentences containing right-dislocated single-word constituents, as in (20a), but is not present in the sentences in (20b), with two-word right-dislocated constituents. In (20a) we see H tone spread into the single-word dislocated phrase from the lexical final H tone of the preceding word. In (20b), where the dislocated phrase is binary, there is no such H tone spread.

(20) Right dislocations in Xitsonga

a. Single-word right dislocations

yá-nwá ↓má:tí, ngúlú:ve 'it drinks water, the pig'

[24] The spreading of H tone from the final tbu of the first object noun into the first two tbu's of the second noun in (19i) shows that the low-ranked faithfulness constraint DEP-IO (A) in (12), which blocks tone spreading, is violated when the H tone spreads between words. By contrast, the length of a multi-tbu H tone span in PO that originates and ends within the same noun root is lexically contrastive in Xitsonga. See discussion in footnote 22. Smith (2001) argues that this sort of fact points to the existence of a special tonal faithfulness constraint for nouns, one that would be confined to the nominal word.

sm9-drink cl6-water pig

ú-vóná ho:↓sí, dókódé:la ‘s/he sees the chief, the doctor’
3sg-see. chief doctor

b. Two-word right dislocations

yá-nwá ↓má:tí, nguluve yi-n↓tsó:ngó ‘it drinks water, the small pig’
sm9-drink cl6-water pig cl9-small

ú-vóná ho:sí, dokodela ↓ló:-ntshwá ‘s/he sees the chief, the new doctor’
3sg-see chief doctor cl1-new

This case of H tone spread is of particular interest because there are good reasons—syntactic and phonological—to think that the right-dislocated constituent is not even part of the same clause as the material that precedes. As Kisseberth (1994) points out, penultimate lengthening in Xitsonga (indicated by a colon following a vowel) appears on the pre-final syllable of the final word of any sentence.[25] Note that penultimate lengthening appears in all the example sentences of this paper. Importantly, Kisseberth observes, penultimate lengthening appears in a word that precedes a right-dislocated phrase, and it also appears in the last word of the right-dislocated phrase itself. He suggests this distribution of penultimate lengthening indicates a nested clause structure, in which the right-dislocated constituent is adjoined at the right edge of a clause and is also daughter of a clause. We adopt that syntactic analysis here, accompanied by the assumption, reflected in (21), that clauses are spelled out as intonational phrases (ι) in PI, and that penultimate lengthening is defined with respect to ι in PO, as illustrated in (21) and (22).

(21) MSO $[_{\text{clause}}\,[_{\text{clause}}\,[_{\text{VP}}\,[_{\text{V}}$ *ya nwa* $[_{\text{NP}}$ *ma ti* $]\,]\,]_{\text{clause}}\,[_{\text{NP}}$ *nguluve* $]\,]_{\text{clause}}$
sm9 drink cl6 water pig

PI (ι (ι (φ yá$^{/H\backslash}$nwá$^{/H\backslash}$ (φ ma tí$^{/H\backslash}$)))ι (φ nguluve))ι

[25] This is the only vowel length contrast in PO in Xitsonga.

PO (ι (ι (φ yá/H\nwá /H má:H\tí/H))ι ngúlú:H\ve)ι ‘it drinks water, the pig’

(22) MSO [clause [clause [VP [V *ya nwa* [NP *ma ti*]]]clause [NP *nguluve* [*yi ntsongo*]]]clause
sm9 drink cl6 water pig cl9 small

PI (ι (ι (φ yá/H\nwá/H\ (φ ma tí/H\)))ι (φ nguluve (φ yi ntsó/Hngó H\ )))ι

PO (ι (ι (φ yá/H\nwá /H má:H\tí/H))ι (φ nguluve yi ntsó:/Hngó H\ ))ι

‘it drinks water, the small pig’

Given the nested ι structure in these sentences, it’s no surprise that the penultimate syllable of each of the two ι would show penultimate lengthening.

Now let’s turn to the pattern of ‘between phrase’ H tone spread that’s observed in these example sentences. In both of these sentences our focus is on the rightward spread of the lexical H tone of the noun root *ti* ‘water’ which is final in the lower clause. In (21), the right-dislocated phrase consists of the single lexically toneless noun *nguluve* ‘pig’, and we see that H tone spreads from the ι-final *ti* into *nguluve*, which in PO has the status of a prosodic word (ω) but not of a φ. In (22), on the other hand, the H tone of *ti* is blocked from spreading into the following φ, which consists of *nguluve* and a following modifier. This effect of the binarity of a φ on H tone spread is by now a familiar sort of pattern. Our solution to the puzzle of the effects of constituency on H tone spread has been to characterize the sensitivity of H tone spread to prosodic structure in terms of markedness constraints on the tone-edge relation, namely NONFINALITY(φ) and CRISPEDGELEFT(φ,H). The latter alone will block H from spreading into the φ of the binary right dislocation in (22).

It should be pointed out that the structure-sensitivity of the phenomenon of H tone in Xitsonga can *not* be characterized in terms of a prosodic structure ‘domain’ for tone spread, where the notion ‘domain’ is taken to be a prosodic constituent (whether ω, φ, or ι (as in Selkirk 1980))

and/or the terminal string of that constituent. In the cases of H tone spread in Xitsonga there simply isn't a designated constituent type—whether phonological or morphosyntactic—inside of which and/or across which H tone freely spreads. What we have seen is that, in Xitsonga, H tone can spread rightward out of an intonational phrase (ι), but not into what follows, if what follows has the status of a φ (as opposed to a mere ω). Domain-based accounts, like that of Pak (2008) on domain-based leftward H-tone spreading in Luganda, for example, apparently need to be revisited. On the view we have taken here and in Selkirk and Lee (2015), any supposed 'domain effect' on tonal feature spreading is captured by phonological markedness constraints on tone-edge combinations. So it is simply an accident of a constraint-ranking-based OT typology that the constraints CRISPEDGERIGHT(φ,H) or CRISPEDGERIGHT(ι,H), which would exist in the universal constraint repertoire, do not happen to block rightward H tone spread out of a φ or a ι in (the variety of) Xitsonga reported on here.

3. No spell-out of Fnc-headed phrases in Xitsonga

In the current section we look at patterns of H tone spread from the verb in two morphosyntactic configurations within the VP. These patterns of H tone spread provide evidence for constituency mismatches between PO and MSO that don't have an obvious explanation in the phonology *per se*. Instead, the generalization seems to be that, in Xitsonga, only morphosyntactic phrases of MSO that have a lexical category head are spelled out as a φ in the input representation (PI) for the phonology. Since information about the categorial makeup of words and morphemes is available only in MSO, in the context of the MSO-PI-PO architecture it would be the spell-out module of Xitsonga that is responsible for these mismatches.

3.1 φ formation in double object sentences in Xitsonga

In the double object construction, when the first object phrase consists of a single noun which is lexically toneless, a H tone will spread from the verb into that noun, stopping at the pre-final syllable of the noun, as shown in (23). Given the analysis of H tone spread in section 2, this pattern of H spread in the verb-noun sequence means that the verb and the noun together form a φ in PO, where H tone spread beyond the pre-final tbu of the noun is ruled out by NONFINALITY(φ,H).

(23)

a. verb [*noun1*]$_{NP1}$ [*noun2*]$_{NP2}$

PI: vá xavela munhu ti nguvu
3p-buy-appl someone cl10-cloth

PO: ((váxávélá múnhu)φ tingu:vu)φ
'They are buying clothes for someone'

b. verb [*noun1*]$_{NP1}$ [*noun2*]$_{NP2}$

PI: ndzi nyíka xi koxa nyama
1s-give cl7-old.woman meat

PO: ((ndzinyíká xíkóxa)φ nya:ma)φ
'I am giving an old woman meat.'

