Annali Online di Ferrara - Lettere
Vol. 1 (2007) 1/13

CARLO CECCHETTO

Some preliminary remarks on a 'weak' theory of linearization

1. Introduction

One traditional research topic in syntax is the question of linearization of hierarchical structures. Extensive research stemming from the seminal work by Noam Chomsky in the fifties has shown that natural languages have a hierarchical organization, but, since utterances take place in time and not instantaneously, hierarchical structures must be mapped in a linear sequence of words. Linearization can be seen as a process in which a bi-dimensional structure (as illustrated by the traditional branching trees) is conflated in a mono-dimensional one (a linear sequence). Clearly, this process is rule-governed. It is not the case that anything goes, probably because the linearized structure is the input for language acquisition, namely a child has to infer some properties of his/her own language based on a linear input. Generative grammar has systematically dealt with the question of linearization at least from the seventies, when what would get the name of X-bar theory was first elaborated. X-bar theory claims that the each phrase adheres to a universal format according to which the Head (a lexical item) can be attached to a Complement (itself a phrase). To the unit formed by Head and Complement, a further phrase, called Specifier, is attached (adjuncts can further complete the structure).

According to a more or less standard version of X-bar theory, the only significant point of systematic cross-linguistic variation in phrase structure is the linear order between Head and Complement, namely languages would come in two big varieties: the [Spec [Head-Compl]] type, illustrated by Italian and English, and the [Spec [Compl-Head]], illustrated by Japanese and Turkish. According to this view, linear order is directly encoded in syntax and a child acquiring a language must set the value of the Head parameter for his/her target language. As for Specifiers, the question was less extensively investigated, but at least prototypical Specifiers created by a movement operation (say, Spec,CP or Spec,IP) seemed to be universally left branching, so no Spec Parameter, akin to the Head Parameter, was part of the standard syntactic toolbox[1].

The picture that I have roughly summarized, which was a point of contact between generative theory and the typological tradition stemming from Greenberg's work (cf. Greenberg [1966]),

[1] However, X-bar theory assumes that the basic X-bar skeleton can be enriched by the presence of other phrases that have the role of adjuncts, which can be attached more or less freely either to the left or to the right. So, some instances of rightward movement and some rightward branching structures were allowed, but, crucially, not so for canonical cases of movement (*wh*-movement and A movement).

prevailed until the early nineties when Richard Kayne proposed a revised theory of linearization. Kayne's theory (cf. Kayne [1994]) is based on Linear Correspondence Axiom (LCA), a condition that, roughly summarized, states that asymmetric c-command is mapped into linear precedence. The theory based on LCA allows a radical simplification of X-bar theory, so it has been adopted as a guideline by many researchers. However, it also met opposition in the field, partly because the way LCA is formulated does not leave a space for the Head Parameter, since all languages are predicted to be Head initial (prior to re-arrangement that can lead to the superficial Complement-Head order). LCA also blocks rightward movement of any type.

Chomsky, in minimalist work (cf. Chomsky [1995] and following work), proposes that the basic structure building operation (Merge), which creates a set by merging two syntactic objects, is insensitive to linear order and that linearization is decided only at the PF interface, i.e. in phonology. So, strictly speaking, in core syntax it does not make sense to ask whether a Head precedes or follows its Complement. Therefore a theory of linearization would not sit in the core of the system but in a more peripheral position (at the PF interface).

In this squib, I will approach the question of linearization by reflecting on some cases that seem recalcitrant to a systematic treatment, namely the fact that recent work on sign languages has shown that specifiers can be linearized on the right and the existence of disharmonic languages in which the order of Head and Complement is not stable across categories. I will propose that starting from these "exceptions" can indeed be very useful because discovering why "exceptions" are tolerated might allow us to identify the "rules" that govern linearization in general. As is usual practice in other disciplines, you start from "pathology" in order to discover regularities.

2. The Linearization of Specifiers

Although a syntactic category can either be base generated in or moved to a Specifier position[2], in the latter case it is easier to detect the linear position of (the material in) the Specifier. The reason is that in most cases the material that is base generated in a Spec position moves to a higher functional position (i.e. the external argument of a transitive verb that moves to Spec,IP) or the material surrounding it does, so the initial linear order configuration is often obscured[3].

If one focuses on Specifier positions created by a movement operation, the common wisdom, as I already said, is that Specifiers are linearized on the left. This is particularly clear for cases of *wh* movement, since in the overwhelming majority of the languages that have been studied up to recent

[2] In this section, I use the term Specifier as a rather informal label to refer to the position that is identified as such in X'-bar theory. Assuming Bare Phrase Structure theory, the notion of Specifiers is not a primitive.
[3] Of course, if linearization is a PF phenomenon, it does not even make sense asking how the material that will move is linearized in its base position.

years, *wh* phrases are found either *in situ* (namely, they don't move) or at the left edge of the sentence. Cases of languages in which *wh* phrases systematically occur at the right edge of the sentence are extremely rare (cf. Haider [1997]; Petronio – Lillo Martin [1997] for discussion on this). However this picture, which might have been adequate until a decade ago, is seriously challenged by recent syntactic investigation of sign languages. In fact, in sign languages *wh* elements can and in some cases *must* systematically move to the right edge of the sentence. I refer to Cecchetto – Geraci – Zucchi (2007) for data presentation. They focus on Italian Sign Language (LIS) but also mention other sign languages (American Sign Language, Indo-Pakistani Sign Language, Israeli Sign Language, Sign Language of the Netherlands a.o.) and some typological work that shows that phrases naturally sit in the right periphery (Zeshan [2004]).

