

The Syntactic Derivation of Complex Wh-Words in Wolof

Harold Torrence
UCLA

1.0. Introduction

Wolof is a West Atlantic language spoken in Senegal and Gambia, with basic SVO word order and noun classes.¹ It presents an interesting analytical problem in that there are two basic ways of forming Wh-questions, with what I will call “*an*-forms” or “*u*-forms”:

- (1) **k-an** (mo)-o dóór xale yi *an*-Form
class-an 3sg-*a* hit child the
“who hit the children?”
- (2) **k-u** dóór xale yi *u*-Form
class-*u* hit child the
“who hit the children?”

The “*an*-forms” are composed of a class marker followed by *-an*, as seen in (1). The “*u*-forms”, on the other hand are composed of a class marker followed by *-u*, as in (2). There are no known interpretive differences between (1) and (2).² In both (1) and (2), the class marker that appears is *k*-. The *k*-class is the default human noun class. When this combines with either *-an* or *-u* the clause can be translated into English with “who” or “which one (human)”. Wh questions with either *an*-forms or *u*-forms can be answered by a cleft or not. The clefted form is the most natural answer:

- (3) Bubəkër mó-ó dóór xale yi Cleft
bubakar 3sg-*a* hit child the
“it’s Bubakar who hit the children”

¹ I would like to thank my consultants Maryam Sy, Seynabou Ndoeye, and Mustapha Djigo for the many hours spent providing me with the data discussed in this talk. I would also like to thank Hilda Koopman, Dominique Sportiche, Tim Stowell, Leston Buell, Marcus Smith, and Eric Jackson for their comments.

² There is a slight preference for the *an*-forms to be D-linked.

2.0. Distribution of *u*-forms

The *u*-forms can be used to form Wh-questions for subjects, objects, adjuncts, etc.³

- (6) a. **k-u** sàcc gato bi? Subject
 class-u steal cake the
 “who stole the cake?”
- b. **l-u** Isaa sàcc? DO
 class-u isaa steal
 “what did Isaa steal?”
- c. **n-u** Isaa sàcc-e gato bi Adjunct
 class-u isaa steal-manner cake the
 “how did Isaa steal the cake?”

The *u*-forms cannot be used with PPs or genitives:

- (7) a. *ci **f-u** xale yi di fo-we -*u*- + PP
 prep class-u child the di play-loc
 “at where are the children playing?”
- b. ci **f-an** l-a xale yi di fo-we -*an* +PP
 prep class-an xpl-*a* child the di play-loc
 “at where are the children playing?”
- c. *xac-u-(b) **k-u** mo-o-la màtt -*u*- + Genitive
 dog-u-class class-u 3sg-*a*-you bite
 “whose dog bit you?”
- d. xac-u-(b) **k-an** mo-a-la matt -*an* + Genitive
 dog-u-class class-an 3sg-*a*-you bite
 “whose dog bit you?”

The *u*-forms can also co-occur in relative clauses with indefinite heads:

- (8) di-na-a wax ak k-u ñëw
 imperf-na-1sg speak with class-u come
 “I will speak with whoever/someone who comes”

³ There is no *u*-form for ‘why’ or ‘how many’. There is speaker variation as to whether there is an *u*-form that corresponds to ‘when’, *b-u*.

3.0. Differences between *u*-forms and *an*-forms

There are a number of syntactic differences between the *u*-forms and *an*-forms. Overall, *u*-forms have the distribution of clitics (Kayne 1975). This is clearly manifested in the fact that the *an*-forms can stand by themselves, the *u*-forms cannot:

- (9) k-enn ñëw-na
 class-one arrive-neutral
 “someone arrived”
- a. k-an ? b. *k-u
 class-an class-u
 “who?” “who?”

Further, the *an*-forms and non-Wh DPs can be coordinated with *ak*, *u*-forms cannot:

- (10) a. **Isaa** ak **woto bi** l-a xale yi dàq Non-Wh DP
 isaa and car the xpl-*a* child the chase
 “it’s Isaa and the car that the children chased”
- b. **k-an** ak **l-an** l-a xale yi dàq? *an*-Form
 class-an and class-an xpl-*a* child the chase
 “who and what is it that the children chased?”
- c. ***k-u** ak **l-u** xale yi dàq *an*-Form
 class-u and class-u child the chase
 “who and what did the children chase?”

Nor can the *u*-forms be coordinated with the clausal coordinator *te*:

- (11) ***k-u** te **l-u** xale yi dàq *u*-Form
 class-u and class-u child the chase
 “who and what did the children chase?”

