Word order in contemporary French: a functional view
Martin C. Harris
University of Salford
Word order in contemporary French: a functional view

In Harris (forthcoming) I argued that earlier analyses of French word order by Lambrecht, Ballard and myself were all inadequate, in that they failed to capture the full variety of alternative word orders current in the contemporary language, competing for favour in (presumably) different pragmatic circumstances. It is not my intention to repeat that paper here, but rather simply to summarize without the supporting evidence its conclusions, as a prelude to a presentation of the position in French, both synchronic and diachronic, from within the functional framework expounded by Simon Dik (1978; 1980). The paper will of necessity be tentative, partly because the situation in contemporary French really does seem to be extremely complex, with many questions as to the frequency and functions of particular structures still to be answered.

Standard (literary) French is often said to be an almost canonical SVO language, having evolved to this stage via a topic-initial (TVX) period (to which we shall return later) from an earlier SOV order in Latin. (Harris 1978: Ch. II describes this process in detail.) However, during at least the present century (and maybe before: the detailed research work remains to be done), two major and related changes have taken place, at least in the spoken language and written representation thereof. Firstly, left- and right-dislocated structures are now widely attested, with up to three constituents occurring before or after the "core" sentence. (See Bauche 1920: 156 for an early discussion; Grevisse 1980: para. 280 n.7, para. 313 is uncharacteristically uninformative on the structures in question.) Secondly, the inherited set of atomic pronouns has been reinterpreted as a set of bound affixes preposed to the verb, in such a way that forms which, in the standard language, may still serve as the sole exponent of the grammatical categories Subject, Object etc. are now frequently to be interpreted in less formal registers as case markers co-referential with the "real" exponents of these categories elsewhere in the sentence. Thus a paradigm example of "right-dislocation" such as:

(1) Je le lui ai donné moi le livre à Pierre
would be analysed (leaving aside for a moment the important question of what is and what is not within the Predication) as:

(2) je - le - lui - ai donné moi le livre à Pierre
clitic-clitic-clitic- verb subject object Prep Phr (Ind Obj)

In Harris (forthcoming) I conclude, after synthesizing much available data derived from corpora of spoken utterances and from the analyses of native speakers, that there are two types of "left-dislocation" (one type being only infrequently attested) and one type of "right-dislocation". The major type of "left-dislocation" is a classic instance of the topicalization process, the nominal or nominals that are topicalized being placed in Theme position, analysed by all observers as being outside the Predication. This initial position behaves exactly like the topic slot in so many languages: the nominal or nominals are syntactically (though not of course pragmatically) independent of the framework imposed by the verb within the Predication, and carry no overt case marking (i.e. no preposition, in the case of French) (cf. Dik 1978: 135). If the verb selected in due course by the speaker is one which permits the attribution of a syntactic function to the left-moved nominal, then an appropriate clitic pronoun is generally found:

(3) La plage i-faut y-aller quand i-fait chaud
     the beach it-is necessary to-it to go when it-is hot
     Theme Predication

(Note the absence of the 'dative' case-marker à before la plage, but the presence of a 'dative' pre-verbal clitic y within the Predication.)

The major type of "right-dislocation" differs formally from the structure just discussed in that the nominal(s) dislocated - occurring in Tail position - must be coreferential with a clitic position which has already occurred within the Predication, that is, the Predication must be complete in itself and only nominals whose function in a given sentence is one for which a clitic exists in French can occur as Tails. Equally, overt case marking (i.e. the use of a preposition, if appropriate) is obligatory (cf. Dik 1978: 155). Contrast with (3) above an example such as (4):

...
(4) I-faut y-aller quand i-faut chaud, à la plage

where the 'dative' clitic y is again attested, but where la plage is now marked with the 'dative' preposition à.

