The Interaction of state, prosody and linear order in Kabyle (Berber)
Grammatical relations and information structure

Amina Mettouchi
Ecole Pratique des Hautes Etudes, Paris, and CNRS-LLACAN

The aim of this paper is to show how, starting only from forms belonging to various domains (morphology, syntax, and prosody), it is possible to compute grammatical relations and information structure constructions in Kabyle spontaneous speech.

The study is based on recordings made in the field, transcribed, translated, and annotated with Praat and Elan-CorpA. The methodology consists of systematically retrieving the sequences containing a verb and looking for the presence of a noun (and its inflection) within the prosodic group of the utterance, or outside, as well as studying the linear orders involved. This non-aprioristic methodology reveals the close interaction between grammatical relations and information structure in Kabyle. The study provides evidence to support the claim that the encoding of grammatical relations on nouns is a by-product of information structure constraints in Kabyle.

Keywords: prosody, information structure, grammatical relations, Kabyle, Berber

1. Introduction

The aim of this paper is to show how, starting only from forms belonging to various domains (morphology, syntax, and prosody), it is possible to compute grammatical relations and information structure constructions in Kabyle spontaneous speech.

The study is based on data recorded in the field, transcribed, translated, and annotated with Praat and Elan-CorpA.

1. The hypotheses in the present paper were first presented in June 2011 at the 14th International AfroAsiatic Linguistics Conference in Turin.

DOI 10.1075/cilt.339.14met
© 2018 John Benjamins Publishing Company
The methodology consists of systematically retrieving the sequences containing a verb, and looking for the presence of a noun (and its inflection) within the prosodic group of the utterance, or outside, as well as studying the linear orders involved.

From such forms, it is possible to ask fundamental questions about the nature of grammatical relations and the structures associated with various information structure values in Kabyle.

The hypothesis underlying the paper is that the marking of grammatical relations on nouns is a byproduct of the information structure constraints described in §2.

1.1 General information about Kabyle

Berber languages are spoken in Northern Africa, in a zone delimited by the Atlantic Ocean to the west, the Mediterranean to the north, the oasis of Siwa (Egypt) to the east, and the southern borders of Mali and Niger to the south. Those languages constitute a family within the Afroasiatic phylum. Well-known members of the family are Kabyle (spoken in Northern Algeria), Tashelhiyt (Shilha) (spoken in Southern Morocco), and Tamashek and Tahaggart (also called Tuareg), spoken in Southern Sahara. Kabyle is spoken by about four million people in the north of Algeria. The variety investigated in this paper is a western one, spoken in the village of Ait Ikhlef, close to the town of Bouzeguene. I have collected all the data on fieldwork between 2007 and 2011.4

In Kabyle, as in all Berber languages, a minimal predication consists of a verb and its bound personal pronoun, or a non-verbal predicate. In this paper I focus on verbal predicates. In addition to this core, the clause may contain noun phrases and prepositional phrases, as well as adverbs. Within noun phrases, modifiers follow the modified constituent. The language has two genders and two numbers, marked on pronominal affixes and clitics to verbs, nouns, and prepositions, as well as on adjectives (a subclass of nouns) and on nouns. It also has two states, marked on nouns.

1.2 Relevant coding means

This section describes the formal means that come into play for the encoding of grammatical relations and information structure.5

---

4. My warmest thanks go to the speakers I recorded in the village over the years.

5. Examples are transcribed morphophonologically according to the following rules: IPA symbols are used whenever they consist of a single character. Otherwise, they are replaced by
1.2.1  *The state distinction*

Nouns have two forms, the absolute (ABS, traditionally called ‘état libre’ (‘free state’)), and the annexed (ANN, traditionally called ‘état d’annexion’ (‘annexation state’)). E.g., man = argaz in the absolute, wrgaz in the annexed; girl = taqift in the absolute, tajift in the annexed.

Table 1. Illustration of state alternation in Kabyle

<table>
<thead>
<tr>
<th></th>
<th>Masculine sg</th>
<th>Masculine pl</th>
<th>Feminine sg</th>
<th>Feminine pl</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annexed</td>
<td>w-rgaz “man”</td>
<td>j-rgz-n “men”</td>
<td>t-qif-t “girl”</td>
<td>t-qift-in “girls”</td>
</tr>
<tr>
<td>Absolute</td>
<td>a-rgaz “man”</td>
<td>i-rgz-n “men”</td>
<td>t-a-qift-t “girl”</td>
<td>t-i-qift-in “girls”</td>
</tr>
</tbody>
</table>

The state distinction in itself does not code grammatical roles (Galand 1964; Chaker 1988). Evidence for this fact is developed in Mettouchi & Frajzyngier (2013), where the states are thus defined: the annexed state indicates that the noun provides the value (in the logical sense) for the variable of the function grammaticalized in a preceding constituent (Mettouchi & Frajzyngier 2013). Such functions are diverse. The absolute state is simply the default form of the noun, and as such it has no overall function of its own (Mettouchi & Frajzyngier 2013). Within various structures, the states contribute to the creation of grammatical meaning, thus being the backbone of the grammar of Kabyle.

1.2.2  *Prosody*

An intonation unit is a segment of speech that has a coherent intonation contour (Chafe 1994), and is delimited by its boundaries (Cruttenden 1997), which bear a ‘boundary tone’ (Pierrehumbert & Hirschberg 1990). In Kabyle, Intonation Units are marked by one or more of the following cues:

6. In this table, only underlying forms are given, so that the morphology of the state alternation is clearer. Syllabification rules which result in schwa insertion are reflected in the examples throughout the paper.

All rights reserved
Main external cues:
1. final lengthening;
2. initial rush;
3. pitch reset;
4. pause;
5. creaky voice.

