Johanna Kuningas & Jaakko Leino

Word Orders and Construction Grammar

Abstract

Word order phenomena are versatile in the world's languages. Fixed word-order languages as well as free word-order languages have constructions which allow deviations from the so-called "basic word order". This paper sketches out a holistic approach to word order, based on *Construction Grammar*. This framework allows us to combine morphosyntactic, semantic, and pragmatic features in a unified description, and to bring forth the fact that word order as a linguistic phenomenon may be associated with a wide range of motivating factors.

1. Word order in Construction Grammar

Word order phenomena are numerous, and concern—in one way or another—all languages. Even languages with fixed word order have constructions which permit word orders other than the so-called "basic word order" (in a Greenbergian sense; e.g. 1966). The most commonly evoked and the cross-linguistically most widespread variations are due to information structure. For example dislocations and cleft sentences are instances of this kind of word order variation. In addition, there are naturally more language or language family specific word order phenomena which may be related to semantics or morphosyntax.

We believe that all the orders or, as we prefer, construction types are equally important and need to be integrated in a grammatical description. However, a description without any indication of how or under which constraints the order in question occurs, is without interest. *Construction grammar* (henceforth CxG) allows the constraints or motivations to enter into the description. If necessary, even several different kinds of constraints can be integrated in the same description.

In our study, we take a look at some word order phenomena in Kabyle,[1] English, and Swedish, and we propose a description of these phenomena in the CxG framework.

2. Why Construction Grammar?

Why do we need CxG for describing word order phenomena? After all, there is a wide range of literature on word order typology, including very useful accounts on word order phenomena in Kabyle, English, Finnish, and Swedish. Are we merely re-inventing the wheel and claiming that CxG is entitled to the honor for this wonderful novel invention?

What makes CxG useful for our research topic is the fact that it enables us to point out explicitly certain (both universal and language-specific) semantic and pragmatic phenomena and connect these, within the grammar of the language in question, to the formal structure. The possibility to bring together morphosyntactic structure, word order, semantics, and information structure in a formal account of conventional units of the grammar is essential to a proper understanding and description of word order phenomena. Our aim is certainly not to re-invent, neither to discard, the findings of Greenbergian word order typology, nor any other previous account. Rather, our aim is to integrate existing knowledge about word order into a more comprehensive and coherent account of grammar.

We wish not only to point out the basic word order and other notable word order patterns of a given language, but also to make some suggestions about what *motivates* different word order patterns. Central to this aim is the notion of *construction.* A construction is, briefly, a conventionalised combination of form and meaning; it is any linguistic unit, no matter how big, as long as it is conventionalised in the language. Every word is a construction, every grammatical "rule" or template is a construction, and so forth. Or, as Fried and Östman (2004: 18) put it, "A construction is an *abstract,* representational entity, a conventional pattern of linguistic structure that provides a general blueprint for *licensing* well-formed linguistic expressions." A crucial feature of constructions is that they are holistic: they express several linguistic features simultaneously. In other

---

[1] Kabyle (*taqbaylit*) is a Berber language spoken in Northern Algeria. Kabyle belongs to the Afro-Asiatic language group, and it is classified typologically as having a VSO basic constituent order.

words, a construction is a (conventionalized) overall pattern of various facets: formal, semantic, functional, and—notably—information-structural.

3. Word order and information structure

In Construction grammar, word order phenomena are described as a part of a certain type of constructions; these constructions could be labelled *ordering constructions*. In ordering constructions, the specific order of the parts of that construction is combined with a certain sense; some other features of the construction often "complete" this sense.

An example par excellence of these ordering constructions are constructions motivated by information structure. One such construction is the *Kabyle Left-Topic Construction*. In the Kabyle Left-Topic Construction, the noun initial order, with a dislocated sentence structure, is combined with the Topic Promotion function.[2]

Let us start with some examples (data collected by J. Kuningas in 2000).[3]

(1) *bien sûr tilawin n   tmurt   euh... msakint... xeddment di berra,*
     of course women of country euh   poor   work.they in outside

     xeddment deg uxxam. degmi ara d- kkrent   d   lehbir.
     work.they in   house   When REL PCL- get.up.they AUX chore
     'Of course, the women of the country… poor women… work outside as well as inside. Once they get up, it's the chores (that begin).'

(2) *neṭṭat, assehfeḍ-is,   s   weʕakk°az.*
     She   education-her with stick

     a i- d- taf   sexsreɣ   aɣrum-nni a(d) i-tewwet.
     PV me-PCL she.finds have.spoiled.I bread- the PV   me-she.hits
     'Of course she educated with a stick. When she discovered that I had spoiled the bread, she hit me.'