There is also a minor type of "left-dislocation" in which overt case marking does occur, presupposing of course that the whole utterance has been pre-planned; such structures are by that very token likely only in more formal registers. Larsson (1979), for instance, cites (her (21)):

(5) De la résistance, il n'y avait rien à en dire

Somewhat confusingly, Larsson labels this structure "Topicalization" and the vastly more common structure discussed earlier "left-dislocation". That apart, hers is a perceptive analysis set within an illuminating comparative framework. Her "left-dislocation" structure is clearly the primary means of marking Topics in contemporary French, with the minor pattern just discussed being marginal in comparison (though none the less interesting for that). (For a brief discussion of sentences like (5) above, see Bally 1965: 67. Bally concludes: 'On voit combien ces tours nous rapprochent des inversions simples de la phrase liée'.)

The movement of nominal elements to the left and/or right of the core sentence is of course anything but unusual in cross-linguistic terms. What is perhaps not so common is that both structures are very frequently attested in the spoken language, to the point where the 'canonical' order is clearly a minority pattern: indeed, recent work by Lambrecht, as yet unpublished, suggests that archetypal SVO sentences are in fact extremely rare in spoken texts. At least three questions therefore arise: what are the currently preferred functional patterns in French, what (so far as we can tell) is their history, and what their likely future? In attempting to answer these questions, we shall (as we indicated earlier) try to situate them within the context of Dik's work, particularly Dik (1980).

The treatment of left-dislocated structures seems to be very much more straightforward than that of right-dislocated structures, and we shall accordingly start with that. It was noted briefly at the start of this paper that, earlier in the history of French, there was a lengthy
period when the language was a strong V2 (V2s) language (Dik 1980: 152). Specifically, the preferred pattern was for precisely one element (Subject, complement, adverb) to precede the verb ('to occur in P1 position'), the choice of such element being pragmatically rather than syntactically governed (i.e. primarily for reasons of topicalization or anaphora). The operation of (R3) can be clearly seen at this stage, in that, if the pre-verbal position would otherwise have remained unfilled, an unstressed Subject pronoun was almost invariably used to prevent this: by the literary period, instances of actual verb-initial structures are extremely rare in prose texts, V2 generalization having taken place. (For a concise summary of the position of Old French, cf. Price 1971: 145-9; for a detailed survey, see Foulet 1968: paras 446-60.) There was thus a situation, in 'standard' Old French, in which the verb was second to either an adverb or to a Topic (instances of the latter of course very often being also the Subject), or simply to a (non-topicalized) Subject pronoun inserted by the operation of (R3). In other words, the stage was set for the passage, so often described elsewhere, of TVX to SVX as the unmarked order in French, a development largely concluded by the time the classical language was codified: hence the designation of French as an SVO language.

But the very change just described, reserving initial position within the Predication for the (grammatically defined) Subject, precluded using the position for any sort of distinctive difference (Dik 1980: 167) and thus opened the door for the re-emergence of a pragmatically-controlled Topic slot, in left-dislocated absolute initial position, that is as Theme. Although this structure has, as was indicated above, only been fully documented within the spoken language this century, it seems reasonable to suppose that its emergence was in fact earlier, hand in hand with the progressive rigidification of the SV ... order. Specifically, as non-Subject nominal elements came to be excluded from pre-verbal position within the Predication, we may presume that they, when topicalized, were fronted in an absolutely familiar way, Theme thus becoming the appropriate slot for topicalized terms. This process would then have extended from non-subject nominals pragmatically linked with the Predication (whether or not these were in addition a syntactic complement of some kind) to include Subjects also, when explicitly topicalized. In the case of Subjects always, and in the case of other complements generally, a co-referential clitic
is found prefixed to the finite verb, as we have seen. In this way, sentences such as (6), (7) and (8) became widespread:

(6) Ton Père, Marie l'attend
(7) Marie, elle attend ton père
(8) Ton père, Marie, elle l'attend

There are as yet, despite Dik's expectations (1980: 158-9), few signs of the absorption of these topicalized elements within the Predication, that is, within the P1 position which would then rival or replace Theme for this purpose.

