

The union spell-out mechanism

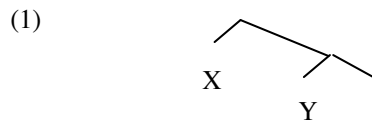
Peter Kinyua Muriungi
*Chuka University College - Kenya*¹

Abstract

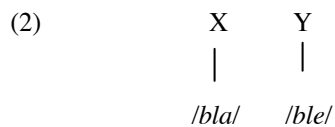
This paper argues that grammar has a union spell-out mechanism that dictates that contiguous heads be spelled out by a single morpheme/phrase if there exists a morpheme/phrase in the lexicon with a collection of the features of the contiguous heads.

1. The nature of union

Suppose there are two contiguous features X and Y in the verbal functional sequence (Henceforth fseq, following Starke (2001)), (1).



Suppose these features can sometimes get realized separately as *bla* (for X) and *ble* (for Y) in the overt syntax, (2).

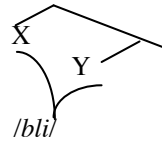


Suppose also that the two features X and Y can be realized by a single lexical item *bli*.

¹ I would like to thank Peter Svenonius, and Michal Starke for their comments on this paper. All the errors in this paper however are my responsibility.

THE UNION SPELL-OUT MECHANISM

(3)



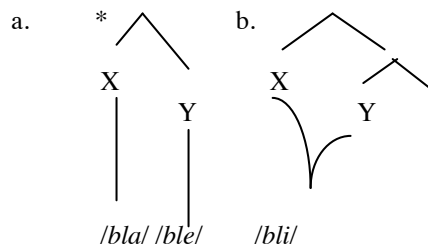
I argue in this paper that grammar might provide a mechanism where the spell-out system in (3), blocks the spell-out system in (2). The cause of this blockage might lie with a spell-out mechanism — *the union mechanism*. This mechanism is informally stated as follows:

(4) *The Union Mechanism (to be revised)*

Spell out of contiguous heads with a single morpheme wins over spell out of such heads by separate morphemes if there exists a single morpheme in the lexicon with a superset of the features of the contiguous heads.

The lexical item *bli* has a superset of the features that both *bla* and *ble* has — the features X and Y. It therefore wins over and consequently blocks insertion of *bla* (which only spells out X) and *ble* (which only spells out Y) (cf. (5-a) and (5-b)).

(5)



The union mechanism has its antecedent a long debate on blocking effects in grammar for which see e.g. Aronoff (1976), Kiparsky (1973), Poser (1992), and more recently Embick and Marantz (2008).

2. The Kiitharaka data that unionizes

To concretize the working of the union mechanism, consider the following three verbs from Kiitharaka (Bantu, SVO, Kenyan), paying attention to the three morphemes *îk*, *ûr* and *ûk* (all in bold):

(6)

- a. kun- **îk** -a
'X covers Y'
- b. kun- **ûr** -a
'X uncovers Y'
- c. kun- **ûk** -a
'Y gets uncovered'

Following our intuitions about the meaning changes in the verbs in (6), it plausible to claim that the morpheme *îk* conveys transitivity, the morpheme *ûr* transitivity and reversiveness and the morpheme *ûk* just mere reversiveness. What is crucial to note here is that there is a single morpheme *ûr* that can spell out two features at a time, the features for transitivity and reversiveness.

Consider in this light, the verb *kuama* ‘bend by self’ which has an obligatory suffix *am*, a suffix that usually conveys a meaning of being in some position or state. This verb forms a simple reversible (one that is not transitive) with the reversible morpheme *ûk*, (7).

- (7) kuam- **ûk**-a
 ‘Y reverts from a bending position by self’

In order to transitive, this verb uses *i*, a transitive causative morpheme, (8).

- (8) kuam- **i** -a
 ‘X bends Y’

Let us now turn to the formation of transitives that are also reversives. How does the verb *kuama* behave? There are two variants we would expect given the inventory of morphemes in Kîtharaka. One, the transitive reversible could be formed step by step by addition of the simple reversible morpheme *ûk* and the transitive causative morpheme *i*. Alternatively the transitive reversible could be formed by use of the single morpheme *ûr* that encodes both transitivity and reversiveness. The data facts show that this latter strategy is the right one, (9).

