Wolof *Wh*-Movement at the Syntax-Morphology Interface

Contents

1 Introduction

One of the important questions generative linguists try to answer is how diverse languages truly are. Proponents of universal grammar tend to believe that, when properly unpacked, a lot of apparent differences between languages arise from operations not having to do with narrow syntax, and that in fact most languages have a very similar underlying structure. The layers to be unpacked are various processes that take place at the interface of syntax with other components of the grammar. This paper focuses on the interface of syntax with morphology.

In the last several decades, with the rise of Distributed Morphology, the boundary between syntax and morphology is treated as much more permeable than by lexicalist approaches to word formation. In a theory of the syntax-morphology interface in which the post-syntactic component of the grammar is a separate module, with its own principles and constraints, the phonological reflection of syntactic structure is complicated and the expected output of syntactic derivations can be obscured, creating the appearance of syntactic complexity and peculiarities where in reality there are none. By looking more closely at the types of post-syntactic processes that occur in languages, we arrive at the conclusion that a lot of the complexity that we may be tempted to blame on language-particular syntactic behavior is in fact attributable to the interaction of the syntactic and the post-syntactic components. This is due to various operations occurring in post-syntax, affecting the output of syntax by deleting features, or even entire nodes, by allowing a single node to be realized in more than one morphological position, or by collapsing multiple nodes into one. The operations of the post-syntactic component are no less constrained than those of the syntactic component, and just as we see uniformity in syntactic processes, we also expect to see cross-linguistic similarities in morphological patterns. This view is defended at length by Arregi and Nevins (2012), who attribute morphological processes to universal or language-specific markedness constraints in the post-syntax.

This paper contributes to this line of research, by investigating the seemingly complex syntactic structure of A′-movement in the Niger-Congo language Wolof, and argues that its syntax is, in a sense, completely unremarkable – there is nothing out of the ordinary about A′-movement constructions in Wolof. There is also nothing completely unique to Wolof in the post-syntactic component of its grammar. The processes which I argue are at work in its CP-layer are found in

similar morpho-syntactic configurations (specifically, structurally adjacent elements in agreement relations) in various languages. It is thus the interaction of two orderly behaving modules—the syntactic and the post-syntactic one—that creates the appearance of disorder. In the remainder of the introduction, I give an overview of the data and a summary of the analysis.

A′-movement in Wolof exhibits several morphosyntactic effects. What makes it especially interesting, and highly relevant for the discussion of the syntax-morphology interface, are two facts. First, all A′-extraction effects in Wolof surface on C, but with two different versions of the complementizer, *(l)a* and CM-*u*. *(L)a* occurs in every intermediate C position between the extraction site and the final landing site, similarly to *aL* in Irish (McCloskey 2001, 2002). It also exhibits a subject/non-subject asymmetry, a type of the English *that*-trace effect (Perlmutter 1971) as in (1)-(2), surfacing as *a* in subject extraction and as *la* in non-subject extraction.[1]

(1) *Subject question with (l)a*[2]
K-an **a** jox Musaa téere bi?
CM-Q C_{WH} hand Musa book DEF.SG
"Who handed the book to Musa."

(2) *Object question with (l)a*
K-an **l-a** Musaa gis?
CM-Q *l*-C_{WH} Musa see
"Who did Musa see?"

CM-*u* does not exhibit the same effects as *(l)a*, but reflects agreement in φ-features with the covert *wh*-operator in its specifier:

(3) *Subject question with* CM-*u*
K-u jox Musaa téeré bi?
CM-C_{WH} hand Musa book DEF.SG
"Who handed the book to Musa."

(4) *Object question with* CM-*u*
Y-u Musaa gis?
CM.PL-C_{WH} Musa see
"What(pl) did Musa see?"

The second crucial observation about Wolof is that the two versions of the complementizer are for the most part in complementary distribution: they both occur in *wh*-questions, as in (1)-(2) and (3)-(4), but only one of the versions of C_{WH} is allowed in all other A′-constructions. For example, relative clauses are only formed with CM-*u*, and exhaustive focus constructions only with *(l)a*. The near-complementary distribution of the two versions of C_{WH} gives the impression that different syntax underlies the distinctions in their CP-layer. The goal of this paper is to argue against this view, and show that all surface differences between the two seemingly different A′-movement constructions can be explained as simply that – surface differences, resulting from post-syntactic

[1] Unless otherwise noted, all the data in this paper come from my own field work with native speakers of Wolof in Chicago, Paris, and Saint Louis, Senegal.

[2] Abbreviations: AFF = affirmative, AUX = auxiliary, CM = class marker, DEF = definite determiner, DIST = distal, FUT = future tense marker, IMPERF = imperfective marker, INDEF = indefinite determiner, PERF = perfective marker, PRED.FOC = predicate focus, PROX = proximal, PST = past tense marker, Q = question word.

processes. Specifically, I argue that C_{WH} in Wolof has the same relevant featural content in all A′-movement constructions, meaning that all extraction effects—complemetizer agreement in Wh- and φ-features and the subject/non-subject asymmetry—are present in the syntax of each of them. Their overt representation, however, is obscured by post-syntactic operations.

The analysis consists of two parts. First, I make a specific claim about the origin of all A′-extraction effects in Wolof: I propose them all to be reflexes of agreement between C and some element in the clause. This is an obvious analysis of the occurrence of the complementizer in all C positions between the extraction site and the final landing site (Agree between C and the extracted phrase in the Wh-feature) and of the occurrence of φ-features of the extracted phrase on C. The subject/non-subject asymmetry is a more controversial type of effect. However, when its seemingly complicated surface syntax is analyzed as the morphological reflex of agreement, it feeds into the view of syntax of a simple, cross-linguistically uniform system. An analysis along these lines is proposed in Pesetsky and Torrego (2001) for the English *that*-trace effect. Specifically, C_{WH} is an agreeing complementizer, with uninterpretable Wh-, φ- and T-features. The uninterpretable Wh- and φ-features are valued via Agree under c-command between C and the *wh*-phrase. The uninterpretable T-feature on C is related to the subject/non-subject asymmetry, which is analyzed as a T-to-C asymmetry. This feature can be valued via movement of T (which is then realized as *l*-) to C, or the movement of the subject to Spec,CP (which in this analysis has a valued *u*T). Pesetsky and Torrego (2001) predict that C is morphologically more complex in non-subject extraction than in subject extraction, but in languages which they survey both C and T are never overt in A′-movement. In Wolof they are, therefore the analysis presented in this paper significantly strenghtens their proposal and explains part of the surface complexity as a result of the nature and morphology of agreement.

The second part of the analysis tackles the question of the distribution of CM-*u* and *(l)a* by illustrating that the surface visibility of a particular A′-extraction effect correlates with the overtness of the φ-feature in the C head and its specifier. *Wh*-questions, the only A′-movement construction which can surface with either of the allomorphs of the complementizer, show us that the question with *(l)a* obligatorily has an overt question word in its specifier (*k-an*), as in (1)-(2), while the question with CM-*u* obligatorily has a null specifier, seen in (3)-(4) (Torrence 2005, 2012a,b). Since both CM-*u* and the question word CM-*an* contain an overt φ-feature, their complementary distribution indicates that if an overt marker of a φ-feature shows up in C, it cannot show up in Spec,CP,

and vice versa. I argue this to be a consequence of a morphological Obligatory Contour Principle constraint, which prohibits the φ-feature to occupy two adjacent nodes (here the head C and its specifier), and triggers a repair: one of the nodes containing the φ-feature is deleted. This can either be the φ-feature in C, or the phrase in Spec,CP. The Wolof data give support for a particular type of post-syntactic Impoverishment, called Obliteration, which does not delete only offending features, but targets entire nodes (e.g. in Arregi and Nevins 2007, 2012). The difference in surface forms of C (CM-*u* vs. *(l)a*) is the result of post-syntactic morphological rules, which realize C as -*u* when it is adjacent to the φ-feature, and as *a* in all other environments. In other words, *a* and -*u* are allomorphic realizations of C, crucially dependent on the presence or absence of φ (CM) in C. If either C or Spec,CP contain content irretrievable in the CP-layer, the deletion of that element is blocked, which results in a particular construction always surfacing with only one allomorph of C. The distribution of the two variants of C is therefore a surface effect—the result of the OCP constraint and allomorphy—and not the product of underlying syntactic differences.

A′-movement effects in Wolof illustrate the tension between the principles of the syntactic and the post-syntactic component, and the morpho-syntactic consequences of their interaction, particularly well. The syntax of A′-extraction involves agreement, resulting in adjacent featural identity. OCP-type constraint is a typical post-syntactic phenomenon that militates against adjacent identity in particular morphosyntactic configurations, as identified in various languages (e.g. Nevins 2007, 2012; Ackema and Neeleman 2004). Such effects and their repairs are known to be language-, even dialect-specific (Arregi and Nevins 2012). The analysis I propose for the morphosyntax of Wolof A′-movement is therefore not meant to apply in this configuration in all languages. However, OCP-effects are expected to arise precisely in the type of rich agreement configurations as is the CP-layer in Wolof A′-movement.

This paper offers not only a unified analysis of A′-extraction effects and maintains a unified syntax of A′-extraction in Wolof, but crucially offers a principled account for the distribution of different shapes of the CP-layer in different instances of A′-movement. What it ultimately shows is that careful examination of apparent syntactic irregularities in various languages must involve the cross-linguistic exploration of phenomena that occur in other modules of the grammar, and of their behavior at the interface with syntax. Even though some of these processes may appear to be unique to particular languages, by uncovering their uniform cross-linguistic properties, the apparent

syntactic differences that they cause can be easily explained without endangering our view of syntax as a cross-linguistically fairly uniform system.

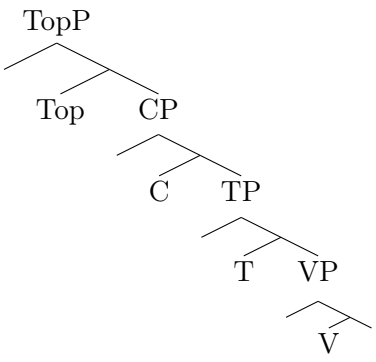
The paper is organized as follows. Section 2 lays out the relevant details of Wolof clause structure and §3 presents the syntax of A′-extraction and justifies a unified syntactic treatment of all A′-movement in Wolof. The syntactic and the post-syntactic analyses of the morpho-syntax of A′-movement effects are laid out in §4 and §5, respectively. The central section for the post-syntactic analysis of the differences in the CP-layers of the two construction types is §5; however, it rests on the syntactic analysis proposed in §3 and §4. The paper is concluded in §6.

2 Wolof clausal structure

In this section, I present arguments for the clause structure I assume in the rest of the paper, focusing on two parts of the clause. First, we briefly familiarize ourselves with the position and function of complementizer-like elements in Wolof, since the A′-movement complementizer is one such element. Second, I address the status of pronominal subjects, because their position signals that the element to their left or right, depending on the presence or absence of A′-movement of the subject, is a complementizer.

Wolof belongs to the Atlantic branch of the Niger-Congo language family, most widely spoken in Senegal, but also in the Gambia and Mauritania. For the purposes of the present discussion, I assume that Wolof sentences have the basic structure in (5).

(5) *Wolof clausal structure*



Wolof is an SVO language, as shown in the neutral, affirmative sentence in (6):

(6) *Affirmative sentence*
Xale yi jox na-ñu Musaa téeré bi.
child DEF.PL give C_{AFF}-3PL Musa book DEF.SG
"The children gave Musa the book."

The basic word order is changed in A′-movement structures, such as focus constructions, in which the focused element is fronted:

(7) *Object focus*
Téeré bi l-a xale yi jox Musaa.
book DEF.SG *l*-C_{WH} child DEF.PL give Musa
"The children gave the $BOOK_{FOC}$ to Musa."

Aside from word order, (6) and (7) differ in two more ways. First, in both sentences there is an element in the clause called a *sentential particle* – *na* in the affirmative sentence (glossed as C_{AFF}), and *la* in the object focus sentence (glossed as C_{WH}). Second, the *subject marker* follows the particle *na* in (6) (*ñu*, 3PL), but is absent in (7).[3] I discuss the properties of complementizers and subject markers in the remainder of this section.

Sentence particles are complementizer-like elements, whose presence is obligatory in order for the sentence to receive temporal interpretation; sentences which lack particles also lack overt tense morphology, and their temporal interpretation depends on context (Njie 1982).[4] There are a number of different particles: A′-extraction particle (traditionally considered subject/complement focus), imperative, affirmative, obligative and negative imperative/obligative particles, and four different temporal modality particles, which express reference to a past event, near or remote in time, or reference to a hypothetical or expected future event. Sentence particles are located in a projection which takes the TP as its complement, and are in complementary distribution. Dunigan (1994) therefore assumes that all sentential particles occupy a single position in the clause, which she terms the Sigma phrase, following Laka (1990). As this paper only explores A′-movement constructions, I am not concerned with the status of other sentential particles; the important fact to bear in mind is that they do not co-occur with each other. I argue that the particles which occur in A′-movement are allomorphs of the *wh*-complementizer located in C.

[3] Another difference is the position of the verb. This is not relevant for the present purposes.

[4] The temporal interpretation of sentences in Wolof is rather complex, and depends on the clause type, the presence or absence of overt tense morphology, and verb type. For detailed discussion, see Dunigan 1994.

We are next concerned with the position of subject markers. In (6), a 3PL subject marker *ñu* follows the clausal particle *na*, but it is altogether absent in (7). Wolof subject markers are differently analyzed in the literature—as either agreement morphemes or pronominals—due to the fact that they are obligatory in addition to a lexical subject in some constructions, as in the affirmative sentence in (6), and in complementary distribution with lexical subjects in others, as in the focus sentence in (7). Dunigan (1994) and Russell (2006) convincingly argue that when lexical subjects are at the left edge of the clause and subject markers follow the clausal particle, as in (6), the structure has an obligatorily left-dislocated (topicalized) lexical subject and a resumptive subject pronoun in Spec,TP. I follow this analysis and consider subject markers to be pronominal elements.

In all A′-movement constructions lexical subjects and subject pronouns are located next to the sentential particle, i.e. the A′-movement complementizer. In case of non-subject extraction this position is to the right of the complementizer, as in (8) and (9).

(8) *Subject marker and lexical subject in object focus*

a. Aali l-a **góor gi** gis.
Ali *l*-C_{WH} man DEF.SG see
"The man saw ALI_{FOC}."

b. Aali l-a-**∅** gis.
Ali *l*-C_{WH}-3.PL see
"He saw ALI_{FOC}."

(9) *Subject marker and lexical subject in non-subject relative*

a. góor g-u **Aali** gis
man CM-C_{WH} Ali see
"the man who Ali saw"

b. góor g-u **mu** gis
man CM-C_{WH} 3SG see
"the man who he saw"

In subject extraction, the lexical subject or the subject marker are left-adjacent to the complementizer, due to the fact that they by hypothesis A′-move to its specifier, as in (10) and (11).

(10) *Subject marker and lexical subject in subject focus*

a. **Aali**-a (>Aalee) gis xale bi.
Ali-C_{WH} see child DEF.SG
"ALI_{FOC} saw the child."

b. **Mu**-a (>moo) gis xale bi.
3SG-C_{WH} see child DEF.SG

"S/HE$_{FOC}$ saw the child."

(11) *Lexical subject in a subject relative*
xale b-i gis Aali
child CM-C$_{WH}$ see Ali
"the child who saw Ali"

No non-clitic element can ever intervene between the complementizer and the lexical subject or the subject marker.[5] I therefore posit that the lexical subject/subject marker is located in the specifier of TP, which C takes as its complement.

In addition to sentence particles, Wolof also possesses a subordinating complementizer *ni* 'that', which can co-occur with sentence particles, in both non-extraction and extraction constructions, as in (12).

(12) *Subordinating complementizer* ni *in Wolof*

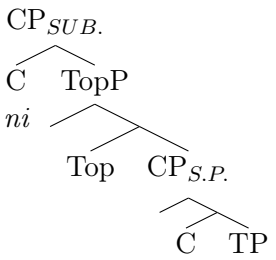
a. Gëm na-nũ **ni** Aali gis na-∅ Musaa
believe C$_{AFF}$-3PL that Ali see-3SG C$_{AFF}$-3SG Musa
"They believe that Ali saw Musa."

b. K-an l-a-nũ gëm **ni** l-a Aali gis?
CM-Q *l*-C$_{WH}$-3PL believe that *l*-C$_{WH}$ Ali see
"Who do they believe that Ali saw?"

The proposed analysis assumes that, if *(l)a* and CM-*u* are treated as complementizers, we have to allow for multiple CP layers in Wolof. In this sense, Wolof would be similar to Korean, which distinguishes between mood markers signaling clause type (statements, questions, etc.), obligatory in every clause, and a subordinating particle, which introduces embedded clauses. Bhatt and Yoon (1992) propose that the category "Comp" be dissociated into two distinct categories – one that indicates clause-type (MOOD), and one that indicates subordination (SUBORDINATORS) – which some languages would conflate, and some keep separate. Following this type of an analysis, Wolof would belong to the second category, distinguishing sentential particles and a subordinating particle *ni*. Another possible route is to assume a Split CP along the lines of Rizzi (1997), where *ni* would be a high complementizer (Force), and *(l)a* a low complementizer (possibly Foc). I do not take a specific stand on this issue, but assume two C heads in Wolof, a higher one which hosts only the subordinating complementizer *ni*, and a lower one which hosts all sentence particles, as in (13).