Main internal cues:
1. declination;
2. tonal parallelism, or isotony.

On the basis of those cues (see also Izre’el & Mettouchi 2015), a one-hour corpus composed of 20 minutes of dialogue and 40 minutes of monologue (narrative) was segmented into intonation units, regardless of any other cue, syntactic, semantic, or other.

1.2.3 Linear orders
Linear ordering concerns not only noun phrases and verbs, as it also comes into play for other units, for instance clauses. In this paper, I consider linear ordering as a series of formal means, crucially depending on the existence of a reference point (Frajzyngier & Shay 2003: 60–62), which is overtly and unambiguously coded for this function. This reference point differs depending on the language.

In Berber, the verb is a salient potential reference point, as it is formally always affixed by a dedicated pronominal paradigm. It cannot therefore be confused with any other element of the clause.

Another reference point that I consider here is the prosodic boundary between intonation units (see § 1.2.2). As a discontinuity in the flow of speech, it constitutes a salient potential reference point, regardless of its functional values, which depend on the boundary tone, and other prosodic cues.

In this paper, I focus on the ordering of noun phrases with respect to the verb: before or after the verb; as well as the relative order of noun phrases when this is formally computable. I also take into account the position relative to the prosodic boundary: before or after the prosodic boundary.

1.2.4 Traditional analysis of the structure of the clause in Berber
Since (Galand 1964), the bound pronoun affixed to the verb has been considered as the actual subject of the Berber clause, coreferent noun phrases being expansions of this pronoun.
(1)  
\textit{j-hma}  
\begin{verbatim}
3SG.M-be\_warm:PFV
\end{verbatim}  
“It is/was warm.”

The lexical item coreferent with the subject pronoun is either (Galand 1964) the ‘indicateur de thème’ (‘theme indicator’) (i.e., the noun is in the absolute state and appears before the verb); or the ‘complément explicatif’ (‘explanatory complement’) (i.e., the noun is in the annexed state and appears after the verb).

(2)  
\textit{adjazin j-hma}  
\begin{verbatim}
tagine:\textsc{abs}\_\textsc{m}\_sg 3SG.M-be\_warm:PFV
\end{verbatim}  
“The tagine (cooking pot) is hot”

(3)  
\textit{j-hma u\textit{adjazin}}  
\begin{verbatim}
3SG.M-be\_warm:PFV tagine:\textsc{ann}\_\textsc{m}\_sg
\end{verbatim}  
“The tagine is hot”

The term ‘indicateur de thème’ is not limited to the subject, as the noun preceding the verb can also be coreferent with pronouns other than the subject:

(4)  
\textit{aksum j-\textit{čča}=t}  
\begin{verbatim}
meat:\textsc{abs}\_\textsc{m}\_sg sbj\textsc{3sg.m}\_eat:pfv=\textsc{absv}\textsc{3sg.m}
\end{verbatim}  
“He ate the meat” (‘object’)

(5)  
\textit{argaz-nni t-mmut tm\textsc{țiřut}=is}  
\begin{verbatim}
man:\textsc{abs}\_\textsc{m}\_sg-cns sbj\textsc{3sg.f}\_die:pfv woman:\textsc{ann}\_\textsc{f}\_sg-\textsc{poss}3sg
\end{verbatim}  
“That man, his wife died” (‘possessor’)

The ‘complément explicatif’ can also be coreferential with pronouns other than the subject in Kabyle:

(6)  
\textit{j-\textit{čča}=t w\textit{aksum-nni}}  
\begin{verbatim}
sbj\textsc{3sg.m}\_eat:pfv=\textsc{absv}\textsc{3sg.m} meat:\textsc{ann}\_\textsc{m}\_sg
\end{verbatim}  
“He ate the meat”

And the coreferent pronoun can be affixed to a noun, and not only to the verb:

(7)  
\textit{t-mmut tm\textsc{țiřut}=is wargaz-nni}  
\begin{verbatim}
sbj\textsc{3sg.f}\_die:pfv woman:\textsc{ann}\_\textsc{f}\_sg-\textsc{poss}3sg man:\textsc{ann}\_\textsc{m}\_sg-cns
\end{verbatim}  
“His wife died, that man.” (‘possessor’)

In this traditional analysis, only two structures are taken into account:

‘Indicateur de thème’ – V  
V – ‘Complément explicatif’
Those structures have essentially been used to argue against the attribution of the subject role to noun phrases (Galand 1964). Some publications have gone further in their endeavor to study the information structure of the constructions actually encountered in Kabyle spontaneous speech (Mettouchi 2007a, 2007b, 2008; Kuningas & Leino 2006), but no systematic study of all possible combinations has been undertaken to date.

1.2.5 Constituent order analysis

The present paper adopts a non-aprioristic methodology: I only take into account the formal means involved (state, prosodic boundary, linear order), without assumptions concerning the function of each noun before the structures are examined. Consequently, a wide array of constructions are actually investigated. The potential structures are listed below. Note that the abbreviation \( V_{sbj} \) represents the minimal predication composed of a verb and its obligatory personal affix (and possibly other bound pronouns).

\[
\begin{align*}
&[V_{sbj}] \\
&[V_{sbj} \text{NP}_{abs}] \\
&[V_{sbj} \text{NP}_{ann}] \\
&[V_{sbj} \text{NP}_{ann} \text{NP}_{abs}] \\
&[V_{sbj} \text{NP}_{abs} \text{NP}_{ann}] \\
&[\text{NP}_{abs} V_{sbj}] \\
&[\text{NP}_{abs} V_{sbj} \text{NP}_{abs}] \\
&[\text{NP}_{abs} V_{sbj} \text{NP}_{ann}] \\
&\text{NP}_{abs} [V_{sbj} (\text{NP}) (\text{NP})] \\
&[V_{sbj} (\text{NP}) (\text{NP})] \text{NP}_{ann}
\end{align*}
\]

When the NP does not bear the indices “ann” or “abs” in the list above, it means that both states are possible in this position. The state opposition is an independent coding means whose function is not to mark grammatical relations or information structure (cf. §0.2.1). Indication of the state borne by the nouns, however, is relevant for distinguishing among some structures: for instance between \([V_{sbj} \text{NP}_{abs}]\) and \([V_{sbj} \text{NP}_{ann}]\), which do not have the same information structure value.