- (9) a. *kuam-**ûk-i** -a
 ‘X takes Y from a bending position’
 b. kuam- **ûr**- a
 ‘X takes Y from a bending position’

The fact that a single morpheme *ûr* blocks a combination of *ûk* and *i* is puzzling. But the puzzle remains only when we fail to recognize that grammar has union.

The verb *kuama* ‘bend’ is not the only one that behaves in this way. There are a number of other verbs that have the *am* suffix with the same properties. These verbs form transitives with the *i* causative morpheme. Transitives that are also reversible are formed with a single morpheme *ûr*, but never with a combination of the transitive causative morpheme *i* and reversible *ûk*. I list the paradigms with these verbs below.

- (10) a. maama
 ‘lie’
 b. ma-am-**i**-a
 ‘X lies Y somewhere’
 c. ma-am-**ûr**-a
 ‘X takes Y from a lying position’

THE UNION SPELL-OUT MECHANISM

- d. *ma-am-**ûk-i**-a
- (11) a. rûûng-am-a
'stand'
b. rûûng-am-**i**-a
'X helps Y to stand'
c. rûûng-am-**ûr**-a
'X takes Y from a standing position'
d. *rûûng-am-**ûk-i**-a
- (12) a. ind-am-a
'stoop'
b. ind-am-**i**-a
'X makes Y stoop'
c. ind-am-**ûr**-a
'X takes Y from a stooping position'
d. *ind-am-**ûk-i**-a
- (13) a. thend-am-a
'crooked/bent'
b. thend-am-**i**-a
'X makes Y crooked/bent'
c. thend-am-**ûr**-a
'X takes Y from a crooked/bent position'
d. *thend-am- **ûk-i** -a

3. Implementing union – some crucial assumptions

In order for union to survive, grammar must be endowed with a number of other mechanisms which I detail below

3.1 Late insertion

The job of syntax is to concatenate features, for example [transitive] and [reversible].² It is only after syntax is done with its job that the syntactic features are given phonological reality in a mechanism called *spell-out*. Replacement of nodes with phonology therefore occurs late, at a juncture different from when mere features with no phonology are combined. This is an architectural view that was entertained in some early work in generative semantics (McCawley (1968)) and taken to be standard in much work in distributed morphology (Halle and Marantz (1993), Marantz (1997a), Marantz (1997b), Harley and Noyer (1999)). Coming back to our context, *ûr*, the transitive reversible morpheme, blocks *ûk+i* at the point of spell-out. Blockage does not occur at the point of syntactic bundling –

² In syntax, merge is the concatenating operation (cf. Chomsky (1995)).

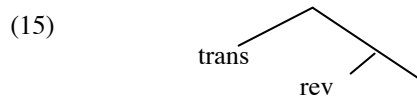
– syntax does not manipulate things with phonology, for example *ûr* and *ûk+i*. All syntax sees are the features [transitive] and [reversive].

3.2 Hierarchy in syntax and in the lexicon

Syntax manipulates features, and these features are replaced (during spell-out) by a morpheme from the lexicon. The lexicon therefore is the storehouse of the items with phonological content which replace syntactic features. Consider our context again. We have two features in the syntax – [transitive] and [reversive]. Furthermore we know that these features are not a mere bundle - they also express a scope asymmetry. If, as usually assumed, scope asymmetries reflect structural asymmetries, then [transitive] and [reversive] must merge at different heights of the syntactic tree. Let us establish the scope of [transitive] and [reversive]. If [transitive] scopes over [reversive], we expect a reading where an action can be undone, without necessarily having been caused before.

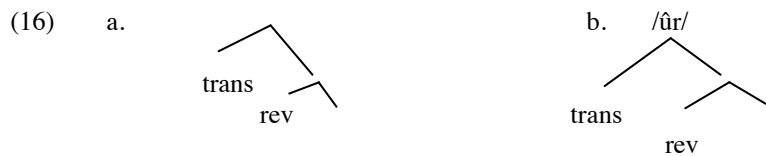
- (14) Maria n-a-bung- *ûr*-ir-e cati
 Maria f-sal-button-rev-perf-fv 9.shirt
 ‘Maria did an event that led the shirt to be in an unbuttoned state (the shirt was made by the machine buttoned).’