The *an*-forms can appear in situ (and be interpreted as echo), the *u*-forms cannot.⁴

⁴ In other cases, an “in-situ” Wh is fine as a non-echo question, this is never the case with the *u*-forms.

- (12) togg-al-na-ñu xale yi ceeb non-Wh DP
 cook-ben-netural-3pl child the rice
 “they cooked the children rice”
- (13) togg-al-na-ñu xale yi **l-an** *an*-Form
 cook-ben-na-3pl child the class-an
 “they cooked what for the children?” (Echo only)
- (14) *togg-al-na-ñu xale yi **l-u** *u*-Form
 cook-ben-na-3pl child the class-u
 “they cooked what for the children?”

Thus, while the *an*-form yields an echo question only in (13), the presence of an *u*-form in (14) is simply unsayable.

In the simple case (but see (56)), the *u*-forms are compatible only with a “relative” CP, while the *an*-forms are incompatible with a relative CP:⁵

- (15) **k-o-o-fa** jox-oon tééré bi *u*-Form + Relative CP
 class-u-2sg-loc give-past book the
 “who did you give the book to there?”
- (16) ***k-an-a-fa** jox-woon tééré bi *an*-Form + Relative CP
 class-an-2sg-loc give-past book the
 “who did you give the book to there?”

Only one *u*-form per clause is allowed, while multiple *an*-forms may occur:

- (17) ***k-u l-u** jox-oon xale yi class-u class-u
 class-u class-u give-past child the
 “who gave what to the children?”

⁵ A “relative CP” is the type of CP which is found in relative clauses and in some clauses with left peripheral adverbs. The template for relative CPs is:

(i) XP-*u*-subj.clt-obj.clt-loc.clt DP_{subj} V-neg-past DO

Neither arguments nor adjuncts can co-occur with the clitic in this construction unless dislocated. The span from *-u-* to loc.cltic is ATR harmonic to XP. The cleft template is very similar:

(ii) XP *l-a*-subj.clt-obj.clt-loc.clt DP_{subj} V-neg-past DO

Neither arguments nor adjuncts can co-occur with the clitic in this construction, unless dislocated. The span from *l-* to loc.clt is ATR harmonic to XP.

(18) **k-an l-an** l-a jox-oon xale yi class-an class-an
 class-an class-an xpl-a give-past child the
 “as for who, what is it that he gave to the children?”
 Echo Question only

(19) **jigéén yi tééré bi** l-a-ñu jox-oon xale yi DP DP
 woman the book the xpl-a-3pl give-past child the
 “as for the children, it’s the woman who they chased”

For cases like (17), neither the order of the *u*-forms, nor the combination of subject and non-subject Wh(, nor would the presence of argument and adjunct Whs) make any difference.

When *an-* and *u-*forms co-occur in a single clause, the *u*-form must precede the *an*-form:

(20) **l-u ñ-an** jox-oon xale yi class-u class-an
 class-u class-an give-past child the
 “who(pl) gave what to the children?”

(21) ***ñ-an l-u(-ñu)** jox-oon xale yi class-an class-u
 class-an class-u-3pl give-past child the
 “who(pl) gave what to the children?”

An-forms can pied pipe an NP, (interrogative) *u*-forms cannot:

(22) **g-ën góór** l-a-ñu gis
 class-an man xpl-a-3pl see
 “which man did they see?”

(23) **góór g-u-ñu** gis
 man class-u-3pl see
 * “which man did they see?”
 “a man who they saw”

As (23) shows, when a non-Wh NP is pied piped, the interrogative interpretation is impossible.

The *an*-forms can be clefted, the *u*-forms cannot be:

- (24) **k-an** (mo)-o dóór xale yi *an*-Form
 class-an 3sg-*a* hit child the
 “who is it who hit the children?”
- (25) **Isaa** l-a xale yi dóór Non-Wh DP
 isaa xpl-*a* child the hit
 “it’s Isaa that the children hit”
- (26) ***k-u** l-a-ñu dóór *u*-Form
 class-u xpl-*a*-3pl hit
 “who did the children hit?”

4.0 Islands

While the *u*-forms and *an*-forms display a number of differences, they are both island-sensitive. This is true for both strong and weak islands, suggesting that both are derived by movement⁶:

Adjunct Island

- (27) xale bi dem-na laata Bintë togg-al Móódu laax
 child the leave-neutral before bintë cook-ben moodu laax
 “the child left before Binta cooked Moodu laax”
- (28) ***l-ani** l+a xale b-i dem [laata Bintë togg-al Móódu t_i]
 cl-an l+a child cl-def go before binte cook-ben moodu
 “what did the child go before Binte cooked Moodu?”
- (29) ***l-u_i** xale b-i dem [laata Bintë togg-al Móódu t_i]
 cl-u child cl-def go before binte cook-ben moodu
 “what did the child go before Binte cooked Moodu?”