The question of right-dislocated structures is more complex. In a recent analysis, William Ashby showed that right-dislocation of the Subject occurs in 17% of the tokens of his large sample of spoken French (Ashby 1982: 34); other more impressionistic reports suggest a higher figure. Be that as it may, sentences with one or more elements in Tail position are not at all infrequent in contemporary spoken French. There seems little controversy over the origin of this structure. A familiar use of the Tail slot is as an 'afterthought' in a very broad sense, that is, to clarify or elaborate upon an element less fully expressed earlier in the utterance, particularly when that initial exponent was a personal pronoun. Clearly, the Predication must be syntactically complete, so that Tails can occur only co-referentially with elements which have an exponent — generally a clitic — within the core sentence. What is less clear is the pragmatic circumstances under which this structure is appropriate, whether it is being grammaticalized and if so what precisely this would mean in this case.

My own initial analysis, which I first proposed in Harris (1976) and have repeated several times subsequently, was that this process could be adequately described in syntactic terms. Specifically, I suggested that sentences such as (1) above indicate the emergence of a VSO order in contemporary spoken French. The passage of a Subject-initial language was discussed as early as 1974 by Theo Vennemann (Vennemann 1974) within the context of typological drift; exactly similar assumptions underlie much of the discussion in chapters 7 and 8 of Dik (1980). The principal arguments for this 'syntactic' hypothesis are that only elements which
have a syntactically definable (and in fact 'cliticizable') role within the core sentence can occur in right-dislocated position (the 'original' Subjects etc. now being downgraded in status to bound affixes) and that many instances of right-dislocation occur in contexts with no apparent special pragmatic value and without 'comma intonation', thus:

(9)  (je) (ne) sais pas moi
(10) On y va nous (à Paris)

These arguments still seem to me to bear considerable weight. However, Lambrecht has argued strongly (1981 and personal communication) that there are (at least) two important ways in which right-dislocated constituents generally differ from 'normal' Subjects. Firstly, they share with Topics the characteristic that they may not be indefinite and they must be pragmatically recoverable; and secondly, even when they are otherwise apparently intonationally integrated within the core sentence, they cannot receive even slight stress. Both these points are well-taken and cannot lightly be set aside.

What, then, seems to be the most adequate analysis of the position of right-dislocated nominals in contemporary spoken French? Firstly, an original "afterthought" structure occurring in Tail position has come to be used far more frequently than its former limited pragmatic function would warrant; indeed, it now seems to be appropriate in precisely that range of pragmatic contexts where left-topicalization is appropriate, but with the absolute syntactic constraints not found in respect of Topics the Theme slot. (The question immediately arises as to how frequent cross-linguistically Topic-final order is: the obvious iconic explanation of Topic-initial order is clearly not available in this case.) As a second stage in the process, no doubt facilitated by the changed status of clitic pronouns, it has become not infrequent for all the lexical content of the sentence to be right-dislocated, subject always to the pragmatic constraint of recoverability already discussed. (This leads to sentences like (i) and (ii) in note 8, in a sense the obverse of (8) above.) Put in other terms, we may say that a pragmatically limited but nevertheless large sub-group of Subjects, Objects and certain complements may now occur in the Tail position.
As a third, as yet very hesitant, stage in the process, we may find the weakening of the pragmatic and/or intonational constraints mentioned by Lambrecht, a development by Givón (1979: 208ff) as the passage from "pragmatic" discourse structures to "grammaticalized" syntactic structures and pointing towards an eventual VSO or VOS order. Thus, to take an example on which all seem to be agreed, the shorter version of (10) above can be uttered when there has been no prior reference to 'us', and nous can bear contrastive stress within a unified intonation pattern, for instance thus:

(11) On y va nous: et vous?

Of course, personal pronouns have a particular deictic value, which may weaken the first point - the same may be said of proper names - but nevertheless it is hard to resist the conclusion that there is at least incipient Tail-grammaticalization at work here, along the lines of similar instances presented in some detail by Dik (1980: 6.4.4), with the clitic pronouns clearly being reanalysed (as we have already suggested) as Subject markers, Object markers etc.¹⁰ as in many of the languages he discusses there. (For an interesting discussion of 'frozen' Object pronouns, cf. Givón 1979: 239ff.) The Tail position is not after all the obvious one for a Topic slot, and the fact that French is a phrase-stress-final language can only facilitate the absorption within the Predication of originally right-dislocated elements. We therefore reject as excessively categorical Dik's claim (1980: 144) that it is 'obviously incorrect' to call:

(12) (=his 40) Il a acheté le livre, ton frère

a VOS structure, not least because, as he himself observes (1980: 148), à propos of an example similar to our (11) above, the use of personal pronouns to 'expand the information given in the predication' is 'obviously vacuous'.

