



University of
Massachusetts
Amherst

Word Order and WH-Movement in Basque

Item Type	article;article
Authors	Azkarate, M.;Farwell, D.;de Urbina, J. Ortiz;Saltarelli, M.
Download date	2025-01-14 22:07:01
Link to Item	https://hdl.handle.net/20.500.14394/36380

Word Order and WH-Movement in Basque

M. Azkarate, D. Farwell, J. Ortiz de Urbina and M. Saltarelli
University of Illinois, Urbana

0. Introduction

Inversion (post-verbal subjects) and pro-drop (null subjects) distinguish languages like Italian and Spanish from languages like French and English. This surface syntactic variation is reduced to the selection of the pro-drop/null subject parameter in the core grammar from a limited number of parameters available in Universal Grammar. Thus, the core grammar of Italian and Spanish incorporate an optional rule of syntax which assigns the Inflection constituent to the Verb-Phrase. If the rule applies, only an empty category may occupy the structural subject position, as predicted by the theory of government and binding (Chomsky 1981).

This paper deals with "free" word order (of which subject inversion is an instance) and the occurrence of empty categories (of which subject pro-drop is an instance) in Basque, within a Universal Grammar framework conceived as a parametrized system. It will be shown that when the data of Basque is included along with English and Italian in a cross-linguistic comparison of surface variations in word order and pro-drop the analytical perspective of the issue is significantly altered. In fact, it will be argued that the definition of pro-drop language as Inflection lowering is too narrow, and that the surface phenomena of inversion and pro-drop follow from independent parameters. The analysis of Basque proposed in this paper finds independent support in that it distinguishes between bounded and unbounded hypotheses of WH-movement. The interaction of intervening clauses requires a successive cyclic application of the rule of WH-movement.

The paper is organized as follows. Free word order phenomena are discussed in section 1. Section 2. presents the interaction between word order and focus (galdegaia) in clauses with extracted question-word. In section 3. a preliminary analysis of Basque as a free word order language is outlined.

1. "Free" word order phenomena.

In affirmative main clauses, all permutations of the major categories are possible except for those which would affect the order or contiguity of the elements in the so-called "nuclear verb phrase;" that is, the sequence of elements beginning with the focus position (galdegaia) immediately before the verb and the verb complex itself composed of the participle and the auxiliary(1).

- (1)(a) galdegaia + verb ('synthetic verbs')
(b) galdegaia + participle + auxiliary ('compound verbs')

Thus, given a sentence such as (2) where the direct object, liburua (the book), is understood as being the element in focus, the possible distributions of the non-nuclear elements can be

Azkarate, Farwell, Ortiz de Urbina and Saltarelli

represented by (3) where X is any unordered subset of elements of Mireni, egon-gelan including the null set, and Y is any unordered subset of elements taken from the complement of X in $\{ \text{Jonek, Mireni, egon-gelan} \}$. This would reflect the fact that such sentences as (4) are all well-formed, while those in (5) are ungrammatical.

- (2) Jon-ek Miren-i egon-gela-n liburu-a
 John-erg Mary-dat living room-def(loc) book-def(abs)

irakur-ri d-io
 read-1prt 3s(0)-[prs]-[have]-3s(d)-[3s(S)]
 (John read the book to Mary in the living room.)

- (3) . . . X . . . liburua irakurri dio . . . Y . . .

- (4)(a) Jonek egon-gelan liburua irakurri dio Mireni.
 (b) egon-gelan Jonek liburua irakurri dio Mireni.
 (c) liburua irakurri dio Jonek Mireni.
 (d) liburua irakurri dio Mireni Jonek.
 (e) liburua irakurri dio.

- (5)(a) *Jonek liburua irakurri Mireni dio.
 (b) *Jonek irakurri liburua dio Mireni.
 (c) *Jonek dio liburua irakurri Mireni.

Specifically, in (5a) the indirect object, Mireni, has broken the continuity of the nuclear verb phrase while in (5b) and (5c) the relative order of the elements within the nuclear verb phrase has been altered.

In general, any of the major categories of a given clause is a possible candidate for the galdegaia or focus position so that in place of the direct object, liburua, in the above examples, we might have found the indirect object, Mireni, the subject, Jonek, or the locative adverb, egon-gelan, or even the verb itself. In each case, the freedom to permute is as open as in the case just outlined.