Noun phrases following prepositions have not been taken into account here because they are unambiguous as far as function is concerned: indirectly affected

---

7. The square brackets [ and ] are used to represent the presence of prosodic boundaries, which are transcribed / or // in actual examples taken from my recordings.

8. Prepositional phrases have not been included, but another study in progress, on linear ordering of PPs, shows that their presence has no effect on the information structure values of the combinations listed in this section.
argument is introduced by preposition *i*, and instrumental complement is introduced by preposition *s*, while locative complements are introduced by prepositions *g*, *sag*, *yər*, among other.

2. Information structure

In investigating the functions of those structures, five constructions were found, each consisting of one or more structures and having one function:

- \[V_{sbj} (N_{abs})\], which subsumes the following:
  \[V_{sbj}\]
  \[V_{sbj} N_{abs}\]
- \[V_{sbj} N_{ann} (N)\], which subsumes the following:
  \[V_{sbj} N_{ann}\]
  \[V_{sbj} N_{ann} N_{abs}\]
  \[V_{sbj} N_{abs} N_{ann}\]
- \[N V_{sbj} (N)\] which subsumes the following
  \[N_{abs} V_{sbj}\]
  \[N_{abs} V_{sbj} N_{abs}\]
  \[N_{abs} V_{sbj} N_{ann}\]
- \[N_{abs} [V_{sbj} (N) (N)]\]
- \[V_{sbj} (N) (N) N_{ann}\]

The exact functions of those constructions will be investigated one by one, in main and independent clauses.

2.1 Function of \([V_{sbj} (N_{abs})]\)

This construction is the default one in terms of syntax (it is in itself a full-formed clause) and information structure. It has no marked value but builds on previous context without any shift or change of perspective. All examples involving only a verb and its bound subject pronoun, possibly followed by a noun in the absolute within the same intonation unit, were encountered inside a subtopic in a narrative or a conversation.

Subtopics are discourse-level topics that rank lower than a basic-level topic but higher than a sentence topic. A topic is “an aggregate of coherently related events, states and referents that are held together in some form in the speaker’s semi-active consciousness” (Chafe 1994: 121). Chafe implies that a subtopic should encompass more than an intonation unit, since an intonation unit is generally associated with a “focus of consciousness” related to “active” information (1994: 29).
In this discursive context, using the verb (with its obligatory subject pronoun), possibly followed by a noun in the absolute – which is never coindexed with a preceding pronoun – (such as lbir, “well”, the object of i-xdəm, “he made”) is the unmarked informational function, namely (sub-) topic continuity: the protagonist is the same, and the narrative is carried forward. This function is very well described in the general literature across languages (Givón 1983; Chafe 1994; Lambrecht 1994); for Kabyle it has been analyzed in Mettouchi 2008. Even if there are several protagonists, only bound pronouns are used (cf. Mettouchi 2005, 2007b).

Once noun phrases other than the direct object – i.e., noun phrases coreferent to a bound pronoun – appear, the information structure value is changed. This shows that the mere presence of a noun phrase coreferent to a bound pronoun, in a pronominal-argument language such as Kabyle, has information structure value. Which value this is depends on the position of that noun with respect to the verb and the prosodic boundary, as well as on the state of the noun (absolute or annexed).

2.2 Function of \([V_{sbj} \, NP_{ann} \, (NP_{abs})]\)

This construction is realized as three different structures, depending on the presence of two NPs or one. The important factor is the presence of a noun in the annexed state in the position after the verb (possibly separated from it by another noun, in the absolute), within the same intonation unit as the verb.

- \([V_{sbj} \, NP_{ann}]\)
- \([V_{sbj} \, NP_{ann}, NP_{abs}]\)
- \([V_{sbj} \, NP_{abs}, NP_{ann}]\)

The three structures promote an event or a state to topic status: both the main participant and the predicate are expressed, and even if the participant is known or
mentioned, the relationship it holds with the predicate is presented as new and is about to be developed in the following intonation units. This structure was analyzed as ‘sentence-focus’ or ‘thetic’ in Mettouchi 2008, because newness was not limited to the referent, but encompassed the state of affairs itself, represented by the association of a verb (and its bound subject pronoun) and a noun in the annexed state. That analysis is true at the level of the sentence, but it doesn’t provide information about the role of the construction in discourse.

An examination of my corpus show that in discourse the structure is used to introduce a new episode in a narrative or a new subtopic in a conversation. Sometimes, as in the following example, it is mentioned as a disclosure:

(9) sp3:  
\textit{anda lla-nt tleta t\textsuperscript{3}q\textsuperscript{fi}n} //
where be:PFPV-sbj\textsuperscript{3}PL.F three girls:ANN //
“Where are those three girls?”
sp2:  
\textit{jahi / t\textsuperscript{3}mmut lwiza} //
interjection sbj\textsuperscript{3}SG.F-die:PFPV Louisa //
“Actually, Louisa died.”