⁶ NB All of these island violations can be repaired with a resumptive pronoun. See Torrence (forthcoming).

Coordinate Structure Island

(30) ***I-ani** l-a-ñu jend a-y nen ak t_i *an-Form*
 cl-an xpl-a-3pl buy indef-cl egg and
 “what did they buy eggs and?”

(31) ***I-u**-ñu jend a-y nen ak t_i *u-Form*
 cl-u 3pl buy indef-cl egg and
 “what did they buy eggs and?”

Wh Island

(32) ***I-ani** l-a Dudu xam ndax t_i l-a-a jënd *an-Form*
 cl-an l-a dudu know whether l-a-1sg buy
 “what does Dudu know whether I bought?”

(33) ***I-u** Dudu xam ndax t_i l-a-a jënd *u-Form*
 cl-u dudu know whether l-a-1sg buy
 “what does Dudu know whether I bought?”

There is a Wolof-specific diagnostic which is also relevant with respect to movement constraints. Wolof has an applied suffix, *-al*, that alternates with a preposition *ak*. Simplifying somewhat, the applied suffix is obligatory when the applied object undergoes A' movement, and is impossible otherwise.⁷ In a non-Wh case, with the applied object in situ, the preposition is obligatory:

(34) jàngalekat yi daje-na-ñu *(**ak**) Isaa
 teacher the meet-neutral-3pl with isaa
 “the teachers met with Isaa”

The applied suffix is impossible when the applied object has not undergone A'-movement:

(35) *jàngalekat yi daje-**el**-na-ñu Isaa
 teacher the meet-appl-neutral-3pl isaa
 “the teachers met with Isaa”

But, the suffix must be present when the applied object undergoes A'-movement:

⁷ This is similar to what happens in Vata (Koopman 1984).

- (36) **Isaa** l-a jàngalekat yi daje-*(**el**)
 isaa xpl-*a* teacher the meet-appl
 “it’s Isaa that the teachers met with”

The applied suffix cannot be present when the applied object is cliticized, whether P is present or not or whether the clitic is doubled by the applied object⁸:

- (37) *jàngalekat yi daje-**el**-na-ñu-ko (ak) (Isaa)
 teacher the meet-appl-na-3pl-him with isaa
 “the teachers met with Isaa”

Given this pattern of behavior, I take it that the applied suffix is a diagnostic for A'-movement. It is therefore telling that the applied suffix must be present when an *an*-form is clefted or when an *u*-form is present:

- (38) **k-an** l-a jàngalekat yi daje-*(**el**) *an*-Form
 class-an xpl-*a* teacher the meet-appl
 “who is it that the teachers met?”

- (39) **k-u** jàngalekat yi daje-*(**el**) *u*-Form
 class-u teacher the meet-appl
 “who did the teachers meet?”

A natural conclusion to draw from both the island facts and the distribution of the applied suffix is that both the *an*-forms and *u*-forms are derived by movement.

5.0. Decomposing the *u*-Form

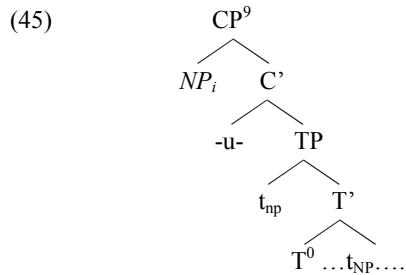
As was noted in the first section, the *u*-forms are composed of both a class marker and an *-u-* element. Taking a closer look at these will be necessary for the analysis to come. Class agreement is widespread in Wolof and it is clear that the class marker in the *u*-form indicates agreement with a noun, but is not itself a noun.

- (40) sàcc-nga tééré **b-u**-nu jàng Overt Noun
 steal-2sg+na book class-u-1pl read
 “you stole a book that we read”

⁸ The applied suffix is impossible even if the clitic has been restructured out of its originating clause.

6.0. Analysis

Given the constellation of properties examined, I analyze the *u*-forms as arising by movement of a silent NP, NP_i , into the left periphery of a clause with *-u-* in the C-domain. That is, *k-u* is not an underlying constituent, but is put together by movement in the course of the derivation, as schematized below:



As always, the class agreement marker is the spellout of the spec-head agreement relation between the silent NP and the head of CP, *-u-*.¹⁰ The distributional properties follow straightforwardly from such an analysis. Recall that the *u*-forms cannot be used independently:

- (46) **k-u*?
 class-u
 “who?”

