Finite Control in Kabyle*

Jurij Božič
McGill University

SUMMARY

The present paper discusses embeddings that occur under verbs such as [jaːrəð] "try" and [jivʁa] "want" in Kabyle Berber. We will show that these constructions differ from embeddings under verbs such as [jinːɛd] "say" according to several properties, viz. tense, mood and the interpretation of the embedded subject "gap". We will also see that the discussed construction hosts a fully inflected embedded verb but it still passes all the standard diagnostics for Control. This means that Kabyle Berber has an instance of Finite Control.

RESUME

Cet article traite d l'intégration syntaxique qui a lieu avec des verbes tels que [jaːrəð] 'essayer' and [jivʁa] 'vouloir' en berbère kabyle. Nous démontrons que ces constructions diffèrent de celles employées avec des verbes tels que [jinːɛd] 'dire' selon plusieurs propriéties: le temps, le mode et l'interprétation de la position vide du sujet intégré. Nous démontrons également que l'intégration syntaxique exige un verbe intégré entièrement fléchi bien que les tests standards suggèrent qu'il s'agit d'une structure de contrôle. Ces observations permettent de conclure qu'il existe une structure de contrôle fini en berbère kabyle.

1 INTRODUCTION

The traditional claim in generative syntax is that the null anaphor PRO can only occur in embedded clauses that are uninflected — in other words, they are *non-finite*. However, a recent strand of research (Ferreira 2004, 2009; Landau 2004; Boeckx et al. 2010) has shown that Control may be formed in *inflected* contexts: it seems that, in several languages, PRO can occur in embedded

* All the data presented in this paper were elicited from Sadia Nahi, a native speaker of Kabyle. Thanks to Sadia for sharing her language with me!

McGill Working Papers in Linguistics, 26.1, 2020

clauses in which the verb is fully inflected. These constructions have been termed *Finite Control*. What the trigger of Control could be in these contexts is currently a matter of great debate. The present paper identifies one such construction in a dialect of Berber, viz. Kabyle, which was collected to contribute to the cross-linguistic knowledge of this phenomenon.

The present paper will focus on the embeddings under verbs such as [jaːrəð] "try" and [jivʁa] "want". I will show that the application of certain fairly standard diagnostics suggests that these embeddings are instances of Finite Control. The set of diagnostics inspects: the covert status of the embedded "gap", the argumental status of the embedding, the interpretation of embedded idioms, the interpretation of the embedded gap under VP-ellipsis, and finally the dependent nature of *tense* in the embeddings. Most of the data will involve *Subject Control*. However, we will see that the verb [jivʁa] can also form *Object Control*, which is unexpected given pre-existing claims on this construction (Ouali 2019).

2 BASIC DESCRIPTION OF EMBEDDING IN KABYLE

Before we start investigating the instances of Finite Control in Kabyle, let us consider a basic embedding under the verb [jinːɛd] "say":

(1) jinːɛd rəzqɛj [(bəlːi/ðɛkin) θtʃɛ kinza jimɛkija]
said:3.SG.M Razqej that ate:3.SG.F Kinza food
'Razqej said that Kinza ate food.'

Verbs such as [jinːɛd] seem to embed CPs just like the verb *say* does in English. The presence of the CP-layer is indicated by the complementizers [bəlːi] and [ðɛkin]. Either of them can occur (of course, not at the same time), but they are optional, as these embeddings can be formed without their overt realization. At first blush, it seems that these embeddings represent the usual, fully finite and non-deficient declarative embedding in Kabyle, and we will show that this is truly the case momentarily.

We now consider embeddings under verbs such as [jaːrəð] "try" and [jivʁa] "want", which, in English and many other languages, embed uninflected Control clauses:

(2) jaːrəð rəzqɛj$_i$ [ɛð=irˤoħ _____$_i$ saxːam]
tried:3.SG.M Razqej IRR=go.AOR.3.SG.M to-home
'Razqej tried to go home.'

(3) jivʁa rəzqɛj$_i$ [að=iʁarˤ _____$_i$]
wanted:3.SG.M Razqej IRR=study:3.SG.M
'Razqej wanted to study.'