The preceding discourse was structured around the subtopic of a woman’s number of children. When it was clear that the two speakers hadn’t come up with the same number of children, Speaker 2 added the crucial information that a daughter had died. Since the statement is sad, it is soon replaced by another a piece of information: the names of the other two daughters, presented with the same structure:

(10) sp2:  
\textit{t\textsuperscript{3}lla nadija / f\textsuperscript{adi}la }\textsuperscript{:::} //
sbj\textsuperscript{3}SG.F-be:PFPV Nadia / Fadila HESIT //
“There is Nadia, Fadila…”
sp1:  
\textit{f\textsuperscript{adi}la d nadija d tin\textsuperscript{gg}gura} //
Fadila ASSOC Nadia COP last:ABS.PL.F
“Fadila and Nadia are the youngest ones.”

This structure is used typically in \textit{wh}- questions with presupposed information, as in Speaker 3’s question above, or for new information (regardless of the activation state of the referent itself).

Sometimes, a noun in the absolute also appears in the structure: either after the noun in the annexed state ($[V_{sbj} N_{ann} N_{abs}]$) or before ($[V_{sbj} N_{abs} N_{ann}]$)
In the following example, the father, prompted by his sons, is sending his younger son to steal a beautiful carpet from an ogress:

\[(11)\]  
\[ ja\-mna=jas \quad zra\-n \quad wajtma\-s \]
\[ sbj3sg.m\-say:pfv=dat3sg \quad see:pfv=sbj3pl.m \quad brother:ann.pl\-kin3.sg \]
\[ ta\- Zarbit \quad ar \quad jomma \quad Nu\-3a / \]
\[ carpet:abs.sg.f \quad to \quad mother:ann.f.sg \quad Nu\-3a \]

“The father said his brothers had seen a carpet at Mother Nu\-3a’s.”

This new piece of information is the basis of his demand to his son, that he should go and steal it from the ogress.

In the following example, taken from the same tale, the brothers come to the father with the news that the ogress has a hen, whose eggs heal all sorts of illness.

\[(12)\]  
\[ t\-s\-ta \quad tajazi\-t \quad jomma \quad Nu\-3a / \]
\[ sbj3sg.f\-possess:pfv \quad hen:abs.sg.f \quad mother:ann.f.sg \quad Nu\-3a / \]

“All three structures belong to the same construction, as they all have the same informational value; the difference is first between intransitive and transitive predication, and second, among the latter, between the default order, which is \([V\_sbj\_N\_ann \_N\_abs]\) if both nouns have equal weight but which becomes \([V\_sbj\_N\_abs \_N\_ann]\) if the noun in the annexed state is heavier (in terms of information status or length).

Note that, apart from construction 1.1, this construction is the only one that can be found in a dependent clause (in particular in relative clauses, but also in complement clauses). This is in keeping with the fact that the construction globally construes the event or situation, it does not comment on a topic, or reactivate a referent.

2.3 Function of \([N\_abs \_V\_sbj \_N]\)]

Within the same prosodic unit, a noun can appear before the verb. This noun is always in the absolute state (the annexed state is only used after a grammatical morpheme encoding a function, for which the noun in the annexed state is the variable (Mettouchi & Frajzyngier 2013).

This construction, which has not been described for Kabyle yet, is used as a background for further developments, when a salient preceding situation is recapitulated, so that the listener grasps the whole situation and its importance for the current discourse.

In the following example, all aspects of the situation have already been introduced, mostly through \([V\_sbj\_N\_ann \_N\_abs]\) structures. Most referents have been previously mentioned, as is shown by the suffix -\_mni, which marks shared reference (Mettouchi 2006, 2011: 482) (often through previous mention).
In the first line, the noun in the absolute is the subject, as in the third line, but in the second line, it is an object and is taken up by the absolutive pronoun = t. What is important here is that one argument appears before the verb but not separated from it by a prosodic boundary. The construction allows the reduction of a long and complicated story into its salient characteristics.

In conversation, similar recapitulations occur, as in the following example, where the speaker takes up information scattered in the preceding context – where various brothers of her grandfather’s, as well as the grandfather himself, were said to have married a number of women – and then goes on to comment on the genealogy of the family:

(14) ʒəddi j-uy sətti /
  grandfather:ABS.SG.M  SBJ3SG.M-take:PFV  grandmother:ABS.SG.F /
  “My grandfather married my grandmother.”

This construction must not be confused with the following one, where the noun in the absolute preceding the verb is also before the prosodic boundary that precedes the verb.

2.4 Function of $N_{ABS} \left[ V_{sbj} (N) (N) \right]$

This construction is characterized not by the internal structure of the intonation unit but by constituent ordering with respect to the verb (possibly preceded by a particle or auxiliary, as in the following example with negation), and prosodic boundary. The argument preceding the prosodic boundary is taken up by a bound pronoun in the clause. Such constructions always imply a shift in perspective or

---

9. Note that preverbal particles (modal, aspectual, or negative) have fixed positions, and nothing else than a string of clitics can separate the particle from the verb. This makes the particle an alternative reference point to the verb for the calculation of linear ordering.
contrast with previous expectations, as in the following example. They are not topic-promoting devices, where a referent is promoted from non-active state to active state as argued in Kuningas & Leino (2006); indeed it is another structure which has this function in Kabyle: \([V_{\text{subj}} (N) (N)] N_{\text{ANN}},\) treated in §1.5.