But this grouping of the verb and the first object into a prosodic phrase φ goes counter to the assumed morphosyntactic phrase structure for such cases. According to current understanding of the morphosyntax of double object constructions in Bantu (Baker 1989, Carstens 2005, Baker et al. 2012, van der Wal 2015), the first object does *not* form a constituent with the verb in MSO. Instead the first object forms a phrasal constituent with the second object in MSO, and the verb lies outside of that phrase. The S-P constituency mismatch here needs explanation.

Sentence (23a) illustrates a double object sentence with a verb stem in MSO that is composed of the verb root *xav-* 'buy' and the applicative morpheme *-el*. The direct object, which appears last in (23a), is an argument of the low-positioned verb at an earlier stage of the syntactic derivation; the higher applicative head is responsible for introducing the object phrase that appears first. As is generally the case in the morphosyntactic derivation of sentences in Bantu, the verb

raises from its original low position in the sentence structure, adjoining to successively higher functional heads, including the applicative, along its way to its initial position in the extended verbal projection. This derivation is reflected in (24).

(24) $[_{VP}\ [_{V}\ sm\text{-}..verb_i\text{-}appl_j..]\ \ldots[_{\boldsymbol{applP}}\ [_{NP}[_{N}\ noun1\ ]_{N}]_{NP}\ [\ \varnothing_{i,j}\ [_{\boldsymbol{VP}}\ \varnothing_{i}\ [_{NP}[_{N}\ noun2\ ]_{N}]_{NP}]_{\boldsymbol{VP}}\ ]\ ]_{\boldsymbol{applP}}]_{VP}$

Verb raising yields an MSO in which a V-headed morphosyntactic word is the daughter and morphosyntactic head of that highest phrase of the structure, which we label here as VP, since it is headed by a word that contains the lexical root V.[26] The ∅ symbols indicate the positions occupied by the verb in the course of verb-raising: the verb's original position ($\varnothing_i$); then as it raises by head movement, the position of adjunction to the applicative head [$\varnothing_{i,j}$]; and then the movement of the $verb_i\text{-}appl_j$ combo through ever higher functional heads, ending in a derived word structure [$sm..verb_i\text{-}appl_j..$] in which the highest functional head of the verbal extended projection, the subject marker, occupies initial position.

So, the two objects are both contained within the applP in MSO and the derived verb lies higher in the sentence structure, outside of the applP. How then to explain the radical mismatch between the phrasal constituency for MSO that's posited in (24) and the φ constituency of PO that's testified to by the verb-to-noun1 pattern of H tone spread in (23)? What changes need to be made in our current analysis of phrasal constituency spell-out and/or the phrasal phonology *per se* that would yield the PO representation ((verb noun) noun) in (23) on the basis of the MSO representation [*verb* [[*noun*] [*noun*]]] in (24)?

[26] The italicization of *VP* and *ApplP* indicates that these structures no long dominate their heads.

Our hypothesis is that, in the grammar of Xitsonga, the constraint for spelling out the morphosyntactic phrasal constituency of MSO as φ constituency in PI is not MatchPhrase, (2), but rather the more specific version $\text{MatchPhrase}_{\text{LEX}}$, (25):

(25) $\text{MatchPhrase}_{\text{LEX}}$

> For every Phrase in the morphosyntactic output representation MSO that is headed by a word containing a lexical category root (N, V, A) there is exactly one φ in the input phonological representation PI that spells out that Phrase phonologically.

$\text{MatchPhrase}_{\text{LEX}}$ would spell out just three phrases of the MSO (24) as φ in PI—the two object noun phrases and the upper VP headed by the verb.[27] This yields the tripartite φ structure in PI shown in (26) in which the verb and the two object phrases are sisters: (verb (noun)φ (noun)φ)φ. It would be up to the phonology to derive the left branching φ structure ((verb noun)φ noun)φ in PO from the φ with three daughters in PI. This is a welcome result, because BINARITY(φ) and STRONGSTART (φ),[28] which are independently motivated types of phonological markedness constraints of the phonology *per* se, can in that case form part of an explanation for the presence of the observed mismatch in PO.

(26) Deriving PO from MSO in two steps

MSO $[_{\text{VP}}\ [_{\text{V}}\ \textit{sm-verb}_i\textit{-appl}_j]\ \ldots[_{\textbf{\textit{applP}}}\ [_{\text{NP}}[_{\text{N}}\ \textit{noun}]_{\text{N}}]_{\text{NP}}\ [\varnothing_{\text{i,j}}\ [_{\textbf{\textit{VP}}}\ \varnothing_{\text{i}}\ [_{\text{NP}}[_{\text{N}}\ \textit{noun}]_{\text{N}}]_{\text{NP}}]_{\textbf{\textit{VP}}}\ ]\ ]_{\textbf{\textit{applP}}}]_{\text{VP}}$

SPELL-OUT (by $\text{MatchPhrase}_{\text{LEX}}$)

PI $(\varphi_1\ (\omega$ sm verb appl $)\omega\quad (\varphi_2\ (\omega$ noun $)\omega)\varphi_2\quad (\varphi_3\ (\omega$ noun$)\omega)\varphi_3\ )\varphi_1$

[27] Modern syntax has vastly expanded the repertoire of functional heads of phrases beyond the set of functional categories that are typically spelled as separate words in the common orthographical and printing practices, as with determiners, verbal auxiliaries, complementizers, and the like. Functional heads in Bantu that end up as part of a larger word in MSO and are spelled typically spelled as part of the word include the applicative head mentioned here, the noun class marker, and the tense marker.

[28] Selkirk (2011), Elfner (2012), Bennett et al. (2016).

PHONOLOGY

PO (φ1 (φ4 (ω sm.verb.appl)ω (ω noun)ω)φ4 (ω noun)ω)φ1

As we saw in sections 1.3 and 2, the phonological markedness constraint BINARITY(φ) plays an important role in Xitsonga in accounting for the 'deletion' of a φ containing just one ω. BINARITY(φ) can moreover be held responsible for the insertion of the new φ4 in PO that groups together the verb and the first object as in (26), as well as for the deletion of the non-binary φ of both object noun phrases in that sentence.

Of course, a phonological analysis of the relation between PI and PO in (26) requires more than the mere recognition that the constraint BINARITY(φ) is playing a crucial role here. There are two issues yet to be addressed. The first issue concerns the creation of the new, 'inserted', φ4 in PO. The presence of a new φ in the output, one that is not present in the input, would incur a violation of a faithfulness constraint of the DEP family, which penalizes the presence of elements in the output that do not correspond to elements of the input (McCarthy and Prince 1995, 1999). The prosodic faithfulness constraint DEP(φ), hereby proposed, would penalize the presence in PO of a φ that is not present in the input:

(27) DEP(φ) ('No insertion of φ')[29]

Any constituent (node) φn of PO must also be present in PI.

[29] Compare Kratzer and Selkirk (2020), where the formulation of DEP(φ), identifies the prosodic constituent referred to in terms of its terminal string, rather than directly, 'by name.'

The constraint-ranking BINARITY(φ) >> DEP(φ) in (28a) in the grammar of Xitsonga would ensure the possibility of the 'insertion' of a φ in order to satisfy the BINARITY(φ) markedness constraint:

(28) a. BINARITY(φ) >> DEP(φ) [drives insertion of φ]

b. BINARITY(φ) >> MAX(φ) [drives deletion of φ]

Recall that we already have evidence from section 2 for the ranking BINARITY(φ) >> MAX(φ) in (28b). This ranking licenses the 'deletion' of φ from the PO of both the unary object phrases.