In Kabyle, verbs of this type embed clauses that are fully *inflected*: notice that [ɛð=irˤoħ] and [að=iʁarˤ] are fully inflected and agree with the matrix subject. In the embedded clause, however, we also have a "gap", which appears to be coreferential with Razqej, or perhaps involves binding.

Analyzing this embedded gap is the core of this paper.[1] We will ultimately argue that it is indeed an instance of PRO, but for now we will refer to it as just the "gap".

A key similarity between (1) and (2)-(3) is that they all host an embedded verb that is *fully inflected*. However, it is important to note that, unlike (1), (2)-(3) type constructions have a fairly restricted embedded domain: the verb is always in the *aorist* form preceded by the particle [að], which is claimed to encode *irrealis* in the literature (Ouali 2019: 4).[2] The aorist in general is claimed to occur in embedded "non-finite" contexts in the Berber literature (Bentolila 1981; Boukhris 1998), which presumably refers to other, non-agreement related deficiencies in this domain. In general, it seems that these embedded domains have a fixed or "deficient" *mood* specification; in other words, they could perhaps be understood as subjunctives of sorts.

3 ANALYZING PROPERTIES OF THE EMBEDDED DOMAINS

In this section, I will show that the embeddings in (2)-(3) display key differences in their internal properties when compared to the standard embedding in (1). We have already stated that there is no difference between the two constructions in terms of embedded verb inflection. We have also stated that (2)-(3) seem to be limited to *irrealis* mood, which is not the case for (1)-type constructions. Below, we consider two other differences: the tense-status of the embeddings and their argumental status.

The first clue that verbs such as [jaːrəð] embed a different type of clause than in (1) is that the two complementizers that can occur in (1) cannot occur under [jaːrəð]:

(4) jaːrəð rəzqɛj [(*ðekːɪn/*bəlːi) ɛð=irˤoħ saxːam]
tried:3.SG.M Razqej that IRR=go.AOR.3.SG.M to-home
'Razqej tried to go home.'

This is further confirmed by the observation that the two types of construction have *different* tense properties. Embeddings under [jinːɛd] "say" seem to host an entirely independent T(ense)-operator:

(5) jinɛjid rəzqɛj <u>ɛsːɛgi</u> [ðekːɪn kinza ɛ=tsəts pidza
told.me:3P.SG.M Razqej today that Kinza IRR=eat.AOR.3.SG.F pizza
<u>ɛzəka</u>]
tomorrow
'Razqej told me today that Kinza will eat pizza tomorrow.'

A standard way of testing the presence of a T-operator in the embedded domain is to endow the sentence with a *conflicting* pair of temporal adverbs (Landau 2004): the matrix clause needs to host

[1] It should be noted that (2) and (3) appear to pattern in the same way in the relevant respects. However, most of the examples in this paper will be with embeddings under the verb [jaːrəð] "try" and only few will involve the verb [jivʁa] "want", for reasons of space and also lack of time in elicitation sessions. Future work should, of course, attempt to elicit entire paradigms with all verbs that display the pattern of (2) and (3).

[2] In my elicitation sessions, I got the impression that this particle may exhibit clitic-like behaviour, which is why I am notating it as such in the examples throughout this paper. Whether it is a true clitic is unclear.

one adverb and the embedded clause the other which must conflict with the matrix one. This is done in (5), where the adverbs are underlined. The embedded domain hosts a temporal adverb that conflicts with the interpretation of the matrix one, which means that there must be an *independent T-operator* in the embedded clause that allows this. In short, the embedded domain under [jinːɛd] "say" is fully inflected and also fully tensed. But the same does not seem to be true of the domain under [jaːrəð] "try":

(6) jaːrəð rəzqɛj ɛsːɛgi [ɛð=irˤoħ saxːam]
tried:3.SG.M Razqej today IRR=go.AOR.3.SG.M to-home
'Razqej tried to go home today.'

(7) *jaːrəð rəzqɛj ɛsːɛgi [ɛð=irˤoħ saxːam ɛzkɛ]
tried:3.SG.M Razqej today IRR=go.AOR.3.SG.M to-home tomorrow
intended: 'Today, Razqej tried to go home tomorrow.'