[5] Russell (2006) argues that object and locative clitics in Wolof undergo syntactic movement to Spec,TP, tucking under the subject. Then phonological reordering takes place, positioning all clitics on the left edge of the TP, as they are enclitics on the C-domain. Therefore, the subject clitic precedes all other clitics, since it also undergoes phonological reordering. The lexical subject, however, appears to the right of the object and locative clitics.

(13) *Two CP layers in Wolof*



Subject markers and some allomorphs of C are clitics (Russell 2006), and they undergo various vowel coalescence and/or morpho-phonological fusion processes, depending on the element they are adjacent to (Ka 1987; Ndiaye 1995). For example, the 3^{rd} person singular subject marker is $\emptyset$ when right-adjacent to *la*, as in (8b), and *mu* when right-adjacent to CM-*u*, as in (9b). Adjacent word-final vowel *i* and the C allomorph *a* become *ee*, shown in (10a), and adjacent *u* and *a* become *oo*, in (10b). We will not concern ourselves with the details of these changes. In all examples, both the underlying form and the surface form of the subject marker and the sentence particle are provided.

This section offered a basic description of those elements of Wolof clausal structure that are relevant to the present discussion. Additional clarifications are provided throughout the paper, as necessary. The remainder of the paper is concerned with the syntactic (§3 and §4) and post-syntactic (§5) analysis of the CP-layer in A′-movement in Wolof.

3 The Syntax of A′-extraction

Torrence (2005) shows that constructions with CM-*u* and those with *(l)a* both involve A′-movement.[6] Both forms signal agreement in the Wh-feature, and the use of one or the other allomorph correlates with Case agreement (subject/non-subject asymmetry marking) and φ-agreement, respectively. At first glance, their distribution appears to be fairly simple: some A′-movement constructions occur only with *(l)a*, and some only with CM-*u*, suggesting that they may be syntactically or featurally distinct. In *wh*-questions, however, both variants of C$_{WH}$ can appear, with no difference in meaning or use. More importantly, even in *wh*-questions, there are cases in which CM-*u* and *(l)a* are in complementary distribution. I therefore argue that the difference between them is not syntactic, or featural, but that the clue to the explanation of their distribution lies in an important

[6]They are sensitive to islands, exhibit reconstruction effects, and pass a Wolof-specific A′-movement test. For details, see Torrence 2005, 2012a.

distinction between the CP-layers of the two types of A′-movement constructions, having to do with the overtness of the phrase in Spec,CP, and the φ-feature in C. I offer a post-syntactic analysis of the surfaces differences in the CP-layer in §5. This section argues that constructions with one or the other complementizer variant have identical syntax, and that the two versions of C_{WH} are allomorphs. In 3.1, I discuss the basic distribution of CM-*u* and *(l)a*, and in 3.1-3.4, I present evidence for a unified syntax of the two structures.

3.1 Basic distribution of CM-*u* and *(l)a*

I propose that C_{WH} in Wolof, in all constructions, has the feature specification in (14): it has uninterpretable Wh-, T-, and φ-features. I follow Pesetsky and Torrego (2001) in treating the EPP feature as a subfeature of features. In this case, all uninterpretable features have an EPP subfeature, meaning that they must be valued via Agree and Move.

(14) *Feature specification of C_{WH}*

$$\begin{array}{c} C_{WH} \\ \left[\begin{array}{c} u\mathrm{Wh}_{\mathrm{EPP}} \\ u\mathrm{T}_{\mathrm{EPP}} \\ u\varphi_{\mathrm{EPP}} \end{array}\right] \end{array}$$

C_{WH} occurs in a variety of sentence types: *wh*-questions, exhaustive focus constructions, relative clauses, comparatives and copular sentences with two DPs (Double-DP copular sentences). Aside from *wh*-questions with a simple *wh*-word, which can contain either *(l)a* or CM-*u*, in all other structures the two allomorphs are in complementary distribution. Table 1 summarizes the distribution of *(l)a* and CM-*u*. In this paper, I only discuss representative examples from each of the two categories: *wh*-questions, exhaustive focus constructions and relative clauses.[7]

[7]Other syntactic differences aside, focus constructions, comparatives and copular clauses all have the surface shape of the CP-layer in common: they all have an overt DP in Spec,CP. In focus structures it is the exhaustively focused constituent, in comparatives the DP which is the target of comparison, and in Double-DP copular sentences the predicate DP. Since there is no obvious unifying featural factor for these three structures, I argue that the fact that they all contain *(l)a* is the result of post-syntax. See §5 for details.

		(l)a	CM-**u**
HIGHEST C	**wh-questions**	√	√
	exhaustive focus	√	*
	comparatives	√	*
	Double-DP copular clauses	√	*
	relative clauses	*	√
INTERMEDIATE C		√	*

Table 1: The distribution of CM-*u* and *(l)a*

Examples in (15) illustrate *wh*-questions with *(l)a*. The two examples also show the subject/non-subject asymmetry that this complementizer exhibits: it surfaces as *a* in case of subject extraction and as *la* in case of non-subject extraction.[8]

(15) a. *Subject question with (l)a*
K-an **a** gis Musaa?
CM-Q C_{WH} see Musa
"Who saw Musa?"

b. *Object question with (l)a*
L-an **l-a** Musaa gis?
CM-Q *l*-C_{WH} Musa see
"What did Musa see?"

Questions completely equivalent in meaning to those in (15a) and (15b) can also be formed with CM-*u*, as in (16a) and (16b) (see §3.5 for details on the semantic equivalence of the two structures).

(16) a. *Subject question with* CM-*u*
K-u gis Musaa?
CM-C_{WH} see Musa
"Who saw Musa?"

b. *Object question with* CM-*u*
L-u Musaa gis?
CM-C_{WH} Musa see
"What did Musa see?"

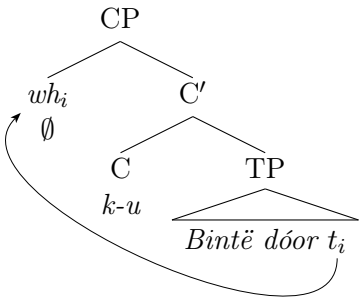
Wh-questions with CM-*u* can be posed about subjects, objects, locatives, manners, and instrumentals, as long as the question corresponds to one simple *wh*-phrase (*who, what, how*) (Torrence 2005, 2012b). The CM-*u* complementizer can be formed with any of the thirteen noun class markers, requiring the answer to contain an item from that noun class. Both CM-*u* and *(l)a* can also occur in embedded questions.

Noun class marking on CM-*u* may suggest that it is a type of a question word, due to its similarity with the question word CM-*an*, which occurs in questions with *(l)a*. However, Torrence (2005, 2012a,b), presents extensive evidence in favor of treating CM-*u* as a complementizer with a null *wh*-phrase in its specifier, as in (17).

[8]Wolof is a noun class language, like other Atlantic languages, and class membership is usually indicated on DP elements other than the noun, such as articles and demonstratives (Torrence 2005, 2012a,b). The class marker morpheme is glossed as CM.

The syntax of CM-*u* (Torrence 2012b, p.1157-1158)

(17) a. K-u Bintë dóor?
φ-C_{WH} Binta hit
"Who did Binta hit?"

b. 

I follow Torrence in considering CM-*u* to be a complementizer, however, I argue in §5 that the obligatorily empty Spec,CP is the result of post-syntactic processes, and not a selectional property of CM-*u*.I do not discuss all of Torrence's evidence for treating CM-*u* as a complementizer; the interested reader is directed to Torrence 2005, 2012a,b. My goal here is to establish a syntactic parallel between constructions with CM-*u* and *(l)a* in support of an analysis according to which constructions with *(l)a* have the same syntax as those with CM-*u*. In the remainder of this section, I give three arguments for this claim. First, we shall see that allomorphy is independently needed to account for different forms of CM-*u* in relative clauses, therefore employing it to account for all A′-movement structures extents an already existing analysis. Next, pied-piping data show that the *wh*-phrase can pied-pipe a preposition with both allomorphs, suggesting their syntactic equivalence, but that pied-piping of other material (which must be overt) is only possible with *(l)a*, pointing to the fact that the two structures differ in which material they allow to be overt in Spec,CP, and not in their syntax. And finally, long distance extraction is possible only with *(l)a* in intermediate C_{WH} position(s), which is a strong argument in favor of it being an A′-movement complementizer. In the final part of this section, I entertain alternative analyses and show that they cannot account for the distribution of CM-*u* and *(l)a*.

3.2 Allomorphy is independently needed

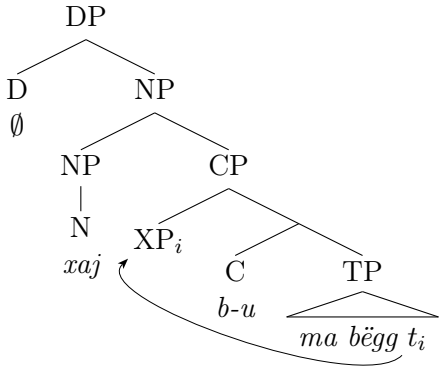
As it turns out, an allomorphy-based analysis is already needed to account for different forms of CM-*u* that occupy C in Wolof relative clauses.[9] I assume that Wolof relative clauses, as in (18a), have the syntax in (18b): a relative operator co-indexed with the head noun is located in Spec,CP,

[9] CM-*u* also occurs in temporal clauses and conditionals, which are a type of relative clause (Torrence 2012a). In this paper I restrict the discussion to relative clauses, assuming that the analysis extends to temporal and conditional clauses.

and the relative CP is the complement of D. I justify this syntax in §5.4.

The syntax of Wolof relative clauses

(18) a. Xaj **b-u** ma bëgg.
dog CM-C_{WH} 1SG like
"*a dog that I like*"

b. 

In addition to noun class, in relative clauses C_{WH} encodes definiteness and proximity, having three allomorphs – CM-*u*, CM-*i* and CM-*a*. Determiners[10] have similar forms in Wolof, *a*-CM (or in some dialects *u*-CM) being the indefinite one, and CM-*i* and CM-*a* the definite ones, with the former denoting a spacially proximal entity, and the latter a distal one. The indefinite determiner is usually omitted.

(19) a. *Indefinite determiner*
a-b xaj
INDEF-CM.SG dog
"*a dog*"

b. *Definite proximal determiner*
xaj **b-i**
dog CM.SG-DEF.PROX
"*the (proximal) dog*"

c. *Definite distal determiner*
xaj **b-a**
dog CM.SG-DEF.DIST
"*the (distal) dog*"

The vowels in the complementizer have the same meaning: CM-*u* is an indefinite relative marker, and CM-*i* and CM-*a* definite ones, denoting that the head of the relative clause is proximal or distal, respectively, in space, time, or discourse (Torrence 2012a).[11] It must also be pointed out that the determiner of the DP containing the head noun and the relative marker normally do not co-occur, shown in (20).[12]

[10] Wolof is for the most part a head-initial language. Interestingly, it has a mixed determiner system, where the indefinite determiner is pre-nominal, and definite determiners post-nominal.

[11] The three versions of this complementizer differ in their distribution. Only CM-*u* can be used in questions, and only CM-*i* in certain free relative constructions (Caponigro and Heller 2007). In this paper, I am disregarding these distributional facts, though the definiteness of the complementizer plays an important role in my analysis, discussed in §5.

[12] According to Torrence (2012a,b), the determiner can optionally surface on the edges of the relative clause, as in (i). (example from Torrence 2012a, p.103).

(20) *Relative clauses and definite determiners in Wolof*

a. (*a-b) xaj **b-u** ma bëgg
INDEF-CM dog CM-C_{WH} 1SG like
"*a dog that I like*"

b. xaj (*b-i) **b-i** ma bëgg
dog CM-DEF.PROX CM-C_{WH} 1SG like
"*the dog (here) that I like*"

c. xaj (*b-a) **b-a** ma bëgg-oon
dog CM-DEF.DIST CM-C_{WH} 1SG like-PST
"*the dog (there) that I liked*"[13]

It appears that CM-*u/i/a* in relative clauses signals not just agreement in noun class, but also in definiteness and proximity. It could be proposed that a morpho-phonological process (fusion or deletion) ensures that the the determiner and the complementizer do not both surface next to each other, due to their phonological similarity, and that what surfaces is the determiner, carrying the definiteness feature. While this could explain examples (20b) and (20c), where the determiners and the complementizer are next to each other and have identical phonological form, it cannot account for the absence of the indefinite determiner in (20a), which is prenominal, and which differs in form from the complementizer.[14] I therefore propose that C in relative clauses, in addition to the φ-feature, carries the definiteness and proximity features. It obtains these features not via agreement with the *wh*-phrase in Spec,CP, but via agreement with D of the external head: D agrees with its complement, the CP, and the features percolate down to CP's head. In §5 I make a tentative proposal for a mechanism which regulates the cooccurrence of D and C in relative clauses.

The data from relative clauses show that allomorphy is independently needed to account for three different variants of the φ-agreeing C_{WH}, which encode definiteness and proximity. Therefore, one of the benefits of the present account is that it represents the extension of an existing analysis.

(i) góór g-i ñu gis (g-i)
man CM-*i* 3PL see CM-DEF.PROX
"*the man that they saw*"

In the variety of Wolof analyzed in this paper, the occurrence of the determiners on the edges of the relative clause is not possible. My analysis does not hinge on the exact position of the determiner, as I assume that a mechanism different from the OCP_{φ} is involved in regulating the cooccurrence of C and D. For details, see §5.4.

[13] In my data, CM-*a* is always accompanied by the past tense morpheme *-oon* on the verb, indicating that the state or event is removed in time from the moment of speaking.

[14] It also cannot account for Torrence's data in footnote 12, in which the determiners can optionally surface on the edges of relative clauses.

3.3 Pied-piping

An important piece of evidence for the parallel treatment of CM-*u* and *(l)a* comes from pied piping. First, not all phonologically overt material is banned from Spec,CP of structures with the allomorph CM-*u*. The locative preposition *ci* can be pied-piped by the *wh*-phrase both in constructions with CM-*u* and in those with *(l)a* (Torrence 2012a):[15]

(21) *Preposition pied-piping in Wolof*[16]

a. {Ci fan}/{Fan ci} la ñu teg téeré bi?
{P-LOC where}/{where P-LOC} *l*-C_{WH} 3PL put book DEF.SG
"On what did they put the book?"

b. Ci l-u ñu teg téeré bi?
P-LOC CM-C_{WH} 3PL put book DEF.SG
"On what did they put the book?"

These examples show that both the overt *wh*-phrase in (21a) and the null *wh*-phrase in (21b) can pied-pipe material to Spec,CP. Crucially, not all material can be pied-piped with both C allomorphs. We have seen that *wh*-questions can contain both variants of C_{WH} (see (15) and (16)), but this is only the case when the *wh*-phrase consists of one word. Questions with complex *wh*-phrases can only be formed with the complementizer *(l)a*, as shown in (22). CM-*u* is banned.[17]

(22) *Questions with a complex wh-phrase can only contain* (l)a

a. Jaaj-u **k-an** **l-a** Aali gis?
mother-of CM-Q *l*-C_{WH} Ali see
"Whose mother did Ali see?"

b. *Jaaj-u **k-u** Aali gis?
mother-of CM-C_{WH} Ali see

c. **B-an** xale **l-a** Faatu gis?
CM-Q child *l*-C_{WH} Fatou see
"Which child did Fatou see?"

d. ***B-an** xale **b-u** Faatu gis?
CM-Q child CM-C_{WH} Fatou see

[15] This is the only preposition that behaves in this way in Wolof. Other prepositional elements, *ag* 'with' (which is also a conjunction) and *ngir* 'for' cannot pied-pipe or be stranded in A′-movement, but are replaced with applicative/benefactive suffixes on the verb (Torrence 2012a).

[16] In *wh*-questions, *ci* can be both a preposition and a post-position. The fact that it can precede or follow the *wh*-word *fan* 'where', but cannot follow CM-*u*, is additional evidence that CM-*u* is not a question word.

[17] Torrence (2012a) reports that there is variation in whether CM-*u* can occupy C_{WH} in these types of questions. I have found no variation amongst speakers from Saint Louis and Dakar, but I address Torrence's data in §5 in more detail and show that my analysis can easily be extended to account for them as well.

Examples in (21) and (22) show that not all material is banned from Spec,CP of CM-*u*, but only the *wh*-DP. I take the fact that constructions with both allomorphs behave identically with respect to preposition pied-piping as another indication that they are syntactically equivalent. I argue in §5 that the reason for the ungrammaticality of complex *wh*-phrases with CM-*u* lies in the tension between the requirement that the phrase in Spec,CP containing φ-features must be deleted when C is realized as CM-*u*, and the fact that complex *wh*-phrases cannot be deleted (due to their non-recoverability). Such questions can therefore only contain *(l)a*. If we assume that structures with CM-*u* and *(l)a* are syntactically equivalent, and their surface differences the result of post-syntactic processes, we have a way of understanding the pied-piping facts, which otherwise remain unexplained.

3.4 Long-distance extraction

Another property of *(l)a*, one that strongly favors an analysis which treats it as an A′-movement complementizer, is that it obligatorily occurs in long-distance movement; extraction out of an embedded clause that contains a different sentence particle is not possible (Dunigan 1994). The example in (23b) illustrates an attempt at extraction out of an embedded clause with the verb-focus particle *da*. Extraction is equally ungrammatical with CM-*u* and *(l)a* in the matrix clause, if the embedded clause retains the particle *da*. Long distance extraction out of the sentence in (23a) is only possible if *(l)a* occupies the embedded C_{WH}, as in (23c).