The second question is why the inserted binary φ appears at the left edge of the higher φ rather than at the right edge? A prosodic markedness constraint of the STRONGSTART family provides the answer to this question regarding φ organization. This constraint family is independently motivated in Selkirk (2011) and crucially exploited in Bennett et al. (2016). (29) gives the constraint in a form that is specific to φ.

(29) STRONGSTART (φ)

A prosodic constituent φ must begin with a leftmost daughter constituent π_n which is not lower in the prosodic hierarchy than the constituent π_{n+1} that immediately follows:

* $(_{\varphi}\ \pi_n\ \pi_{n+1}\ \ldots.)_{\varphi}$, where π_n is lower in the prosodic hierarchy than π_{n+1}

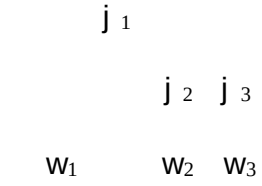
STRONGSTART(φ) and BINARITY(φ) together impose the left-branching φ structure in PO of the optimal candidate in the double object construction, illustrated in (26). Let's examine the optimality-theoretic tableau in (30) to confirm that the proposed system of ranked phonological marked and faithfulness constraints on prosodic structure has the result that the optimal candidate (30d)—i.e., the PO of (26)—is selected as optimal. Given that the labelled bracketings in the

tableau are not easy to read, the prosodic constituent structures of the four candidates under consideration in (30) are also provided in tree-form immediately below the tableau.

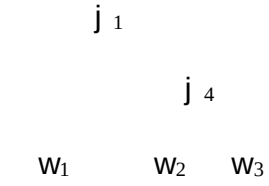
(30) Double object construction with two unary NPs

PI	(φ1 ω1vá/H\xavela (φ2 ωmunhu)φ2 (φ3 ωtinguvu)φ3)φ1	BINARITY (φ)	DEP (φ)	MAX (φ)	STRONG START(φ)
PO	a. (φ1 ω1vá/Hxávélá H\ (φ2 ωmunhu)φ2 (φ3 ωtinguvu)φ3)φ1	*φ1*φ2!*φ3			*ω1
	b. (φ1 ω1vá/H-xávélá H\ (φ4 ωmunhu ωtinguvu)φ4)φ1		*φ4	*φ2*φ3	*ω1!
	c. (φ1 ω1vá/H-xávélá ωmúnhú ωtingú H\vu)φ1	*φ1!		*φ2*φ3	
	☞ d. (φ1 (φ5 ω1vá/H-xávélá ωmún H\hu)φ5 ωtinguvu)φ1		*φ5	*φ2*φ3	

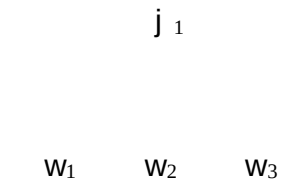
(30a')



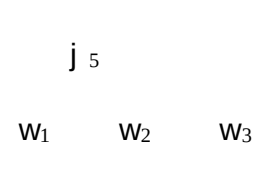
(30b')



(30c')



(30d')



(30a) is the candidate that is fully faithful to the input PI representation, but its violation of STRONGSTART(φ) and BINARITY(φ) (the latter multiple times) is fatal. BINARITY(φ) drives the deletion of the φ for the single-word objects in the surface representations of the three candidates

(30b,c,d), which incur the noted violations of the lower-ranked MAX(φ). BINARITY(φ) also drives the unfaithful φ insertion that incurs a violation of the lower-ranked DEP(φ) violation in candidates (30b) and (30d): BINARITY(φ) over DEP(φ) and MAX(φ) categorically rules out the fully faithful ternary candidate (30a), as well as the ternary (ω ω ω) structure of (30c). It falls to STRONGSTART(φ) to drive the attested positioning of the epenthetic φ at the left edge of the higher φ in the optimal candidate (30d). STRONGSTART(φ) is violated in the right-branching φ structure in (30b), as well as in the totally faithful tripartite structure in (30a)).

This phonological account of the unfaithful positioning of an epenthetic binary φ at the left edge of the higher φ in the optimal candidate (30d) relies on the new and otherwise untested φ-specific markedness constraint STRONGSTART(φ). Looking back to section 2, we need to account for the violation of STRONGSTART(φ) in the optimal PO of a sentence with a transitive verb followed by a single binary object, as in ($(_{\varphi 1}$ (ω v<u>á</u>tísá) $(_{\varphi 2}$ (ω nguluve) (ω yi-<u>ntsó:ngó</u>) $)_{\varphi 2}$ $)_{\varphi 1}$ 'They bring a small pig' from Figure 4 in section 1.3. Importantly, in Figure 4 the noun and following adjective form a φ in the input representation PI, and they do so in PO as well. The maintenance of this binary φ in PO would be the consequence of satisfying the PI-PO faithfulness constraint MAX(φ). Our conclusion is simply that MAX(φ) is higher ranked than STRONGSTART(φ) in the grammar of Xitsonga:

(31) MAX(φ) >> STRONGSTART(φ)

This ranking completes our account of the phonology of φ structure in Xitsonga.

To sum up, by assuming that the spell-out of MSO phrases as φ in PI is accomplished by the interface constraint $\text{MatchPhrase}_{\text{LEX}}$ in Xitsonga, the applP of the MSO of the double object

construction in (26), which lacks a lexical category head, will not correspond to a φ in the phonological representation PI that is input to the phonology *per se.* Crucial to the successful derivation of the attested PO by the phonology *per se* is the assumption that the φ in PI that spells out a VP with the double object construction in MSO is ternary branching. With a PI containing only φ's that correspond to Lex-headed phrases, and a phonology *per se* that consists of the appropriate ranking of independently motivated prosodic markedness and faithfulness constraints, a simple and insightful analysis can be offered for the attested left-branching φ-structure and associated patterns of H tone spread of the optimal candidate for PO in double object constructions.

Our proposal is that the interface constraint $\text{MatchPhrase}_{\text{LEX}}$ is key to understanding the φ-structure of the double object construction in Xitsonga and other languages like it. An alternative is the proposal that only morphosyntactic phrases that have heads which are phonologically overt are spelled out as φ in phonological representation, as put forward in Kalivoda (2018). In the double object construction in Xitsonga there is a phonologically overt applicative head, but, as we saw in (24), in the course of head movement/verb raising in the syntactic derivation, it is raised out of its position as head of the applicative phrase in the double object construction, leaving the applicative phrase in MSO with only the trace of a head. So it is indeed possible that, lacking an overt head, the applicative phrase does not 'count' when phrases of MSO are spelled out as φ in phonological representation.

But there are three reasons for preferring an appeal to the presence of a lexical category head over an appeal to the absence of a phonologically overt head in the formulation of the MatchPhrase constraint that spells out an MSO phrase as φ. The first is based on evidence from Xitsonga to be reviewed in the next section: MSO phrases which have a functional head that *is* phonologically overt still do *not* map onto φ in phonological representation. $\text{MatchPhrase}_{\text{LEX}}$ can

account both for this fact, and for the fact that the applicative phrase in Xitsonga, which lacks on overt functional head, is also not spelled out as a φ. The second reason for making an appeal in Match constraints to the property of being a lexical category is that the functional/lexical distinction clearly plays a role in spelling out word-level constituents cross-linguistically. Functional category items that head morphosyntactic phrases tend not to have the status of prosodic word (ω) in the output phonological representation, while lexical category items do (see, e.g., Selkirk 1996). This is additional independent evidence for relying on an appeal to the lexical vs. functional category status of a morphosyntactic unit in the spelling out of prosodic constituent structure in PI. A third reason is that MatchPhrase in our current conception is a constraint of the spell-out module, which simply produces a φ in PI which corresponds to a morphosyntactically specified phrase of MSO. The proposal by Kalivoda (2018) extends the power of interface Match constraints to allow for an examination of the terminal string of the PI representation. As we've seen, this is not a necessary move.