While the matrix clause can support a temporal adverb, as in (6), ungrammaticality results as soon as the embedded clause contains a conflicting adverb, as in (7). This suggests that these embeddings *lack a T-operator*; in other words, they are *tense-deficient*. Ouali (2019: 16) notes that this is generally true of such embeddings in Tamazight.

Next, we will show that the clauses embedded under [jaːrəð] are true complements and not adjuncts. A fairly typical adjunct can be introduced with the complementizer [iwəkin] "so that" in Kabyle. Consider the following example:

(8) juʁəd rəzqɛj θɛkarosθ ɛsːɛgi [iwəkin ɛtsərˤoħ kinza
bought:3.SG.M. Razqej car today so.that IRR=go.AOR.3.SG.F Kinza
ɛzkɛ saxːam]
tomorrow to-home
'Razqej bought a car today so that Kinza will go home tomorrow.'

Here, the embedded clause is an optional part of the sentence. It seems to be a regular tensed CP since it can contain a temporal adverb that conflicts with a matrix one. Such adjuncts can usually be fronted, and this is also true in Kabyle:

(9) [iwəkin ɛð=ɛrˤoħ ɛzkɛ saxːam] juʁəd rəzqɛj
so.that IRR=eat.AOR.3.SG.M tomorrow to-home bought:3.SG.M Razqej
θɛkarosθ ɛsːɛgi
car today
'So that he will go home tomorrow, Razqej bought a car today.'

However, the same is not true of embeddings under [jaːrəð], as shown below:

(10) *[ɛðirˤoħ saxːam] jaːrəð rəzqɛj
IRR=go.AOR.3.SG.M to-home tried:3.SG.M Razqej
intended: 'To go home, Razqej tried.'

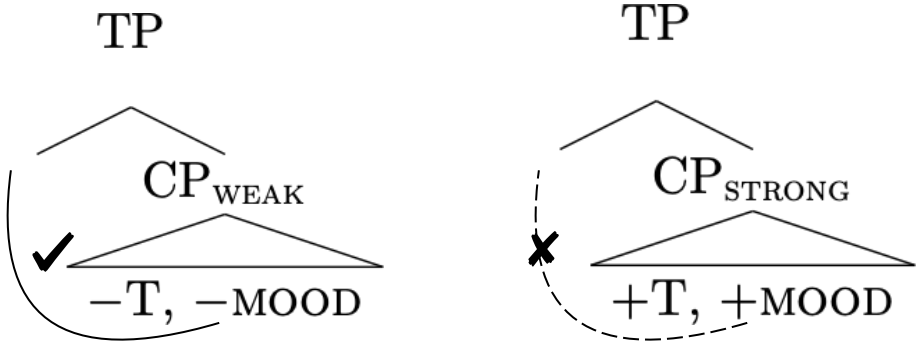
The embedded clause cannot be fronted in (10), which shows that the embedding is a true complement here, but probably just an adjunct in (8)-(9).

The major internal differences between embeddings under [jin:ɛd] "say" and [ja:rəð] "try" are then the following:

(11)	***Embedding under***	jin:ɛd "say"	ja:rəð "try"
	T-deficient	✗	✓
	Mood-deficient	✗	✓
	Inflected	✓	✓

While both embedded domains host an inflected verb, they are distinguished by their tense and mood-properties. In the terminology of Chomsky (2001), embeddings under [jin:ɛd] "say" represent *complete* and hence *strong phases*, and the embeddings under [ja:rəð] "try" represent *incomplete* and hence *weak phases*. While strong phases will spell-out and be inaccessible for further operations, weak phases will delay spell-out:

(12) *Weak vs. Strong embedded phases*



Because weak phases delay spell-out, they "open up" the embedded domain for further interactions with the matrix clause, such as Control formation, Raising, etc. (Boeckx et al. 2010; Gallego 2010). This means that the embeddings under [ja:rəð] are an ideal candidate for one of these cross-clausal operations. In the following sections, we will show that a Control-relation is formed across the weak phasal boundary in question.