(23) *Verb focus particle and A′-extraction*

a. Moodu xam ni Faatu da-fa gis gainde.
Modu know that Fatou $C_{PRED.FOC}$-3SG see lion
"Modu knows that Fatou SAW a lion."

b. {*L-an l-a}/{*L-u} Modu xam ni Faatu da-fa gis?
{CM-Q *l*-C_{WH}}/{CM-C_{WH}} Modu know that Fatou $C_{PRED.FOC}$-3SG see
Intended: "What does Modu know that Fatou SAW?"

c. {L-an l-a}/{L-u} Moodu xam ni **l-a** Faatu gis?
{CM-Q *l*-C_{WH}}/{CM-C_{WH}} Modu know that *l*-C_{WH} Fatou see
"What does Modu know that Fatou saw?"

Similarly, CM-*u* cannot occupy the C of an embedded relative clause; only *(l)a* is allowed:

(24) *Relative Clauses*

a. film b-i ñu bëgg
movie CM-C_{WH} 1PL like
"the movie we liked."

b. film b-i mu wax-oon ni l-**a**-ñu bëgg
movie CM-C_{WH} 3SG say-PAST that *l*-C_{WH}-1PL like
"the movie that s/he said we liked"

c. *film b-i mu wax-oon ni b-i ñu bëgg
movie CM-C_{WH} 3SG say-PAST that CM-C_{WH} 1PL like

In this respect, *(l)a* is similar to the Irish complementizer *aL*, and provides evidence for the cyclic nature of A′-movement.[18]

The following picture emerges from comparing CM-*u* and *(l)a*. Both complementizer variants occur in A′-movement constructions. In one type, *wh*-questions, both versions are possible, with no difference in meaning or contexts of use. In other structures only one is allowed, which correlates with the overtness of the *wh*-phrase in Spec,CP. This is especially significant in *wh*-questions with complex *wh*-phrases, which in the dialect discussed in this paper cannot be formed with CM-*u*. Furthermore, *(l)a* obligatorily occurs in intermediate C positions in long-distance extraction, which is strongly reminiscent of the behavior of A′-complementizers in successive cyclic movement. The fact that certain dialects allegedly also allow the occurrence of CM-*u* in intermediate Cs, as discussed in footnote 18, strengthens the proposal that they are allomorphs of the same complementizer.

3.5 Alternative analysis

In this section, I have argued that CM-*u* and *(l)a* are allomorphs of the same C_{WH}, and that the constructions in which they occur are syntactically identical. A common assumption in the

[18] Torrence (2012b) claims that the complementizer CM-*u* exhibits the same behavior, i.e. that it occupies intermediate C positions in long-distance extraction, shown in (i). He only reports this for questions. Furthermore, he gives examples of *mixed chains*, in which CM-*u* and *(l)a* can alternate along the path of A′-movement, as in (ii) (example from Torrence 2012b, p.1173):

(i) **K-u** Kumba wax ne **k-u** Isaa defe ne **k-u** Maryam di dóór?
CM-C_{WH} Kumba say that CM-C_{WH} Isa think that CM-C_{WH} Maryam AUX.FUT hit
"Who did Kumba say that Isa thought that Maryam will hit?"

(ii) **K-u** Kumba wax ne **l-a** Isaa defe ne **k-u** Maryam di dóór?
CM-C_{WH} Kumba say that *l*-C_{WH} Isa think that CM-C_{WH} Maryam AUX.FUT hit
"Who did Kumba say that Isa thought that Maryam will hit?"

I have not been able to confirm this data. For my consultants from Dakar and Saint Louis, Senegal, CM-*u* in intermediate C_{WH} positions is ungrammatical. However, taking Torrence's data into account can only strengthen my proposal to treat *(l)a* and CM-*u* as allomorphs of C_{WH}. Furthermore, the analysis I propose in §5 also accounts for his data.

literature, however, is that the two constructions are syntactically distinct; in particular, that structures with *(l)a* are clefts (Kihm 1999; Torrence 2005, 2013a,b). In this section, I show that there is no evidence in Wolof to justify this position. In particular, *wh*-questions with *(l)a* do not exhibit properties of clefts. First, there are no semantic differences between structures with *(l)a* and those with CM-*u* in *wh*-questions, the construction in which they can both occur. Specifically, questions with *(l)a* are not associated with a non-cancelable existential presupposition, as is the case in clefts. Second, all positions in which *(l)a* occurs cannot be associated with focusing/exhaustivity, also a hallmark of a cleft construction. Third, there is no evidence to support a bi-clausal analysis of structures with *(l)a*, a crucial property of clefts. And finally, in *wh*-questions, there is no difference in which constituents can form questions with CM-*u* and which with *(l)a*, something that we often see in languages that employ multiple syntactic strategies for forming *wh*-questions.

We begin by comparing the semantic properties of structures with *(l)a* to those of clefts in French and English. It was already mentioned in §3.1 that speakers make no semantic difference between the two types of questions (this is also noted in Torrence 2005, and Torrence's subsequent work). As a point of comparison, consider the difference between French cleft-questions and in situ questions. Clefts are associated with an existential presupposition, seen from the infelicity of the question-answer pair in (25). Namely, the cleft question carries a non-cancelable existential presupposition that there exists something that the person does in life, and since *nothing* which cancels such a presupposition, it is not a felicitous answer to (25a). The in-situ strategy, in (26), does not exhibit the same effect – the existential presupposition can easily be canceled (Shlonsky 2012):

(25) *Cleft question in French*

a. C'est quoi que tu fais dans la vie?
it's what that you do in life
"What is it that you do in life?"

b. #Rien.
"Nothing."

(26) *In-situ question in French*

a. Tu fais quoi dans la vie?
you do what in the life
"What do you do in life?"

b. Rien
"Nothing"

In Wolof, neither the question with *(l)a* nor the question with CM-*u* are associated with a non-cancelable existential presupposition; both (27a) and (28a) can be answered with *"Nothing"*:

(27) *Wh-question with (l)a*

a. L-an **l-a** Musaa gis?
CM-Q C_{WH} Musa see
"What did Moussa see?"

b. Dara.
"Nothing."

(28) *Subject question with* CM-*u*

a. **L-u** Musaa gis?
CM-C_{WH} Moussa see
"What did Moussa see?"

b. Dara.
"Nothing."

Similarly, in English cleft question (29) the existential presupposition (that there exists someone that Moussa saw) cannot be canceled, making the insertion of *if anyone* infelicitous. In a non-cleft *wh*-question in (30) this effect is not observed:

(29) *Cleft wh-question*

a. Who was it that Moussa saw?

b. #Who was it, if anyone, that Moussa saw?

(30) *A′-movement wh-question*

a. Who did Moussa see?

b. Who, if anyone, did Moussa see?

In Wolof, the existential presupposition can be canceled via insertion of *if anyone* both in questions with *(l)a* and in those with CM-*u*:

(31) *Canceling the existential presupposition in la-question*

a. K-an l-a Musaa gis?
CM-Q *l*-C_{WH} Moussa see
"Who did Moussa see?"

b. K-an, s-u dee am na-∅ kenn, l-a Musaa gis?
CM-Q CM-C ? have C_{AFF}-3SG someone, *l*-C_{WH} Moussa see
"Who, if anyone, did Moussa see?"

(32) *Canceling the existential presupposition in u-question*

a. K-u Musaa gis?
CM-C_{WH} Moussa see
"Who did Moussa see?"

b. S-u dee am na-∅ kenn, k-u Musaa gis?
CM-C ? have C_{AFF}-3SG someone, CM-C_{WH} Moussa see
"If anyone, who did Moussa see?"

Semantic differences between structures with CM-*u* and *(l)a*, that we would expect to see if one of the constructions was a cleft, are not found in Wolof.

Another argument for treating structures with *(l)a* as clefts comes from the fact that in some cases they involve focusing. It has been claimed in the literature that *(l)a* is a focus marker (Dunigan 1994; Russell 2006). The syntactic parallel between focus constructions and questions would not be surprising. It has been observed that languages which have a designated focus position tend to move their *wh*-phrases to that position as well (Horvath, 1986), and comparatives, which obligatorily contain *(l)a*, are also claimed to involve focusing (Reglero, 2006; Merchant, 2009). In that view, the occurrence of movement which resembles focus movement in languages such as Hungarian (Horvath 1986; É. Kiss 1998) in these structures is not unexpected. If this were the complete list of environments in which CM-*u* and *(l)a* occurred, it would be reasonable to assume that *(l)a*, in addition to having a Wh-feature, also has a focus feature associated with it, triggering A′-movement of the focused constituent to its specifier, and that CM-*u* is the *elsewhere* A′-movement complementizer. However, *(l)a* occurs in environments which cannot be straightforwardly argued to involve focusing. First, it obligatorily occurrs in every intermediate C between the extraction and the final landing site, as discussed in §3.4, which makes it difficult to argue that every C along the path of A′-movement has a focus feature, or, for that matter, that the only way to A′-extract a constituent out of any clause-type is to focus it.[19] Second, *(l)a* occurs in a structure which is not obviously related to focus: copular sentences with nominal predicates. In an information-structurally neutral[20] predicational copular sentence, the nominal predicate is located in the specifier of *(l)a*, as in (33):

(33) *Copular sentence in Wolof*
Osmaan ndongo **l-a**-∅.
Osman student *l*-C_{WH}-3SG
"*Osman is a student.*"

I do not discuss copular clauses in this paper, because their syntax is more complex (see Author to appear(a), to appear(b)), but taking a closer look at them actually gives support to an analysis

[19] But see Torrence (2013b) on a proposal for successive-cyclic clefting in Wolof.

[20] By *information-structurally neutral*, I mean an out-of-the-blue or a broad sentence focus context (e.g. as a response to a question *What happened?*).

according to which *(l)a* is C_{WH}. There is evidence that even in sentences such as (33), *(l)a* is the same C_{WH} as in other A′-movement constructions. I only mention the most striking fact here: in copular sentences, *(l)a* behaves as a sentence particle. The example in (34) shows a copular sentence in a discourse narrative context. Such sentences differ from ordinary Wolof sentences in that they do not obligatorily contain a sentence particle, which, as was shown in §2, is a class of complementizer-like elements that the A′-movement complementizers belongs to (Dunigan 1994; Robert 1991). The absence of *(l)a* in such structures would be unexpected, if it was itself not a sentence particle. Second, copular clauses with nominal predicates can contain other sentence particles, depending on their information-structure. In such cases, *(l)a* does not occur, as with the verb focus particle in (35). And most importantly, copular sentences can contain CM-*u* (to the exclusion of *(l)a*) in those types of environments that are predicted by my analysis: matrix *wh*-questions, which can also be formed with CM-*u*, as in (36), and in relative clauses, as in (37).

(34) *Copular sentence in a narrative context*
Usmaan di sàcckat.
Ousman IMPF thief
"Ousman is a thief."

(35) *Copular sentence with VF*
Usmaan dafa di (>dafay) sàcckat.
Ousman C_{VF} IMPF thief
"It's that Ousman is a thief"

(36) *Copular sentence as a question*
K-u di (>kuy) sàcckat?
CM-C_{WH} IMPF thief
"Who is a thief?"

(37) *Copular sentence in a relative clause*
góor g-u di (>guy) sàcckat
man CM-C_{WH} IMPF thief
"a man who is a thief"

The fact that *(l)a* behaves as a sentence particle in copular sentences with nominal predicates is a strong argument in favor of it being the same A′-movement complementizer as it is in other structures in which it occurs. This casts serious doubt on the possibility that it obligatorily involves (exhaustive) focus of the element in its specifier, as is the case in clefts.

Another crucial property of clefts is bi-clausality. Evidence of bi-clausality includes a copular verb in the higher clause, an expletive subject, and relativizers/complementizers which usually occur in relative clauses in the language.[21] Wolof does not have one verb that would be the equivalent of the English *be*, nor an overt expletive, making the search for evidence of bi-clausality more difficult. It does, as we have seen, have a complementizer that is always overt in relative clauses. The complementizer usually occurring in relative clauses, CM-*u*, does not occur in structures with *(l)a*,

[21] See, for example, Potsdam (2009) and Potsdam and Polinsky (2011) for tests of bi-clausality of A′-movement constructions in Austronesian languages.

which is unexpected if they are indeed clefts:

(38) *Relative clause*

nit k-u Musaa gis
person CM-C_{WH} Moussa see
"A person who Moussa saw."

(39) *(L)a question*

K-an l-a (*k-u) Musaa gis?
CM-Q *l*-C_{WH} (*CM-C_{WH}) Moussa gis
"Who did Moussa see?"

The structures containing *(l)a* that are under discussion here are also not pseudoclefts. Wolof does possess a pseudocleft structure:

(40) *Pseudoclefts in Wolof*

$\tilde{\text{N}}$-i lekk tangal yi xale yi l-a.
CM-C eat sweets DEF.PL child DEF.PL *l*-C_{WH}
"Who ate the sweets were the children."

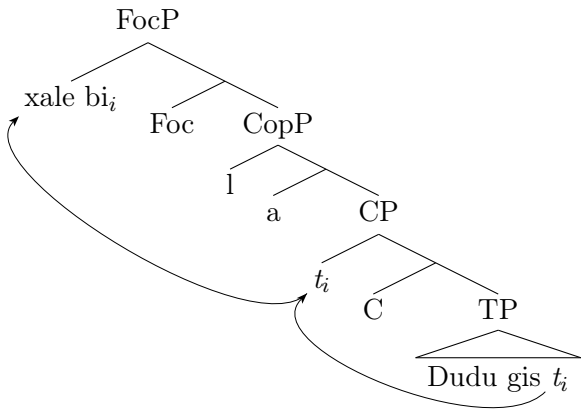
The sentence in (40) shows that a pseudocleft in Wolof contains an overt complementizer in the free relative (*ñi*). No such elements occurs in A′-movement constructions with *(l)a*.

Torrence (2005, 2012a,b, 2013a,b) (see also Kihm 1999) calls the constructions containing *(l)a clefts*, though he does not completely assimilate them to the bi-clausal English cleft-type (*"It's the child that Dudu saw"*). He treats *(l)a* as a raising predicate, occurring in a functional projection CopP above CP, as in (41a) and (41b). Torrence analyzes the *a/la*-asymmetry as a *that*-trace effect, claiming that in subject extraction, the subject raises to Spec,CopP, before moving up to Spec,FocP. In that situation, Cop takes a TP as a complement, and not a CP, in order to avoid a *that*-trace effect, which would occur if the subject moved from Spec,TP to Spec,CP. In non-subject extraction, no such violation occurs (since the subject does not move), so Cop takes a CP as a complement, and *l-* is an expletive, satisfying the EPP requirement of Cop.

(41) *(l)a as a copula* (Torrence 2005, p. 271)

a. Xale bi l-a Dudu gis?
child DEF.SG *l*-C_{WH} Dudu see
"Dudu saw the CHILD?"

b.



My analysis is in fact an implementation of Torrence's. In the following section, I also argue that the *a/la*-asymmetry is a type of the *that*-trace effect; however, I show that, by assimilating it to a T-to-C asymmetry, along the lines of Pesetsky and Torrego (2001), we can explain it without assuming additional functional projections (FocP or CopP), and syntactic differences which offer no insight into the distribution of CM-*u* and *(l)a*.

A final indication that structures with CM-*u* and *(l)a* are not two distinct strategies for forming *wh*-questions comes from comparing Wolof with languages which do have multiple strategies. Crucially, those strategies differ in some way. For example, different strategies might be used for questioning different constituents, as is the case in some Austronesian languages. Seediq (Aldridge 2002, 2004) uses clefts or *wh*-in-situ for argument *wh*-questions, but only *wh*-in-situ for adjunct *wh*-questions. Tagalog (Richards 1998; Aldridge 2002, 2004) uses clefts for argument *wh*-questions, but focus fronting for adjunct *wh*-questions. There is no such division of labor in Wolof: all constituents can be questioned both in a structure with CM-*u* and *(l)a*. The only restriction is in the content of Spec,CP, as discussed in §3.3, which receives a straightforward explanation in the post-syntactic analysis I propose in §5.

Given the semantic equivalence of questions with *(l)a* and CM-*u*, the lack of arguments to support the claim that structures with CM-*u* and *(l)a* are syntactically distinct, and their parallel syntactic behavior, I conclude that the differences in the surface shape of their CP-layer are not of semantic or syntactic provenance.

3.6 Summary

In this section, I have presented arguments for treating *(l)a* as an A′-movement complementizer, on a par with CM-*u*. I show that it behaves like a complementizer in obligatorily occupying all intermediate C positions along the movement path in long-distance extraction, and that it is in complementary distribution with other complementizer-like elements, including CM-*u*. Ultimately, my analysis is an implementation of Torrence's, with a clear advantage: by treating the structures with CM-*u* and those with *(l)a*, I am able to offer an analysis which captures the distribution of the two allomorphs in different A′-movement constructions. This is the goal of the following two sections.