3.2 The spell-out and phonology of noun class markers in Xitsonga

A well-known characteristic of the noun in Bantu languages is its membership in one of a large set of noun classes. In Xitsonga the noun *-lungu* 'European', for example, belongs to class 1. As in Bantu more generally, Xitsonga nouns of class 1 are associated with a pair of class markers, one for the singular and one for the plural. A class 1 noun in the singular takes the corresponding marker *mu-*: *mu-lungu*. The distinct class marker *va-* appears with a class 1 noun in the plural: *va-*

lungu 'Europeans'.[30] In Bantu, noun class is reflected in morphosyntactic agreement. Within the noun phrase, a class marker agreeing with the head noun appears preceding any modifier of the noun. Moreover, the choice of the subject marker that forms part of the verb reflects the noun class of the head of the subject noun phrase. In traditional accounts noun class markers are thought of as prefixes (Welmers 1973, Greenberg 1977, Mufwene 1980), but in Xitsonga they do not behave phonologically as if they form part of the same word as the following noun.

Noun class markers in Xitsonga are for the most part lexically toneless,[31] and so would not in general block the rightward spread of lexical H tone from a preceding word into the noun that follows them. The central puzzle about noun class markers that we must deal with in this section concerns these additional facts:

First, a H tone that is final in a preceding word may *always* spread into a following toneless noun class marker, and second, H tone spread from the preceding word into *both* the toneless noun class marker *and* the following noun happens only when the noun forms part of a single-word phrase (as in (32)), but never when the noun forms part of a phrase consisting of two or more words (as in (33)).

(32) H tone spread into noun class marker and noun of a single-noun phrase

PI ni lá$^{H\backslash}$nguta mu lungu
1sg-look.at *cl1-European*

PO (φ ni lá$^{/H}$ngútá mú lú:$^{H\backslash}$ngu)φ

'I look at a European'

(33) H tone spread into noun class marker but not into noun of a multi-word phrase

[30] Misleadingly, the singular is referred to as a class 1 marker, the plural is referred to as a class 2 marker. (There is no singular noun class 2.) In similar fashion, the singular and plural class markers for Class 3 nouns are referred as class 3 and class 4 markers, respectively. And so on.
[31] One noun class prefix has a H tone (the class 2a, *vá-*).

PI ni lá$^{H\backslash}$nguta mu lungu ló nkúlú
1sg-look.at cl1-European cl1- big

PO (φ ni lá$^{/H}$ngútá mú$^{H\backslash}$ (φ lungu $^{\downarrow}$lónkúlú)φ)φ
'I look at a big European.'

The phonological output representations (PO) of the sentences in (32) and (33) are shown with the φ-structure that would account for the observed pattern of H tone spread in the terms of the phonological account presented in section 2. In the PO of (32) the verb, noun class marker and noun are all contained in the same φ; H tone spreads from the verb through the sentence, where its failure to spread to the final tbu is blocked by NONFINALITY(φ). In the PO of (33), the class marker is followed by a φ that contains the noun and following modifier; H tone spreads from the verb through the noun class marker, but it is blocked from continuing into the φ-initial noun stem by CRISPEDGELEFT (φ,H). What's significant here is the mismatch between the φ structure of the PO of (33) and a commonly assumed morphosyntactic structure of MSO in which the noun class marker is a prefix within the noun, and that noun and its modifier form a Phrase. With an MSO like this, the noun class marker should be grouped with the noun and following modifier in the φ structure of PO. It would *not* undergo the inter-word H tone spread observed in (33).

The PO of sentences containing various other types of noun + modifier construction in (34) show the same H tone spread from the verb into just the class marker alone that hints at a mismatch between PO and MSO constituency. The superscript notation $^{/H...H\backslash}$ makes clear the span of the H tone spread.

(34) Class marker followed by various noun + modifier constructions

i. noun + numeral

a. ni lá$^{/H}$ngútá vá$^{H\backslash}$ lungu va mbi:rhí 'I look at two Europeans'
1s look.at cl2 European cl2 two

b. vá/H xávísá tíH\ mbhongolo ti nhá:rhu ‘They sell three donkeys’

3p sell cl10 donkey cl10 three

ii. noun + associative phrase (headed by associative/possessive particle *-á*)

a. ú/H tísá tíH\ nyama t á mbhóngó:lo ‘He brings lots of donkey meat’

3s bring cl10-meat cl10 assoc donkey

b. ni lá/Hngútá má H\ -sangu m á ndzúlámí:so ‘I look at sleeping mats of Ndzulamiso’

1s look.at cl6 sleeping.mat cl6 assoc Ndzulamiso

iii. noun + relative clause

a. ni lá/Hngútá váH\ lungu la va tirha:ka ‘I look at Europeans who are working’

1s look.at cl2 European rel cl2 work

b. <u>vá</u>/H tísá tí H\-nguluve le ti nwa:ka ‘They bring pigs that drink’

3p bring cl10 pig rel cl10 drink

Facts like these from Xitsonga about H tone spread into the noun classifier but not into a noun followed by a modifier were first reported by Cole-Beuchat (1959) and Beuchat (1962), and subsequently in Herbert (1992) and Kisseberth (1994). Based on such facts, Herbert proposed a syntactic analysis in which the noun class marker is external to the syntactic constituent containing the noun and its modifier:[32]

(35) The noun class marker is sister to an N-root-headed phrase in Xitsonga (Herbert 1992)

[va [[nhwana] [va [khúme]]]] ‘10 girls’

cl2 girl cl2 ten

We will build here on Herbert’s proposed structure for the ‘noun phrase’ in (35). Our version of the proposal is (i) that the class marker *va* for a noun root like *nhwana* in (35) would have the

[32] Kisseberth (1994) also suggests a function word status for the class marker.

status of a functional category head of a class marker phrase (clP), as shown in the MSO of (36) below, and (ii) that the noun stem and following modifier together form a phrase. We are not in a position to offer a more detailed hypothesis about the structure of what follows the head noun in MSO. Note that in (35) the numeral is immediately preceded by a classifier formative that agrees in noun class with that of the head noun, but we are unfortunately not able at this point to identify the structural position of noun classifiers introduced by agreement. We do know that there is no possible evidence from H tone spread for assuming that the agreeing classifier that precedes a following modifier is the head of a classifier phrase taking a modifier complement. Recall from the discussion in the first paragraph of section 2.2 that, for quite independent reasons, there are no possible cases of H tone spread from a head noun into a following modifier in Xitsonga. So we leave the constituents of the modifier phrase unlabeled in the MSO, PI and PO representations in (36).

According to our proposal in the previous section concerning the nature of the Match constraint that spells out MSO phrase structure as φ in Xitsonga, it is the $\text{MatchPhrase}_{\text{LEX}}$ version in (25) that is responsible for generating prosodic φ structure in PI. The full MSO-PI-PO derivation in (36) shows that $\text{MatchPhrase}_{\text{LEX}}$ spells out no φ in PI that would correspond to the class-marker-headed clP of MSO. This is exactly the φ structure for PI that, submitted to the ranking of phonological markedness and faithfulness constraints of the phonology *per se* proposed above in 3.1, would give rise to the PO representation in (36), in which H tone from the verb spreads into the classifier morpheme, but no farther.