It should be noted that the representation in (12) only displays the basic distinction between strong and weak phasal domains. It is not entirely clear whether the embedded domain under [ja:rəð] is a full CP at all; it is just as possible that it is merely a TP — or perhaps an even smaller domain. This is something that future work will need to flesh out. For now, it is only relevant to remember that the embedding under [ja:rəð] seems to be tense and mood-deficient, which contrasts with embeddings under verbs such as [jin:ɛd].

4 ANALYZING THE EMBEDDED "GAP"

Now that we have established that domains embedded under [jaːrəð]-type verbs are deficient, we must determine what types of constructions they form. To do this, we must determine what the status of the embedded "gap" is. It seems that the gap is obligatorily bound by the matrix DP, it also must be covert, as it cannot host a full DP:

(13) *jaːrəð rəzqɛj [ɛtsərˁoħ **kinza** saxxam]
tried:3.SG.M Razqej IRR=go.AOR.3.SG.F Kinza to-home
intended: 'Razqej tried for Kinza to go home.'

(13) contains an overt full DP in the embedded gap-position, where the DP does not refer to the matrix subject. In principle, we have three possibilities for analyzing this gap:

(14) *Analyzing the gap*
Hypothesis A: the gap is occupied by *pro* (prolepsis)
Hypothesis B: the gap is occupied by PRO (Control)
Hypothesis C: the gap is occupied by a deleted copy (Raising)

We will show that PRO must be occupying the embedded gap in this construction. To do this, we must first rule out the remaining options.

4.1 AGAINST A *pro*-ANALYSIS

Kabyle is a *pro*-drop language, which means that arguments do not need to be overt in all situations. This makes determining the status of the embedded gap a little harder than in a non-*pro*-drop language. *pro* usually occupies the positions of subjects, but it can be made overt in a context where the subject contrasts with some other individual:

(15) (nətsa) ɛjθəts pidza
(he) eats:3.SG.M pizza
'He is eating pizza.'

This is an expected property of *pro*-drop languages. If the gap is occupied by *pro*, then we expect it to also follow this pattern, but this is not the case:

(16) *jaːrəð$_i$ rəzqɛj [**nətsa**$_i$ ɛð=irˁoħ saxːam]
tried:3.SG.M Razqej he IRR=go.AOR.3.SG.M to-home
intended: 'Razqej tried he to go home.'

As soon as we introduce an overt pronoun in the embedded clause, ungrammaticality follows — even if we set up a contrastive context for the embedded clause. However, it is possible that we are dealing with a type of *prolepsis*, where some operator requires an embedded *pro* that must stay covert, as in the analysis of prolepsis proposed by Salzmann (2017). But even in such cases, the

hosted *pro* can corefer with different referents and is not subject to strict binding from the matrix clause.

A good diagnostic for showing the distinction between an embedded *pro* and some other element in such cases is the interpretation of the gap under VP-ellipsis (Hornstein 1999; Boeckx et al. 2010). Consider the interpretation of the gap in the elided constituent in (18), with (17) being in the background:[3]

(17) jin:ɛd rəzqɛj$_i$ [bəl:i jittʃa *pro*$_i$ pidza]
said:3.SG.M Razqej that ate:3.SG.M pizza
'Razqej said that he ate pizza.'

(18) kinza$_j$ ðɛʁən ~~. . . [bəl:i . . . *pro*$_{i/j}$ pidza]~~
Kinza too/even
'Kinza did so, too.'

Notice that the embedded clause in (17) contains a standard case of *pro* that can be made overt if needed (not shown here). In the elided part of (18), when uttered after (17), the elided embedded *pro* can refer to either Razqej or to Kinza. In other words, (18) is ambiguous between a *strict reading* of the gap (where the gap refers to Razqej) and a *sloppy reading* (where the gap refers to Kinza). This is expected of prolepsis.

However, once the same diagnostic is applied to the cases involving verbs such as [ja:rəð], the ambiguity disappears:

(19) ja:rəð rəzqɛj$_i$ [að=ittʃ ____$_i$ pidza]
tried:3.SG.M Razqej IRR=eat:3.SG.M pizza
'Razqej tried to eat pizza.'