In both (1) and (2) we have an instantiation of the Kabyle Left-Topic Construction. The second line, in both cases, is added in order to give a

---

[2] The Left-Topic Construction is understood as a syntactic structure, in which a referent is promoted from a non active state to an active state. For the Kabyle Left-Topic Construction, see Kuningas 2004.

[3] AUX: auxliliary of predication, PCL: particle, here: particle indicating a movement towards the speaker.

minimum of context. In (1), the dislocated part consists of *tilawin n tmurt* 'women of the country', in (2), the situation is somewhat more complicated, as we have two dislocated elements, *neṭṭat* 'she' and *asseḫfeḍ-is* 'her education'. The dislocated part is followed by a predication, which in (1) is verbal and in (2) consists of a prepositional phrase. Sentences without a verb are very common in Kabyle. The non-verbal predicate very often consists of a preposition followed by a noun, as in our example, but many other types of non-verbal predicates also exist (cf. Chaker 1991, ch. 8). In the present paper, however, we shall concentrate on the description of the dislocated part (see below).

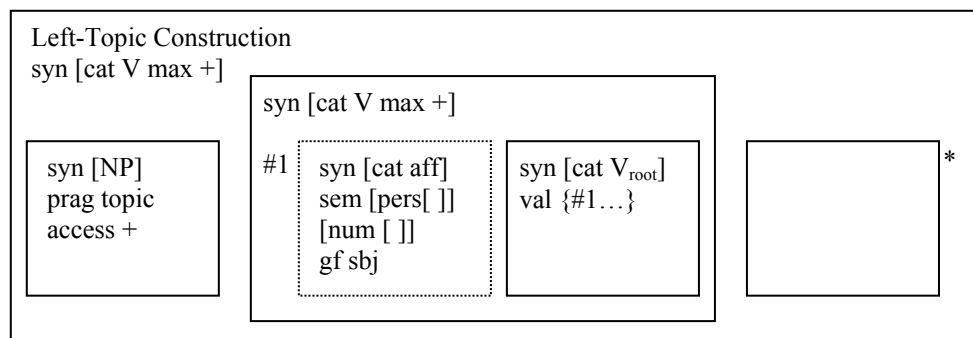A box diagram for the Kabyle Left-Topic Construction with one topic element would be as follows.

Left-Topic Construction
syn [cat V max +]

syn [cat V max +]

syn [NP]
prag topic
access +

#1   syn [cat aff]
sem [pers[ ]]
[num [ ]]
gf sbj

syn [cat V$_{root}$]
val {#1…}

*

Figure 1.

In the description proposed in figure 1, word order is an essential feature of the construction and its component parts. It is described with the discourse function which it conveys; this is done with the aid of information structure categories such as *topic*, *accessibility*, and *identifiability* (these categories are to be understood here as presented in Lambrecht 1994).

On the left, there is the topic expression, the referent of which must be accessible (access +). The topic expression is separated from the predication by an intonation break. This information is not made explicit in the formalism, because it is considered intrinsic to sentence initial topic expressions in Kabyle.

syn [*cat V max* +] means we are dealing with a predicate phrase. In Figure 1, the *syn* [*cat V max* +] is instantiated by a verbal predicate consisting of a verb root and a personal affix, but it could, as well, be replaced by a non-verbal predicate. (In the latter case, there would be, instead of the personal affix and the verb root, information concerning the

non-verbal predicate construction.) The [*cat V max* +] in the case of Kabyle stands for (any kind of) predicate phrase.

The unification variable *#1* in the predicate phrase box, to the left of the affix, means that the root cannot occur without the personal affix subject. The dotted line means that the personal affix may be on either side, or on both sides, of the root depending on the grammatical person. *val* stands for valence. The Kleene star (*) after a box indicates that there may be zero or more of those constituents.

Notice that (1) and (2) are attested examples. The transcription is not phonetic, merely phonological; nevertheless, pauses and hesitations do appear. The box diagram (in Figure 1) does not display these prosodic features, not even the apposition *msakint* "poor", because they are not considered as a part of the Left-Topic Construction. If we wanted to describe the utterance, all these factors should and could be included by using other constructions which contribute to the sentence.

Let us now examine example (2) in more detail. Can CxG account for this "more complicated" case as well?

As mentioned above, here too, we are dealing with an instantiation of the Left-Topic Construction. The essential difference between the examples (1) and (2) is that in (2), there seem to be two topic elements: the first one of them has as its referent an "already known" discourse feature, whereas the referent of the second one has not yet been introduced. (The issue whether or not the two dislocated elements should be analysed as topics will be discussed in the forthcoming PhD thesis of J. K. See also Kuningas 2004. For the purposes of this paper, we do analyse them as topics.)