We are now left with two puzzles to solve. First, why is it that when we see φ-agreement on C_{WH}, we never see the subject/non-subject asymmetry, and vice versa? Second, what mechanisms determines which allomorph of C_{WH} is observed in which context, i.e. why can C_{WH} in questions with a simple *wh*-phrase surface as either of the allomorphs, in relative clauses only as CM-*u*, and in questions with complex *wh*-phrases, focus constructions and all embedded C positions only as *(l)a*? In §5 I argue that the shape of the CP-layer in Wolof is the result of the interface between the syntactic and post-syntactic component of the grammar. In particular, I propose that a markedness constraint is at play in Wolof, a morphological Obligatory Contour Principle, prohibiting adjacent φ-features in C_{WH} and its specifier. This configuration arises because C_{WH} agrees in φ-features with the extracted phrase. In my analysis, φ-agreement occurs in all A′-movement constructions, but due to the OCP-triggered deletion, we only observe it when the deletion targets the phrase in Spec,CP and the complementizer surfaces as CM-*u*. In the following section, I make explicit the morpho-syntactic details of the agreement patterns in A′-movement in Wolof, as the post-syntactic analysis depends on the assumptions about the syntax.

4 The Morphosyntax of A′-movement Effects in Wolof

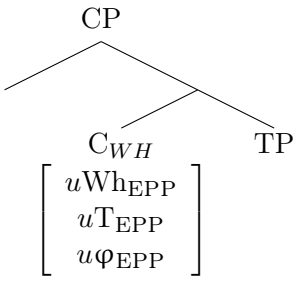
In the previous section, I argue that all A′-movement constructions share the same syntax. The goal of this section is to show that this syntax is quite unremarkable, and that part of the surface differences between structures with CM-*u* and *(l)a* comes from the agreement relations between C_{WH} and other elements of the clause. The other part, which in particular environments obscures some

of the agreement effects, comes from post-syntactic processes discussed in §5; crucially, no surface differences are the result of extraordinary syntactic behavior.

The important contribution of the analysis presented in this section is its treatment of Wolof's A′-extraction effects—complementizer agreement in φ-features and a subject/non-subject asymmetry—as the result of the same mechanism, that of agreement, building on a proposal by Pesetsky and Torrego (2001) who thus analyze the English *that*-trace effect. Their analysis predicts that C in non-subject extraction is more complex than C in subject extraction. The examination of the Wolof subject/non-subject asymmetry provides direct support for their analysis.

My proposal of a unified syntax assumes that every C_{WH} has at least some of the same features, in particular those specified in (42).[22]

(42) *Feature Specification of C_{WH} in Wolof*



The analysis according to which all A′-extraction effects are the result of agreement reinforces and motivates the post-syntactic analysis I propose in §5. Namely, agreement results in featural identity, and in spec-head agreement configurations the identical features are adjacent. Various dissimilations in different modules of the grammar militate precisely against adjacent identity in particular morphosyntactic configurations, as identified in various languages (e.g. Nevins 2007, 2012; Ackema and Neeleman 2004). We therefore expect OCP effects to arise precisely in the type of rich agreement configurations as is the CP-layer in Wolof A′-movement. Since this paper argues that the shape of the CP-layer in A′-movement in Wolof is the result of the interaction between syntax and post-syntax, in this section I lay out the details of the syntactic derivations and the resulting morphology.

The complementizer's uninterpretable features must be valued via Agree and Move, and the A′-extraction effects we observe are the morphological reflex of agreement relations that C_{WH} es-

[22] C_{WH} in different A′-movement constructions can of course also have additional features. In the analysis I propose they do not play a role in the pronunciation of the CP-layer.

tablishes with different elements of the clause. One effect is agreement between C_{WH} and the extracted phrase in φ-features, to check *u*φ on C_{WH}, displayed as a class marker on the complementizer allomorph CM-*u*:

(43) a. *Subject question with* CM-*u*
K-u gis Musaa?
CM-C_{WH} see Musa
"Who saw Musa?"

b. *Object question with* CM-*u*
L-u Musaa gis?
CM-C_{WH} Musa see
"What did Musa see?"

This type of agreement occurs long-distance; in (44), it is present at the top of the dependency. In the variety of Wolof discussed in this paper, CM-*u* cannot occur in intermediate positions of long-distance extraction, though Torrence (2005, 2012a,b) reports that some dialects allow this (see footnote 18).

(44) *Long-distance extraction with* CM-*u in final position*
K-u ñu gëm ni l-a Musaa xalad ni l-a Aali gis?
CM-C_{WH} 1PL believe that *l*-C_{WH} Musa think that *l*-C_{WH} Ali see
"*Who do we believe that Musa thinks that Ali saw?*"

As (44) shows, intermediate C_{WH} positions must be occupied by the allomorph *(l)a*. This variety of C_{WH} does not show φ-agreement, but exhibits a subject/non-subject asymmetry – it surfaces as *a* in subject extraction, as in (45a), and as *la* in non-subject extraction, in (45b). The asymmetry is local: as can be seen from (46), it only occurs at the bottom of the dependency, on the C_{WH} local to the extraction site.

(45) a. *Subject focus with (l)a*
Aali **a** (>Aalee) gis Musaa
ali C_{WH} see Musa
"ALI_{FOC} *saw Musa.*"

b. *Non-subject focus with (l)a*
Musaa **l-a** Aali gis
Musa *l*-C_{WH} ali see
"*Ali saw* $MUSA_{FOC}$."

(46) *Long-distance extraction with (l)a*
Aali **l-a**-a gëm ni **l-a** Musaa xalad ni mu **a** (>moo) leen gis
Ali *l*-C_{WH}-1SG believe that *l*-C_{WH} Musa think that 3SG C_{WH} 3PL.OBJ see
"*I believe that Musa thinks that* ALI_{FOC} *saw them.*"

I follow Pesetsky and Torrego (2001) and analyze the *a/la*-asymmetry as a T-C asymmetry, and

propose an agreement-based analysis for this phenomenon in Wolof. In subsection §4.1, I justify the view that A′-movement effects are all agreement effects. My analysis is restricted to Wolof, and it remains to be seen to what extent it can be applied to similar effects in other languages. In §4.2, I give the details of the analysis.

4.1 A′-movement Effects as Agreement

A′-movement and morpho-syntactic phenomena resulting from it have given rise to much discussion in the field since the seminal works on the topic in the late 60's and early 70's, beginning with Ross 1967. A lot has been uncovered regarding the nature of this type of movement, most notably that the seemingly distant relationship between an A′-moved element and the extraction site in fact consists of a series of local connections, first proposed by Chomsky (1973) and substantiated by investigation of languages with overt morphological evidence for the locality of movement. One such language is Irish, in which an A′-movement complementizer (*aL*) marks each intermediate C position between the extraction site and the final landing site (McCloskey 2001, 2002). Another A′-movement effect reflects agreement between the complementizer and the *wh*-operator and/or the trace in its specifier, usually in φ-features, as, for example, in Kinande (Schneider-Zioga 1995, 1996, 2000, 2007). That these effects involve agreement between C and the extracted phrase is not controversial; that the third A′-extraction phenomenon in Wolof, the subject/non-subject asymmetry, is also a reflex of agreement is less agreed upon. This is the claim I am making, following Pesetsky and Torrego (2001) and their analysis of the *that*-trace effect in English. Specifically, I argue that C_{WH} has an uninterpretable T-feature which can be checked either by T-to-C movement, in non-subject extraction, yielding *la*, or by the moved subject itself, which carries a T-feature in the form of Nominative case, resulting in the C_{WH} surfacing as *a*. The crucial part of this analysis is that it predicts additional morphological material in C in case of non-subject extraction. This is exactly the situation we observe in Wolof. In my analysis, *l-* that occurs in C_{WH} in case of non-subject extraction is the morphological realization of T in C.

An important cross-linguistic property of A′-extraction effects, one that motivates an agreement-based analysis for all of them, is that agreement between C and the extracted phrase in the Wh-feature and/or the φ-feature occurs long-distance or in a cyclic fashion along the entire path of A′-movement, while the various subject/non-subject asymmetries are always local, limited to the clause

in which the extracted phrase originates.[23] In Wolof, φ-agreement occurs only in final landing sites of A′-movement, never in intermediate positions (but see footnote 18). This is nothing unusual; just because a language has an agreeing complementizer, it does not mean that this agreement will show up in every C position. In Kinande, for example, the occurrence of a complementizer in intermediate positions is optional, but if the complementizer does occur, it exhibits φ-agreement (Schneider-Zioga 1995, 1996, 2007). In Chamorro (Chung 1998), the complementizer in A′-extraction alternates depending on the features of the operator in Spec,CP.[24] Crucially, only the highest C alternates in occurrences of A′-movement; the form of intermediate C's is determined by considerations which govern complementizer choice in cases of non-extraction. The fact that agreement between C and the operator in its specifier is not always cyclic can have various explanations: for example, A′-movement in Chamorro may not be cyclic. In §5, I attribute the absence of φ-agreement in intermediate C positions in Wolof to post-syntactic mechanisms that determine C selection.

The presence or absence of a locality effect follows straightforwardly from the nature of agreement. Consider, for example, local extraction schematized in (47), where the C head has an uninterpretable feature *u*F, valued via agreement with a DP in C's c-command domain. The DP can either have a matching interpretable feature *i*F, or a matching uninterpretable feature *u*F.

(47) *Local extraction*

a. $[_{CP}$ C[*u*F] ... DP[*i*F]]
b. $[_{CP}$ C[*u*F] ... DP[*u*F]]

C can check its *u*F against the corresponding feature in both (47a) and (47b); in the former case, the DP carries a matching interpretable feature, and in the latter case a valued uninterpretable feature, which, upon being checked by an *i*F on a different head, is not immediately deleted but stays available for further operations within the same phase (Pesetsky and Torrego 2001). In sum, both interpretable and valued uninterpretable features of a DP can check uninterpretable features on C in local extraction.

The situation in long-distance extraction, illustrated in (48), is different.

[23] This goes for all types of subject/non-subject asymmetries in A′-movement: those similar to the *that*-trace effect, but also anti-agreement effects, which on the surface look quite different, and are clearly tied to agreement (Ouhalla 1993, 2005; Henderson 2007, 2013).

[24] Whether it is a noun phrase or a prepositional phrase, whether the operator is a null relative operator, and whether or not it denotes a location in space or time.

(48) *Long-distance extraction*

a. [$_{\text{CP}}$ C[*u*F] ... [$_{\text{CP}}$ DP[*i*F] ...]]

b. [$_{\text{CP}}$ C[*u*F] ... [$_{\text{CP}}$ DP[~~*u*F~~] ...]]

The DP in the embedded CP can check *u*F on the higher C in (48a), but not in (48b). This is the result of feature life-span: I assume that, while interpretable features are available throughout the entire derivation, uninterpretable features, once valued, must be deleted at the end of the phase in which they are generated (Pesetsky and Torrego 2001). The *u*F on the DP therefore cannot participate in agreement outside its own phase. Thus the uninterpretable feature of C in (48b) must be checked by some other means. In other words, only interpretable features can participate in long-distance agreement relations.

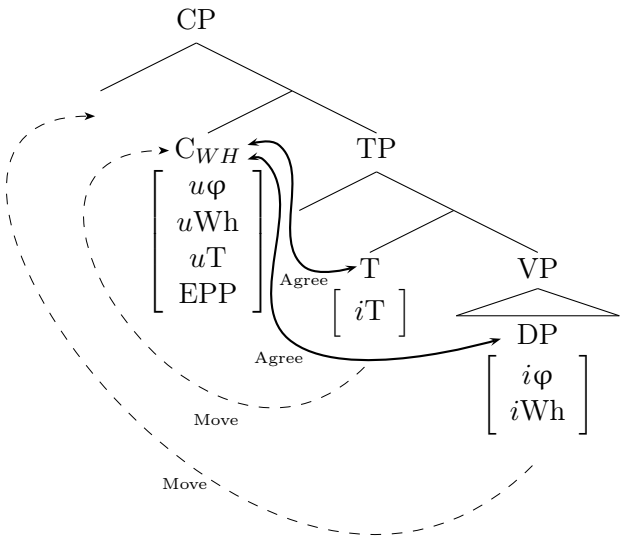
Having set up the stage for a unified treatment of A′-extraction effects in Wolof, in the remainder of this section I lay out the details of the resulting morphosyntactic structures.

4.2 The Subject/Non-subject Asymmetry in Wolof

The most important novelty of my analysis is the treatment of the *a*/*la*-asymmetry as T-C agreement, following Pesetsky and Torrego (2001) (P&T). The key assumption that the analysis is based on is that C_{WH} has an uninterpretable T-feature, which can be checked either by Head Movement of T to C, or by movement of the subject to Spec,CP. The reason why the subject can check *u*T on C_{WH} is that it carries Nominative Case, which Pesetsky and Torrego (2001) argue is *u*T on D.

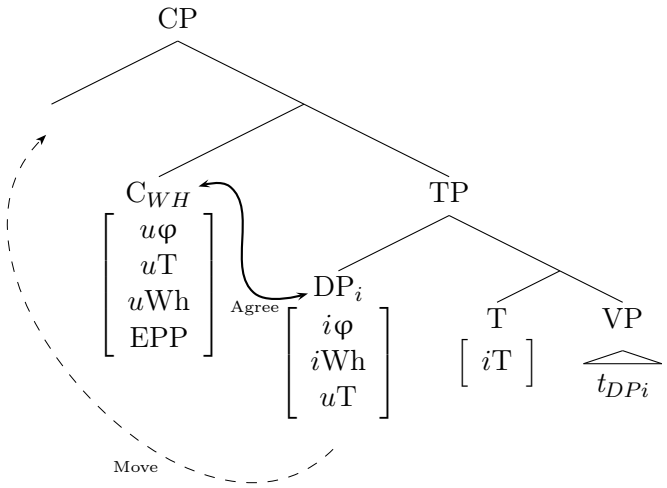
The tree in (49) illustrates agreement relationships that C_{WH} enters into in non-subject extraction. In order to value its uninterpretable φ- and Wh-features, it establishes agreement with the non-subject phrase in its c-command domain, which has interpretable φ- and Wh-features. This phrase also moves to Spec,CP, because C_{WH} has an EPP feature. In order to check its uninterpretable T-feature, C_{WH} establishes Agree with T, which then undergoes Head Movement to C_{WH}.

(49) *C's Agreement Patterns in Non-Subject Extraction*



In case of subject extraction, as in (50), the agreement pattern is somewhat different. C$_{WH}$ agrees with the subject phrase in order to check [uφ] and [uWh], but since the subject phrase also carries a valued [iT], it can value [uT] on C$_{WH}$.

(50) *C's Agreement Patterns in Subject Extraction*



P&T's unified analysis of the T-to-C asymmetry and the *that*-trace effect in English rests on two assumptions: (i) T-to-C movement is motivated by an uninterpretable T feature (uT), with an EPP feature,[25] on C, and (ii) Nominative case is uT on D. The relevant principles for the analysis are the following:

[25]Meaning that the uninterpretable feature must be valued via both Agree and Move.

1. Attract Closest (Chomsky, 1995): only the closest constituent can be attracted.
2. Head Movement Generalization: the movement from a complement to the nearest head is always realized as head movement.
3. Principle of Minimal Compliance (Richards, 1997): a constituent that is farther away may be extracted, if an element that complies with Attract Closest has already moved.

I proposed that C_{WH} in Wolof, in addition to having an uninterpretable Wh- and φ-features, also has a *u*T feature. Adopting P&T's assumption that nominative case is *u*T on D, we expect the sentence in (51a) to have the structure in (51b), before the movement of the focused phrase:

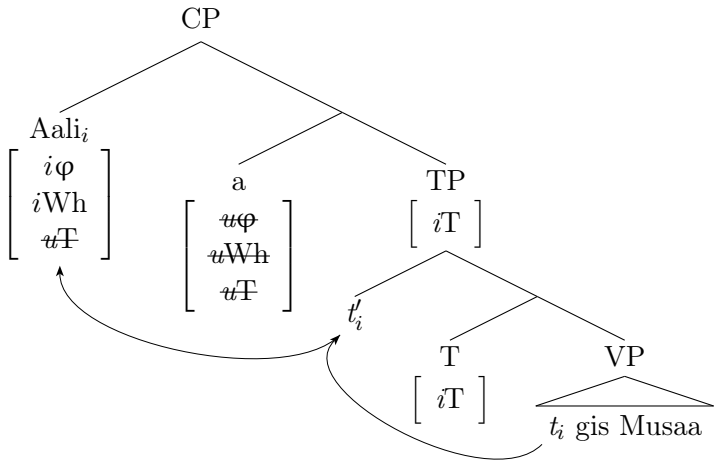
(51) *Subject extraction*

a. Aali **a** (>Aalee) gis Musaa.
Ali C_{WH} see musa
"ALI_{FOC} *saw Musa.*"

b. [$_{C}$ $a_{uT,uWh,u\varphi}$] [$_{TP}$ $Aali_{uT,iWh,i\varphi}$ *i*T [$_{VP}$ gis Musaa]]

The subject in (51b) is already attracted to Spec,TP by T's need to check its uninterpretable φ-features. *u*T on the subject is also marked for deletion by agreement with *i*T on T; however, this feature may remain undeleted until the end of the CP cycle, and be accessible to further operations. P&T explain the lack of T-to-C movement in subject extraction in the following way. C bears *u*Wh and *u*T, with an EPP feature. TP and its nominative specifier both count as the closest constituent to C, so, in principle, C can choose to delete its *u*T feature by attracting TP (realized as Head Movement), or by attracting the specifier. Since the nominative phrase already moves to Spec,CP to check *u*φ and *u*Wh, *u*T can also be deleted. The Economy Condition prevents unnecessary movement to take place, and bans T-to-C.

(52) *Subject extraction as uT valuation*



On the other hand, if a non-subject is extracted, as in (53), the extracted constituent has only the iWh and $i\varphi$ features. The structure after C has merged with TP is shown in (53b).