In negatives, as with affirmatives, any ordering of the major categories is possible except those which would affect the contiguity or relative order of the elements within the nuclear verb phrase as expressed in (6).

- (6)(a) ez (neg) + verb + galdegaia
 (b) ez + auxiliary + galdegaia + participle

The possible distributions for the negative counterpart of (2), then, may be represented as (7),

- (7) . . . X . . . ez dio liburua irakurri . . . Y . . .

where X and Y are as described with respect to (3) above. Again, this reflects the fact that the sentences in (8) are grammatical while those in (9) are not.

Azkarate, Farwell, Ortiz de Urbina and Saltarelli

- (12)(a) Mikelek esan dio Jonek Mireni liburua irakurri
Michael-erg tell-1prt

dio-la
-COMP

(Michael told him that John read the book to Mary.)

- (b) Mikelek galde-tu dio Jonek Mireni liburua irakurri
ask-1prt

dio-n
-COMP

(Michael asked him if John read the book to Mary.)

In (12a) the embedded clause is marked by the complementizer -ela which roughly corresponds to the "that"-complementizer in English while the embedded clause in (12b) is marked by the complementizer -n which correspond to the WH-complement in English. The non-finite types are represented by the sentences in (13).

- (13)(a) Jon-i Mireni liburua irakur-tze-a gusta-tzen
John-dat read-nom-def(abs) please-2prt

z-a-io

3s(S)-prs-[e] -3s(D)

(Reading the book to Mary is pleasing to John.)

- (b) Joni liburua nola irakur-ri ahaz-tu zaio
how read-1prt forget-1prt
(John forgot how to read the book.)

- (c) Jon ikus-i d-u Mireni liburua
John[-abs] see-1prt 3s(0)-[prs]-have-[3s(S)]

irakur-tzen

read-2prt

(He saw John reading the book to Mary.)

The complement in (13a) is marked by the -tze nominalizer, the one in (13b) by the "first" (past) participle and that in (13c) by then "second" (present) participle. Each of these complements may serv a variety of syntactic roles and, in doing so, is as open to permutation as any non-sentential constituent. Each complement is furthermore, a potential candidate for galdegaia.

In affirmative subordinate clauses, regardless of their syntactic form, we found the same freedom of permutation as in main clauses provided that they neither (a) are in focus positions nor (b) bear non-verbal morphology (i.e., case or number markers). Thus, while the sentences in (14) are well-formed, those in (15) are ungrammatical.

Word Order and WH-Movement in Basque

- (14)(a) Mikel-ek esan dio liburua irakurri dio-la Jonek
 Michael-erg tell-[1prt] -COMP
 Mireni
 (Michael told him that John read the book to Mary.)
- (b) Mikelek galde-tu dio liburua irakurri dio-n Jonek
 ask-1prt -COMP
 Mireni.
 (John asked him if John read the book to Mary.)
- (c) Jon ikus-i du liburua irakur-tzen Mireni
 John-[abs] see-1prt read-2prt
 (He saw John reading the book to Mary.)
- (d) Jon-i ahaz-tu z-a-io nola irakur-ri
 John-dat forget-1prt 3s(S)-prs-[be]-3s(D) how read-1prt
 liburua
 (John forgot how to read the book.)
- (15)(a) *Mikelek liburua irakurri dio-la Jonek Mireni esan dio
 -COMP
 (Michael told him that John read the book to Mary.)
- (b) *Mikelek liburua irakurri dio-n Jonek Mireni galdetu
 -COMP
 dio
 (Michael asked him if John read the book to Mary.)
- (c) *liburua-ak Mireni irakur-ri-ak Jonek
 book-def.pl(abs) read-1prt-def.pl(abs)
 dir-a
 3p(S)-prs-[be]
 (The books are read to Mary by John.)
- (d) *erabak-i du bihar irakur-tze-a Mireni
 decide-1prt tomorrow read-nom-def(abs)
 liburua
 (He decided to read the book to Mary tomorrow.)

In (15a) and (15b), the embedded clause is in the preverbal focus position. In (15c) and (15d), the embedded clause bears non-verbal morphology.

The negative counterparts to the examples presented in (12) and (13) above are even further constrained in terms of word order. In fact, the negative counterparts to (13) simply do not allow any element to appear behind the verb at all. Thus the sentences in (16) are ill-formed.