(20) kinza$_j$ ðɛʁən ~~. . . [. . . ____$_{*i/j}$ pidza]~~
Kinza too/even
'Kinza did so, too.'

This diagnostic shows that the construction in (19) *cannot* involve prolepsis. Specifically, in (20), the elided gap can *only* refer to Kinza, which means that only a *sloppy reading* is available, but not a *strict one*, where the gap could reference Razqej. This means that (20) either involves Raising or Control, since both Control and Raising yield only sloppy readings under ellipsis.

Before we move on with discussing the distinctions between Control and Raising, let us briefly consider a possible objection to the ellipsis diagnostic used above. One might wonder whether it is problematic that ellipsis elides so much structure in (17), crucially involving the Control verb and with it possibly something like a prolepsis 'operator', so that the potential embedded *pro* cannot get a strict reading here. However, this seems fairly unproblematic: as ellipsis is usually claimed to

[3] The usual way to employ this diagnostic is to give the full sentence, such as (14), as the first element in a coordination, and the sentence with ellipsis as the second element of the same coordination. However, the consultant found ellipsis very marginal in coordinations, but found it very natural if (14) was uttered, then followed by a question such as "What about Kinza?", which was in turn followed by (15).

be a PF-phenomenon — and in some sense it needs to be, since the LF-side of the elided constituent remains intact for interpretation — it is unlikely that any actual syntactic/semantic content would get deleted here. This means that the distinction between the sets of readings in (14)-(15) and (16)-(17) must lie in the differing syntactic configurations.

4.2 AGAINST RAISING BUT FOR CONTROL

We have now ruled out the analysis that posits a *pro* in the embedded clause under verbs such as [ja:rəð]. In this section, we will show that the gap cannot be a deleted copy of a Raised DP either. A standard diagnostic that teases apart Raising from Control is the interpretation of embedded idioms:

(21) *Raising*
The cat seems [~~the cat~~ to be out of the bag].
Idiomatic meaning: ✓

(22) *Control*
The cat tried [PRO to be out of the bag].
Idiomatic meaning: ✗

We can test two idioms to show that the discussed construction does not involve Raising. First, consider the idiom *"Foks ate Razqej*", which has the meaning "Razqej was late":[4]

(23)

rəzqɛj	jɪtsɛθ	foks
Razqej.OBJ	ate:3.SG.M	Foks

'Foks ate Razqej.' (lit.)
'Razqej is late.' (idiom.)

As soon as the idiom in (23) is embedded under the verb [ja:rəð], with the subject *Foks* occurring in the matrix clause, the idiomatic meaning is lost:

(24)

ja:rəð	foks	[ɛð=jittʃ	rəzqɛj]
tried:3.SG.M	Foks	IRR=eat.AOR.3.SG.M	Razqej

'Foks tried to eat Razqej.' (lit.)
#'Razqej tried to be late.' (idiom.)

The same effect can be observed with other idioms. Consider the following one:

(25)

θɛrɣʲɛɣʲi	θɛsɛw
shook:3.SG.F	my.liver

'My liver shook.' (lit.)
'I had a premonition of something.' (idiom.)

[4] *Foks* is the name of a specific dog.

The idiom "*My liver shakes"* is uttered when the speaker has a premonition of something. When this idiom is embedded under the verb [ja:rəð], the idiomatic meaning is again completely lost:

(26) θa:rˤəðˤ θɛsɛw [ɛ=tsɛrɣʲɛɣʲi]
tried:3.SG.F my.liver IRR=shake:3.SG.F
'My liver tried to shake.' (lit.)
#'I tried to have a premonition of something.' (idiom.)

This shows that raising cannot be involved. It seems therefore that the construction that we are dealing with is truly a *Control construction*. It passes all the relevant diagnostics for Control: it involves obligatory binding of a covert element in the embedded clause, it only yields *sloppy* readings under ellipsis, and it does not preserve the interpretation of idioms when they are "split" between the matrix and embedded domains.

