Proceedings of the 20th International AEDEAN Conference

P. GUARDIA and J. STONE, editors

UNIVERSITAT DE BARCELONA
Facultat de Filologia

INFORMATION PACKAGING IN ENGLISH AND CATALAN

Enric Vallduví
Universitat Pompeu Fabra

Communication in a speaker-hearer interaction is not simply an unstructured sequence of propositions. Given that transfer of knowledge from speaker to hearer is a defining element in communication, it is not surprising that speakers, in Prince's (1986:208) words, 'form their utterances so as to structure the information they are attempting to convey [...] in accordance with their beliefs about the hearer: what s/he is thought to know, what s/he is expected to be thinking about'. In other words, speakers 'package' information in different ways in view of their assumptions about the hearers' knowledge and attentional state. Hence the term **information packaging**.

In every language there is an array of truth-conditionally equivalent sentences which differ only in the way the information they convey is packaged. English, for instance, may express the proposition that the boy does not eat chocolate with the alternatives in (1), among others, and Catalan can express this propositional content through at least the options in (2) (capitals indicate sentential nuclear stress) :

(1)   a. The boy doesn't eat CHOCOLATE.
        b. The boy doesn't EAT chocolate.
        c. The BOY doesn't eat chocolate.
(2)   a. No en menja el XIQUET, de xocolata.
        b. No menja XOCOLATA, el xiquet.
        c. De xocolata no en MENJA, el xiquet.

These alternatives are nevertheless not thoroughly identical, as shown by the fact that they cannot be freely interchanged in discourse. Contrast, for instance, the felicity of (3b) as an answer to (3a) with the infelicity of its truth-conditional equivalent (3c) in that same context (the symbol '#' indicates pragmatic infelicity):

(3)   a. On són, els ganivets?
          'Where are the knives?'
        b. Els vaig deixar al CALAIX.
          'I left them in the DRAWER.'
        c. #Els hi vaig DEIXAR, al calaix.
          'I LEFT them in the drawer.'

Sentences (3b) and (3c) encode the same proposition. They differ not in terms of truth-conditions but rather in terms of information packaging. It is these different ways of packaging information which are responsible for the existence of the alternatives in (1) and (2) and for the contrast in (3).

There is increasing awareness of the large degree of crosslinguistic diversity involved in

the structural realisation of information packaging. Whereas English and other languages mainly exploit intonation for informational purposes, in languages like Catalan syntax plays the primary role. This is illustrated by the contrasts in (4) and (5):

(4)   a. The boss CALLED.
        b. The BOSS called.
(5)   a. L'amo ha TRUCAT.
        b. Ha trucat l'AMO.

Sentences (a) and (b) are truth-conditionally equivalent but differ in their felicity conditions. This difference is structurally reflected through an intonational contrast in English but through a syntactic contrast in Catalan: in English the position of nuclear stress varies, while string order is maintained (SV). In Catalan, the position of nuclear stress is maintained (clause-final) and string order varies. This paper provides a comparative analysis of the structural realisation of information packaging in Catalan and English.

How do speakers go about packaging utterances? What are the informational primitives that underlie the diversity in sentential form information packaging is meant to account for? In trying to answer these questions several proposals have been put forward: the terms focus, ground, topic, comment, given, new, theme, and rheme, among others, all refer to proposed informational primitives.

Information packaging is seen in Vallduví 1992, 1995 and Vallduví & Engdahl 1996 as instructions for information update. The sentences in (4), for instance, have the same propositional content but encode different instruction-types, i.e. different ways of indicating how the meaning of the sentence updates the hearer's information state. Instruction-types correspond to different **focus-ground** partitions. The **focus** is defined as the actual update potential of a sentence S, i.e. the only contribution that (according to the speaker) S makes to the information state of the hearer at the time of utterance. Since all sentences have some update potential, they all have a focal segment. The **ground**, in contrast, is already subsumed by the input information state and acts as an 'anchor' for the focus: it indicates how the information update is to be carried out. Sentences have a ground only if the context warrants its use, i.e. if the anchoring is (thought by the speaker to be) required. The ground is further divided into **link** and **tail**. Link and tail each contribute in their own way to the anchoring role of the ground. Links indicate *where* the focus should go in the input information state: they establish a particular locus-of-update in the input information state. A tail indicates *how* the focus fits there: the presence of the tail indicates that a nondefault mode of update is (in the speaker's eyes) required at that point in discourse.

In order to see exactly how the elements in the ground carry out their task, something must be said about the structure of information states. For current purposes, let us view an information state as a file-like data structure (cf. Heim's 1982 files in File Change Semantics). Files are collections of file cards. Each file card has a number of records or conditions written on it listing attributes and relations about the entity it denotes. The content of these file cards is updated during communication. Information packaging reflects the way the speakers take into account their assumptions about the structure of the hearer's information state in order to optimise information update.