In the following example, what is important is the fact that the presupposition concerning the stepmother’s relationship to her husband’s daughters (built throughout the previous episode: she promised she would take care of them and love them dearly) is rejected.

\[
(15) \quad \text{aj argaz tura jassi-k-agi} / \vspace{2mm}
\text{VOC man:ABS.SG.M now daughter:ABS.PL-KIN2SG.M-PROX} / \vspace{2mm}
\text{“My husband, now those daughters of yours,} \vspace{2mm}
\text{ur zaddy-ay ara jid-sont} // \vspace{2mm}
\text{NEG dwell:IPFV-SBJ1SG POSTNEG COM-PREP3PL.F //} \vspace{2mm}
\text{I’m not living with them!”}
\]

Similarly, in the next example, the father had given his seven wives apples. They all gave birth to a normal boy, except the last one, who had eaten only half an apple:

\[
(16) \quad \text{tin iwumi ja-fka akka mnzfs} / \vspace{2mm}
\text{the_one:SG.F to_whom SBJ3SG.M-give:PFV thus half} \vspace{2mm}
\text{“The one to whom he had given half} \vspace{2mm}
\text{n tətəftaht} / \text{t-urw=ədd} / \vspace{2mm}
\text{GEN apple:ANN.SG.F / SBJ3SG.F-give_birth:PFV=PROX human_being:ABS.SG /} \vspace{2mm}
\text{an apple, she gave birth to a human being,} \vspace{2mm}
\text{qqar-n=as} / \text{atmar mnzfs} // \vspace{2mm}
\text{say:IPFV-SBJ3PL.M=DAT3SG / Aṣmar half //} \vspace{2mm}
\text{whom people called Amar the Midget,} \vspace{2mm}
\text{t-urw=it=idd} \text{d mnzfs} // \vspace{2mm}
\text{SBJ3SG.F-give_birth:PFV=ABS3SG.M=PROX COP half //} \vspace{2mm}
\text{it was half a child she had given birth to.”}
\]

I suggest to call those structures ‘contrastive comments’, since they go against a presupposition about the topic which was built in the preceding context.
2.5 Function of $[V_{\text{sbj}} (N) (N)] N_{\text{ANN}}$

Those structures are characterized by the presence of a noun in the annexed state\(^{10}\) after the prosodic boundary. Here again, the internal composition of the prosodic unit is not important. It is the state of the noun and the position after the boundary, which formally identify the construction. This type of construction is often called 'right-dislocated' in the literature. I will not use that term because I make no assumptions concerning possible underlying structures.

All examples involving such a construction are used to activate a referent that had lost its active (and even semi-active) status. This reactivation is generally associated with further continuation of the discourse with the activated referent as topic. In the following example, the house is reactivated after having stayed unmentioned for a few intonation units, and then the following subtopic starts (the girls set to explore the house, which is described in details).

(17) $t$-ufa $d$ amfif $n$ wədrar //  
   sbj3sg.f-find:pfv cop cat:abs.sg.m gen mountain:ann.sg.m //
   “She found it was the Mountain Cat
   $i=t$ izədyon / wəxxam-nni //  
   rel.real=absv3sg.m dwell:pfv:relsbj.pos / house:ann.sg.m-cns //
   who inhabited it, the house.”

2.6 Synthesis on information structure

The study of formal sequences based on position relative to the verb, and to a prosodic boundary, provides a number of structures which can be grouped together on the basis of function.

Construction $[V_{\text{sbj}} (N_{\text{abs}})]$, without any surrounding nouns belonging to the clause outside the prosodic group of the verb, marks (sub-)topic continuation. Typically, those structures are realized as sequences of verbs with their obligatory person affix, possibly complemented by nominal direct objects.

$[V_{\text{sbj}}]$

$[V_{\text{sbj}} N_{\text{abs}}]$

---

\(^{10}\) If the noun is in the absolute state and after the prosodic group of the verb in Kabyle, it cannot be related to the preceding verb, but has to be construed as starting a new clause. This is due to the function of the state distinction (see Mettouchi & Frajzyngier 2013).
Construction $[V_{\text{subj}} \ N_{\text{ann}} \ (N)]$ has no surrounding nouns belonging to the clause outside the prosodic group of the verb either, but there is at least one noun after the verb, and it bears the annexed state. Typically, those are structures where the nominal subject follows the verb within the same prosodic unit. A nominal direct object can be present too, and in that case both relative orders of nouns are possible, without any change in the function of the construction. The function of this construction is topicalization (promotion to topic status) of an event or a state, in a thetic perspective. This construction is used to present situations or events as a whole as new, regardless of the activation status of the referents themselves: they inform the listener, providing a statement that triggers a new subtopic.

$[V_{\text{subj}} \ N_{\text{ann}}]$
$[V_{\text{subj}} \ N_{\text{ann}} \ N_{\text{abs}}]$
$[V_{\text{subj}} \ N_{\text{abs}} \ N_{\text{ann}}]$

Construction $[N_{\text{abs}} \ V_{\text{subj}} \ (N)]$ has no surrounding nouns belonging to the clause outside the prosodic group of the verb either, and its defining feature is the presence of a noun in the absolute state just before the verb and still within the prosodic group containing the verb. Another noun can appear after the verb. This construction recapitulates salient elements of a situation which have been exposed before at some length. It is a summary, which provides backgrounding for the following discourse.

$[N_{\text{abs}} \ V_{\text{subj}}]$
$[N_{\text{abs}} \ V_{\text{subj}} \ N_{\text{abs}}]$
$[N_{\text{abs}} \ V_{\text{subj}} \ N_{\text{ann}}]$

Construction $N_{\text{abs}}[V_{\text{subj}} \ (N) \ (N)]$ is characterized by the presence of a noun in the absolute before the prosodic boundary opening on the prosodic unit containing the verb. This noun has to bear the special continuative (rising) boundary tone that links it to the following sequence; otherwise it would not be interpreted as a topic but as belonging to the previous clause. This construction is binary, as it involves a topic and a comment. The topic is a referent reactivated from the previous discourse, but the important element here is the comment: it goes against a presupposition about the topic that was built in the preceding context. This is why I propose to label the function of this construction ‘contrastive comment’.