(36) Two steps in the derivation of the constituency of PO from MSO: one mismatch may lead to yet another

MSO [$_{VP}$ [$_{V}$ *ni languta*]$_{V}$ [$_{clP}$ *mu* [$_{NP}$ [$_{N}$ *lungu*]$_{N}$ [*lo- nkulu*]]$_{NP}$]$_{clP}$]$_{VP}$

1sg look.at *cl1* *European* *cl1* *big*

SPELL-OUT (with MatchPhrase$_{LEX}$)

PI (φ_1 (ω ni lánguta)ω mu (φ_2 (ω lungu) (ló nkúlú))φ_2)φ_1

PHONOLOGY

PO (φ_1 (ω(ω nilángútá) mú)ω (φ_2 (ω lungu) (ló nkú:lú))φ_2)φ_1

Crucially absent from the PI in (36) is a φ that would correspond to the classifier phrase clP in MSO. MatchPhrase$_{LEX}$ is responsible for this mismatch. It delivers a PI for the MSO in (36) in which the functional head *mu* is 'stray', grouping with neither the following noun phrase nor the preceding verb. This constituency mismatch would be further compounded if the constraint system of the phonology *per se* were indeed to result in the prosodic adjunction of the class marker to the preceding verb in PO, as shown in (36). In some languages there is clear phonological evidence that the functional head of a post-verbal phrase in MSO is enclitic to the preceding verb in PO (see below for discussion). In Xitsonga the phonology of H tone spread testifies only to the absence in PO of the left edge of a φ preceding the class marker.

The predicted representation for the PO in (36) nonetheless commits to an encliticized status for the class marker. This is because the high-ranked phonological markedness constraint

BINARITY(φ) would rule out a PO for (36) that is identical to the non-binary structure in PI, in which the class marker has both the status of sister to the verb and of sister to the following phrase. We see this in the tableau in (37) where the φ structure of the PI is what MATCHPHRASE$_{LEX}$ would deliver, given that it only spells out Lex-headed phrases of MSO as φ in PI, as in (36).

(37) Prosodic structure for the class marker in PO with PI supplied by MATCHPHRASE$_{LEX}$

PI	$(\varphi_1(\omega_1$ ni lá$^{/H\backslash}$nguta) mu $(\varphi_2\,(\omega_2$ lungu) $(\varphi_3\,(\omega_3$ lo nkúlú) $)\varphi_3\,)\varphi_2\,)\varphi_1$	BINARITY(φ)	DEP(φ)	MAX(φ)	STRONG START (φ)
PO	a. $(\varphi_1\ \ \omega_1$nilá$^{/H}$ngútá mú$^{H\backslash}$ $(\varphi_2$ lungu $(\varphi_3$ lonkúlú $)_{\varphi 3}\,)_{\varphi 2}\,)_{\varphi 1}$	*φ_1!*φ_3			*ω_2
	b. $(\varphi_1\ \ \omega_1$nilá$^{/H}$ngútá mú$^{H\backslash}$ $(\varphi_2$ lungu lonkúlú $)_{\varphi 2}\,)_{\varphi 1}$	*φ_1!		*φ_3	
	☞ c. $(\varphi_1\,(\omega_4\ (\omega_1$nilá$^{/H}$ngútá) mú$^{H\backslash})_{\omega 4}$ $(\varphi_2$ lungu lonkúlú $)_{\varphi 2}\,)_{\varphi 1}$			*φ_3	*ω_4
	d. $(\varphi_1\ \omega_1$ni-lá$^{/H}$ngútá$^{H\backslash}$ $(\varphi_5$ mu $(\varphi_2$ lungu lonkúlú $)_{\varphi 2}\,)_{\varphi 5}\,)_{\varphi 1}$		*φ_5!	*φ_3	*ω_1, **mu**

Candidates (37a) and (37b) are nonoptimal because of the violation(s) of high-ranked BINARITY(φ). In both, the φ_1 that spells out the VP as a whole immediately dominates three prosodic constituents. The medial 'stray' syllable expressing the class marker is sister in prosodic structure to both the ω_1 spelling out the verb and the φ_2 that spells out the NP complement of the class marker phrase. In the optimal candidate (37c), that BINARITY(φ) violation is removed through the encliticization of the class marker to the verb. The optimal candidate (37c) does show a single violation of STRONGSTART(φ), which is due to the φ_1-initial ω_4 preceding a φ within the higher φ. That initial ω_4 consists of the ω formed by the verb and the class marker which has encliticized to it. As for candidate (37d), here the class marker is left-adjoined to the following φ,

with which it forms a (new) φ in violation of DEP(φ). (37d) also shows two violations of STRONGSTART(φ). Therefore (37c) emerges as optimal.

Let's now put the prosodic structure of a functional head like the class marker of Xitsonga in cross-linguistic perspective. There is evidence from ω-internal-phonological patterns in a variety of languages that a nonlexical (functional) head which in the syntax c-commands an immediately following phrase must be analyzed as enclitic to a preceding verb or to some other preceding constituent in the phonological representation. The well-known case of post-verbal noun phrase determiners forming a phonological unit with the preceding verb in Kwakwala is reported by Boas (1947) and analyzed by Anderson (1984). Based on word-level sound patterns in the Bantu language Kukuya (Paulian 1974), Hyman (1987) shows that 'prefixal' elements including noun class markers and prepositions must be analyzed as forming part of the same prosodic word as the lexical item that precedes, typically a verb. Klavans (1985) reports an analogous S-P mismatching case from Nganhcara (Smith and Johnson 1979). Certain pre-nominal determiners in Shanghai Chinese must also group in a prosodic word with the verb that precedes (Selkirk and Shen 1990). Parker (1999) brings to light similar facts from Chamicuro. And see Werle (2009) on the enclitic, proclitic vs. free ('stray') status of functional heads in Bosnian-Serbian-Croatian. It would be interesting to see whether other prosodic structure-sensitive phenomena of sentence phonology in these languages point to a role for the spell-out constraint $\text{MATCHPHRASE}_{\text{LEX}}$ in the language, which, as in Xitsonga, would explain the enclitic status of these functional heads in PO.

Noun class markers positioned before noun stems are pervasive in the Bantu languages, of course, and in many Bantu languages phonological evidence appears to show that the noun class marker is prefixal, in the sense of forming a phonological unit with the noun stem that follows, parallel to what we have seen in Xitsonga with the subject marker and following verb stem (see

Myers 1987, Herbert 1992 for discussion). Given the MSO-PI-PO model of the syntax-phonology interface proposed here, there are a variety of potential sources for the grouping of the functional classifier head with what precedes or follows. The morphosyntax itself could produce a 'prefixal' status for the classifier in MSO, the result of noun-raising comparable to the verb-raising in Bantu. Alternatively, spell-out could conceivably call for the procliticizing incorporation of the c-commanding classifier head into the ω that corresponds to the noun in the phonological input representation PI. And of course the phonology *per se* could be given responsibility for the decidedly non-prefixal status of the classifier in Xitsonga, or in languages like those cited above, where the phonological patterning testifies to the enclitic status of the functional head of the following phrase. As we saw above, this is a natural consequence of the 'stray' status for the classifier that would result from the effect of the constituency interface constraint MATCHPHRASE$_{LEX}$, which gives spell-out only to phrases of MSO that are headed by lexical category items.