Since the reading in (14) is possible, we know that [transitive] scopes over [reversive] semantically. Syntactically, [transitive] must merge above [reversive]. We therefore have the hierarchy of functional projections in (15):



According to union, a single lexical item *ûr* can replace, at spell-out, the (whole) syntactic tree in (15). Given that the replaced node is syntactically complex (with at least two maximal projections - transitiveP and reversiveP), it is also highly likely that the morpheme in the lexicon is also syntactically hierarchically complex, with a structure exactly like the one in the syntax, except that in addition, it has some phonology (Starke class lectures, Abels and Muriungi (2008)).

I illustrate the structure in the syntax and in the lexicon in (16) ((16-a) for syntax, (16-b) for lexicon).



3.3 The superset principle

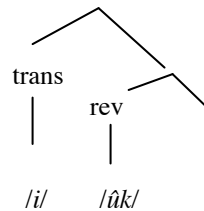
Since the trees in the lexicon and trees in the syntax are substantially similar (modulo the phonology), then lexical insertion reduces to matching trees in the lexicon, with those in the syntax, under a spell-out principle, the superset principle (cf. Starke, class lectures; Caha (2007)):

(17) *The Superset Principle*

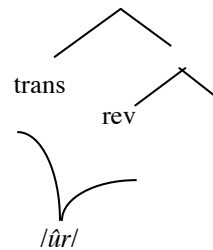
Insert a tree in the lexicon for a (sub) tree in the syntax, if the features of the tree in the lexicon are a superset of the features of the (sub) tree in the syntax. When lexical items compete for insertion, insert the tree with the least unused features. Do not insert a tree from the lexicon if it does not contain (a) feature(s) in the syntax.

Turning to our context, given the presence of the features [transitive] and [reversible] in the syntax, a single lexical item *ûr* can be inserted since it has both the features [transitive] and [reversible]. Note however, given the nature of the superset principle, insertion of *ûr* for [transitive] and [reversible] will be blocked, given the presence of more specific morphemes: the morpheme *i*-causative that only has a [transitive] feature, and the morpheme *ûk* which has only the feature [reversible]. All things being equal therefore, and given the nature of the superset principle, spell-out by a combination of *uk+i*, (18-a) should win over spell-out by a single morpheme – *ûr* (18-b).

(18) a.



b.



3.4 Ironing the conflict between union and superset

Given the superset principle (whose elsewhere condition requires more specific morphemes to win over morphemes with more junk (i.e. unused features), the effects of union (which prefers spell-out of contiguous morphemes with a single morpheme) will never be felt. This is because at the point when [reversible] is merged, the morpheme with less junk (*ûk*), will inserted. Similarly, when [transitive] is merged, the morpheme with less junk (*i*-causative) will be inserted. Insertion of *ûr*, an amalgam of [transitive] and [reversible] will be disallowed. In order for the effects of union to be felt, we have to require that the union mechanism applies before the superset principle.

- (19) Apply union before minimize junk (the elsewhere of the superset principle)

Borrowing from Michal Starke, we can claim that union type of spell-out overrides a principle requiring one to minimize junk.

3.5 Phrasal spell-out

The morpheme *ûr* replaces two features ([transitive] and [reversive]) that occupy different heights of the syntactic tree and therefore project phrasal categories of different types— transitive a transitiveP, and reversive a reversiveP. Evidently therefore, the spell-out of these features with a single morpheme must be a case of phrase spell-out— a case of a morpheme targeting a non-terminal (McCawley (1968), Weerman and Evers-Vermeul (2002), Neeleman and Szendrői (2007), Caha (2007), Abels and Muriungi (2008), Starke class lectures). We therefore differ from those accounts that only limit spell-out to terminal nodes (for a recent account of only terminal node spell-out, see Embick and Marantz (2008)).

3.6 Spell-out for constituents

Union requires that the features that get a single exponent are contiguous in fseq. We can translate this to mean that the features that get a single exponent are a constituent. The indication that it is spell-out of constituents that is relevant is reinforced by the fact that two (or more features) are spelled out by a single morpheme. The nature of this spell-out brings to mind an old test for constituency —only constituents can be substituted. A pronoun substitutes a noun phrase— a constituent. Expanding this view of constituency somewhat, a single morpheme will replace/substitute a sequence of syntactic features because those features are a constituent.