As far as the ordering of S and O is concerned, we recall that, within the Tail, this was pragmatically governed; there is as yet in general no syntactic rule governing the choice of VSO or VOS in any
particular instance. As we have already observed (note 8), Dik's own theory might seem to predict VOS in the case of French (though his illustration is much less complex than the situation under discussion here), whereas if and when grammaticalization is complete, S preceding O is overwhelmingly more probable on general theoretical grounds. And all of this assumes that we are indeed dealing with S's and O's, something which is, as Lambrecht has shown, not at all beyond dispute. And I leave unanswered one of the most interesting questions of all: why did French so significantly increase, at first in Tail position, its use of right-dislocated structures for Topic-like purposes, when it was already making massive use of the (much more expected) left-dislocated Theme position for a like purpose, particularly since only the Theme position, being free of syntactic constraints, is able to handle all topicalized nominals, whatever their pragmatic links with the predicate? (Dik labels this all-embracing pragmatic function 'scene-setting'; 1980: 143.)

In sum, then, French has, in the period since the emergence of the written language, seen two important changes in respect of the ordering of the major constituents of the sentence. Firstly, we find that a Topic-initial language gives way to a Subject-initial language ('a V2s language gives way to a V3 language'), this change taking place within the Predication. This development, in no way surprising cross-linguistically, was complete before the codification of the standard language, and is thus taken to represent the canonical word order. At some subsequent period, however, - and in reality probably only very shortly afterwards - a new method of Topic marking came into use, that is, the use of Theme position, outside the Predication and marked by 'comma intonation'. This structure is now so widely found that one might legitimately claim that spoken French is once again a Topic prominent language. This change, however, has not received any 'official' recognition, although Topic-initial sentences are, as Ashby's data clearly show, a major sentence type when talking about a given referent.

At the same time, less predictably, French has made greatly increased use of what was presumably originally a very minor sentence type employed to express 'afterthoughts'. The pragmatic function of this sentence type has clearly expanded significantly, to the point of being very similar to that of Topic marking, but beyond that there is some evidence of
grammaticalization, on both the pragmatic and the intonational levels (cf. Dik 1978: 156). A context in which the speaker is thinking of someone (hence the referent is 'given' for him), utters a sentence incorporating an anaphoric pronoun, and then realizes that the referent is, or may be, unknown to the hearer, causing him to add new information at the end, thus:

(13) Elle est très jolie, Claire, n'est-ce pas?

is clearly a grey area in terms of givenness, and might be the type of structure in which 'subjecthood' comes to outweigh 'objecthood'. We have also suggested that the phrase-stress nature of contemporary French may be a relevant factor. Whether one sees these right-dislocated elements in syntactic or pragmatic terms or (most plausibly) as sitting uncomfortably between the two, the widespread use of VXX structures cannot be doubted. This fact also is in general passed over by prescriptive and pedagogical grammars.

It seems appropriate to note at this point just how 'free' the word order of contemporary spoken French actually is. We have discussed left- and right-dislocated structures, but it should not be forgotten that, in addition, these may co-occur. Dik, for instance, discusses (1980: 143):