3.3 Prosodic structure formation in the spell-out and phonology modules

This section has examined syntax-phonology constituency mismatches seen in the Xitsonga double object construction and in the class marker construction. For both cases, the proposal is that the syntax-phonology interface constraint MATCHPHRASE$_{LEX}$ spells out phrasal constituency in MSO as φ constituency in PI, creating mismatches, and that in the phonology *per se* a ranking of the phonological constraints BINARITY(φ), MAX(φ), DEP(φ) and STRONGSTART(φ) results in mismatches between PI and PO.

From a typological perspective, the theory of constituency spell-out that is advanced here predicts the possibility (but not the necessity) that Match constraints may introduce different classes of mismatches between the constituency of MSO and the constituency of PI, depending on whether it is MatchPhrase$_{LEX}$ or the general MatchPhrase that's at play in a language. And indeed, Elfner (2012, 2015) makes a strong case that in Irish it is the general version of the MatchPhrase constraint—an interface constraint that holds of *all* phrases, whether Lex-headed or not—that determines the φ-structure of PO. Irish is a verb-initial language where the verb precedes the subject and the object of the sentence. In the morphosyntax, the subject and the object are constituents of a same Fnc-headed phrase, i.e., one that does not have a lexical head. Yet, Elfner argues, this Fnc-phrase must be assumed to be spelled out as a φ, in order to account for the predictable distribution of edge tones that appear at the left edge of the subject phrase in PO (and in other contexts as well).[33] So, from the point of view of constituency spell-out, Irish would be a 'MatchPhrase language', one in which *all* phrases of MSO are spelled out as φ, and so no phrasal mismatches would be produced in PI. But mismatches do arise in the phonology *per se* in Irish. In the PO representations of Irish sentences, Elfner (2012, 2015) and Bennett et al. (2016) report, there are no nonbinary φ. At present the facts suggest that in Irish only phonologically-driven MSO-PO mismatches are found, produced by the now-familiar constraint ranking BINARITY(φ) >>MAX(φ) of the phonology *per se*.

It seems, then, that languages may differ in whether it is MatchPhrase in general, or only the more specific MatchPhrase$_{LEX}$, that is responsible for spelling out phrases of MSO. How then

[33] Modern Irish, like Standard British and American English, is a language without lexical tone in which there is a phonologically predictable distribution of nonlexical, non-meaning-bearing tones in declarative pragmatically neutral sentences.

to formally characterize within the MSO-PI-PO model the difference between the grammars of Xitsonga and Irish in the spelling out phrases of MSO as φ in PI? Is it a matter of the relative ranking of these two varieties of Match constraint in the spell-out module of the different languages? Recall that in section 1.2 (and in Kratzer and Selkirk (2020)) it was suggested the ranking of spell-out constraints DephraseGiven >> MatchPhrase has central responsibility for the attested prosody of [G]-marked constituents in PO in Standard American and British English. We suggest here, though, that placing the notation '>>' for constraint ranking between a pair of spell-out constraints should simply be understood as meaning that the higher-ranked constraint takes precedence over the lower-ranked one. A much smaller font size could be used to show this. Our informal proposal, then, is that MatchPhrase >> $\text{MatchPhrase}_{\text{LEX}}$ in Irish, and that $\text{MatchPhrase}_{\text{LEX}}$ >> MatchPhrase in Xitsonga.

It is important to keep in mind that all spell-out constraints of the MSO-PI-PO model we assume involve only a relation between MSO and PI. As a result, the role and nature of constraint ranking in the spell-out module is likely to be very different from that in the phonology *per se.* Within the spell-out module the spell-out constraints generate an output representation (PI) whose primitive elements, which are phonological in kind, are entirely distinct from those of the input representation (MSO), which are morphosyntactic in kind. Moreover, as currently conceived, there are no constraints in the spell-out module that hold exclusively of PI. In other words, as a grammatical system, the spell-out module has properties that are quite different from those of the phonology module, where both PI and PO representations are strictly phonological in character, where there is a set of faithfulness constraints on the relation between PI and PO representations, and where a set of phonological markedness constraints on PO embodies the ideal of phonological well-formedness.

4. Predictions of the MSO-PI-PO model of the syntax-phonology relation

Let's now put at the forefront the MSO-PI-PO model of the 'P-side' of grammar that we have been assuming: (i) the phonology *per se* defines a phonological output representation (PO) on the basis of a phonological input representation (PI), (ii) a generative syntax defines a (morpho)syntactic output representation (MSO), and (iii) the sole interface between phonology and the morphosyntax is between MSO and PI. Given this architecture, the only potential impact of morphosyntactic constituent structure in MSO on the prosodic constituent structure of PO is through the intermediary representation PI. The MSO-PI-PO model allows no direct impact of the output of the morphosyntax module on the output of the phonology module.

The set of syntax-phonology interface constraints that collectively spell out MSO as the phonological representation PI necessarily exploits a mixed syntactic and phonological vocabulary. The interacting constraints of the phonology *per se*, however, exploit only the vocabulary of phonological representation, whether they are markedness constraints on PO or faithfulness constraints on the relation between the phonological representations PI and PO. The ordering of the spell-out and phonology modules of the P-side derivation, and the limits on the nature of the representational information available at each step, provide a quite restrictive theory of these mismatches, and at the same time open up the possibility of genuinely new insights into the MSO-PO relation.

In committing to the spell-out of morphosyntactic constituent structure as prosodic constituent structure in PI, the MSO-PI-PO model commits to other predictions concerning the relation between syntax and phonology, in particular concerning the linearization of words and

phrases. In this final brief section we sketch out a research hypothesis involving linearization of the words and phrases of a sentence that is informed by the particular assumptions of the MSO-PI-PO model, and could lead to new understanding of this aspect of the syntax-phonology relation. The hypothesis is that linearization phenomena in any language may be diverse in kind, determined respectively by the morphosyntax, by spell-out, or by phonology—and produced in that order in the derivation. Linearization in the morphosyntax would be based on purely morphosyntactic principles. In spell-out and in the phonology *per se*, as well, the representational properties that play a role in defining possible linearizations are those available to the constraint types of the respective modules. Linearization in spell-out would involve 're-ordering' of formatives or constituents with specific morphosyntactic properties into new positions that are characterized in terms of the phonological properties of PI. Linearization within the phonology module would be free of any direct reference to morphosyntactic properties. Rather, it would involve a change between 'word order' in PI and 'word order' in PO that would result from the ranking of purely phonological markedness and faithfulness constraints in the phonology *per se.*

The existence of a class of phonological movement phenomena was proposed in Agbayani and Golston (2016), and in other earlier work of theirs.[34] The conclusion to the Agbayani and Golston (2016) paper reads:

> "We have argued that the classical phenomenon of hyperbaton in Latin is a case of phonological movement, a species of movement that is strictly prosodic in that it moves prosodic constituents to the edges of other prosodic constituents. Because it applies in the

[34] Agbayani and Golston (2010ab), Golston (1995).

> phonological component of the grammar, it is sensitive to the Obligatory Contour Principle but insensitive to syntactic island conditions. Though it appears to be motivated by the need to mark discourse prominence effects, it also has the property of being vacuous with respect to LF effects in relation to binding and scope. Thus, it is movement that is entirely syntax-free. The possibility of phonological movement opens up an intriguing parallelism between phonology and syntax that has, until recently, been largely unexplored. The diagnostics for phonological movement presented in this paper may be employed to uncover other cases of post-syntactic phonological movement crosslinguistically."