- (16)(a) *Jon ikusi du ez entzu-ten irakasle-ari
 neg listen-2prt teacher-def(dat)
 klase-an
 class-def(loc)
 (He has seen John not listening to the teacher in class.)
- (b) *Joni ahaztu zaio ez ahaz-tu nola gauza-k
 forget-1prt how thing-def.pl.(abs)
 (John has forgotten how not to forget things.)

Nevertheless, the freedom to postpose in the finite clauses as in (17) is perfectly acceptable provided that the conditions mentioned above are met.

- (17)(a) Mikelek esan dio Jonek ez dio-la liburua irakurri
 neg -COMP
 Mireni
 (Michael told him that John hasn't read the book to Mary.)
- (b) Mikelek galdetu dio Jonek ez dio-n liburua irakurri
 -COMP
 Mireni
 (Michael asked him if John hasn't read the book to Mary.)

2. Question-word formation and focus interpretation.

As pointed out above, question-word questions take the same form as statements with the exception that the focus position must be filled obligatorily by the question-word. Otherwise, they exhibit the same word order restrictions as statements. With respect to question-word questions formed over some element in an embedded clause, there are two possible strategies. By the first of these strategies, the appropriate question-word is focussed before the embedded verb, the entire embedded constituent appearing as the galdegaia of the matrix verb. Thus, the questions in (18) are grammatical while those in (19) are ill-formed.

- (18)(a) Mikel-ek Mireni liburua nor-k irakurri diola esan dio?
 Michael-erg who-erg
 (Michael told him that who read the book to Mary?)
- (b) Jon-i Mireni zer irakur-tze-a gusta-tzen zaio?
 John-dat what-[abs] read-nom-def(abs)please-2prt
 (It is pleasing to John to read what to Mary?)
- (c) Jon-ek liburua nor-i irakur-tzen ikas-i du?
 John-erg who-dat read-2prt learn-1prt
 (John has learned (how) to read the book to whom?)

Word Order and WH-Movement in Basque

(19)(a) *Mikelek esan dio Mireni liburua nori irakurri diola?
 (Michael told him that who read the book to Mary?)

(b) *Joni zer Mireni irakurtzea gustatzen zaio?
 (It is pleasing to John to read what to Mary?)

In (19a) the embedded question does not appear in the preverbal focus position while in (19b) the embedded question is improperly formulated, the question-word not being in focus position. Because of the requirement that the embedded question appear in preverbal position, it is not surprising to find that the permutation of the elements outside of the nuclear verb phrase is not possible, this being the very position where postposing was not allowed.

The second strategy for the formulation of question-word questions with respect to elements in an embedded clause involves the focussing of the appropriate question-word before the matrix verb. In this case, the embedded clause with which the question-word is associated must appear after the matrix verb. This strategy is reflected by the questions in (20), which are well-formed alternatives to those in (18).

(20)(a) zer esan dio Mikelek, irakurri diola Jonek Mireni?
 what-[abs]
 (What did he see John reading to Mary?)

Note that whereas the "clause internal" strategy is extremely free in terms of the syntactic types and functions which the relevant clause may have, the "cross-clause boundary" strategy is limited to the types of complements shown in (20), functioning as absolutive arguments of a certain class of verbs.

In all of the acceptable examples in (20), the embedded verb has appeared immediately after the matrix verb. In fact, in the two dialects we looked at there was a marked tendency to place the embedded verb in clause initial position when such questions were formulated, and in one of these dialects, the initial position was obligatory. The speaker of this dialect, then, would not accept a question of the type (21)

(21) zer esan dio Mikelek, Mireni irakurri diola Jonek?
 (What did Michael tell him that John read to Mary?)

as the dative object, Mireni, intervenes between the matrix material and the lower verb. Of greater interest, however, is that in both dialects, particularly in the one which accepts (21) as well-formed, there is no galdegaia in the embedded clause. Rather the fronted question-word is interpreted as the lower clause galdegaia.

Finally, it should be noted that the "cross-clause boundary" focussing strategy can operate over various levels of embedding as can be seen in (22a) and (22b), which correspond to the two dialects mentioned.

- (22)(a) zer esa-ten du Haritz-ek, uste du-ela Mikelek,
 tell-2prt Haritz-erg believe-[2prt]-COMP
 irakurri dio-la Mireni Jonek?
 -COMP
 (What does Harry say that Michael believes that John read
 to Mary?)
- (b) zer esaten Haritzek, Mikelek uste duela, Jonek Mireni
 irakurri diola?
 (What does Harry say that Michael believes that John read
 to Mary?)