We noted that the *u*-form has the distribution of a clitic. Therefore, we predict that it will not be possible to use it by itself, given its prosodic properties. This is reinforced because the *u*-form always appears as the first element in the clitic complex (as in (56)). Note that the *u*-forms cannot be like Romance clitics because, for example, only one *u*-form per clause is allowed. Another possibility is that the relative CP simply cannot be elided. We also saw that the *u*-forms cannot be coordinated with *ak* or with *te*:

- (47) ***k-u** ak **l-u** [FinP xale yi dàq *u*-Form
 class-u and class-u child the chase (= 0)
 “who and what did the children chase?”

⁹ Although *-u-* is written in as the head of a unitary CP, it is not the highest head in the left periphery of the clause.

¹⁰ There could be an additional Agr projection which mediates agreement; this is not important for the discussion here, and will therefore be ignored.

- (48) ***k-u** te **l-u** [FinP xale yi dàq *u*-Form
 class-u and class-u child the chase (=11)
 “who and what did the children chase?”

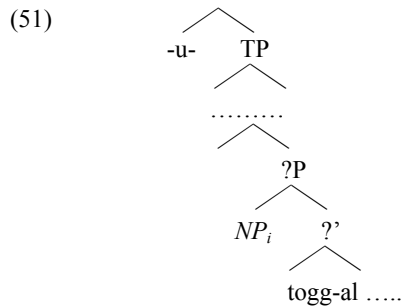
It was concluded that the *u*-forms are not DPs, therefore (47) is bad because *ak* only coordinates DPs, and not CPs. If correct, this would constitute an argument for the proposed analysis of *class-u* as an agreeing C⁰. However, coordination is also blocked by the stipulation that *class-u* is a clitic. (48) can be explained if we find a way to explain Right Node Raising (i.e. the string in (48) must be derived from two full CPs.) That is, RNR or ellipsis would leave no phonological material for *class-u* to lean on. (The properties of Right Node Raising and ellipsis are not understood in Wolof.) Note that when appropriate phonological material is present, coordination is possible:

- (49) [CP **k-u-ñu** dàq te [CP **l-u-ñu** dàq
 class-u-3pl chase and class-u-3pl chase
 “who and what did they chase?”

Another property was that the *u*-forms cannot appear in situ:

- (50) ***togg-al-na-ñu** xale yi **l-u** *u*-Form
 cook-ben-neutral-3pl child the class-u (=14)
 “they cooked what for the children?”

Consider the partial derivation of (50) below:



In the analysis here, *-u-* is merged in the C-domain, not as an argument in the v/VP region. Thus, *-u-* does not appear “in situ”, but only in the left periphery.

It was also noted that the *u*-forms are only compatible with a relative CP. The *an*-forms are, in the simple case, incompatible with a relative C:

(52) **k-o-o-fa** jox-oon tééré bi *u*-form and Relative CP
 class-u-2sg-loc give-past book the
 “who did you give a book to there?”

(53) ***k-an-a-fa** jox-oon tééré bi *an*-form and Relative CP
 class-an-2sg-loc give-past book the
 “who did you give a book to there?”

Again, according to the present analysis, *-u-* is an element in the C-domain. As is commonly the case, a complementizer is compatible with a clause with a particular kind of left periphery. It is not surprising that the *u*-form is only compatible with a relative CP.

Interestingly, the *an*-form is compatible with a relative CP *only if a u-form is also present*:

(54) ***l-an-a** xale yi lekk foofu
 class-an child the eat there
 “what did the children eat there?”

(55) **l-u** ñ-an lekk foofu
 class-u class-an eat there
 “who(pl) ate what there?”

The fact that an *an*-form can appear in a relative CP only when an *u*-form is also present follows from *-u-* being a complementizer that attracts a particular type of NP/DP. The *an*-form is simply of the wrong type to satisfy the selectional properties of *-u-*. Let me also note also that in (55), although the *u*-form and *an*-form appear to be adjacent, if present, clitics will intervene between them. (*ñ-an* is the subject below.):

(56) **n-u-leen-ko-fa** ñ-an togg-e-woon démb
 class-u-them-it-there class-an cook-manner-past yesterday
 “how did who cook it for them there yesterday?”

From (56) we see that the *an*-form occupies a lower position than an *u*-form and that there is “space” for clitics between the two.

One interesting difference was that only one *u*-form per clause is allowed:

(57) ***k-u l-u** jox-oon xale yi class-u class-u
 class-u class-u give-past child the
 “who gave what to the children?”