(53) *Object extraction*

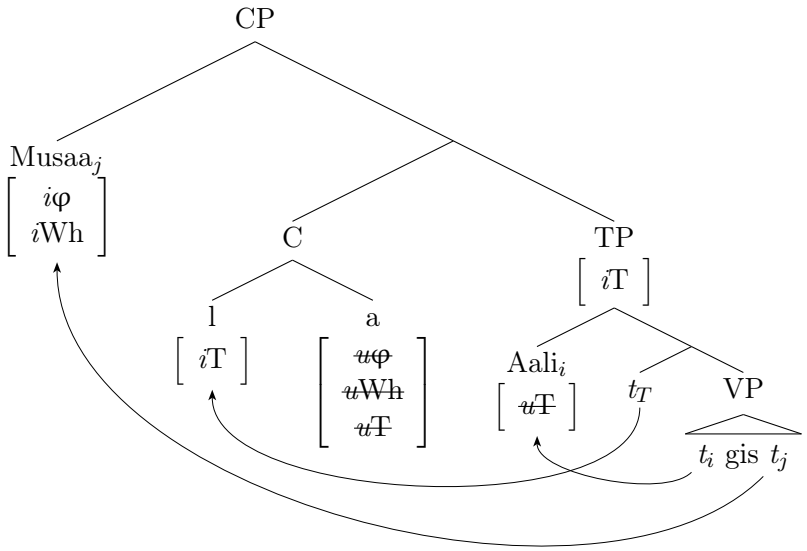
a. Musaa **l-a** Aali gis
musa *l*-C$_f$ ali see
"*Ali saw MUSA$_{FOC}$.*"

b. [$_{\text{C}}$ a$_{uT,uWh,u\varphi}$] [$_{\text{TP}}$ Aali$_{uT}$ iT [$_{\text{VP}}$ gis Musaa$_{iWh,i\varphi}$]]

The closest constituent that can value an uninterpretable feature on C is the TP or its specifier, so Attract Closest forces C to delete one of its uninterpretable features (uT) by attracting the closest constituent. In Wolof, the only option is to attract the TP, which according to the Head Movement Generalization is realized as Head Movement of T to C. This does not value uWh on C, which can only be deleted by attracting the object DP. Therefore, C needs two movement operations to delete all of its uninterpretable features. The key to this proposal is that *l-* is the spell-out of T that has moved to C, as shown in (54).[26]

[26]Overt material can still occur in T, in particular the past tense marker *-oon*. This analysis therefore assumes two things. First, *l-* is the Spell-Out of T in C, and not of T in T. Second, provided that *-oon* is generated in T, it presupposes that one feature can be spelled out in two positions. Multiple expression of a single element in different structural positions, as a result of A-movement, has been proposed, for example, for Clitic Doubling (Anagnostopoulou 2003; Harizanov To appear), and is similar to analyses of resumption in languages where the resumptive pronoun behaves as a phonetically realized trace of movement (Engdahl 1985; Demirdache 1991).

(54) *T-to-C in object extraction*



We now have an account of the *a/la*-asymmetry as a T-to-C asymmetry: the result of the ability of the subject phrase to check uT on C, due to the fact that it carries Nominative Case. Seeing how C in local subject extraction differs from C in any other position by not having *l-* prefixed to it, the parallel between the function that *l-* performs and the function that an extracted subject in Spec,CP performs is hereby accounted for.

A question that needs addressing is why C in sentences like (53) cannot choose between Spec,TP and TP (i.e. its head), to delete uT. In other words, why is the sentence in (55) not a possible way to focus an object?

(55) *Subject movement to Spec,CP in object extraction*
*Musaa Aali **a** (<Aalee) gis
Musa Ali C_{WH} see
intended:"*Ali saw MUSA$_{FOC}$.*"

From the ungrammaticality of (55), the following appears to be true of the C_{WH} in Wolof: (i) the phrase carrying the *Wh*-feature must occupy the specifier of C_{WH}, and (ii) C_{WH} has only one specifier position.[27] In other words, if the subject moved to Spec,CP in order to delete uT, uWh would remain unchecked because no other phrase could move to Spec,CP.[28]

[27]This involves an implicit claim that the number of EPP subfeatures (here all the relevant features have an EPP subfeature) does not necessarily correspond to the number of specifier positions. At this point, this is only a stipulation which needs to be further tested.

[28]See Pesetsky and Torrego (2001) for arguments that this is a possible derivation in English, with consequences for interpretation.

This analysis improves on Pesetsky and Torrego (2001) by finding an example of a language in which both C and T which moves to C are overt. A weakness of their evidence lies in the fact that in the languages they discuss (Standard English, Belfast English, Spanish), C is always phonologically null, and only T is overt. If my analysis is on the right track, the evidence for a T-to-C analysis of the *that*-trace effect is considerably strengthened.

Let us now turn to long distance movement. The occurrence of *(l)a* in C of embedded clauses is straightforwardly explained by assuming that the extracted element passes through the Spec,CP of each embedded clause. If *(l)a* is the spell-out of a complementizer that carries a Wh-feature, its presence in C of every embedded clause is necessary for the A′-moved element to be fronted to the beginning of the sentence. The example in (56) illustrates the extraction of an embedded focused subject.

(56) *Long-distance subject extraction*

a. Aali **l-a**-a gëm ni **l-a** Musaa xalad ni mu-**a** leen gis
ali *l*-C$_f$-1SG believe that *l*-C$_f$ musa think that 3SG.SBJ-C$_f$ 3PL.OBJ see
"*I believe that Musa thinks that ALI$_{FOC}$ saw them.*"

b. $[_{CP}$ [Aali$_{iWh}$]$_i$ l$_{iT}$ a$_{uT,uWh}$ a t_T gëm ni

$[_{CP}$ t$_i$' l$_{iT}$ a$_{uT,uWh}$ Musaa t_T xalad ni $[_{CP}$ mu$_i$ a$_{uT,uWh}$ t$_i$ leen gis]]]

In (56), the subject first moves from inside the VP to Spec,TP, to check $u\varphi$ on T. At the same time, the uninterpretable T feature on the subject is checked and marked for deletion via Agree with T. However, it is not immediately deleted and remains available for further operations within the same cycle. Next, the subject moves from Spec,TP to Spec,CP in order to check both uT and uWh on C. uT on the subject now has to be deleted, since the phase has ended. The C of the next higher clause needs to delete its uT and uWh. The closest constituent that it can attract to delete its uT is TP, resulting in T-to-C, which surfaces as *l-* preceding *a*. The phrase that carries iWh is the extracted subject located in the lower Spec,CP, which is now attracted to the higher Spec,CP. This analysis explains why *l-* occurs in every C, except the one where the local subject is extracted, and accounts for the locality of the subject/non-subject asymmetry – it is only there that the subject can delete the uninterpretable T feature on C.

Another thing to notice is the occurrence of a subject pronoun *mu* in lieu of the extracted subject in the most embedded clause in (56). I assume that this has to do with the phonological status of *a*, which is a clitic and thus cannot stand on its own. I propose that *a* has to lean on an element within its maximal projection. This is supported by the fact that *a* always forms a phonological unit with the element in its specifier, causing phonological changes in the form of vowel coalescence if that element happens to end in a vowel, as discussed in §2 (examples (10a) and (10b)). In long distance extraction, a trace occupies the specifier of the most embedded *a*. In order to provide it with a host, a subject pronoun is pronounced in the position of the trace. The fact that *a* cannot lean on an element outside its maximal projection, for example on the embedding complementizer *ni*, can be related to phases: at the point when *a* is spelled out and searches for a host, *ni* has not been spelled out and cannot provide phonological support for *a*.

In this section I have offered an analysis of the subject/non-subject asymmetry in Wolof A′-movement constructions by analyzing it as a T-to-C asymmetry along the lines of Pesetsky and Torrego 2001. I argue that *l*-, which precedes *a* in all instances, except at the local subject extraction site, is T that has moved to C in order to delete the uninterpretable T feature on C. T-to-C does not occur in local subject extraction due to nominative case being *u*T on D, and as such capable of deleting *u*T on C by moving to its specifier. Since in those cases the subject also deletes *u*Wh on C, T-to-C movement is unnecessary, and banned by the Economy Condition. In case of extraction of any other element, T-to-C movement must take place, because the extracted phrase does not bear nominative case (i.e. *u*T), or is not close enough to be attracted by the complementizer.

In this section, I have investigated A′-extraction effects exhibited by C's two allomorphs, CM-*u* and *(l)a*, and shown that they can be reduced to agreement between C_{WH} and other elements of the clause, and that there the syntactic operations Agree and Move resulting in the A′-extraction effects behave completely orderly. Table 2 repeats the distribution of the two C_{WH} allomorphs in the relevant constructions in the variety of Wolof discussed in this paper.

		(l)a	CM-**u**
HIGHEST C	questions	√	√
	focus	√	*
	relative clauses	*	√
INTERMEDIATE C		√	*

Table 2: The distribution of allomorphs CM-*u* and *a* in Wolof.

Two puzzles remain to be solved. First, why do the subject/non-subject asymmetry and agreement in φ-features not surface simultaneously? Second, what determines their distribution in different constructions, i.e. why can C in questions surface as either of the complementizers, in relative clauses only as CM-*u*, and in focus constructions and all embedded C positions only as *(l)a*? I propose an answer to these questions in the following section.

5 OCP Effects and C_{WH} Allomorphy

In §3 we investigated the similarities in behavior and syntactic properties between CM-*u* and *(l)a*: they occur in A′-movement constructions, they occupy a position immediately dominating the TP, and they cannot co-occur with any other sentence particle, including each other. Furthermore, there is no difference in meaning in the environment in which they can both occur – matrix *wh*-questions. All these properties are consistent with the view that CM-*u* and *(l)a* are allomorphs of the same complementizer, C_{WH}. I propose that C_{WH} in Wolof has the feature specification in (57) and discuss the details of the analysis in §4: the C head needs to establish agreement both in φ-features, and with T, to check its uninterpretable φ- and T-features.

(57) *The feature specification of C_{WH}*

$$\begin{array}{c} C_{WH} \\ \left[\begin{array}{c} u\mathrm{Wh}_{\mathrm{EPP}} \\ u\varphi_{\mathrm{EPP}} \\ u\mathrm{T}_{\mathrm{EPP}} \end{array}\right] \end{array}$$

The analysis in §4 claims that the syntax of A′-movement is simple, and that some of the surface complexity comes from the agreement relations in the CP-layer. However, we have seen that the two allomorphs of C_{WH} do not each overtly exhibit all agreement features: CM-*u* shows Wh- and φ-agreement, and *(l)a* Wh- and T-agreement. If the difference between constructions with CM-*u* and those with *(l)a* is not structural, and does not seem to correlate with a particular feature specification, and if they both encode A′-movement, then the first task of any analysis is to offer a satisfactory explanation of their distribution.[29] This is the purpose of this section, where I uncover another stratum responsible for some of the surface complexity of the Wolof CP-layer—the post-syntactic one—and show that the operations in the post-syntax of A′-movement structures in Wolof

[29] Analyses which assume they are syntactically different, such as Torrence (2005) (and his subsequent work), do not have an account of their distribution.

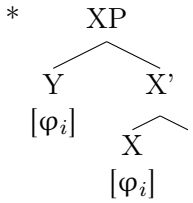
are not unique, or in any way disorderly, but occur in similar morphosyntactic configurations in various languages. When the post-syntactic component is properly described, and its interaction with the syntactic processes understood, Wolof A′-movement looks very ordinary.

In particular, I claim that the surface form of C_{WH} depends on the presence or absence of the φ-feature (i.e. the class marker) in C, which in turn depends on the presence or absence of the φ-feature in Spec,CP. The data show that the φ-feature is overt in only one position in the CP-layer: either in the C head, or in its specifier. If it is overt in C, the complementizer surfaces as CM-*u* and Spec,CP is null; if Spec,CP is overt (and therefore contains the φ-feature), the complementizer surfaces as *(l)a*. Restrictions on the co-occurrence of the *wh*-phrase in Spec,CP and the complementizer in C are well known from English, German, French etc., as the Doubly-Filled-Comp-Filter (Chomsky and Lasnik 1977). The Wolof case is similar, but it is only in constructions where the C_{WH} allomorph overtly exhibits φ-feature agreement that the *wh*-phrase is obligatorily absent from Spec,CP; if the A′-movement structure contains the allomorph *(l)a*, both C and Spec,CP are overt. I argue that the restriction on φ-feature co-occurrence in the specifier-head configuration in A′-movement is the result of two post-syntactic constraints: the Obligatory Contour Principle, which prohibits adjacent identical features, and Recoverability, which prohibits the deletion of unrecoverable material. Wolof also gives support for a particular type of post-syntactic impoverisment, called Obliteration (Arregi and Nevins 2007, 2012; Calabrese 2010; Pescarini 2010), in which entire nodes are deleted from the surface representation. This section is dedicated to spelling out the details of this analysis.

5.1 OCP_{φ}

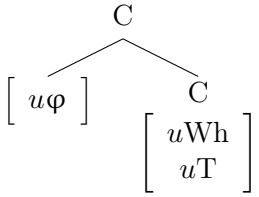
The φ-feature in the CP layer always surfaces overtly only in one place: either in the specifier, or on the complementizer. I pursue the idea that this observation is the key to understanding the distribution of CM-*u* and *(l)a* and propose that there is a morphological Obligatory Contour Principle (OCP) constraint in Wolof which prohibits identical φ-features to surface in a specifier-head configuration, as in (58).

(58) *Morphological Obligatory Contour Principle constraint in Wolof*

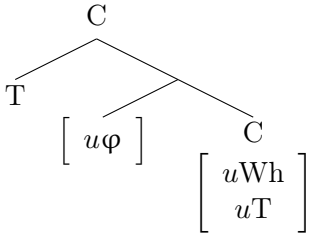


The repair to the OCP violation is to delete the φ-feature node in C, or to delete the entire phrase in Spec,CP. I furthermore propose that φ in C is realized as a separate node in the post-syntax, as CM-*u* is bi-morphemic. In subject extraction, *u*T is valued via subject movement to Spec,CP, but in non-subject extraction, which leads to T-to-C movement (see §4), T is adjoined to C above φ. The structure of C in subject and non-subject extraction is represented in (59) and (60), respectively.

(59) *C in subject extraction*



(60) *C in non-subject extraction*



I also propose the following set of Vocabulary Insertion rules. In the environment of φ, C surfaces as *u*, and as *a* elsewhere. T surfaces as *l* when adjacent to C.

(61) *Vocabulary Insertion Rules, first version*[30]

a. $C_{WH} \rightarrow u / \varphi\_\_$

b. $C_{WH} \rightarrow a$

c. $T \rightarrow l / \_\_C$

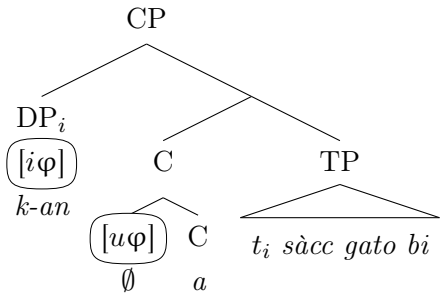
When the extracted element is merged in the final Spec,CP, there are two possible resulting configurations. We start with A′-extraction in questions, in which the complementizer can surface both as *(l)a*, in (62), and as CM-*u*, in (63).

[30]The Vocabulary Insertion rules will need to be made slightly more precise to account for versions of the complementizer in relative clauses.

(62) *Overt φ-feature in Spec,CP*

a. K-an a sàcc gato bi?
CM-Q C_{WH} steal cake DEF
"Who stole the cake?"

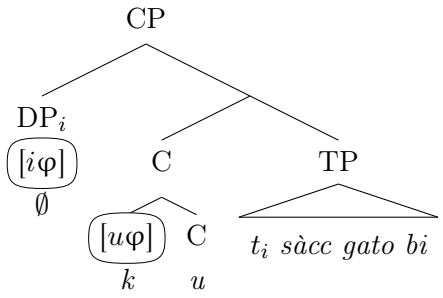
b.



(63) *Overt φ-feature in C*

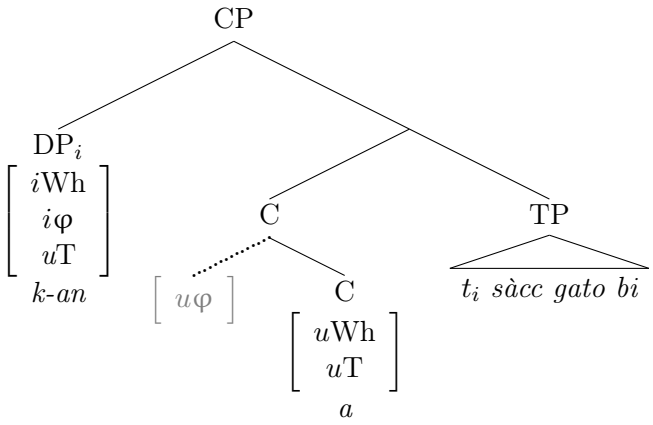
a. K-u sàcc gato bi?
CM-C_{WH} steal cake
"Who stole the cake?"

b.