(14) Ton frère, il l'a donné à Pierre, le livre

(which he analyses, unexceptionably, as Theme, Predication, Tail) (and cf. also Dik 1978:156). This fits in in general with his view (loc cit) that 'terms with only a semantic function have great freedom of expression with respect to the Verb'. We note also how closely French corresponds to Steele's categorization (1978: 610) of 'free word order languages' as exhibiting 'person agreement with the subject of the sentence', often 'transparently related to the independent pronominal forms of the language': clitics co-referential with most of the other major syntactic categories (derived incidentally from exactly the same source) can, it seems to me, only enhance the possibilities for maximally flexible word order. (Morphologization of previously independent constituents is of course a frequent concomitant of the passage from 'discourse' to 'syntactic' structures: Givón 1979: 211.)
And this leads to the final question to be posed in this short survey: what now? Dik is clearly right when he says that more than one functional pattern can co-occur: but do they always have identifiably different functions? In spoken French, the canonical SV... pattern seems to be largely restricted to intransitive verbs, with dislocated or cleft structures used to avoid SVO, the latter, of course, also being fully acceptable in the standard language. Left-dislocated structures are very frequent when the Subject is appropriate for topicalization, with right-dislocation also available in most cases. What criteria determine the use of one rather than another? At present, I have no answer to this question: indeed, in his classic analysis, Charles Bally (1965: 61) suggests explicitly that it is the 'segmentation' itself rather than the precise position of the dislocated element which serves to indicate a topicalized structure. (See also v.Wartburg 1967: 259.) If the pragmatic differences are as slender as they appear, perhaps this in itself militates against the survival of Tail as a second Topic marking slot, and points to the reanalysis of constituents occurring in this position in syntactic terms: recall that only nominals with a specifiable syntactic link with the Predication can occur in this position. On balance, I still feel this to be the most likely outcome, while recognizing the force of some of Lambrecht's objections insofar as one is speaking purely synchronically.

Last of all, what chance is there of any of these changes coming to be recognized as part of the standard language? As long ago as 1968, Marjorie Zengel argued that a significant degree of literacy would retard linguistic change (her actual field of interest being vocabulary). Alongside any possible retardation in the rate of change itself, however, we must consider also the delayed acknowledgement of changes that have occurred, either after standardization but before mass literacy, or even despite literacy (which of course does not and will not inhibit change completely). As a powerful complement to literacy itself — indeed, some would say the more important factor — we must bear in mind the effects of formal schooling and the instillation of linguistic norms as a conservative force. Certainly, the evidence of these very major changes in the most frequent arrangements of the principal constituents of French sentences suggests that the 'canonical' structures can be almost wholly absent from other than formal (largely written) registers, and yet in no way accepted
as being worthy to be described and taught. Indeed, Romaine (forthcoming: ch. 6) clearly shows, on the basis of data presented by Dannequin (1977) relating precisely to the structures under consideration here, that French-speaking children are frequently corrected for using 'dislocated' structures in conversation, even though these are, as we have already observed, on many occasions the most natural mode of expression. Perhaps this resistance to innovation is particularly strong in the case of France, given the peculiarly intimate links between nationhood, culture and language in that country: nevertheless, it is clear that the 'function' of maintaining traditional standards of educational attainment and of literacy is one which interacts with other pressures in determining when or indeed whether a particular incipient change may be consummated.
Notes

*This is a slightly amended version of a paper read at the workshop 'Functional causes of language change and language variation', 19-21 December 1983.

1. But not always. Consider, for instance, an example drawn to my attention by colleagues at the Lille Round Table 'La diachronie hier et demain' (September 1982), attested in a deodorant advertisement:

(1) Les filles, moi, j'aime ... mais j'ai le nez délicat

The precise pragmatic circumstances under which the non-use of a co-referential is acceptable clearly need to be investigated. When the left-dislocated element is a prepositional phrase, then the clitic pronoun can only be omitted if the preposition appears overtly, in which case we are dealing with the 'minor' type of left-dislocation discussed below.

2. This example is drawn from Lambrecht 1981: (example 63). The extent to which my current views have benefited from Lambrecht's published work, and indeed from much personal correspondence, is made clear in Harris (forthcoming).

3. In general, this means Subjects, Direct Objects and prepositional phrases where the relevant preposition is de or à. The scope of the latter is considerably expanded in popular French, thus permitting cliticisation where this would be ruled out in standard French: cf. Lambrecht (1981: 39f).

4. Recall the observation in Dik (1980: 170): 'A language may well have, and many languages do have, different functional patterns side by side, to be used under different conditions'.