In this light, links are argued to designate a specific file card in the input file where information update is to be carried out (akin to Reinhart's 1982 topic and Kuno's 1972 sorting key). The tail further specifies how the update must be effected. In particular, it indicates that focus is not simply added to the file card designated by the link as a new condition (default mode), but rather that focus must complete or alter a condition which is already there and is designated by the tail (nondefault mode). The ground, both link and tail, performs an 'anchoring' role for the focus: it guarantees that the update potential of the sentence is anchored to the appropriate location (from the speaker's perspective) in the input file. If (the speaker assumes) no anchoring is needed, a sentence will have no ground. If only some anchoring is needed, a sentence may have a link but not a tail, or vice versa.

Sentences (6) and (7) illustrate two different instruction-types, link-focus and link-focus-tail, respectively (the F-labelled brackets identify the focus):

(6) a. Tell me about the people in the White House. Anything I should know?
     b. The president [F hates CHOCOLATE].
(7) a. And what about the president? How does he feel about chocolate?
     b. The president [F HATES] chocolate.

Roughly, link-focus instructions designate a locus-of-update and indicate that the update is carried out with an addition of a condition on that locus. Sentence (6), for instance, is an instruction to use the file card denoting the president as a locus of update and add a condition there with the property 'hates chocolate'. Link-focus-tail instructions designate a locus of update too, but in addition they point to a given condition and indicate that the update completes or alters that condition in some way. (7) instructs the hearer to search for a condition of the form *feels-like-about chocolate* on the file card denoting the president and substitute 'hate' for the underspecified predicate. The corresponding linkless instructions occur when no locus of update need be designated because the current one is inherited from discourse.

In English-type languages the prosodic contour of the sentence is 'malleable' in that it allows nuclear stress to shift back and forth over different positions in the clause. Focushood is associated with nuclear stress and focal phrases normally appear in their canonical position. This can be seen in the contrast between (6) and (7) above and in (8):

(8) a. The pipes are [F RUSTY].
     b. The pipes [F are RUSTY].
     c. [F The PIPES are rusty].
     d. [F The PIPES] are rusty.
     e. The pipes [F ARE] rusty.

The pitch accent associated with nuclear stress in English (in capitals in (8)) is described by Steedman 1991, adopting Pierrehumbert's (1980) phonology of intonation, as a level high tone (H*). The H* is associated with one particular word within the focus and is generally followed by a falling boundary tone. Strings that are prosodically identical, like (8a) and (8b) and (8c) and (8d) are informationally ambiguous (see the notion of range of

permissible focus in Chomsky 1971). Of course, these sentences are ambiguous only when used out of context, since context indicates which focus reading is intended. The subject and the predicate in the sentences in (8) appear in situ no matter whether they are focus or ground. The burden of realisation is placed on the intonational dimension.

Links in English are also associated with a particular pitch accent, different from the pitch accent associated with focal/nuclear stress. Steedman defines it as a high tone preceded by a distinctive low level (L+H*). Link-focus sentences, then, are associated with a particular intonational contour L*H*... H* (L%) dubbed by Cohen & 't Hart 1967 a 'hat pattern'. Links may also appear in a left-hand slot (e.g. topicalisation), but syntactic realisation is no substitute for entonational marking, since fronted links must still be associated with a L+H* accent. Fronting, then, is not only optional but also redundant. The sentences in (9), annotated for intonation, illustrates these facts:

(9)   a. Fred ate [the BEANS].
           L+H* LH%   H* LL%
        b. The beans [FRED] hates.
           L+H* LH% H* LL%

Tails, finally, do not appear to be structurally characterised in any way other than by being typically deaccented. English tails, as any other informational element, are free to remain in situ. The position of the direct object tail *chocolate* in (7) illustrates this point.

It appears to be the case, then, that information-packaging interpretation in English is not associated with any particular syntactic manisfestation but rather with prosodic realisation instead. Two sentences with identical syntax may be associated with distinct information-packaging instructions. This contrasts radically with Catalan-type languages, where each instruction-type is necessarily correlated with a different syntactic structure. As in English, an item within the focus (the rightmost one) is marked with nuclear stress. In Catalan, however, it appears that nuclear stress cannot be shifted to the left. Attempting to do so yields ungrammatical strings:

(10)   a. Portarem el gos a EIVISSA.
           'We'll take the dog to IBIZA.'
        b. *Portarem el GOS a Eivissa.
        c. *PORTAREM el gos a Eivissa.