Construction $[V_{\text{subj}} \ (N) \ (N)] \ N_{\text{ann}}$ is characterized by the presence of a noun in the annexed state after the right prosodic boundary of the prosodic unit containing the verb. The annexed state here is the mirror image of the continuative boundary tone on the topic in the absolute in the abovementioned construction: its role is to tie this noun to the clause and to indicate that it does not belong to the following clause. The function of this construction is the reactivation of a participant for topic promotion.
3. Grammatical relations

I have just shown that constructions characterized by word order and prosodic grouping, and to some extent state alternation, had specific informational values in discourse. My hypothesis in this second part is that information structure values imply that in some cases grammatical relations on nouns should be transparent, whereas in other cases, they needn’t be. Indeed, in this section, I show that grammatical relations are not systematically coded on nouns. For this, I am using the same non-aprioristic method as in Part 1, with a query in mind: to what extent are prosodic grouping, state alternation, and word order also involved in the encoding of grammatical relations in Kabyle, at the level of the clause? This way of investigating things makes it possible to show how discourse and clause-level grammar interact in spontaneous speech in Kabyle.

The bound pronouns that are obligatorily affixed to the verb form a special paradigm that codes the main participant in the situation, regardless of semantic role, animacy, topicality, etc. This paradigm can therefore be considered as a subject paradigm. As the verb and its subject affix alone constitute a well-formed sentence, I do not consider them as agreement markers, but as a case of pronominal argument marking (Galand 1964; Mettouchi 2005). Other bound pronominal paradigms are absolutive clitics (objects of transitive verbs and main participant of non-verbal predicates) and dative clitics (affected participant).

All pronominal paradigms in Kabyle are inflected for person, number, and gender (in the second and third person (singular and plural) for the subject and

---

11. A preliminary version of this section was written during a period of collaborative work with Zygmunt Frajzyngier on the state opposition in Kabyle. I am grateful to him for his comments on those hypotheses and for the role his methodology played in the elaboration of my argumentation.

12. Here is the subject pronominal paradigm, with perfective radical -uli-, “go up”, “climb”.

<table>
<thead>
<tr>
<th></th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>uli-ɣ</td>
<td>n-uli</td>
</tr>
<tr>
<td>2m</td>
<td>t-uli-dˤ</td>
<td>t-uli-m</td>
</tr>
<tr>
<td>2f</td>
<td>t-uli-mt</td>
<td></td>
</tr>
<tr>
<td>3m</td>
<td>j-uli</td>
<td>uli-n</td>
</tr>
<tr>
<td>3f</td>
<td>t-uli</td>
<td>uli-nt</td>
</tr>
</tbody>
</table>

13. Clitics are characterized by their ability to climb from the default position after the verb to the position after the negative, modal, or aspectual particle preceding the verb or to the position after the relative marker in relative clauses.
absolutive paradigms, in the second (singular and plural), and third (plural) for the dative paradigm).

(18) tə-wwot=it
    3SG.F-kick.PFV=ABSV3SG.M
    “She kicked it.”

(19) ula=j=it
    EXNEG=ABSV3.SG.M
    “He is not there.”

(20) tə-mmut=as təqift
    3SG.F-die.PFV=DAT3SG.M girl.ANN.SG.F
    “She lost a daughter (lit. a girl died on her).”

In this context, noun phrases are often used as expansions of the bound pronouns. This is true for NPs coreferent to subject affixes, or absolutive ones. Indirect objects are introduced by preposition i, and they won’t be considered here because their role is quite transparent given their prepositional phrase structure. Only one type of NP is not coreferent to any pronoun, and this is the nominal direct object, as will be shown below.

Formally, all we can rely on, for the computation of grammatical relations on nouns, is the state of the noun, its gender-number markers, its position with respect to the verb and the prosodic boundary, and the gender-number markers on the pronominal affixes and clitics.

I will now show the following:

a. the state opposition in itself does not mark grammatical relations;
b. coreference in gender and number between the noun and the bound pronoun, in itself, is not transparent for the encoding of grammatical relations; and
c. word order in itself does not mark grammatical relations.

However, the interaction of state, word order, and prosodic grouping allows the computation of grammatical relations for nouns.

3.1 Grammatical relations are not marked unambiguously by one coding means

3.1.1 The state opposition in itself does not mark grammatical relations
This first statement is demonstrated by the fact that the noun in the annexed state can be a subject or an object. If the same mark is compatible with two grammatical relations that are in principle incompatible, then its function is not to indicate grammatical relations. In the following example, the noun in the annexed state has the same referent as the subject pronoun of the clause:
In the following example, the noun “house” in the annexed state has the same referent as the ‘object’ (absolutive) pronoun: 

(22) \( t\text{-ufa} \quad d \quad amsi\text{id} \quad n \quad w\text{drar} \)

\[ \begin{array}{c}
\text{SBJ3SG.F-find:PFV COP cat.ABS.M.SG GEN mountain.ANN.M.SG} \\
i=t \quad iz\text{dy}on \\
\text{REL=ABSV3SG.M inhabit:PFV:RELSBJ.POS} / \text{house.ANN.M.SG-CNS} \\
\end{array} \]

“She found it was the Mountain Cat who inhabited it, the house.”

3.1.2 Coreference in gender and number in itself does not mark grammatical relations

This second statement is demonstrated by the fact that a noun can have the same features of gender and number than the subject affix, without being coreferent with it. This shows that identity in number and gender (often called ‘agreement’) is not sufficient to mark grammatical role.