Curiously, in this context it is a property of the intervening clauses that, like the most deeply embedded clause with which the question word is associated, they cannot have any element interpretable as galdegaia. In addition, in (22a), which corresponds to that dialect in which the verb is required to appear in clause initial position whenever a question word is extracted from it, the intervening clauses must be verb initial as well. That is, when some element appears in preverbal position in an intermediate clause in this dialect as in (23), the question will be formed.

- (23) zer esaten du Haritzek, Mikelek uste duela irakurri diola
 Jonek Mireni?
 (What does Harry say that Michael believes that John read
 to Mary?)

3. Government, binding and the unordered base hypothesis.

Basque displays a surface structure behavior in maximal S whereby all permutations of major categories are possible, except for the category which is interpreted as being in galdegaia, i.e., the focus position.

Word order variation in maximal S has been treated in the literature in terms of hypotheses which assume an invariant word order in the base plus rules of syntax which account for all possible permutations at S-structure. Recent studies include Katalin Kiss' analysis of Hungarian "free" word order (Kiss 1981) and Chomsky's analysis of "inversion" in pro-drop (null subject) languages like Italian (Chomsky 1981a,b).

The government and binding analysis of subject-verb inversion in Italian, which is normal in sentences with bare verbs like Gianni ha scritto/ha scritto Gianni 'Gianni has written', attributes this apparently free surface order phenomenon to the "pro-drop (null subject)" parameter which is displayed in (24) at S-structure.

- (24)(a) [Gianni' AGR' [ha scritto]]
 S VP

- (b) [PRO' [AGR' ha scritto]] "pro-drop (null subject)"
 S VP

Word Order and WH-Movement in Basque

- (c) [PRO' [[AGR' ha scritto] Gianni']] "inversion"
 S VP VP

At D-structure AGR, a constituent of INFL, is co-superscripted with the NP it governs, namely the subject GIANNI in (24a), which receives the nominative Case in S-structure if it is co-superscripted with and is governed by AGR.

Pro-drop languages are characterized by a rule R which may apply in the syntax, and if so, assigns INFL to VP (25) (Chomsky 1981b).

- (25) Rule R assigns INFL to VP

As a consequence of rule (25), available only in the grammar of pro-drop languages and in accordance with the theory of government and binding, the subject Gianni is governed by AGR in (24a). The "missing" subject in (24b) and the subject inversion structure (24c) must have PRO in subject position as that position is not governed by AGR at S-structure, since rule R lowered INFL into the VP.

Basque, like Italian, falls under the definition of pro-drop language by virtue of the fact that it displays, like Italian, the two major syntactic phenomena of pro-drop and inversion. In fact, Basque has considerably richer evidence in support of the pro-drop parameter. Whereas in Italian pro-drop and inversion is limited to the category subject, in Basque the two phenomena obtain for the subject, the direct object and the indirect object as well. It is this extended inversion process which would account, at least in part, for the apparently free word order in maximal S discussed in the first part of this paper.

Assuming the government and binding analysis proposed by Chomsky for Italian, the pro-drop parameter representation for Basque requires that all three categories (which correspond to subject, direct object and indirect object in accusative languages) must appear in ungoverned position at S-structure. This must be so because the language under consideration allows null subjects as well as null direct and indirect objects (26a) and because inversion of all three said categories is possible (26b).

- (26)(a)(i) Mikel-i eskutitza idatzi d-io-zu (null Ergative
 Michael-D letter-A write A-D-E subject)
 (you write the letter to Michael)
- (ii) Mikel-i idatzi d-io-zu (null Ergative/subject)
 Michael-D write A-D-E (null Absolutive/direct object)
 (you write it to Michael)
- (iii) idatzi d-io-zu (null Ergative/subject,
 write E-D-E null Absolutive/direct object,
 (you write it to him) null Dative/indirect object)

- (b) idatzi d-io-zu zu-k eskutitza Mikel-i (inversion of E-A-D)
(you write the letter to Michael)

The extended pro-drop parameter just exemplified in Basque indicates that the characterization of pro-drop vs. non-pro-drop languages, proposed in view of Italian and English, as the R rule hypothesis lowering INFL into VP, cannot be the correct one. INFL lowering captures the facts of the pro-drop parameter of Italian, limited to the manipulation of the binding properties of the subject which is governed by AGR. As such the R rule hypothesis cannot in principle account for the extended pro-drop phenomena of Basque which involve in addition the category direct and indirect object.