In our analysis, $-u-$ is a C^0 that selects for a TP; therefore there can be only one $-u-$ per clause. In other words, this follows from the selectional properties of heads. The $-u-$ does not select for a CP headed by $-u-$. A case like (57) could naturally arise if more than one $-u-$ were merged (in a single clause), or if $-u-$ projects multiple specifiers. Apparently, neither of these possibilities is borne out. That is, the non-iterability of $-u-$ is what we expect from a complementizer.

Little needs to be said about the u -forms being island-sensitive. Recall the pattern:

Adjunct Island

(58) ***I-*ui*** xale b-i dem [laata Bintë togg-al Móódu t_i
 cl-u child cl-def go before binte cook-ben moodu
 “what did the child go before Binte cooked Moodu?”

Coordinate Structure

(59) ***I-*ui*** ngeen jend a-y nen ak t_i
 cl-u 2pl buy indef-cl egg and
 “what did yall buy eggs and?”

Wh Island

(60) ***I-*u*** Dudu xam ndax t_i l+a-a jënd
 cl-u dudu know whether l+a-1sg buy
 “what does Dudu know whether I bought?”

The u -form is island sensitive because it is derived by movement. The silent NP_i moves from its original merge position into the left periphery. In the cases immediately above, NP_i has crossed an island boundary, thus, we predict them to be ungrammatical. The same reasoning suffices for the applied construction. NP_i undergoes A' -movement and thus the presence of the applied suffix is obligatory when an u -form corresponds to the applied object.

It was seen that the (interrogative) u -form cannot pied pipe an NP:

(61) góór g-u-ñu gis (= (23))
 man class- u -3pl see
 * “which man did they see?”
 “a man who they saw”

If NP_i is merged as the thematic subject, object, etc., it follows from θ -Theory that there cannot be another NP which is assigned that θ -role. In other words, NP_i should be in complementary distribution with overt NPs. We must conclude that cases like (61), which might appear to have NP_i and an overt NP do not. The fact that a clause like (61) with an overt NP cannot be interpreted as a question, only as an indefinite DP is otherwise mysterious. We know that NP_i is compatible with an interrogative interpretation. That is, the only noun present is the lexical NP, not a (silent) Wh-phrase.

Finally, it was shown that the u -forms cannot occur in a cleft:

(62) ***k-u** l-a-ñu dóór (= (26))
 class-u xpl-a-3pl hit
 “who did they hit?”

According to the analysis, NP_i only occurs with $-u-$. The left periphery of a cleft clause is incompatible with the presence of $-u-$ and therefore, incompatible with NP_i . Such co-occurrence restrictions are expected.

We predict that the prohibition on clefting of NP_i will *not* hold, so long as NP_i can eventually reach a $-u-$. Such a situation occurs when, a u -form is extracted from an embedded clause:

(63) **k-u** jigéén ji wax ne l-a-ñu dóór?
 class-u woman the say that xpl-a-3pl hit
 “who did the woman say that they hit?”

What (63) suggests is that the apparent incompatibility of clefting and the u -forms is not due to the u -forms themselves. When there is another clause where NP_i can be licensed, then it can be clefted.

Finally, the fact that when a u -form and an -form co-occur the u -form must precede follows from the analysis here. There is only one “Wh” attracting head in the left periphery and it is higher than the subject position and $-u-$. The C^0 , $-u-$ merges higher than the subject position. When the Wh attractor is merged, NP_i is closer, therefore it always precedes an an -form, even a subject an -form.

7.0. Conclusions:

- Wolof has a class of apparent Wh words, *u*-forms, with a distribution distinct from that another class of Wh words, *an*-forms.
- Careful examination of the *u*-forms reveals that they simply do not have the distribution of NP/DPs, unlike the *an*-forms.
- These differences can be explained if we assume that the phonological class-*u* string is not a Wh-word, but an agreeing complementizer.
- Instead, the *u*-forms result from syntactic movement operations which put a silent NP and a complementizer in very close proximity, which gives the illusion of constituency.
- As a result, the *u*-forms are predicted to have movement properties (e.g., island sensitivity), but at the same time lack the properties of an NP/DP constituent, a seeming contradiction.
- This can be resolved if the *u*-form is derived by movement, but not of the *class-u* string, but the silent NP.

Bibliography

- Kayne, Richard. 1975. *The French Transformational Cycle*. Cambridge, Massachusetts: the MIT Press
- Koopman, Hilda. 1984. *The Syntax of Verbs: From Verb Movement Rules in the Kru Languages to Universal Grammar*. Dordrecht: Foris Publications.