4. Apparent counterexamples to union

We have seen that given two features X and Y that are contiguous in fseq, if these features can get a single exponent at spell-out, then they do – courtesy of the union principle. There are some cases however that seem to be counterexamples to this general trend.

Consider the verb *cuura* ‘hang’. As with some of the verbs we have encountered before, this verb transitivizes with the *i* causative morpheme.

- (20) *cuur- i -a*
‘X hangs Y’

Unlike the verbs we have encountered before however, *cuura* forms the transitive reversive step by step by use of the mere reversive morpheme *ûk* and the simple transitive *i*, (21-a). It cannot use a single morpheme *ûr* that is both transitive and reversive, (21-b).

THE UNION SPELL-OUT MECHANISM

- (21) a. *cuur-ûk-i* -a
 ‘X takes Y from a hanging position’
 b. **cuur-ûr* -a

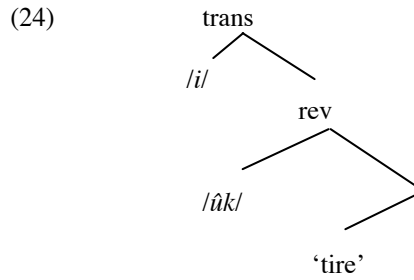
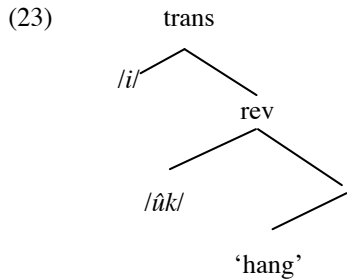
There is another verb *noga* ‘be tired’ that behaves like *cuura*. It forms the transitive with *i*, (22-a), and the reversive transitive with *ûk+i* (22-b), not with *ûr* (22-c).³

- (22) a. *nog-i* -a
 ‘X tires Y’
 b. *nog-ûk-i* -a
 ‘X does work to rest Y’
 c. **nog-or* -a

The data such as in (21) and (22) seem to argue that there is no union mechanism, a mechanism that wins over minimize junk.

I will argue below that we still need the union mechanism, and therefore the requirement for union to take precedence over minimize junk. We have seen that both syntax and the lexicon are structured. The fact that the lexicon is structured opens a possibility for there being linguistic items stored with a complex structure (traditionally called idioms).

Turning to the two exceptions, suppose that these verbs are stored in the lexicon, with the transitive and reversive projections having the phonology of *i* (the causative) and *ûk* (the bare reversive) respectively, i.e. these verbs are idiomatic with the reversive and transitive projection.



Then when we build a syntax, where the verb root is *noga* (‘tire’) or *cuura* (‘hang’), at spell-out, transitive and reversive will be realized as *i* and *ûk* respectively, not as a single morpheme *ûr* (which is both transitive and reversive). The reason for this

³ The morpheme *or* is a phonological variant of the transitive reversive morpheme. The transitive reversive morpheme occurs as *or* when there is an immediately preceding *o*, and as *ûr* elsewhere.

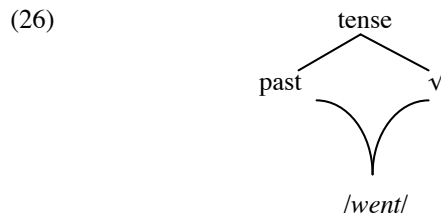
spell-out pattern is that the lexicon has complex idioms, made up of the respective roots+rev+trans, as shown in (23) and (24). Thus, a stored non-compositional form will win over other compositional competitors at spell-out. This type of blockage might itself follow from the nature of union:

- (25) *The Union mechanism (final version)*
 Spell-out of contiguous heads with a single morpheme/phrase wins over spell-out of such heads by separate morphemes if there exists a single morpheme/phrase in the lexicon with a superset of the features of the contiguous heads.