What would be the motivation for the two dislocated topic elements? The topic ('education') that the speaker wants to introduce is not accessible to the hearer and cannot, as such, be coded as a Left-Topic expression. In order for it to appear as a Left-Topic expression, it must first be promoted to the accessible state. Cognitive accessibility presumes that the referent is identifiable. Identifiability is a cognitive notion having to do with the anchoring of a referent in the preceding discourse. Anchoring can be done by different means; in (2), the second topic expression is anchored to the first topic expression by a coreferential possessive suffix (-*is* 'her/his'); by this means, its referent becomes identifiable, and thereby accessible, and can be promoted to the active state.

The first topic is introduced again (we know from its pronominal coding that its referent is considered as active by the speaker) in order to reduce the hearer's processing effort. In fact, in the preceding discourse,

the topic had changed; *asseḥfeḏ-is* 'her education' alone would not have been sufficient in order for the hearer to know whose education the speaker is referring to (especially as in Kabyle the 3rd person possessive suffix does not indicate the gender). At least, it would have taken her a lot more time.

In brief, we are dealing here with a particular case of the Left-Topic Construction, which could be called *Double Topic Construction*. This construction occurs in discourse contexts where 1) the topic to be introduced is not cognitively accessible and 2) the possessive (or demonstrative) anchoring alone would demand too much processing effort. Its motivation is then slightly different from that of the (simple) Left-Topic Construction. In other words, it is a construction of its own right, but belongs to the same construction family.

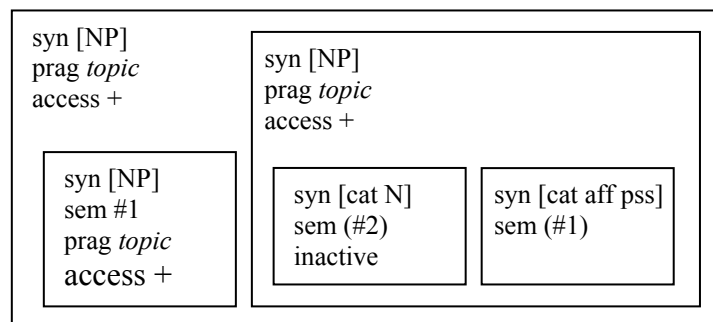Figure (2) shows the analysis for the topic part:

```
┌──────────────────────────────────────────────────────────────┐
│ syn [NP]                                                       │
│ prag topic      ┌────────────────────────────────────────────┐│
│ access +        │ syn [NP]                                    ││
│                 │ prag topic                                  ││
│                 │ access +                                    ││
│  ┌────────────┐ │  ┌──────────────┐  ┌──────────────────────┐ ││
│  │ syn [NP]   │ │  │ syn [cat N]  │  │ syn [cat aff pss]    │ ││
│  │ sem #1     │ │  │ sem (#2)     │  │ sem (#1)             │ ││
│  │ prag topic │ │  │ inactive     │  │                      │ ││
│  │ access +   │ │  └──────────────┘  └──────────────────────┘ ││
│  └────────────┘ └────────────────────────────────────────────┘│
└──────────────────────────────────────────────────────────────┘
```

Figure 2.

The first element and the possessive suffix are co-referential. This is indicated by identical semantic values (unification variable #1). The possessive suffix functions as an anchoring element: it makes the referent of the second dislocated element identifiable and, thereby, accessible to the hearer. This change of activation state is indicated by the attribute *inactive* in the first component (inner box) of the second topic constituent, as well as by the attribute *access +* in the same constituent's upper box.

In order to describe a Left-Topic Construction with two topic elements, we must place this description in the diagram (Figure 1); this is possible because the values [syn [cat NP]], [prag *topic*] and [access +] are the same. As far as the predication is concerned, the description would be slightly different, if the predicate to be described were nominal rather than verbal. Note however that [syn [cat V max +]] covers both cases.

4. Other word order phenomena

Although information structure is central to a number of word order phenomena, it is by no means sufficient for explaining all existing word order variations. CxG may, however, be used for many other kinds of explanations as well. To take a simple example, in the English transitive sentence, word order is used to indicate grammatical functions. The NP before the verb is the subject, and the NP after the verb is the object:

(3)   *Bill saw John ≠ John saw Bill* (Karlsson 1977: 385)

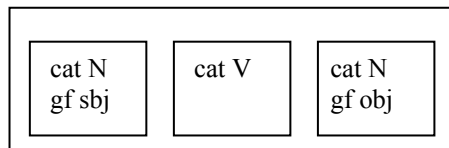| cat N gf sbj	cat V	cat N gf obj

Figure 3.

Different kinds of motivations may also be combined in a single ordering construction. A good example of this is the Swedish V2 constraint, which states that (if the clause does not begin with the subject) the verb must be the second constituent of a sentence:[4]

(4)   *Nu är bilen här* (Karlsson 1994: 156)
        now   is   the car   here
        'Now the car is here.'