From the above it follows that any adequate government and binding theory of Basque must be rich enough to provide at S-structure a scenario where all major categories in maximal S (a) appear in all possible permutations and (b) are ungoverned. Any hypothesis characterizing such a theory must insure that INFL is in \bar{V} and at the same time that none of the major categories is in \bar{V} .

One such viable theory, which we are proposing for Basque, assumes a (partially) unordered base in the form (27).

- (27)(i) $S : X^{n*} \bar{V}$
 (ii) $\bar{V} \rightarrow e$ \bar{V} $e = \text{galdegaia}$
 (iii) $\bar{V} \rightarrow V$ INFL

The first base rule (27i) expands S as a 'mobile' set, following a proposal discussed in Bach (1975), Peterson (1971), and others. The mobile-S hypothesis allows the grammar to provide at D-structure a set of ordered phrase markers defining the possible permutations of the verb and the other categories in S. In the base (27), furthermore all major categories are generated in ungoverned positions.

In Basque, the cyclic nodes are S and \bar{V} . \bar{V} is the binding node. INFL is not a governor. There are two general rules of syntax (28). The rule Move- α will move optionally any major category into the empty

- (28)(i) Move α (into galdegaia)
 (ii) Move WH (from galdegaia to galdegaia)

structural galdegaia position e. Galdegaia must be filled. Operators like interrogative WH-words and negatives must be in galdegaia, the focus position. Extraction of WH-words is possible from galdegaia position through the rule of Move-WH which applies in Basque from galdegaia to galdegaia.

Basque displays morphological case marking on all major categories. There are perhaps as many as fifteen different cases, with Absolutive being the unmarked case. Absolutive, Dative, and

Word Order and WH-Movement in Basque

Ergative are "doubled" in the inflection of the verb by a system of three ordered structural positions which are filled by morphemes resembling clitic pronouns in portmanteau relation to the auxiliary verb (in compound verbs) or the the main verb (in synthetic verbs).

Following a proposal in Osvaldo Jaeggli's dissertation (Jaeggli 1980) in regard to the analysis of "clitic doubling" in Porteno Spanish (29), we assume that clitics are governed by V and cannot be PRO

(29)(i) lo' vimos a Pedro' (we saw Pedro)

(ii) lo' vimos PRO' (we saw him)

(iii) *vimos a Pedro

in languages with clitic doubling. As in the case of pro-drop and inversion, Basque displays a more extensive phenomenon of clitic doubling which involves Absolutive, Dative and Ergative arguments (cf. (26)).

We propose that INFL is expanded in the base as follows, in the case of Basque:

(30) INFL → (Aux) A-D-E

Thus the inflectional morpheme positions are generated in the base under \bar{V} . Consequently, they are governed by V and receive Case. Inflectional morphemes are cosuperscripted at D-structure with any variable in S. Major categories are assigned (or checked for) Case at S-structure. In Basque, therefore, the governors are two, V and P.

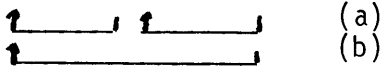
At this point we would like to return to the rule of WH-movement for a brief discussion of its effect on Basque "free" word order and its apparent significance for subjacency.

Questioning out of a clause involves in Basque the extraction of the WH-word from its logical galdegaia into a galdegaia position in a higher clause. In such cases the "free" word order permutation of constituents is restricted for those clauses which intervene between the extracted WH-word and its scope. The only acceptable word order is V-initial, in one strategy (31bi). In another strategy (31bii) there is no restriction on word order in the intervening clauses. Both strategies, however, require that the galdegaia cannot be filled by any of the categories in S. A descriptively adequate description consistent with both strategies would then be one which ensures that the rule of Move- α does not apply in the intervening clauses.

(31)(a) [Jon-ek nor-k esan z-u-en[uste d-u-ela Peru-k[dei-tu
 John-E WH-E say Aux believe Aux-COMP Peter-E call
 Aux-COMP M.-D
 z-io-la-Mikel-i
 (Who did John say that Peter believes called Michael?)