Another case of what appears to be phonological movement has been pointed out by Shih and Zuraw (2017). They report that in Tagalog the normal A-N order of noun and adjective in a phrase is not found when the final segment of the A and the initial segment of the N are both nasal consonants. Rather the opposite order N-A is attested, where the nasal-nasal sequence is avoided. This change in word order could plausibly be understood as a variety of metathesis, a phonological phenomenon in which the order of two phonological units in the input PI is inverted in PO, in order to avoid a violation of a phonological markedness constraint in the output PO representation (Buckley 2011). In Tagalog, the constraint ranking of the phonology would select the word order $(_{n}\omega\ \omega_{n})_{\varphi}$ as the optimal representation in PO for the PI representation $(\omega_{n}\ {}_{n}\omega)_{\varphi}$ if the phonological markedness OCP(nasal) ruling out the *n-n* sequence were to outrank a PI-PO faithfulness constraint which disallows a phonological change in order (metathesis) of two ω daughter constituents of a φ.

The case of phonological pronoun postposing in Modern Irish is treated in detail in Bennett et al. (2016). They argue that a version of the prosodic markedness constraint STRONGSTART is

responsible for the postposing of a weak object pronoun from its original position at the left edge of a φ corresponding to the VP to an enclitic position at the right edge of a lower φ corresponding to a complement phrase within the VP. Beyond this case, it's long been noted that the word order of a sentence may be influenced by 'rhythmic' factors. Presumably, prosodic markedness constraints on the relative positioning of phrasal prominence(s) (and their interaction with prosodic faithfulness constraints) have an ample role in determining word order in PO that departs from PI.

As these examples illustrate, the hypothesis that there are phonology-driven post-syntactic linearization phenomena is a very promising one. What's key in the cases briefly reviewed here is the role for the prosodic constituent structure of phonological representation in delimiting what are possible word order changes between MSO and PO. It seems likely that these particular cases can be seen as phenomena of the phonology *per se*, in which the phonological constraint system introduces changes between the linearization and prosodic structure of PI and PO.

What also needs to be further examined is whether changes in linearization might be introduced by spell-out, between MSO and PI. Because the Match constraints of spell-out produce a prosodic constituent structure in PI, the spell-out of morphosyntactically designated formatives or constituents in MSO could potentially target prosodic structural configurations in PI that would constitute the phonological expression, or exponence, of the designated formative or constituent. The existence of the spell-out constraint DephraseGiven in Standard American and British English suggests that spell-out can be responsible for constituency mismatches between MSO and PI. In another language spell-out might be responsible for mismatches in word order. A [G]-marked constituent of MSO could potentially be spelled out in initial position of some φ that might dominate it in PI, for example.

Clearly, there's much to be learned about phonologically characterizable word order changes in language. The MSO-PI-PO model should be a useful tool in this enterprise.

References

Agbayani, Brian and Chris Golston (2010a) Phonological Movement in Classical Greek. *Language* 86.1: 133-167.

Agbayani, Brian and Chris Golston (2010b) Second-position is first-position: Wackernagel's Law and the role of clausal conjunction. *Indogermanische Forschungen* 115: 1-21.

Agbayani, Brian and Chris Golston (2016) Phonological Constituents and their Movement in Latin. *Phonology* 33: 1-42.

Anderson, Stephen (1984) Kwakwala Syntax and the Government-binding Theory. *Syntax and Semantics* 16: 21-75.

Archangeli, Diana and Douglas Pulleyblank (1994) *Grounded phonology*. Cambridge, MA: MIT Press.

Baker, Mark (1989) Object Sharing and Projection in Serial Verb Constructions, *Linguistic Inquiry* 20, 513-543.

Baker, Mark C., Ken Safir, and Justine Sikuku (2012) Sources of (A)symmetry in Bantu Double Object Constructions. In Nathan Arnett and Ryan Bennett (eds.) *Proceedings of the 30th West Coast Conference on Formal Linguistics*. 54-64.

Bennett, Ryan and Emily Elfner (2019) The syntax–prosody interface. *Annual Review of Linguistics* 2019, 5:1, 151-171.

Bennett, Ryan, Emily Elfner, and James McCloskey (2016) Lightest to the Right: An Apparently Anomalous Displacement in Irish. *Linguistic Inquiry* 47:2, 169-234.

Beuchat, P.-D. [Cole] (1962) Additional notes on the tonomorphology of the Tsonga noun. *African studies* 21: 105-122.

Boas, Franz (1947). Edited by Helene Boas Yampolsky, with the collaboration of Zellig S. Harris. 'Kwakiutl Grammar (With a Glossary of the Suffixes)'. *Transactions of the American Philosophical Society*. American Philosophical Society. 37 (3): 201–377. doi:10.2307/1005538. JSTOR 1005538.

Bobaljik, Jonathan D. (2017) Distributed morphology. *Oxford Research Encyclopedia of Linguistics*. https://doi.org/10.1093/acrefore/9780199384655.013.131

Buckley, Eugene (2011) Metathesis. In *The Blackwell companion to phonology. Vol. 3, Phonological processes*. Edited by Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume, and Keren Rice, 1380–1407. Blackwell Companions to Linguistics Series. Malden, MA: Wiley-Blackwell.

Carstens, Vicki (2005) Agree and EPP in Bantu. *Natural Language and Linguistics Theory* 23: 219-279.

Cassimjee, Farida and Charles Wayne Kisseberth (1998) Optimal domains theory and Bantu tonology: a case study from Isixhosa and Shingazidja. In Hyman, Larry Michael and Kisseberth, Charles Wayne (eds.), *Theoretical aspects of Bantu tone*, Stanford: Center for the Study of Language and Information (CSLI), Stanford University. 33-132.

Chomsky, Noam and Morris Halle (1968) *The sound pattern of English*. Harper and Row.

Cole-Beuchat, P.-D. (1959) Tonomorphology of the Tsonga noun. *African Studies* 18: 133-145.

Elfner, Emily (2012) 'Syntax-Prosody Interactions in Irish'. *Open Access Dissertations*. 505. https://doi.org/10.7275/3545-6n54
https://scholarworks.umass.edu/open_access_dissertations/505

Elfner, Emily (2015) Recursion in prosodic phrasing: Evidence from Connemara Irish. *Natural Language and Linguistic Theory* 33 (4): 1169-1208.

Elfner, Elfner (2018) The syntax-prosody interface: current theoretical approaches and outstanding questions. Linguistics Vanguard, 1(open-issue).

Elordieta, Gorka and Selkirk, Elisabeth (2018) Notes on prosodic headedness and tone in Tokyo Japanese, Standard English and Northern Bizkaian Basque. In Bennett et al., Hana Bana: *A Festschrift for Armin Mester and Junko Ito*. Retrieved from https://escholarship.org/uc/item/7q03m9qm

Embick, David and Rolf Noyer (2007) Distributed morphology and the syntax/morphology interface. *The Oxford Handbook of Linguistic Interfaces*, 289-324.

Ghini, Marco (1993) Phi-formation in Italian: A new proposal. *Toronto Working Papers in Linguistics* 12. 41–78.

Golston, Chris (1995) Syntax outranks phonology: Evidence from Ancient Greek. *Phonology* 12. 343-368.

Greenberg, Joseph H. (1977) Niger-Congo noun class markers: Prefixes, suffixes, both or neither. In *Studies in African Linguistics*, supplement 7, 97-106. Papers from the 18th Conference on African Linguistics.

Halle, Morris and Alec Marantz (1993) Distributed morphology and the pieces of inflection. In K. Hale and S.J. Keyser eds. *The View from Building 20: Linguistics Essays. in Honor of Sylvain Bromberger*. Cambridge MA: MIT Press. 111-176.

Halle, Morris and Alec Marantz (1994) Some key features of distributed morphology. In A. Carnie, H. Harley, and T. Bures (eds.), *MITWPL 21*. 275-288.

Harley, Heidi and Rolf Noyer (1999) Distributed morphology. *Glot international* 4(4): 3-9.