This constraint may be captured with the following simple construction:

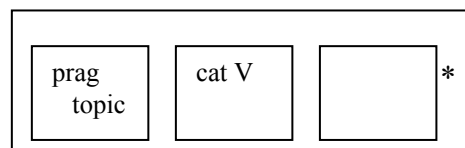| prag topic	cat V		*

Figure 4.

The function of this construction is twofold. On the one hand, it states (much like the Kabyle constructions above) that the first element is the topic. On the other hand, it makes a rather mechanistic structural point:

---

[4] Cf. *Bilen är här nu, Här är bilen* etc. (the verb is the second constituent), but **Nu bilen är här* (ungrammatical since the verb is the third constituent).

there is exactly one element in the sentence before the verb–namely, the topic–and therefore, the verb is the second constituent.

In the case of information structure, we may say that the motivation of the word order phenomena is, broadly speaking, pragmatic. In the case of grammatical functions, the motivation is obviously syntactic. In addition, we find word order phenomena which have a semantic motivation; one such phenomenon is the effect of word order on the scope interpretation of locative adverbials (cf. Huumo 1995a, b). And, as the Swedish V2 constraint exemplifies, combinations of different motivations also occur. Therefore, the system that we use for describing word order phenomena must be versatile enough to be able to capture all these different kinds of underlying reasons for seemingly similar word order variations.

As we pointed out in section 2, a crucial feature of constructions is that they are *holistic:* they express several linguistic features simultaneously. It is precisely this feature which makes CxG so well-suited for a comprehensive analysis of different word order patterns.

5. Conclusion

In CxG, word order is an inherent property of constructions. It is always motivated. Be this motivation semantic, information structural or merely syntactic, it can be captured in the description by establishing different kinds of ordering constructions.

Let us add that a description that includes the motivation is interesting, not only for its own worth, but also because it allows some interesting cross-linguistic observations. As an example, we could give the striking resemblance between the Kabyle Left-Topic Construction and the Spoken French equivalent (cf. Leino & Kuningas 2005). Given this, and the fact that French is an SVO language while Kabyle is a VSO language, we can draw the conclusion that surprisingly similar word orders occur in languages which are classified as entirely different according to the traditional word order typology. Without a framework that allows us to combine morphosyntactic, semantic, and pragmatic features in a unified description, such similarities might go unnoticed.

References

Chaker, Salem (1991) *Manuel de linguistique berbère* I. Alger: Bouchène.

Fried, Mirjam & Jan-Ola Östman (2004) Construction Grammar: A thumbnail sketch. In Fried & Östman (eds.) *Construction Grammar in a Cross-L*anguage Perspective, pp. 11–86. Constructional Approaches to Language 2. Amsterdam & Philadelphia, PA: John Benjamins.

Greenberg, Joseph H. 1966. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph H. Greenberg (ed.) *Universals of language,* 2nd edition, pp. 73–113. Cambridge, MA: The MIT Press.

Huumo, Tuomas (1995a) Word order and the scope of locatives and possessives. *Word* 46.3: 315–332.

—— (1995b) Ongella Elmeri huitaisi Anselmia ongella. Lauseaseman vaikutuksesta eräiden adverbiaalityyppien tulkintaan. *Virittäjä* 99.1: 45–70.

Karlsson, Fred (1977) Syntaktisten kongruenssijärjestelmien luonteesta ja funktioista. *Virittäjä* 81.4: 359–391.

—— (1994) *Yleinen kielitiede*. Helsinki: Yliopistopaino.

Kuningas, Johanna (2004) Structuration informationnelle – étude de deux constructions à thématisation en Kabyle. In Jocelyne Fernandez-Vest & Shirley Carter Thomas (eds.) *Structure informationnelle et particules énonciatives: Essai de typologie*, pp. 167–188. Grammaire & Cognition 1. Paris: L'Harmattan.

Lambrecht, Knud (1994) *Information Structure and Sentence Form: Topic, focus and, the mental representations of discourse referents*. Cambridge Studies in Linguistics 71. Cambridge: Cambridge University Press.

Leino, Jaakko & Johanna Kuningas (2005) La structure informationnelle et l'ordre des mots avec les exemples du français, kabyle et du finnois. In Jocelyne Fernandez-Vest (ed.) *Les langues ouraliennes aujourd'hui : approche linguistique et cognitive / The Uralic Languages Today: A linguistic and cognitive approach*, pp. 587–598. Bibliothèque de l'École des hautes études. Sciences historiques et philologiques 340. Paris: H. Champion.

Contact information:

Johanna Kuningas
Department of General Linguistics
University of Helsinki
jkuningas(at)luukku(dot)com

Jaakko Leino
Department of Finnish language and literature / researchers
P.O. Box 4
FI-00014 University of Helsinki
jaakko(dot)leino(at)helsinki(dot)fi