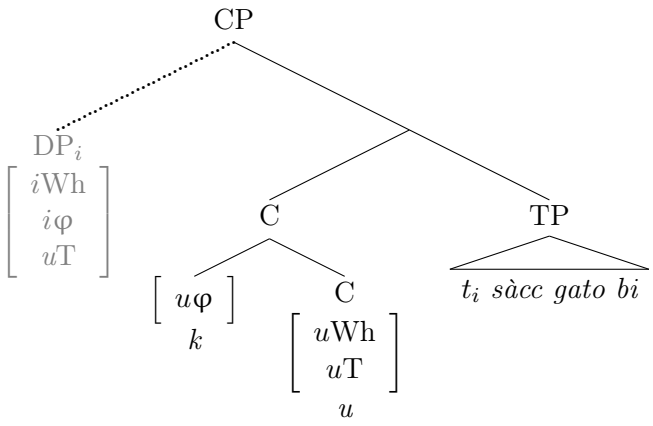


The φ-feature in these constructions occurs in two places – in C, and in the *wh*-word in its specifier – however, it is overt either in one or in the other position. I argue that this is the result of an Impoverishment rule, which militates against structures offending the OCP_{φ} constraint in (58) by deleting one of the nodes containing the φ-feature. One option is to delete the φ-feature in C, which is realized as a separate node in post-syntax. In that case, as per the Vocabulary Insertion rules, C surfaces as *a*, illustrated in (64a). The *wh*-phrase in this case is overt and has the form CM-*an*. The second option is to delete the *wh*-phrase. In that case φ in C is present and conditions allomorphy: the complementizer is realized as *u*, shown in (64b). The examples illustrate subject extraction, when *u*T is valued by the subject phrase itself, as explained in §4.

(64) a. *Subject a-question*



b. *Subject u-question*



Impoverishment was initially proposed as a rule that deleted features (Bonet 1991, 1995), and in some approaches it is quite constrained.[31] There are, however, proposals in the literature that argue that post-syntactic deletion can target not only features, but also entire nodes (Arregi and Nevins 2007, 2012; Calabrese 2010; Pescarini 2010). For example, Arregi and Nevins (2007, 2012) show that Basque dialects systematically avoid the combination of 1*st* and 2*nd* person clitics (characterized as a dissimilation in the [Participant] feature), but there is great dialectal variation in how this violation is repaired. Some dialects delete only the person feature, changing, for example, 2*nd* person singular ergative clitic to the 3*rd* ergative, which results in a less marked feature-value. There are, however, also instances of entire clitics being deleted from the surface representation. Such deletion, which results in the removal of the entire node, is referred to as Obliteration. In Wolof, where the entire phrase is deleted from Spec,CP, gives support for this type of impoverishment. Additional evidence for the existence of Obliteration in Wolof comes from dialectal variation in *wh*-questions with a complex *wh*-phrase in Spec,CP, discussed in §5.2.

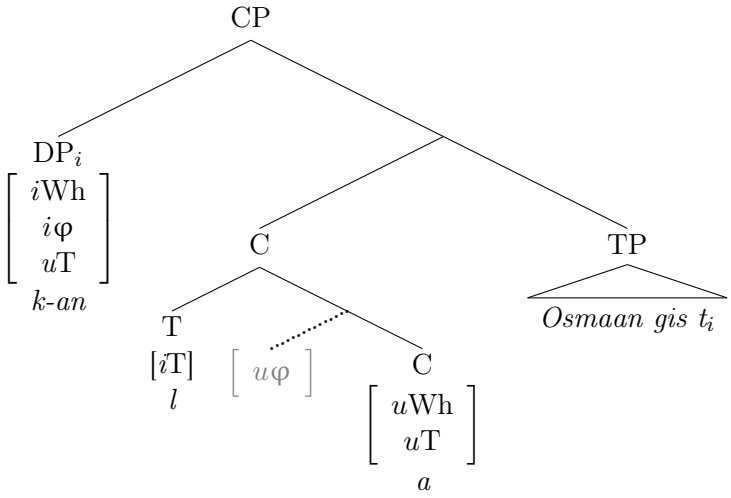
Let us next look at non-subject extraction. When a non-subject moves to Spec,CP, uT is valued by T-to-C movement, in which case T adjoins to C. If the OCP$_\varphi$-repair deletes the node containing the φ-feature in C, as in (65), T is realized as *l*. I consider this to be the result of universal adjacency conditions on contextual allomorphy (Embick 2010): *l* is a contextual realization of T, triggered by its adjacency to C. If, on the other hand, the OCPφ-triggered repair deletes the phrase in Spec,CP, φ surfaces in C and as a result T is not overt, as it is not adjacent to C, as illustrated in (66).

[31] For example, Noyer (1997) argues that, if two features, A and B, are targeted by a constraint such as *[A, B], only the feature lower on the feature hierarchy can delete.

(65) *Non-subject a-question*

a. K-an l-a Osmaan gis?
CM-Q *l*-C$_{WH}$ Osman see
"Who did Osman see?"

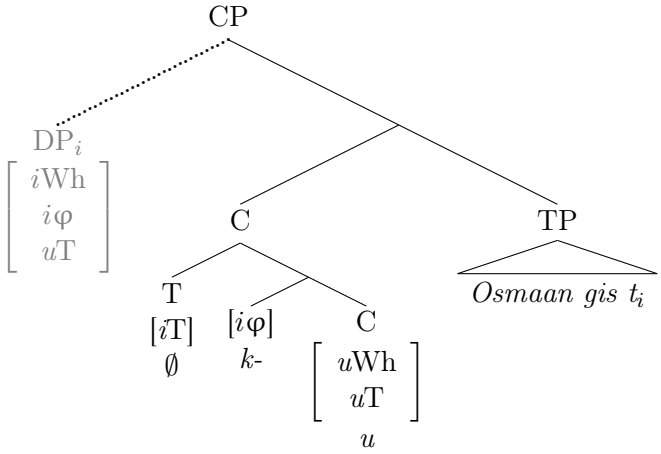
b.



(66) *Non-subject u-question*

a. K-u Osmaan gis?
CM-C$_{WH}$ Osman see
"Who did Osman see?"

b.



Under the analysis developed here, we do not need to stipulate that the complementizer CM-*u* selects an empty operator, and that the empty operator only occupies the specifier of CM-*u*. A post-syntactic analysis offers a natural explanation for the occurrence of the φ-feature only in one position in the CP-layer, attributing it to an Impoverishment rule with the purpose of avoiding a marked structure banned by a morphological OCP$_\varphi$.

OCP_{φ} was first proposed in phonology (Leben 1973; Goldsmith 1976) as a constraint that prohibits adjacent identical elements at the melodic level. It was initially an analysis for tonal dissimilation in African tone languages, but was later extended to account for all kinds of dissimilation phenomena concerning adjacent segments or features on the same (autosegmental) tier (e.g. McCarthy 1986). Similar phenomena (referred to as morphological dissimilations, haplologies, repetition avoidance, morphological OCP, etc.) have been identified to occur in the mapping between syntax and phonology, prohibiting adjacent identity (in form and/or content of morphemes) in particular morphosyntactic configurations. For example, Nevins (2007) analyzes the spurious *se* in Spanish as the result of dissimilation of adjacent clitics both bearing the feature [-Participant]. Ackema and Neeleman (2004) (also Benmamoun and Lorimor (2006)) argue that the post-syntactic allomorphy rules delete identical features of terminal nodes contained within the same prosodic domain in subject-verb agreement in languages such as Dutch and Arabic, resulting in *weakened* agreement on the verb: in Dutch the verb does not agree with the second person subject and exhibits first person agreement, which is explained as the dissimilation in the feature addressee [Add], and in Arabic the agreement is in person and gender, but not in number. Crucially, both these accounts rely on some type of structural adjacency. Similarly, in Wolof an OCP_{φ} markedness constraint prohibits φ-features in adjacent nodes (in a Specifier-Head configuration), as in (58).

We now have an account of the two versions of *wh*-questions with simple *wh*-phrases, which are the only A′-movement construction which can occur with either of the A′-complementizer allomorph, CM-*u* and *(l)a*. In all other constructions, only one of the allomorphs is allowed. In the following sections, I offer an account of the distribution of CM-*u* and *(l)a* in other constructions.

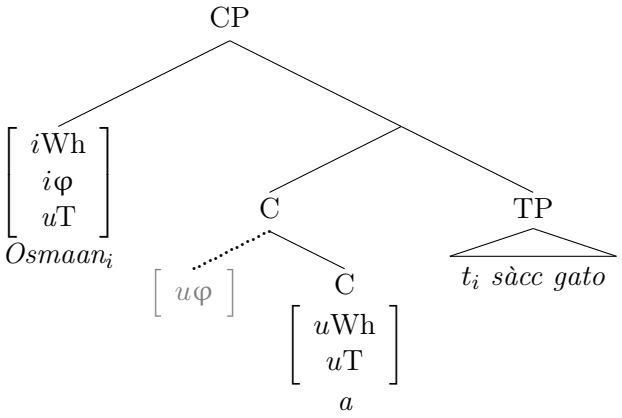
5.2 Recoverability

The only option of avoiding an OCP_{φ} violation in focus constructions is to delete the φ-feature in C, causing those structures to always surface with *(l)a*, as in extraction of a focused subject illustrated in (67), and of a focused non-subject in (68).

(67) *Subject focus*

a. Osmaan$_i$ a t_i sàcc gato.
Osman$_i$ C_{WH} t_i steal cake
"OSMAN$_{FOC}$ stole a cake."

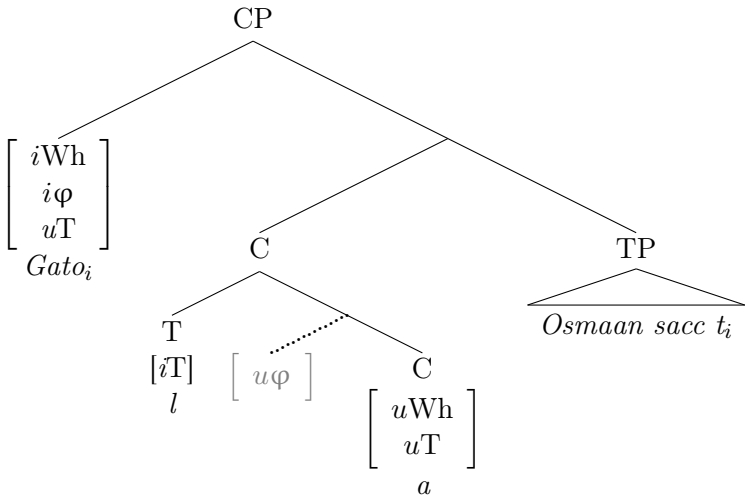
b.



(68) *Non-subject focus*

a. Gato l-a Osmaan sàcc.
cake *l*-C$_{WH}$ Osman steal
"Osman stole a CAKE$_{FOC}$"

b.



I argue that the obligatoriness of *(l)a* in all constructions in which Spec,CP is occupied by a lexical DP that is not a simple *wh*-phrase is due to an important constraint on deletion – Recoverability. Namely, a node can only be deleted if there is no unrecoverable material that gets deleted along with it. The notion of deletion under recoverability dates back to Chomsky and Lasnik (1977), a similar constraint is proposed by Pesetsky (1998), in an OT-approach to the pronunciation of complementizers, and in Ackema and Neeleman (2004), the suppression (i.e. deletion, impoverishment) of a morphosyntactic feature is also subject to a notion of recoverability: the target of the rule and the terminal mentioned in the rule's context must agree. The notion of recoverability under agreement is particularly applicable to the case under discussion: the φ-feature is deleted from one of the nodes in an agreement configuration. Focus structures can only contain the allomorph *(l)a*

because the DPs in their Spec,CP contain irretrievable material and can therefore not be deleted, whereas the featural content of the *wh*-operator can easily be retrieved from C_{WH}, which contains all of the same features.

Recoverability also explains the pied-piping data from §3.3: why complex *wh*-phrases can occur only in the specifier of *(l)a*, as in (69).

Complex wh-questions can only contain (l)a

(69) a. B-an xale l-a Faatu gis?
CM-Q child *l*-C Fatou see
"Which child did Fatou see?"

b. *B-an xale b-u Faatu gis?
CM-Q child CM-C Fatou see

The obliteration of the phrase in Spec,CP is blocked, because it contains irretrievable material: the noun phrase. The deletion of the specifier node would result in the deletion of the entire DP, as in (70).

(70) *Obliteration of Spec,CP with a complex wh-phrase*
*B-u Faatu gis?
CM-C_{WH} Fatou see
intended: "Which child did Fatou see?"

The ungrammaticality of (70) is in my analysis the result of the target of the Impoverishment rule, which is the whole phrase in Spec,CP containing the offending feature. However, there seems to be dialectal variation in the amount of material from a *wh*-phrase that is allowed to remain in Spec,CP. According to Torrence (2012a), for some speakers a null *wh*-phrase can pied-pipe an NP to Spec,CP in questions:

(71) *Null wh-phrase in a question with* CM-*u*
%Picc **m-u** xale y-i dáq?
bird CM-C_{WH} child CM.PL-DEF chase
"Which bird did the children chase?"

All speakers from Saint Louis and Dakar that I have consulted judge this to be ungrammatical as a question. However, my analysis can be easily extended to account for this data. Dialectal variation in Impoverishment rules is not unusual, as mentioned in the case of Basque (Arregi and Nevins 2007, 2012). Similarly, in the variety of Wolof that Torrence reports, Obliteration would target only the *wh*-word, and not the entire phrase in Spec,CP.

There is a case in which the specifier of CM-*u* can contain some overt material. Recall from §3.3 that the locative preposition *ci/si* can be pied-piped to Spec,CP, and both CM-*u* and *(l)a* can occupy C_{WH} (Torrence 2012a):

(72) *Preposition pied-piping in Wolof*

a. {Ci fan}/{Fan ci} la ñu teg téeré bi?
{P-LOC where}/{where P-LOC} *l*-C_{WH} 3PL put book DEF.SG
"On what did they put the book?"

b. Ci l-u ñu teg téeré bi?
P-LOC CM-C_{WH} put book DEF.SG
"On what did they put the book?"

This shows that the Obliteration rule targets only the wh-DP, which contains the φ-features. Other elements, such as prepositions, can be overt in Spec,CP of CM-*u*. Again, variation in the amount of material targeted by a particular rule is common in these types of morphological repairs: in C_{WH}, only the φ-feature (albeit contained in its own node) is deleted, in Spec,CP, the whole wh-phrase containing the φ-feature must go, but a preposition in the PP which got pied-piped by the *wh*-phrase is not in danger – still only the *wh*-phrase deletes.

I have now offered an analysis for the occurrence of both allomorphs in *wh*-questions, and the obligatoriness of *(l)a* in focus constructions. The distribution of CM-*u* and *(l)a* in A′-movement constructions in Wolof is repeated here in Table 3.[32]

		(l)a	CM-**u**
HIGHEST C	questions	√	√
	focus	√	*
	relative clauses	*	√
INTERMEDIATE C		√	*

Table 3: The distribution of allomorphs CM-*u* and *(l)a* in Wolof.

In the remainder of this section, we look at very different data, which may seem to contradict the proposed analysis. First, relative clauses are only possible with the complementizer CM-*u*, meaning that we have to explain why the relative operator, which should contain all retrievable material, must obligatorily be deleted in relative clauses; in other words, why relative CPs do not behave like

[32]Recall from Table 1 that two other A′-movement structures, comparative constructions and copular sentences with nominal predicates, also contain *(l)a*. As it was mentioned in footnote 7, other syntactic differences aside, their CP-layers are identical in that they involve movement of a non-*wh* DP to Spec,CP. Recoverability therefore prohibits their deletion, and such constructions also surface only with *(l)a*.

interrogative CPs. The situation in long distance extraction is quite the opposite: since only the complementizer *(l)a* can occur in intermediate C positions, it is φ in C that must delete. This is also surprising, because the feature content of a copy/trace is not expected to differ from that of the extracted phrase. Intermediate positions should thus behave just as final landing positions with respect to OCP_{φ} and Recoverability. My analysis can explain both cases.

5.3 Intermediate traces

In long-distance extraction in Wolof, *(l)a* obligatorily occupies all intermediate C positions. In the variety of Wolof discussed in this paper, this is the only option, meaning that the OCP_{φ} violation in a non-final CP-layer can only be avoided by deleting the φ-feature in C, regardless of the nature of the element in Spec,CP. I argue that this is the result of the timing of the Spell-Out, which in those dialects precedes movement out of the spelled out domain. In successive-cyclic movement the derivation in which the phrase in Spec,CP is deleted crashes, because there is nothing left to be attracted by the higher C. As was mentioned in footnote 18, Torrence (2005, 2012a,b) also claims that there are dialects in which CM-*u* can occupy intermediate C positions. While I have not been able to confirm such data, my analysis can be modified to account for it by timing movement before Spell-Out in those dialects.

It is fairly standard to assume that Spell-Out happens cyclically, as a result of the merger of particular heads (*phase heads*), that trigger Spell-Out of their complement (a *phase*) (Chomsky 2000). The notion of phases is particularly relevant in considerations of long-distance movement, as it is mostly assumed that such movement can only happen through specific positions on the edges of phases, which results in cyclicity effects in long-distance extraction. We have seen that Wolof is one of the languages that provides evidence for such movement, since it requires the A′-movement complementizer to occupy every intermediate C position along the path of extraction. Since the analysis of complementizer allomorphy proposed here is a post-syntactic one, meaning that the described processes take place during Spell-Out, we need to consider the predictions it makes for long-distance movement.

In much of the literature on long-distance extraction, it is assumed that a phase is impenetrable to further movement (Chomsky 2000). I follow Fox and Pesetsky (2005) and propose that movement can happen either before or after Spell-Out, provided that the linear order within a phase established

during Spell-Out is preserved after movement. This crucially means that some syntax can happen after post-syntax. Let us examine the predictions of this proposal. Consider the sentence in (73).