Hayes, Bruce (1995) *Metrical stress theory*. Chicago: University of Chicago Press.

Herbert, Robert K. (1992) Phonological words and syntactic constituents in XiTsonga. *South African Journal of African Languages* 12(1): 17-24.

Hyde, Brett (2011) 'Extrametricality and nonfinality', in Marc van Oostendorp, Colin Ewen, Beth Hume, and Keren Rice (eds.) *A Companion to Phonology*, 1027-1051. Blackwell.

Hyman, Larry (1987) Prosodic domains in Kukuya. *Natural Language and Linguistic Theory* 5(3): 311-333.

Hyman, Larry M. (2009) Penultimate Lengthening in Bantu: Analysis and Spread. *UC Berkeley Phonology Lab Annual Report* 195-209.

Ito, Junko and Armin Mester (1992/2003) Weak Layering and Word Binarity. In Honma, Takeru, Masao Okazaki, Toshiyuki Tabata and Shin-ichi Tanaka (eds.), *A New Century of Phonology and Phonological Theory. A Festschrift for Professor Shosuke Haraguchi on the Occasion of His Sixtieth Birthday*. 26-65.

Ito, Junko and Armin Mester. (1999) Realignment. In R. Kager, H. v.d. Hulst, and W.Zonneveld, eds. *The Prosody-Morphology Interface*. 188-217. Cambridge: Cambridge University Press.

Ito, Junko, and Armin Mester. (2020) Match Theory and Prosodic Wellformedness Constraints. In Zhang, Hongming, and Youyong Qian, eds. *Prosodic Studies. Challenges and Prospects.* London and New York: Routledge. pp. 252-274.

Kalivoda, Nicholas (2018) Syntax-Prosody Mismatches in Optimality Theory. UCSC PhD Dissertation.

Kisseberth, Charles W. (1994) On domains. In Jennifer Cole and Charles Kisseberth (eds.), *Perspectives in Phonology*. 133-166. Stanford: CSLI,

Klavans, Judith L. (1985) The Independence of Syntax and Phonology in Cliticization. *Language* 61: 95-120.

Kratzer, Angelika and Elisabeth Selkirk (2020) Deconstructing information structure. *Glossa: A Journal of General Linguistics*, *5*(1), 113. DOI: http://doi.org/10.5334/gjgl.968

Ladd, Robert D. (1980) *The Structure of Intonational Meaning: Evidence from English.* Bloomington, IN: Indiana University Press.

Leben, William (1973) *Suprasegmental Phonology*. PhD dissertation, MIT. Distributed by Indiana University Linguistics Club.

Lee, Seunghun J. (2014) Domains of H tone spreading and the noun class prefix in Xitsonga, *Southern African Linguistics and Applied Language Studies* 32(1): 21-34, DOI: 10.2989/16073614.2014.925215

Lee, Seunghun J. and Elisabeth Selkirk (2014) Structural restrictions on H tone spread in Xitsonga. ACAL 45. University of Kansas. Lawrence, KS. April 17-19, 2014

McCarthy, John J. (2001) *A Thematic Guide to Optimality Theory*. Cambridge University Press.

McCarthy, John J. and Alan Prince (1995) Faithfulness and reduplicative identity. In Jill Beckman, Laura Walsh Dickey and Suzanne Urbanczyk (eds.), *Papers in Optimality Theory*. *University of Massachusetts Occasional Papers* 18. 249–384. Amherst, Mass.: Graduate Linguistic Student Association. [Rutgers Optimality Archive 60, http://roa.rutgers.edu]

McCarthy, John J. and Alan Prince (1996) Prosodic morphology. In John Goldsmith (ed.) *The Handbook of Phonological Theory.* Blackwell Publishing.

McCarthy, John J. and Alan Prince (1999) Faithfulness and identity in prosodic morphology. In René Kager, Harry van der Hulst and Wim Zonneveld (eds.), *The Prosody-Morphology Interface*. 218-309. Cambridge University Press.

Myers, Scott P. (1987) Tone and the structure of words in Shona. Doctoral dissertation, University of Massachusetts, Amherst. https://scholarworks.umass.edu/dissertations/AAI8805958

Myers, Scott (1997) OCP Effects in Optimality Theory. *Natural Language and Linguistic Theory* 15, 847–892 https://doi.org/10.1023/A:1005875608905

Mufwene, Salikoko (1980) Bantu noun class prefixes: Inflectional or derivational? In J. Kreiman and O.E. Ojeda (eds.) CLS Working Papers, 246-258. Chicago.

Pak, Marjorie (2008) *The postsyntactic derivation and its phonological reflexes*. Doctoral dissertation, University of Pennsylvania. Downloaded 8/3/2020 from: http://www.marjoriepak.com/

Parker, Stephen (1999) On the behavior of definite articles in Chamikuro. *Language* 75. 552-562.

Paulian, Christiane (1974) *Le Kukuya: Langue teke du Congo,* S.E.L.A.F., Paris.

Prince, Alan and Paul Smolensky (1993/2004) Optimality Theory: Constraint Interaction in Generative Grammar. Wiley.

Selkirk, Elisabeth (1980) Prosodic domains in phonology: Sanskrit revisited. In Mark Aronoff and Mary-Louise Kean (eds.) *Juncture*. Saratoga, CA: Anma Libri.

Selkirk, Elisabeth (1996) The prosodic structure of function words. In James L. Morgan and Katherine Demuth (eds.) *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*, 187-214. Psychology Press.

Selkirk, Elisabeth (2009) On clause and intonational phrase in Japanese: The syntactic grounding of prosodic constituent structure, *Gengo Kenkyu (Journal of the Linguistics Society of Japan)*, 136, 1-39. Special Issue on Linguistic Interfaces ed. by Haruo Kubozono), 2009.

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

Selkirk, Elisabeth (2017) Syntactic Constituency Spell-Out through MATCH Constraints, talk delivered at SPOT Workshop, UC Santa Cruz November 18, 2017.

Selkirk, Elisabeth and Lee, Seunghun J. (2015) Constituency in sentence phonology: An introduction. *Phonology* 32(1): 1-18.

Selkirk, Elisabeth and Tong Shen. (1990) Prosodic domains in Shanghai Chinese. In Sharon Inkelas and Draga Zec (eds.) *The phonology-syntax connection*. 313-338. The University of Chicago Press.

Shih, Stephanie and Kie Zuraw (2017) Phonological conditions on variable adjective-noun word order in Tagalog. *Language: Phonological Data and Analysis* 93: e317-e352

Smith, Jennifer L. (2001) Lexical category and phonological contrast. In *Papers in experimental and theoretical linguistics 6: Workshop on the Lexicon in Phonetics and Phonology*. Edmonton: University of Alberta. 61-72.

Smith, Ian and Steve Johnson (1979) The cliticization of case-markers in Nganhcara. *Paper presented at the Australian Linguistic Society, Newcastle*.

van der Wal, Jenneke (2015) Bantu Syntax. *Oxford Handbooks Online*. DOI: 10.1093/oxfordhb/9780199935345.013.50

Volk, Erez (2011) Mijikenda tonology. Tel Aviv University. PhD Dissertation.

Welmers, W. (1973) African language structures. Berkeley: University of California Press.

Werle, Adam (2009) Word, phrase, and clitic prosody in Bosnian, Serbian, and Croatian. PhD Dissertation. UMass Amherst.

Yip, Moira (1988) The Obligatory Contour Principle and phonological rules: A loss of identity, *Linguistic Inquiry* 19(1): 65–100.

Yip, Moira (2002) *Tone*. Cambridge University Press.