5. Beyond Kiitharaka

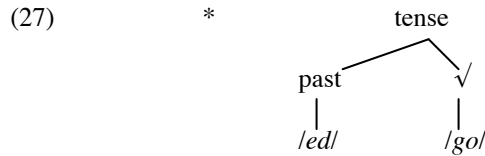
5.1 Irregular past tense in English

It might be that union is also responsible for a number of irregular formations that we see in languages like English (cf. also Starke, class lectures for related argumentation). Just to illustrate, we know independently that the English lexicon has the verb *go* (We have sentences like *I want to go.*). We also know that the English lexicon has the past tense morpheme *-ed* (and its various allomorphs). The question then is why does *went*, block *goed* when it comes to formation of the past tense? This is the effect of union– the verb *went* spells out both the root, and the past tense nodes, as shown, (26), and therefore blocks the compositional spell-out in (27) (Starke, class lectures).⁴



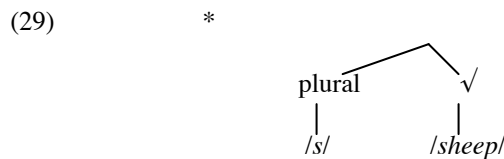
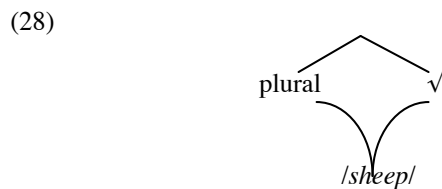
⁴ In the early stages of language acquisition, children overgeneralize the rule of past tense formation and produce past forms such as *holded* (cf. Cazden (1972)). We could hypothesize that at this stage, the children have not established the correct ordering of the superset and the union principles.

THE UNION SPELL-OUT MECHANISM

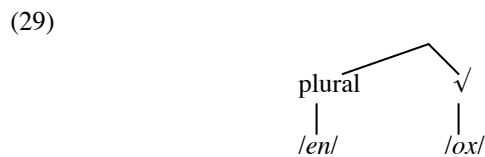


5.2 Irregular plurals in English

We can extend the same story to irregular plurals in English. *Sheep*, (28) blocks *sheeps*, (29), because there is a complex item *sheep* stored in the lexicon, spanning both the features of the root *sheep*, as well as the plural node (Starke, class lectures).



Oxen, will block *oxe+s* because already in the lexicon, there is a complex item *oxen*, and the root *ox*, in the context of plural will always be pronounced as *oxen*, by union.

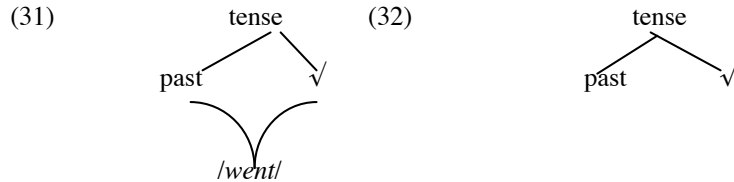


6. Why union: some syntax-brewed speculations

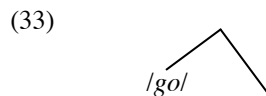
It might be worth the while asking why grammar would allow for a spell-out mechanism such as union, alongside another mechanism, the superset principle. I will suggest two reasons here, both of them very speculative and therefore very tentative.

We have entertained the idea that the lexicon as well the syntax are hierarchically ordered. This means we can have two parallel structures in the syntax

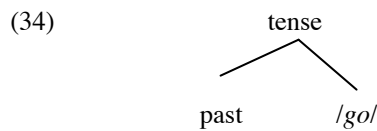
and the lexicon, the two structures differing only in there being some phonology in the lexicon. Consider for example what the structures would be for the verb *went* in English, (31), for the lexicon, and (32) for the syntax.



Recall now what it would require to lexicalize the tree in (32) with the tree in (31), given the superset and union principles. We just take the tree in (32) and insert it/match with the tree in (31). We do not need any movement in the syntax to create a constituent that is like the one in the lexicon – the constituent in the syntax is already like the one in the lexicon. Lexicalization can therefore take place *in situ* – by simple matching. In contrast, consider what would be required if the root and the past tense node were to be lexicalized by two different morphemes, the root by *go* and past tense by *-ed*. First, since the root, by standard accounts is the bottommost element in the tree, it would be spelled out without any problem by *go*, as shown in (33). The root already is a constituent on its own.⁵



A variety of studies on fseq reveal that tense and other suffixes merge above the verb (cf. Baker (1985), Cinque (1999), Julien (2002), Muriungi (2008)). Thus past will merge above the projection with *go*, (34).