- (b)(i) [Jon-ek nor-k esan z-u-en *[Peru-k uste d-u-ela[dei-tu
z-io-la Mikel-i
(ii) [Jon-ek nor-k esan z-u-en[Peru-k uste d-u-ela[dei-tu
z-io-la Mikel-i

Given a successive cyclic (32a) vs. an unbounded (32b) rule of WH-movement as alternative descriptive devices, the data of Basque and the two

- (32) [WH-V [t-V [t-V
 (a)
(b)

strategies fall out only if one assumes that WH-movement is successive cyclic. This is so because Move- α and WH-movement are mutually exclusive in Basque, given a trace theory and the uniqueness in galdegaia. Thus the two rules of syntax, Move- α and Move-WH are bounded, a characteristic which falls out from subadjacency.

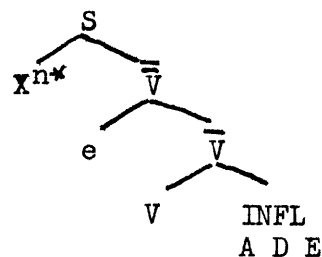
The data of Basque concerning the limitations on galdegaia (focus) interpretation in connection with question-word extraction distinguishes between a bounded and an unbounded theory of WH-movement in this language.

4. Summary and conclusions.

In section 3, an analysis of Basque was outlined which, while in keeping with the concepts of government and binding, requires a description which is at variance with some of the conclusions reached in Chomsky (1981) on the basis of Italian and English. Only if factored as in 3, we claim, can an adequate account be given of phenomena such as "free" word order, pro-drop (null) category, (inflection) doubling, focus (galdegaia) interpretation, and question-word in extraction. Moreover, given the discussion in 3, the data of Basque clearly distinguishes between the rule of Move-WH and Move- α as mutually exclusive. Given the same premise the data of Basque distinguishes furthermore between a successive cyclic and an unbounded analysis of question-word extraction.

Basque sentences, we propose in (27, 30, and 28), conform with the representation (33) where X^{n*} is an unordered set of categories

(33)



Word Order and WH-Movement in Basque

including S, NP, PP with inherent case. e is the focus (galdegaia) structural position. INFL includes (in addition to Tense and other inflectional categories) the three thematic positions Absolutive, Dative, Ergative, which are governed by V. The thematic positions in INFL are co-superscripted uniquely at D-structure with the categories in X^n , in accordance with the binding conditions. There is one movement rule in its general form Move- α , which moves any major category including V in galdegaia (focus) position e (28i). All permutations of the constituents dominated by S are possible, as postulated by (27i). Constituent order is fixed in the domain of other categories (27ii, iii, and 30).

Basque "free" word order in S is characterized in our analysis in terms of the (partially) unordered base parameter. Accordingly, Basque would differ from English by selecting at core grammar the less restrictive initial base rule (27i). Inversion in Italian can be captured either by Rule R (25) or by a variant of the unordered base parameter ($S : NP VP$) with some restrictions. In either case, the categories X^n in S must be ungoverned.

The appearance of null categories (Subject, Direct Object, and Indirect Object) (cf.26) follows from two factors: (a) the category is ungoverned and (b) the category is "doubled" (i.e., co-superscripted) in INFL. Thus in Basque the inflectional positions A-D-E are the thematic positions. Pro-drop, we claim, is independent of the inversion phenomenon or of the more extensive "free" word order found in languages like Basque.

The structural position galdegaia (e in 33) combines the positions of FOCUS and COMP in Basque. This is apparently the only descriptive solution, given the mutual exclusiveness of Move- α and Move-WH observed in relation to question-word extraction (cf. section 2).

References

- Bach, E. (1975), "Order in Base Structures," in Charles N. Li (ed.) Word Order and Word Order Change. University of Texas Press: Austin, 1975. 307-345.
- Chomsky, N. (1981a) Lectures on Government and Binding. Foris Publications: Amsterdam.
- _____. (1981b) "A Note on Non-control PRO*," unpublished manuscript.
- Jaeggli, O. (1980) On Some Phonologically-Null Elements in Syntax, unpublished MIT Ph.D. dissertation.
- Kiss, K. (1981) "Structural Relations in Hungarian, a Free Word Order Language," Linguistic Inquiry 12.2. 185-213.
- Peterson, T.H., (1971) "Multi-ordered Base Structures in Generative Grammar," Chicago Linguistic Society 7. 181-192.