Rather, the association of nuclear stress and focushood in Catalan is attained through the mediation of syntax. Where English uses a prominence shift, Catalan uses a syntactic operation. Let us compare the English sentences in (6) and (7) to their Catalan analogues:

(11)   a. Tell me about the people in the White House. Anything I should know?
        b. Sí. El president [F odia la XOCOLATA].
           'Yes. The president hates CHOCOLATE.'
(12)   a. And what about the president. Anything I should know?
        b. Malament! El president [F l'ODIA], la xocolata.
           'Bad move! The president HATES chocolate.'

The direct object NPs in (7) and (12) are part of the ground. In (7) this informational fact is realised purely through intonation, but in (12) it is realised through syntax. The object NP in the Catalan sentence does not appear in situ, but rather in a right-detachment slot, external to the core clause.

What is going on in Catalan is that nuclear stress is invariably fixed on clause-final position. Since all major constituents remaining within the entonational phrase headed by nuclear stress are interpretated as focus, the only way to prevent a focus interpretation for the clause-final object is to remove it from clause-final position and the 'scope' of the nuclear stress altogether. Detachment is a means to achieve this. The generalisation that can be drawn for Catalan is that focal phrases remain within the core clause whereas ground phrases must be detached. Example (3c) above is infelicitous because the structural encoding of *el calaix* 'the drawer' is incompatible with the interpretation demanded by context (3a). The context requires *el calaix* to be focal and this is possible only in structure (3b).

With one additional assumption, this generalisation is also valid for subjects. Contra the traditional view, we need to assume that Catalan, like other Romance languages is underlyingly VOS (as argued in e.g. Contreras 1991, Vallduví 1993). The canonical position of the subject is a postverbal one and the preverbal subject slot is always a left-detachment slot. The subject in (11) and (12), then, which is part of the ground, is not in a clause-internal position but rather in a left-detachment slot. Focal subjects, in contrast, are necessarily postverbal. This accounts for the facts in (5) above. Sentence (5b), with a focal subject interpretation (as in an out-of-the-blue context), is a VS structure, whereas (5a), where the subject is part of the ground (as when the boss has been under discussion), is an SV structure. Notice also that the Catalan analogue of English (8c) has a postverbal subject as well, as shown in (13a):

(13)   a. [F Deu estar rovellada la CANONADA].
            'The PIPES must be rusty.'
        b. La canonada [F deu estar ROVELLADA].

Sentences (8c) and (13a) are all-focus and contrast with (8b) and (13b), which are link-focus. As discussed above, this informational contrasts correlates with an intonational contrast in English and with a syntactic contrast in Catalan. English sticks to an SV order and plays around with intonation. Catalan sticks to clause-final nuclear stress and plays around with syntax. Two different structural strategies to convey one single informational interpretation.

REFERENCES

Cohen, A. & J. T. Hart. 1967. "On the anatomy of intonation". *Lingua* 19:177-192.

Contreras, H. 1991. "On the position of subjects." S.D. Rothstein (ed.) *Perspectives on phrase structure: Heads and licensing*, pp. 63-79. New York: Associated Press.

Heim, I. 1983. File Change Semantics and the familiarity theory of definiteness. R. Bauerle et al. (eds.) *Meaning, use and interpretation of language*, 164-189. Berlin: Gruyter.

Kuno, S. 1972. "Functional sentence perspective." *Linguistic Inquiry* 3:269-320.

Pierrehumbert, J. 1980. *The phonology and phonetics of English Intonation.* MIT doctoral dissertation.

Prince, E.F. 1986. "On the syntactic marking of presupposed open propositions." *CLS Parasession* 22:208-222.

Reinhart, T. 1982. "Pragmatics and linguistics: An analysis of sentence topics.*" Philosophica* 27:53-94.

Steedman, M. 1991. Structure and intonation. *Language* 67:260-296.

Vallduví, E. 1992. *The informational component.* New York/London: Garland Publishing.

Vallduví, E. 1993. "Catalan as VOS: Evidence from information packaging." In W. Ashby et al. (eds.) *Linguistic perspectives in the Romance languages*, 335-350. Phila.: Benjamins.

Vallduví, E. 1995. "Aspects of information structure in a dynamic setting.". C. Martín-Vide (ed.) *Actes de l'Onzè Congrés de Llenguatges Naturals i Llenguatges Formals*, 145-162. Barcelona: Promociones y Publicaciones Universitarias.

Vallduví, E. & E. Engdahl. 1996. "The linguistic realization of information packaging." *Linguistics* 34:459-519.

[i] Para un estudio en profundidad sobre tipología textual en traducción, véase "Teorías de traducción actuales y su aplicabilidad a la tipología textualî, de Cristina Valdés, 1995.