Suppose we want to lexicalize the past tense node with *-ed*. Given our tree and the assumption that spell-out is only possible for constituents, this is not possible, since there is no constituent made up of only the tense projection. The tense projection also dominates the projection with the root. In order to create this constituent with only tense, we will be required to move the projection with *go*, to a position above

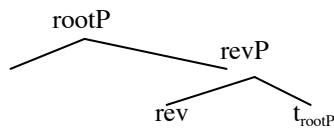
⁵ We are providing a branching phrase marker for the verb *go*, because we take verbs to be internally complex, an idea expounded in greater length in Ramchand's work on verb meaning and the lexicon (Ramchand (2008)).

THE UNION SPELL-OUT MECHANISM

tense as shown in (35).

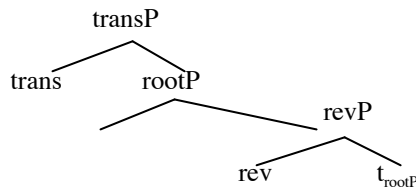
Let us come to the second type of speculation why union consistent spell-out could be preferred. Consider again the differences that could accrue from spell-out of transitive and reversive by using a single morpheme, *ûk* or two morphemes, transitive *i* and reversive *ûk*. Let us start with the structure where the two nodes are spelled out by different morphemes. Assuming that spell-out is only possible for constituents, we will require some movements to create the right configurations. First, in order to create a constituent with only the reversive, we will have to move the root above the reversive, (36).

(36)



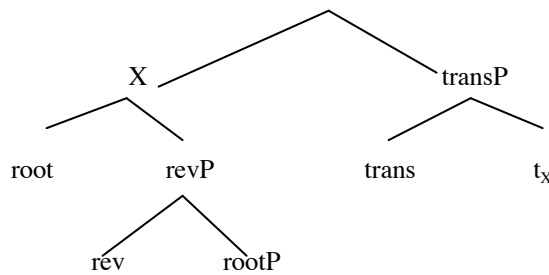
In this structure, revP is a constituent and can be spelled out by *ûk*. Next we can merge the transitive above the landing site of the root, since transitive scopes above reversive, (37).

(36)



In order to create a constituent with only transitive phrase for spell-out purposes, we will have to move the constituent containing both the root and [reversive], above the transitive projection, as shown in (38).

(38)

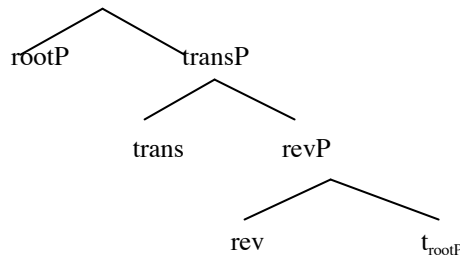


Given the structure above, [transitive] can be spelled out– it is a constituent. What is important to note is that when nodes are spelled out by different morphemes, the

moved chunk increases in size with each step of movement. We will first move the root, then the root+rev, and so on.

Consider however the context where the two nodes, transitive and reversive, are spelled out by the same morpheme. In this scenario, the root moves cyclically, past the reversive and the transitive, (39).

(39)



In other words what is moved does not increase in size with each step of movement.⁶ It might be that a cyclic type of movement is less costly than a roll-up type of movement given that in cyclic movement, the moved constituent does not increase in size. If this is true, then syntax will first exploit a cyclic type of movement, and then check whether there is a single morpheme to spell out the cycled nodes. If there is none, then syntax will resort to the roll-up type of movement. Since cyclic movement precedes roll-up, then the effects of union will be felt. Note that the fact that cyclic movement will be considered before roll-up does not lead to the expectation that there should be more cyclic than roll-up movement in morphology – the lexicon is the determinant!

7. Summary

This paper argues that there might be a union spell-out principle alongside the superset spell-out principle that requires that spell-out of contiguous nodes with a single morpheme/phrase takes precedence over spell-out of such nodes by different morphemes/phrases. I have speculated that this requirement might follow from there being some “in situ” type of spell-out (no re-organization required for spell-out), and movement targeting a small constituent (in the context of cyclic movement).