(73) *Long-distance object extraction in Wolof*
[$_{CP2}$ K-an$_i$ l-a Isaa wax ne [$_{CP1}$ t_i l-a xaj bi matt t_i]]?
[$_{CP2}$ CM-Q *l*-C$_{WH}$ Isaa say that [$_{CP1}$ t_i *l*-C$_{WH}$ dog DEF bite t_i]]
"Who did Isaa say that the dog bit?"

I first consider the option in which movement follows Spell-Out, as this is the variety of Wolof described in this paper. We start at the moment when the Spell-Out of the edge of the embedded CP-layer (CP1) is triggered by a higher phase head,[33] and the OCP$_\varphi$ evaluates the resulting construction. This proceeds in the way explained above: either φ in C or the whole Spec,CP node can delete, resulting in the complementizer surfacing as either *(l)a* or CM-*u*, respectively:

(74) *Deletion of φ in C, Step1*
[$_{CP2}$ C ... [$_{CP1}$ k-an$_i$ l-a xaj bi matt t_i]]
[$_{CP2}$ C ... [$_{CP1}$ CM-Q *l*-C$_{WH}$ dog DEF bite t_i]]
"who did the dog bite?"

(75) *Deletion of Spec,CP, Step1*
[$_{CP2}$ C ... [$_{CP1}$ ∅ k-u xaj bi matt t_i]]
[$_{CP2}$ C ... [$_{CP1}$ ∅ CM-C$_{WH}$ dog DEF bite t_i]]
"who did the dog bite?"

Since in this derivation movement follows Spell-Out, the *wh*-phrase in CP1 is only attracted to the Spec,CP2 after the described post-syntactic processes have taken place. If the φ-feature in C is deleted, as in (74), the *wh*-phrase is present in Spec,CP1 and can move to Spec,CP2, as in (76). The Spell-Out of the CP2 layer proceeds in the same way, with two possible outcomes:

(76) *Deletion of φ in C, Step2*
[$_{CP2}$ {K-an$_i$ l-a}/{K-u} Isaa wax ne [$_{CP1}$ t_i l-a xaj bi matt t_i]]?
[$_{CP2}$ {CM-Q *l*-C$_{WH}$}/{CM-C$_{WH}$} Isaa say that [$_{CP1}$ t_i *l*-C$_{WH}$ dog DEF bite t_i]]
"Who did Isaa say that the dog bit?"

If, however, the phrase in Spec,CP of CP1 is deleted as a result of OCP$_\varphi$ repair, as in (75), there is no *wh*-phrase left in the specifier of CP1 to move to Spec,CP2 in a dialect in which Spell-Out precedes movement. This derivation therefore crashes (in (77)), and the derivation in which φ in C

[33] For simplicity of exposition, I assume that it is the higher C$_{WH}$ head that triggers Spell-Out of the edge of the embedded CP phase.

is deleted is the only one that converges, meaning that, if movement happens after Spell-Out, *(l)a* is the only exponent that can surface in intermediate positions.

(77) *Deletion of Spec,CP, Step2*
*[$_{CP2}$ __ C Isaa wax ne [$_{CP1}$ ∅ k-u xaj bi matt t_i]]
[$_{CP2}$ __ C Isaa say that [$_{CP1}$ ∅ CM-C$_{WH}$ dog DEF bite t_i]]

Let us now investigate the second option, in which movement occurs before Spell-Out. In that case, in the moment of Spell-Out of CP1, Spec,CP contains the copy of the *wh*-phrase which has all of the same features as the phrase which is moved into the higher Spec,CP. Again, deletion of φ in C or the phrase in Spec,CP is governed by Recoverability, meaning that only a *wh*-operator could be deleted, and never a full DP. Unlike in the case in which movement occurs after Spell-Out, in this case the phrase from Spec,CP is already located in the higher Spec,CP, so deleting its copy does not cause the derivation to crash.[34] When movement precedes Spell-Out, intermediate C's behave just like matrix C's. The prediction is that in the dialects in which movement occurs before Spell-Out either *(l)a* or CM-*u*, as in (78), can surface in the intermediate position. However, if CM-*u* can occupy intermediate positions, this should only occur in *wh*-questions, and never in focus constructions, due to Recoverability.

(78) *Successive cyclic movement follows Spell-Out*
[$_{CP2}$ {K-an$_i$ l-a}/{K-u} Isaa wax ne [$_{CP1}$ {∅ k-u}/{t_i l-a} xaj bi
[$_{CP2}$ {CM-Q *l*-C$_{WH}$}/{CM-C$_{WH}$} Isaa say that [$_{CP1}$ {∅ CM-C$_{WH}$}/{t_i l-C$_{WH}$} dog DEF
matt t_i]]?
bite t_i]]
"Who did Isaa say that the dog bit?"

These predictions are confirmed by the data. In the variety of Wolof discussed in this paper, only *(l)a* can occupy intermediate positions of movement, meaning that in that dialect, Spell-Out precedes movement. Torrence (2005, 2012a,b) shows data in which CM-*u* is allowed in intermediate positions, alongside *(l)a* (see footnote 18), and, as expected if this analysis is on the right tract, he only reports such an option for *wh*-questions, never focus constructions. If such dialects do in fact exist, in them movement precedes Spell-Out. This analysis also accounts for the occurrence of what Torrence calls *mixed chains*, where CM-*u* and *(l)a* can interchangeably occupy embedded Spec,CP position, since in every CP, whether embedded or not, either of the two allomorphs of C$_{WH}$ can surface. Crucially,

[34]The fact that a higher copy of the move phrase exists does not affect Recoverability in the post-syntax, which only evaluates the immediate Spell-Out domain.

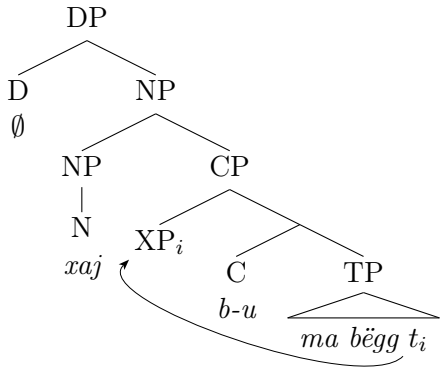
there are no reported dialects in which only CM-*u* occurs in intermediate C positions, and indeed, my analysis predicts that such dialect should not exist.

5.4 Relative clauses

Finally, let us examine the predictions of the analysis for C in relative clauses. I follow the *matching analysis* (Lees 1960, 1961; Chomsky 1965; Sauerland 1998, 2003 among others), and assume that Wolof relative clauses, as in (79a), have the syntax in (79b): a relative operator co-indexed with the head noun is located in Spec,CP, and the relative CP is adjoined to the head NP.

The syntax of Wolof relative clauses

(79) a. Xaj **b-u** ma bëgg.
dog CM-C_{WH} 1SG like
"*a dog that I like*"

b. 

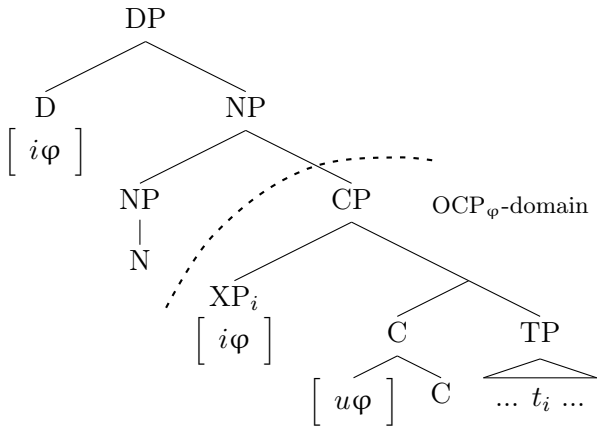
The reason for assuming a head-external representation of Wolof relative clauses is maintaining the generalization that Spec,CP of CM-*u* is always empty, as we have seen is the case in *wh*-questions. Torrence (2005, 2012a,b) advocates a structure for Wolof relative clauses in which the relativized DP occurs in Spec,CP, which would make relative clauses different from questions in that they could have an overt phrase in their specifier. He argues for such an analysis because relative clauses exhibit reconstruction effects, suggesting that the relativized DP starts out inside the clause and moves to Spec,CP, binding its trace. It is, however, not necessary for the relativized noun to be inside the relative clause to account for reconstruction effects. In the here adopted *matching analysis*, an internal head corresponding to the external head is located in Spec,CP, and then deleted under identity with the external head.[35] The representation in (79b) maintains the parallel between questions, which in Torrence's analysis have an empty *wh*-operator in Spec,CP, and relative clauses,

[35] Another possibility which can account for reconstruction effects is the *head raising* analysis (e.g. Brame 1968; Schachter 1973; Vergnaud 1974; Åfarli 1994; Kayne 1994; Bhatt 1999, 2002), according to which the head NP originates inside the relative clause, but is not necessarily located in Spec,CP in the final structure (its final position varies in different analyses). Whether one of the two analyses should be given preference in Wolof is not relevant for the present purposes, and is left for future research.

which along the same lines would have an empty relative operator in Spec,CP.

Since Spell-Out occurs in a cyclic fashion, I propose that the domain of the OCP_{φ} is the CP phase, as in (80), meaning that the φ-feature in D is not taken into consideration in evaluating markedness.

(80) *Wolof relative clause and the domain of OCP_{φ}*



The analysis developed thus far then predicts that the complementizer allomorphy in relative clauses should parallel that of matrix questions: the complementizer should surface as either CM-*u* or *(l)a*; if the former, the operator should be absent, and if the latter, it should be overt. This is, however, not what we observe: in relative clauses, only the allomorph CM-*u* is possible, and the operator is never overt.

Relative CPs, however, are not quite identical to interrogative CPs, and where they differ is in the featural content of C. Recall from the discussion in §3.2 that the φ-feature complex in relative clauses consists of the class feature, the definiteness feature, and the proximity feature. All these features are also contained on the relative operator.[36] In questions, the complementizer allomorph CM-*u* has only one variant – the one with the vowel -*u*. In relative clauses, on the other hand, CM-*u* has in fact three variants: CM-*u*, CM-*i*, CM-*a*, encoding definiteness and proximity of the head noun. The relevant examples are repeated in (81).

(81) *The relative complementizer encodes definiteness and proximity*

a. (*a-b) xaj **b-u** ma bëgg
(INDEF-CM) dog CM-C_{WH} 1SG like
"a dog that I like"

[36] As a result of whichever mechanism ensures co-indexing between the head noun and the relative operator.

b. xaj (*b-i) **b-i** ma bëgg
dog (CM-DEF.PROX) CM-C_{WH} 1SG like
"*the dog (here) that I like*"

c. xaj (*b-a) **b-a** ma bëgg-oon
dog (CM-DEF.DIST) CM-C_{WH} 1SG like-PST
"*the dog (there) that I liked*"

I proposed in §3.2 that C in relative clauses carries the definiteness-proximity feature complex, which it obtains via agreement with D. Crucially, definiteness and proximity are a part of the φ-feature complex. They are only expressed with the complementizer CM-*u*, which can be realized with three different exponents corresponding to indefinite, definite proximal, and definite distal features. The Vocabulary Insertion rules therefore need to be made more precise:

(82) *Vocabulary insertion rules, second version*[37]

a. C[+Wh] → $u/\{\varphi, -Def\}$__

b. C[+Wh] → $i/\{\varphi, +Def, +Prox\}$__

c. C[+Wh] → $a/\{\varphi, +Def, -Prox\}$__

d. C[+Wh] → *a*

e. T → *l*/__C

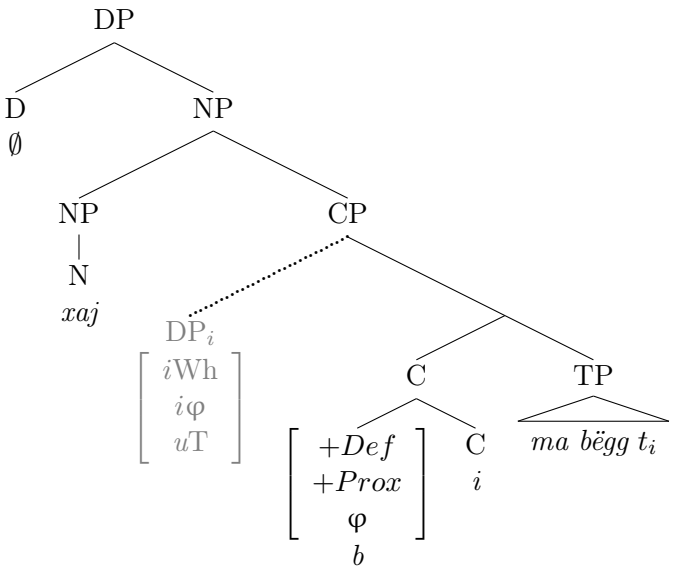
I argue that the reason for the obligatory deletion of the relative operator in Spec,CP lies precisely in the fact that the φ-feature complex in C of relative clauses does not only contain only the class and number features, but also definiteness and proximity. Recoverability therefore prevents the deletion of irretrievable material, protecting the deletion of φ in C. The only option, then, is to delete the operator in Spec,CP, which causes the φ-feature in C to be pronounced and the complementizer to surface as *-u*, *-i*, or *-a*, per the Vocabulary Insertion rules in (82).

Obligatory deletion of Spec,CP in relative clauses

(83) a. xaj b-i ma bëgg
dog CM-C_{WH} 1SG like
"the dog that I like"

[37]Technically, the Vocabulary Insertion rule in (82c) is not necessary, since the exponent for the definite distal complementizer is homophonous with the exponent for the elsewhere condition in (82d). I posit two rules to make the distinction between the two complementizer forms clearer.

b.



And finally, a note on the restriction of co-occurrence of C and D in relative clauses. In the dialect of Wolof that this paper is concerned with, D which never occurs in relative clauses. The fact that the definiteness feature does not surface twice is reminiscent of a similar phenomenon in some Scandinavian languages. In Wolof, the two heads, D and C, agree in φ-features, definiteness and proximity. As a result, the determiner and the complementizer have identical feature specifications. I propose that in such a case only one of the two heads can be pronounced, and that in this configuration in Wolof, it is the lower one. The determiner is therefore deleted. That this analysis is on the right track is corroborated by data in Torrence 2012a,b, where in some dialects the determiner can optionally surface on the edges of the relative clause, reported here in footnote 12. Such variation is expected in the scenario sketched above: in some dialects the expression of identical features in two different heads is prohibited, in others it is not. A similar restriction exists in some Scandinavian languages (Embick and Noyer 2001; Hankamer and Mikkelsen 2002, 2005, i.a.), where the definite feature can occur in two positions inside the DP – as a suffix on the noun, or as an article, depending on the presence or absence of adjectival modifiers. In some languages, for example in Danish, the definite determiner can only be expressed once. Swedish and Norwegian, on the other hand, exhibit the phenomenon of Double Definiteness, whereby the presence of a modifier requires the definite feature to surface both as an article, and as a suffix on the noun. The variation in the expression of definiteness in Wolof is similar: in some dialects it can only be expressed on one head, while other dialects allow the determiner to optionally surface on the edges of the relative

clause. The details of this proposal are left for future research. What is important to stress, is that the deletion of the determiner of the head noun in the relative clause is not the OCP_{φ}-triggered deletion that occurs in the CP-layer, and is presumably handled by a different mechanism.

In this section I argued for a post-syntactic analysis of the A′-complementizer distribution in Wolof, attributing the difference between complementizer form in different A′-movement constructions to the interplay of a morphological OCP_{φ} constraint, which prohibits adjacent identical φ-features, and a Recoverability condition on deletion, which prevents the deletion of nodes that carry irretrievable featural content. This analysis allows us to maintain a uniform account of A′-movement and two extraction effects that occur in Wolof: complementizer agreement and a subject/non-subject asymmetry.

5.5 Issues with Deletion and Recoverability

Before concluding, I address several questions related to the use of Deletion and Recoverability mechanisms in my analysis. A reviewer points out significant tension between the notion of recoverability used here, and the one known from Chomsky and Lasnik (1977). In their influential paper *Filters and control*, they explain the ability to delete the *wh*-words in English relative clauses as a result of their recoverability due to their relation with the relative clause head. Interrogative pronouns in English, bearing no relation with another head, cannot be deleted. In my analysis, however, interrogative pronouns can be deleted. If Chomsky and Lasnik are right about English, and if my analysis of Wolof is on the right track, the difference between the two languages needs to be addressed. Related to this is the question of the phenomena resulting from the constraint known as the Doubly-Filled-Comp Filter (also Chomsky and Lasnik 1977), which offered explanation on the co-occurrence restriction between the *wh*-phrase and the complementizer by requiring the obligatory deletion of the material in C.