(23) \( j\text{-kr\oe}z \quad igor \)

\[ \begin{array}{c}
\text{SBJ3M.SG-plough.PFV field:ABS.M.SG} \\
\end{array} \]

“He ploughed the field” (and not “The field is ploughed”, for which the noun needs to be in the annexed state)

According to Frajzyngier & Shay (2003: 64), agreement “must occur in any clause with a singular nominal or pronominal subject in the clause and it cannot occur if there is no nominal singular subject in the clause”. This definition of agreement clearly excludes what happens in Berber in terms of person-gender-number marking on the verb, since this marking is obligatory, regardless of the presence of a nominal argument in the clause. This brings Frajzyngier & Shay (2003: 64) to the following conclusion: “in many languages, so-called agreement phenomena are actually independent coding means in that they occur regardless of whether the argument that they code appears in the clause”.

I will not pursue in detail the role of person-number-gender marking on the verb, but observe that on the one hand this mark is necessary for the verb to become a clause, a predication, and on the other hand it plays a role in referent-tracking in discourse, and, as shown in Section 1.1, if not accompanied by a coreferential noun in the same clause (in the prosodic group of the verb or in the immediately preceding prosodic unit), has the function of marking continuing topic as far as information structure is concerned.
3.1.3 Word order in itself does not mark grammatical relations

As nouns that are computable as subjects and objects can either precede or follow the verb and as, if they follow the verb, there is no fixed ordering between them, word order in itself is not sufficient to mark grammatical relations.

In (24) the noun in the position before the verb can be interpreted as subject or object:

(24) $\text{aygadaj-nni } j\omega\text{-c}\check{e}a=t$
rat.ABS.M.SG-CONS sbj3M.SG-eat.PFV=ABSV3M.SG
“He ate the rat.” or “The rat ate it.”

In (25)–(26) the nouns following the verb can be interpreted as subject or object, the position just after the verb does not code subject or object exclusively. We have to additionally take into account the state marked on the noun.

(25) $j\omega\text{-swa } w\text{omfif } aijki$
sbj3M.SG-drink.PFV cat.ANN.SG.M milk.ABS.SG.M
“The cat drank milk.”

(26) $j\omega\text{-swa } aijki w\text{omfif}$
sbj3M.SG-drink.PFV milk.ABS.SG.M cat.ANN.SG.M
“The cat drank milk.”

As none of those coding means alone transparently code grammatical relations, we have to hypothesize either that grammatical relations are not relevant for nouns in Kabyle, but only for pronouns, or that they are transparently retrievable, but through the interaction of several coding means.

3.2 The interaction of state, position, prosodic grouping, and gender-number marking

Investigation of the various coding shows that the first distinction is between nouns that are outside of the prosodic group of the verb and nouns that are inside.

3.2.1 Nouns outside the prosodic group of the verb

In the position before the opening prosodic boundary of the prosodic group containing the verb, nouns are in the absolute state and can have any grammatical role.

In the following example, the NP $\text{tamittut n lqbial}$ can be interpreted as subject of the following verb.
(27) tamṭṭut n lqbjl / woman.abs.f.sg gen kabyle_tribe.ann.pl / “The Kabyle woman,
ad=dd t-kkr/ ad t-ruh
POT=prox sbj3sg.f-stand_up:aor POT sbj3sg.f-go:aor
she would stand up, she would go
ad=dd t-zdom /
POT=prox sbj3sg.f-gather_wood:aor /
gather wood (...)”

However other roles also appear in this position: in the following example, the noun ajtma is coreferent with the dative clitic = asn.

(28) ajtma / t-uy-d=asn=idd /
brother:abs.m.pl / sbj2-buy:pfv-sbj2sg=dat3pl.m=prox /
“My brothers, you bought them things.”

The position before the prosodic boundary opening on the prosodic group of the verb is therefore not a coding means for grammatical relations. The referents themselves can be retrieved through coreference in gender and number with one of the pronominal affixes or clitics surrounding the verb. However, ambiguity is always possible if more than one pronoun has the same features of gender and number as the initial noun:

(29) taqšift / t-uy=as=idd taksiwt
girl:abs.sg.f / sbj3sg.f-take:pfv=dat3sg=prox dress:abs.sg.f
“The girl bought her a dress.” or “She bought a dress for the girl.”

Gender-number identity of features between pronoun and noun is therefore not a coding means for grammatical relations either.

In the position before the prosodic boundary, grammatical relations cannot be transparently computed.

In the position after the prosodic boundary closing the prosodic group of the verb, nouns are in the annexed state, and can have any grammatical role.

In the following example, the noun wmyar is coreferent with the possessor pronoun -is on the noun; its grammatical role, if one is to be attributed to it, is possessor.

(30) tə-mmmt tmtṭṭut-is / wmyar-nni // sbj3sg.f-die:pfv woman.ann.f.sg-kin3sg / old_man.ann.m.sg-cns // “His wife died, that man”
In the following example, the noun in the annexed state after the prosodic group of the verb is coreferent with the subject pronoun; its grammatical role, if one is to be attributed to it, is subject.

(31) \texttt{ad=dd hku-ɣ / amk i / īṣif-nt}
    \texttt{pot=prox tell:aor-sbj1sg / how rel.real / live:ipfv-sbj3pl.f}
    \texttt{zik / lxalat n lqbajl-nnəɣ /}
    \texttt{long_ago / woman.ann.pl.f gen kabyle_tribe.ann.pl-poss1pl/}
    “I will tell how they lived in the old days, the Kabyle women.”

Those are not the only grammatical roles that can be found in this position. Others are object, kinship relationship (but not indirect object, which is always preceded by a preposition).

The position after the prosodic boundary closing on the prosodic group of the verb is therefore not a coding means for grammatical relations. The referents corresponding to the pronouns can be retrieved through coreference in gender and number with one of the pronominal affixes or clitics surrounding the verb, but ambiguities in gender and number can always arise.