5. Interestingly, conjunctions do not function like adverbs, in that they do not fill the P1 position. Verbs therefore typically occur in third or final position in subordinate clauses. The position is discussed in detail in Foulet (1968: paras 451-5), where the difficulty of distinguishing rigidly in certain cases between 'adverbs' and 'conjunctions' is noted.
6. '(R3)' is a general rule (Dik 1980: 152) which reads 'X \rightarrow P1'.

7. It was at this time and in this way that atonic Subject pronouns began to develop in French (unlike Spanish or Italian) along the road which led them first to obligatory status when there was no other Subject in surface structure, then to obligatory status tout court, and then ultimately to the status of bound affixes, as described here: see Harris (1978: 111-4). For an indication as to why Subject pronouns were more liable to 'omission' in mains clauses than subordinate clauses, see Foulet (1968: para. 459).

8. Bailard (1982) suggests rather that the emergent order is VOS, a claim I discuss in Harris (forthcoming). Given Dik's claim (1980:136) that 'the Subject is invariably placed before the Object', VOS is perhaps inherently less likely, that is, we might expect:

   (ii) Elle l'attend, Marie, ton père

rather than:

   (iii) Elle l'attend, ton père, Marie

In fact, both structures are frequently attested. Two observations seem apposite. Firstly, provided these nominal elements are seen as part of the Tail rather than the Predication, Dik's (syntactically defined) ordering constraint would not be applicable (1980: 137), as the relative position of S and O within the Tail would be pragmatically governed. (This counter-argument of course becomes progressively weaker as and when grammaticalization takes place, but then the predictions found in Dik (1980: 146) would lead us (given an SVO or VSO starting point) to expect VOS rather than VSO, whereas no such preference appears actually to be found.) Secondly, Dik's constraint does apply specifically to Subjects, whereas we shall shortly examine possible reasons for not regarding these right-dislocated constituents as Subjects in the usual sense.

9. The existence of two slots in which topicalized elements may occur is apparently not in itself unusual. Dik (1980: 131), for instance, proposes the existence of two such slots in Serbo-Croatian, both within the Predication.
10. It is worth observing that popular French also provides an instance of a former clitic developing a function other than that of affixed category marker, parallel once again with changes described in a number of other languages. In this case the pronoun il, inverted to form questions thus, vient-il? has fused with with the final [t] of [vještî] etc., and been reanalysed as a general interrogative marker (diversely orthographed), affixed to the verb. A wide range of examples can be found in Grevisse (1980: para. 1759), including

(iv) vous êtes-t-y prêts?
(v) je savais-ty, moi, pauvre innocent?
(vi) tu veux-t-y que je relâche, oui ou non?

11. I take this opportunity of heavily qualifying, in the light of observations by Knud Lambrecht (personal communication), the claim which both Bailard (1982: 22) and I made about the ambiguity of:

(vii) elle l'aimait bien Marie sa mère
(viii) elle l'aimait bien sa mère Marie

Both are indeed grammatical in an appropriate context, and one can easily construct truly ambiguous sentences to illustrate the respective points she and I were making; this particular example, however, is vitiated by the use of the possessive sa, which virtually imposes, for obvious reasons, the interpretation that Marie is the Subject, who loved her mother. (vii) is thus an instance of a VSO structure, and (viii) of VOS (with the caveats already noted as to whether we are dealing with canonical Subjects and Objects).

12. The frequent overlap in practice between Topics and Subjects does not of course help to resolve this question.

13. Consider, for example, from Dannequin (1977: 75) the child who says 'J s'promène(n)t les oies', to which the teacher responds 'Alors tu me dis "I s'promène(n)t les oies". Il faut mieux dire "Les oies se promènent"'. The force of the 'mieux' in this sentence is clear! I am grateful to Suzanne Romaine for drawing Dannequin's material to my attention.
Bibliography

Wartburg, W.v. (1967)\textsuperscript{8}, \textit{Evolution et structure de la langue française}.
Berne: Francke.

Zengel, M. (1968) 'Literacy as a factor in language change'. In:
J.A. Fishman (ed.), \textit{Readings in the Sociology of Language}.