4.3 AN INSTANCE OF OBJECT CONTROL

While embeddings under [ja:rəð] "want" only yield Subject Control, it seems that embeddings under [jivʁa] "want" (see (3)) can also give rise to Object Control. First, we can demonstrate that the Subject Control pattern with [jivʁa] is truly an instance of Control and not some Raising operation:

(27) θivʁa θɛsɛw [ɛ=tsɛrɣʲɛɣʲi]
wanted.3.SG.M my.liver IRR=shake:3.SG.F
"My liver wanted to shake." (lit.)
'A premonition wanted to come about.' (idom.)

The idiom already used in the previous section cannot be interpreted if it is split between the two clauses in (27), which shows that the matrix subject [θɛsɛw] is binding an embedded PRO. However, the verb [jivʁa] can specify an additional internal *θ-role*; when this occurs, Object Control is formed. Interestingly, Ouali (2019: 16) claims that this construction involves subject-to-object Raising (he states this generally for Tamazight). But the language consultant, who is a Kabyle speaker, rejected the non-compositional interpretations of idioms when embedded in such contexts:

(28) jivʁa rəzqɛj θɛsɛs [ɛ=tsɛrɣʲɛɣʲi]
wanted.3.SG.M Razqej his.liver IRR=shake:3.SG.F
'Razqej wanted his liver to shake.' (lit.)
'Razqej wanted to have a premonition of something.' (idom.)

In other words, cases like (28) seem to be instances of Object Control rather than Raising: if the object [θɛsɛs] raised from the embedded subject position, it should be possible to reconstruct it for the interpretation of the embedded idiomatic phrase. But this is not the case, at least not in the version of Kabyle spoken by our consultant. This either means that there could be dialectal variation at work, or that Ouali (2019) has not considered the interpretation of embedded idioms in this construction.

5 Conclusion

This paper has shown that the embeddings under verbs such as [jaːrəð] "try" and [jivʁa] "want" are instances of Finite Control in Kabyle Berber. I showed that this construction cannot involve an embedded *pro* because the "gap" cannot be made overt under any circumstances, and it fails to give rise to "strict identity" readings under VP-ellipsis. I also demonstrated that it cannot host a deleted copy of a raised DP-argument because idiomatic phrases, when split between the matrix and embedded clauses, never retain their non-compositional readings. The embeddings in question are also tense-deficient, as well as mood-deficient, which renders them *weak phases* (Chomsky 2001), creating the expected environment for Control-formation. This paper contributes to the ongoing investigation of Finite Control patterns in the languages of the world.

REFERENCES

Boeckx, C., Hornstein, N., & Nunes, J. (2010). *Control as Movement*. Cambridge: Cambridge University Press.

Boukhris, F. (2013). *Les clitiques en Berbere Tamazighte: Approche Minimaliste*. Publication de l'institut Royal de la Culture Amazighe Serie: Etudes et Recherches N 38. IRCAM, Rabat.

Chomsky, N. (2001). Derivation by Phase. In Kenstowicz, M. (Ed.), Ken Hale. *A Life in Language*, (pp. 1–52)., Cambridge, MA. MIT Press.

Ferreira, M. (2004). Hyperraising and null subjects in Brazilian Portuguese. *MIT Working Papers in Linguistics 47:* Collected Papers on Romance Syntax, 57–85.

Ferreira, M. (2009). Null subjects and finite control in Brazilian Portuguese. In J. Nunes (Ed.), *Minimalist Essays on Brazilian Portuguese Syntax* (pp. 17–49). Amsterdam/Philadelphia: John Benjamins.

Gallego, A. (2010). *Phase Theory*. Amsterdam/Philadelphia: John Benjamins.

Hornstein, N. (1999). Movement and control. *Linguistic Inquiry*, *30,* 69–96.

Landau, I. (2004). The scale of finiteness and the calculus of control. *Natural Language and Linguistic Theory*, *22,* 811–877.

Ouali, H. (2019). On the syntax of tense and aspect in Tamazight. Handout from talk given at the *Workshop on Amazigh Languages*. McGill University, Montreal, QC. 21-22nd March.

Salzmann, M. (2017). Prolepsis. In M. Everaert & H. van Riemsdijk (Eds.), *The Wiley Blackwell Companion to Syntax, 2nd edition* (pp. 3203–3245). Oxford: Wiley-Blackwell.