In the position after the prosodic boundary, grammatical relations cannot therefore be transparently computed either.

3.2.2 \(N_{abs}\) after the verb within the prosodic group of the verb

This situation is much more constrained than the one described in §2.2.1: the absolute state implies that the noun is not to be interpreted as the variable of a function grammaticalized on the preceding constituent (Mettouchi & Frajzyngier 2013), namely the subject role marked by the bound pronoun on the verb. By default, it is interpreted as the direct object.

(32) \texttt{i-sa taqǥ̀dsit /}
    \texttt{sbj3sgm-possess:pfv herd.abs.f.sg}
    “He had a herd.”

Only two noun phrases can appear after the verb within its prosodic group: one in the absolute state, one in the annexed state. The latter is necessarily the nominal subject (no other grammatical interpretation is possible for such nouns in such position), and the former the nominal object. The relative ordering of the two noun phrases is irrelevant, as state is here sufficient to disambiguate the grammatical role of each NP (cf. §3.1.3). The noun in the absolute can therefore immediately follow the verb, as above, or be separated from it by another noun, as in (33).
If only position was involved, the computation of grammatical relations would not be possible. But since one of the nouns must be in the absolute state and the other in the annexed state, the two roles cannot be confused. The noun in the absolute appears only when the verb is transitive, and it refers to its second argument. It is therefore the object. The nominal object can therefore be defined as a noun in the absolute state following the verb inside the prosodic group of the verb.

No coreference is involved because in Kabyle the nominal object is the only direct complement in postverbal position, which does not corefer to a pronoun, within the prosodic group of the verb.

3.2.3  $N_{ann}$ after the verb within the prosodic group of the verb

Only one noun phrase in the annexed state can occur within the prosodic group of the verb. This noun is always coreferent with the subject affix. If the construction is intransitive, only the noun in the annexed state occurs. If it is transitive, a noun in the absolute may appear, which is computed as an object (cf 3.2.2.

The combination of annexed state and position (following (immediately or not) the verb within the prosodic group of the verb) provides unambiguous instructions for the decoding of the grammatical relation ‘subject’.

3.2.4  $N_{abs}$ before the verb within the prosodic group of the verb

The noun in the absolute is unambiguously the subject if and only if there is no clitic pronoun in the prosodic group of the verb (only the subject affix). Indeed, in the position before the verb, the noun has to be coreferent with a pronoun affixed or cliticized to the verb.
If there are one or two clitics the grammatical relation is no longer transparent and has to be disambiguated thanks to gender-number feature identity. In the following example, based only on coreference in gender and number, the noun can be computed as the affected object, but if the subject affix was feminine, ambiguity would arise and the noun jiwət could be interpreted as the subject as well:

(37) jiwət jo-fka=jas nnafs //
    one:F sbj3sg.m-give:pfv=dat3sg half //
    “He gave half (an apple) to one (of his wives).”

In the position before the verb, the noun can be transparently coded as subject if and only if there are no pronouns cliticized to the verb, only the subject affix.

3.4 Implications

This raises the question of what can be called a function: do we call something a function only if there is a one-to-one relationship between a coding means and a value? In that case only pronominal paradigms mark grammatical relations, and nouns can only be indirectly computed as subjects or objects.

If we accept that a function can be marked by the interaction of several coding means (i.e., a construction), then we can say that grammatical relations are coded on nouns in Kabyle but that this coding is complex and only holds within some structures.

In Kabyle, nominal subjects and objects can only be unambiguously computed within the prosodic group of the verb:

a. A noun is a nominal subject if and only if, within the prosodic group of the verb:
   – the verb has no clitics other than the subject affix AND the noun occurs before the verb, and is in the absolute state; and
   – the noun occurs after the verb (immediately or not) and is in the annexed state.

b. A noun is a nominal object if and only if, within the prosodic group of the verb, the noun occurs after the verb (immediately or not) and is in the absolute state.
As far as information structure is concerned, functions are clearly complex and involve the presence or absence of the noun, its state, and its position relative to the verb and prosodic boundaries.

4. **General conclusion**

Why is it that in a language such as Kabyle, nominal subject and nominal object are only coded in specific environments? I suggest that this has to do with the fact that the language has pronominal argument marking instead of agreement. Mithun (1992: 58) already links that kind of language with pragmatically-based word order, and the appositive role nouns play in those languages, as opposed to pronouns, which bear the primary case relations to the verb. I have shown that in Kabyle nouns are not simply appositive (for instance, nominal objects can appear without an object clitic) but that their presence and ordering code a number of pragmatic functions and, in some cases, syntactic ones.

The question now is why subject and object grammatical relations are relevant only in some contexts when nouns are involved, whereas they are always unambiguously coded by bound pronouns (subject affixes and absolutive clitics). A look at those contexts may provide an answer: structures involving two nominal arguments (or one when the verb is intransitive) within the prosodic group of the verb are either topicalizations of events or states with a thetic perspective or recapitulations of salient elements of a previously narrated episode. In both cases, the construal of the situation is not of the ‘comment on a topic’, or ‘topic continuation’, or ‘reactivation of a referent’ type, but rather implies the role of each participant in the event or state. This type of context therefore calls for disambiguation of grammatical roles, which is what we see in Kabyle. Other structures conveying information of the ‘aboutness’ or ‘topic promotion’ type are less tied to their argument structure; what is important there is referential information and topical information. In those structures, grammatical relations on nouns are not unambiguously retrievable.

It would be interesting to see whether pronominal argument languages with no basic word order and a pragmatically-based ordering in the sense of Mithun (1992) tend to restrict to some constructions only the unambiguous coding of grammatical relations on nouns, leaving other constructions grammatically unspecified regarding the interpretation of nouns as subjects or objects.
References