First, with respect to the discrepancy between English and Wolof in the deletion of interrogative *wh*-phrases, one obvious difference between the two languages is in the fact that Wolof has an overt C_{WH} which agrees in φ-features with the phrase in its specifier. The *wh*-phrase's features really are recoverable from C_{WH} in Wolof. This is not the case in English. Furthermore, there is reason to believe that different types of deletion target the *wh*-phrase in Wolof interrogative and relative clauses and the *wh*-phrase in English relative clauses. Related to the recoverability in relative clauses,

in my analysis the domain of the OCP_{φ} and operations related to it is restricted to the CP-layer (see (80)) and the moment of its Spell-Out (which occurs when a higher phase head is merged). Being characterized as a post-syntactic process, the presence of matching features on D has no bearing on the recoverability within the CP-layer, as the higher phase is not yet spelled out. If the *wh*-phrase in English relative clauses can be deleted precisely because of its identity with the external head, it would appear that deletion of elements can happen during different stages of the derivation, and in different modules of the grammar. If this is correct, then by extension the constraints on deletion and on what counts as recoverable differ as well. The OCP_{φ} constraint I propose only cares about φ-features, and the deletion of the phrase in Spec,CP only happens when φ-features are overtly encoded in C_{WH}. In questions with the allomorph *(l)a*, both C_{WH} and Spec,CP are overt. This also underlines the difference between the situation in Wolof and phenomena related to the Doubly-Filled-Comp Filter – in Wolof, both C_{WH} and Spec,CP can be overt, as long as they do not both overtly encode φ-features. All this suggests that deletion of elements can happen during different stages of the derivation, and in different modules of the grammar. In Wolof, the OCP_{φ} constraint operates on the CP layer, and cares only about φ-features, and so does recoverability related to OCP_{φ}-triggered deletion. An argument in favor of this view comes precisely from Wolof and the dialectal variation in the co-occurrence of determiners and C_{WH} in relative clauses, discussed in the previous section. The very fact that in some dialects the determiner can surface on the edges of the relative clause (see footnote 12), but in others it cannot, while such variation does not exist with respect to the deletion of φ-features in C_{WH} or Spec,CP of relative clauses, points to the fact that a different mechanism from the OCP_{φ}-triggered repair handles the deletion of the determiner.

6 Conclusion

In Distributed Morphology, some aspects of word formation take place in the syntax proper, while other aspects occur in the post-syntactic component, during Spell-Out. Processes that happen at PF are highly constrained and limited to minor manipulations of terminal nodes, such as feature or node deletion. They can nonetheless alter the surface output of syntax, creating the appearance of differences between structures that are syntactically identical. This paper adds to the body of work supporting this view of morphology and syntax interaction by investigating the morphosyntax of

A′-movement in the Niger-Congo language Wolof, and showing that surface distinctions in two A′-movement constructions can be reduced to the interaction of syntactic and post-syntactic processes, in particular agreement and dissimilation.

A′-movement in Wolof results in two seemingly syntactically distinct structures, which differ in the shape of the complementizer and the overtness of the phrase in Spec,CP. Furthermore, each of the complementizer variants exhibits a different agreement-based A′-extraction effect – CM-*u* shows φ-agreement and *(l)a* a subject/non-subject asymmetry, and it obligatorily marks the cyclicity of A′-movement. The two versions of C_{WH} are for the most part in complementary distribution: they both occur in *wh*-questions with a simple *wh*-phrase, but only one variant is allowed in all other A′-constructions, and their distribution does not correspond to any obvious syntactic or semantic differences. I therefore argue that the two variants of C_{WH} are allomorphs, meaning that all extraction effects are present in the syntax of each of them, but not observed on the surface due to post-syntactic processes.

The central part of the analysis presented in this paper lies in identifying the source of the distribution of the two C_{WH} allomorphs in Wolof in different constructions as post-syntactic. Namely, adjacent featural identity is known to trigger post-syntactic dissimilations. In Wolof, dissimilation targets the φ-feature, which, due to agreement, is found in C_{WH} and in Spec,CP. However, an overt φ-feature marker never surfaces both in C_{WH} and in Spec,CP. I argue this to be a consequence of a morphological Obligatory Contour Principle constraint, which prohibits the φ-feature from occupying two adjacent nodes. Repairs of such violations are language-specific, as are the constraints themselves. In Wolof, one of the nodes containing the φ-feature, either in C_{WH} or Spec,CP, is deleted. However, if either of the two contains content irretrievable in the CP-layer, its deletion is blocked, which results in a particular construction always surfacing with only one allomorph of C_{WH}. This paper offers not only a unified analysis of A′-extraction effects and maintains a unified syntax of A′-extraction in Wolof, but crucially offers a principled account for the distribution of different shapes of the CP-layer in different instances of A′-movement.

The more far-reaching contribution of the paper is the view of the grammar that it promotes. The increased interest in lesser known languages which seemingly greatly differ from the Indo-European types in their syntax and morphology raises questions as to the tenability of linguistic universals and the uniformity of syntactic structures and processes. This research shows that, as

long as we carefully examine the root of apparent syntactic irregularities in various languages, we can expect to find that their origin is not in underlying syntactic differences or derivations, but in processes that happen on the interface of syntax and other modules of the grammar. It is only the incomplete understanding of those other modules and their interactions with syntax that casts doubt on the cross-linguistic uniformity of syntactic structure.

References

Ackema, Peter, and Ad Neeleman. 2004. *Beyond morphology: Interface conditions on word formation*. Oxford: Oxford University Press.

Aldridge, Edith. 2002. Nominalization and wh-movement in Seedig and Tagalog. *Language and Linguistics* 3:393–427.

Aldridge, Edith. 2004. Ergativity and word order in Austronesian languages. Doctoral Dissertation, Cornell University.

Anagnostopoulou, Elena. 2003. *The syntax of ditransitives: Evidence from clitics*. Berlin: Mouton de Gruyter.

Arregi, Karlos, and Andrew Nevins. 2007. Obliteration vs. impoverishment in the Basque g-/z-constraint. In *Proceedings of the 30th annual Penn Linguistics Colloquium*, ed. Tatjana Scheffler, Joshua Trauberer, Aviad Eilam, and Laia Mayol, volume 13.1 of *University of Pennsylvania working papers in linguistics*, 1–14. Department of Linguistics, University of Pennsylvania, Philadelphia. Http://ling.auf.net/lingBuzz/000280.

Arregi, Karlos, and Andrew Nevins. 2012. *Morphotactics. Basque Auxiliaries and the Structure of Spellout*. Dordrecht: Springer.

Bayer, Josef, and Ellen Brandner. 2008. On Wh-Head-Movement and the Doubly-Filled-Comp Filter. In *Proceedings of the 26th West Coast Conference on Formal Linguistics*, ed. Charles B. Chang and Hannah J. Haynie, 87–95. Somerville, MA: Cascadilla Proceedings Project.

Benmamoun, Elabbas, and Heidi Lorimor. 2006. Featureless expressions: When morphophonological markers are absent. *Linguistic Inquiry* 37:1–23.

Bhatt, Rajesh. 1999. Actuality entailments of ability modals. In *Proceedings of the 17th West Coast Conference on Formal Linguistics*, ed. Kimary N. Shahin, Susan Blake, and Eun-Sook Kim, 74–87. Palo Alto, CA: CSLI.

Bhatt, Rajesh. 2002. The raising analysis of relative clauses: evidence from adjectival modification. *Natural Language Semantics* 10:43–90.

Bhatt, Rakesh, and James Yoon. 1992. On the composition of COMP and parameters of V2. In *Proceedings of the 10th West Coast Conference in Formal Linguistics*, ed. D. Bates, 41–52. Stanford, CA: CLSI.

Bonet, Eulàlia. 1991. Morphology after syntax. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Bonet, Eulàlia. 1995. Feature structure of Romance clitics. *Natural Language and Linguistic Theory* 13:607–647.

Brame, Michael K. 1968. A new analysis of the relative clause: Evidence for an interpretative theory. Unpublished Manuscript, MIT.

Calabrese, Andrea. 2010. Investigations on markedness, syncretism and zero exponence in morphology. *Morphology* 21:283–325.

Caponigro, Ivano, and Daphna Heller. 2007. The non-concealed nature of free relatives: Implications for connectivity in specificational sentences. In *Direct compositionality*, ed. Chris Baker and Pauline Jacobson, 237–263. Oxford: Oxford University Press.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Massachusetts: MIT Press.

Chomsky, Noam. 1973. Conditions on transformations. In *A Festschrift for Morris Halle*, ed. Stephan R. Anderson and Paul Kiparsky, 232–286. New York, NY: Holt, Rinehart, and Winston.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by step: Essays on Minimalist syntax in honor of Howard Lasnik*, ed. Roger Martin, David Michaels, and Juan Uriagereka, 89–155. Cambridge, Mass.: MIT Press.

Chomsky, Noam, and Howard Lasnik. 1977. Filters and control. *Linguistic Inquiry* 8:425–504.

Chung, Sandra. 1998. *The Design of Agreement: Evidence from Chamorro*. Chicago and London: The Univeristy of Chicago Press.

Demirdache, Hamida. 1991. Resumptive chains in restrictive relatives, appositives and dislocation structures. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Dunigan, Melynda B. 1994. On the clausal structure of Wolof. Doctoral Dissertation, University of North Carolina at Chapel Hill, Chapel Hill, NC.

É. Kiss, Katalin. 1998. Identificational focus versus information focus. *Language* 74:245–273.

Embick, David. 2010. *Localism versus globalism in morphology and phonology*. Cambridge, Massachusetts: MIT Press.

Embick, David, and Rolf Noyer. 2001. Movement operations after syntax. *Linistic Inquiry* 32:555–595.

Engdahl, Elisabet. 1985. Parasitic gaps, resumptive pronouns, and subject extractions. *Linguistics* 23:2–44.

Fox, Danny, and David Pesetsky. 2005. Cyclic linearization and syntactic structure. *Theoretical Linguistics* 31:1–45.

Goldsmith, John. 1976. Autosegmental phonology. Doctoral Dissertation, Massachusetts Institute of Technology. [Published, New York: Garland, 1979].

Hankamer, Jorge, and Line Mikkelsen. 2002. A morphological analysis of definite nouns in Danish. *Journal of Germanic Linguistics* 14:137–175.

Hankamer, Jorge, and Line Mikkelsen. 2005. When movement must be blocked: A reply to Embick and Noyer. *Linguistic Inquiry* 36:85–125.

Harizanov, Boris. To appear. Clitic doubling at the syntax-morphophonology interface: A-movement and morphological merger in Bulgarian. *Natural Language and Linguistic Theory* .

Henderson, Brent. 2007. Anti-agreement and [Person] in Bantu. In *Selected Proceedings of the 38th Annual Conference on African Linguistics: Linguistic Theory and African Language Documentation*, ed. Fiona McLaughlin Masangu Matondo and Eric Potsdam, 171–316. Somerville, MA: Cascadilla Press.

Henderson, Brent. 2013. Agreement and person in anti-agreement. *Natural Language and Linguistic Theory* .

Horvath, Julia. 1986. *Focus in the theory of grammar and the syntax of Hungarian*. Dordrecht: Foris.

Ka, Omar. 1987. Wolof phonology and morphology: a non-linear approach. Doctoral Dissertation, University of Illinois at Urbana Champaign, Urbana, IL.

Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, Massachusetts: MIT Press, (linguistic inquiry monographs 25) edition.

Kihm, Alain. 1999. Focus in Wolof. A study of what morphology may do to syntax. In *The grammar of focus*, ed. Georges Rebuschi and Laurice Tuller, 245–273. Amsterdam: John Benjamins Publishing Company.

Laka, Itziar. 1990. Negation in syntax: On the nature of functional categories and projections. Doctoral Dissertation, MIT, Cambridge, MA.

Leben, William. 1973. Suprasegmental phonology. Doctoral Dissertation, Massachusetts Institue of Technology. [Published, New York: Garland, 1979].

Lees, Robert B. 1960. *The grammar of English nominalizations*. The Hague: Mouton.

Lees, Robert B. 1961. The constituent structure of noun phrases. *American Speech* 36:159–168.

McCarthy, John. 1986. OCP Effects: Gemination and antigemination. *Linguistic Inquiry* 17:207–263.

McCloskey, James. 2001. The morphosyntax of WH-extraction in Irish. *Journal of Linguistics* 37:67–100.

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. Daniel Seeley, 184–226. Oxford: Blackwell.

Merchant, Jason. 2009. Phrasal and clausal comparatives in greek and the abstractness of syntax. *Journas of Greek Linguistics* 9:169–179.

Ndiaye, Moussa D. 1995. Phonologie et morphlogie des alternances en wolof: Implications théoretiques. Doctoral Dissertation, Université de Montréal.

Nevins, Andrew. 2007. The representation of third person and its consequences for person-case effects. *Natural Language and Linguistic Theory* 25:273–313.

Nevins, Andrew. 2012. Haplological dissimilation at distinct stages of exponence. In *The morphology and phonology of exponence*, ed. Jochen Trommer. Oxford University Press.

Njie, Codu Mbassy. 1982. *Description syntaxique du wolof de Gambie*. Dakar, Senegal: Les Nouvelles Editions Africaines.

Noyer, Rolf. 1997. *Features, positions and affixes in Autonomous Morphological Structure*. New York: Garland Press.

Ouhalla, Jamal. 1993. Subject-extraction, negation and the anti-agreement effect. *Natural Language and Linguistic Theory* 11:477–518.

Ouhalla, Jamal. 2005. Agreement features, agreement, and anti-agreement. *Natural Language and Linguistic Theory* 23:655–686.

Perlmutter, David. 1971. *Deep and surface structure constraints in syntax*. New York, NY: Holt, Rinehart, and Winston.

Pescarini, Diego. 2010. Elsewhere in Romance: Evidence from clitic clusters. *Linguistic Inquiry* 41:427–444.

Pesetsky, David. 1998. Some optimality principles of sentence pronunciation. In *Is the best good enough? Optimality and competition in syntax*, ed. P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis, and D. Pesetsky. Cambridge, MA: MIT Press.

Pesetsky, David, and Esther Torrego. 2001. T-to-C movement: causes and consequences. In *Ken Hale: A life in language*, ed. Michael Kenstowicz, 355–426. Cambridge, MA: MIT Press.

Potsdam, Eric. 2009. Austronesian verb-initial languages and *wh*-question strategies. *Natural Language and Linguistic Theory* 27:737–771.

Potsdam, Eric, and Maria Polinsky. 2011. Questions and word order in Polynesian. In *Topics in Oceanic Morphosyntax*, ed. Claire Moyse-Faurie and Joachim Sabel. Berlin: Mouton de Gruyter.

Åfarli, Tor A. 1994. A promotion analysis of restrictive relative clauses. *Linguistic Review* 11:81–100.

Reglero, Lara. 2006. Spanish subcomparatives: the "obligatory gapping stragegy". In *Selected Proceedings of the 9th Hispanic Linguistic Symposium*, ed. Nuria Sagarra and Almeida Jacqueline Toribio, 67–78. Somerville, MA: Cascadilla Proceedings Project.

Richards, Norvin. 1997. What moves where in which language. Doctoral Dissertation, MIT, Cambridge, MA.

Richards, Norvin. 1998. Syntax versus semantics in Tagalog *wh*-extraction. In *UCLA occasional papers in linguistics 21: recent papers in Austronesian linguistics*, ed. Matthew Pearson, 259–275. Los Angeles: University of California, Los Angeles Department of Linguistics.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of Grammar*, ed. Liliane Haegeman, 281–337. Kluwer Academic Publishers.

Robert, Stéphane. 1991. *Approche énonciative du système verbal: le cas du wolof*. Paria: Éditions du Centre National de la Recherche Scientifique.

Ross, John Robert. 1967. Constraints on variables in syntax. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Russell, Margaret A. 2006. The syntax and placement of Wolof clitics. Doctoral Dissertation, University of Illinois at Urbana-Champaign, Urbana, IL.

Sauerland, Uli. 1998. The meaning of chains. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, Massachusetts. Distributed by MIT Working Papers in Linguistics.

Sauerland, Uli. 2003. Unpronounced heads in relative clauses. In *The Interfaces: Deriving and interpreting omitted structures*, ed. Kerstin Schwabe and Susanne Winkler, 205–226. John Benjamins.

Schachter, Paul. 1973. Focus and relativization. *Language* 49:19–46.

Schneider-Zioga, Patricia. 1995. Specifier/head agreement in Kinande. *Cahiers Linguistiques d'Ottawa* 23:67–93.

Schneider-Zioga, Patricia. 1996. An argument in favor of Agreement Phrase. In *ESCOL Proceedings*.

Schneider-Zioga, Patricia. 2000. Anti-agreement and the fine structure of the Left Periphery. *University of California Irvine Working Papers in Linguistics* 6.

Schneider-Zioga, Patricia. 2007. Anti-agreement, anti-locality and minimality. The syntax of dislocated subjects. *Natural Language and Linguistic Theory* 25:403–446.

Shlonsky, Ur. 2012. Notes on wh in situ in French. In *Functional Heads. The Cartography of Syntactic Structures*, ed. Laura Brugé, Anna Cardinaletti, Giuliana Giusti, Nicola Munaro, and Cecilia Poletto, volume 7, 242–252. New York: Oxford University Press.

Torrence, Harold. 2005. On the distribution of complementizers in Wolof. Doctoral Dissertation, University of California, Los Angeles, Los Angeles, CA.

Torrence, Harold. 2012a. *The clause structure of Wolof: insights into the left periphery*. Amsterdam/Philadelphia: John Benjamins Publishing.

Torrence, Harold. 2012b. The morpho-syntax of silent wh-expressions in Wolof. *Natural Language and Linguistic Theory* 30:1147–1184.

Torrence, Harold. 2013a. A promotion analysis of Wolof clefts. *Syntax* 16:176–215.

Torrence, Harold. 2013b. The morpho-syntax of Wolof clefts: Structure and movement. In *Cleft Structures*, ed. Katharina Hartmann and Tonjes Veenstra. Amsterdam: John Benjamins Publishing.

Vergnaud, Jean Roger. 1974. French relative clauses. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, Massachusetts.