References

- Abels, Klaus, and Peter Muriungi. 2008. The focus marker in Kĩtharaka: Syntax and semantics. *Lingua* 118:687–731.
- Aronoff, Mark. 1976. *Word Formation in Generative Grammar*. Cambridge, MA: MIT Press.

⁶ In fact given cyclic (=non-roll up) movement, the moved constituent can only decrease in size. For example, some material can be stranded on the path of movement, as in the case of *wh*-movement stranding quantifiers in [Spec, CPs] (cf. McCloskey (2000)).

THE UNION SPELL-OUT MECHANISM

- Baker, Mark. 1985. The mirror principle. *Linguistic Inquiry* 16:373.
- Caha, Pavel. 2007. Case movement in PPs. *Tromsø Working Papers in Linguistics* 34:239–299.
- Cazden, Courtney. 1972. *Child Language and Education*. New York: Holt, Rinehart, and Winston.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- Cinque, Guglielmo. 1999. *Adverbs and Functional Heads - a Cross-linguistic Perspective*. New York and Oxford: Oxford University Press.
- Embick, David, and Alec Marantz. 2008. Architecture and blocking. *Linguistic Inquiry* 39:1–53.
- Halle, Morris, and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger*, ed. Kenneth Hale and Samuel Jay Keyser, 111–176. Cambridge, MA: MIT Press.
- Harley, Heidi, and Rolf Noyer. 1999. State-of-the-art article: Distributed morphology. *GLOT International* 4:3–9.
- Julien, Marit. 2002. *Syntactic Heads and Word Formation*. New York: Oxford University Press.
- Kiparsky, Paul. 1973. 'Elsewhere' in phonology. In *A Festschrift for Morris Halle*, ed. Paul Kiparsky and Steven Anderson. New York: Holt, Rinehart and Winston.
- Lindblom, Gerhard. 1914. *Outlines of a Tharaka Grammar: With a list of words and specimens of the language*. Archives D'Etudes Orientales, volume 9. Uppsala: J.-A. Lundel.
- Marantz, Alec. 1997a. 'Cat' as a phrasal idiom: Consequences of late insertion in Distributed Morphology.
- Marantz, Alec. 1997b. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. In *Proceedings of the 21st Annual Penn Linguistics Colloquium*, ed. Alexis Dimitriadis and Laura Siegel, University of Pennsylvania Working Papers in Linguistics, 201–225. Philadelphia: University of Pennsylvania.
- wa Mberia, Kithaka. 1993. *Kitharaka Segmental Morphophonology with Special Reference to the Noun and the Verb*. Doctoral Dissertation, University of Nairobi.
- McCawley, James D. 1968. Lexical insertion in a transformational grammar without deep structures. In *Papers from the 4th regional meeting*, 71–80. Chicago: Chicago Linguistic Society.
- McCloskey, James. 2000. Quantifier float and *wh*-movement in an Irish English. *Linguistic Inquiry* 31:57–84.
- McCloskey, James. 2002. Resumption, successive cyclicity, and the locality of operations. In *Derivation and explanation in the minimalist program*, ed. Samuel David Epstein and T. David Seely, 184–226. Malden, MA and Oxford, UK: Blackwell.
- Muriungi, Peter. 2003. *Wh-questions in Kitharaka*. Master's thesis, University of the Witwatersrand, Johannesburg.
- Muriungi, Peter. 2008. *Phrasal movement inside Bantu verbs: Deriving affix scope and Order in Kĩtharaka*. Doctoral Dissertation, University of Tromsø, Norway.
- Neeleman, Ad, and Kriszta Szendrői. 2007. Radical pro-drop and the morphology of pronouns. *Linguistic Inquiry* 38:671–714.
- Poser, William J. 1992. Blocking of phrasal constructions by lexical items. In *Lexical matters*, ed. Ivan A. Sag and Anna Szabolcsi, 111–30. Stanford, CA: Center for Study of Lang. & Information.
- Ramchand, Gillian. 2008. *Verb Meaning and the Lexicon: A First Phase Syntax*. Cambridge University Press: Cambridge.

PETER KINYUA MURIUNGI

- Starke, Michal. 2001. *Move Dissolves into Merge: A Theory of Locality*. Doctoral Dissertation, University of Geneva.
- Weerman, Fred, and Jacqueline Evers-Vermeul. 2002. Pronouns and case. *Lingua* 112:301–338.