In order to explain why *wh* elements can move to the right edge of the sentence in sign languages, Cecchetto – Geraci – Zucchi (2007) elaborate on an idea by Ackema – Neeleman (2002).

Ackema and Neeleman propose that rightward movement is more limited than leftward movement because of processing difficulties that arise with the former. Here is a brief summary of their account. In sentence processing, the parser must build a syntactic representation on the basis of a linear (left-to-right) input (the incoming sentence). In order to reduce the pressure on short-term memory during the analysis of an input string, the parser closes off certain units of already parsed structure and removes them from the short-term memory by treating them as atoms with no discernible internal syntactic structure. As for processing of antecedent-trace dependencies, it is standard to asuume that the parser starts searching for a position to insert a gap only when an item is identified as a moved element[4].

Given what we said, if the trace follows its antecedent (as in leftward movement), in principle it can be inserted in any phrase following the antecedent, since for any such phrase the insertion of the trace may occur while the parser is building up the phrase. However, a trace can precede its antecedent (as in rightward movement) only if trace and antecedent are in the same unit; otherwise, in order to insert a gap, a "closed" (fully analyzed) unit would have to be reopened for inspection when the parser hits the antecedent, something which is either not possible or very costly. In a nutshell, the basic intuition underlying Ackema and Neeleman's approach is that rightward movement requires backward localization of a trace and backward localization is possible only if

[4] Ackema and Neeleman motivate this assumption by observing that in a sentence like (i) there is nothing which suggests that it contains a gap. Thus, positing a trace in the string "you think Mary fixed the bike" in (ii) depends on the presence of the fronted interrogative element *how*.
(i) You think Mary fixed the bike
(ii) How do you think Mary fixed the bike?

trace and antecedent are close enough to belong to the same unit[5]. This approach can derive Ross's (1967) Right Roof Constraint, according to which rightward movement (and only rightward movement) is clause-bound.

In order to extend to LIS (and other sign languages) this type of approach, Cecchetto – Geraci – Zucchi (2007) stress some important features of the grammar of sign languages. In LIS multiple levels of subordination are rarer than in spoken languages: for example, a sentence of a spoken language that presents multiple subordinate clauses is commonly translated in LIS by using a multisentential discourse. Furthermore, when subordinate clauses occur in LIS, they are never center-embedded, but occur in peripheral position in the sentence (cf. Cecchetto – Geraci – Zucchi [2006]). Cecchetto – Geraci – Zucchi (2007) suggested that these properties are related to a fundamental difference between signs and words with respect to short-term memory: the span, namely the number of items that a subject can repeat in the exact order they were presented to her, is significantly lower for signs than it is for words. Indeed, this difference between signs and words is confirmed for LIS and Italian by an experimental study reported by Geraci – Gozzi – Papagno – Cecchetto (2007), where it is shown that, other things being equal, short-term memory for LIS signs is significantly more limited than short-term memory for Italian words. This suggests that avoidance of multiple levels of embedding and of center-embedded structures are due to short term memory limitations: the grammatical options chosen by the grammar of LIS are those that put a light burden on short term memory. The same need may also explain why many functions words of spoken languages (articles, auxiliaries, prepositions, etc.) are not expressed in LIS, and in other sign languages, by manual signs, but by face expressions, spatial orientations and body postures. Even some content words like adverbs are not expressed in LIS by manual signs, but by modifications of

[5] The Right Roof Constraint can be illustrated by Heavy NP Shift in English and by Right Dislocation in Italian. That Heavy NP Shift is clause bound is shown by the contrast between (i) and (ii) (cf. ROCHEMONT – CULICOVER [1990]). (i) is acceptable since the backward localization of the trace is very local, given that the rightward moved NP *the entire book of Revelation* does not escape the embedded clause (*carefully* modifies the verb *study*). In (ii) the backward localization of the trace is not local, given that *the entire book of Revelation* escapes the embedded clause (*carefully* modifies the verb *want*).
(i) John wanted to [study *t* carefully [the entire book of Revelation]]
(ii) *[John wanted to study *t* dearly [the entire book of Revelation]]
A similar pattern emerges by looking at right dislocation in Italian (cf. CECCHETTO [1999]). In (iii) *a Gianni* is right dislocated to the edge of the matrix clause. The backward localization of the trace should take place at long distance (the matrix clause *mi sembra strano* occurs between *a Gianni* and its trace) and this results in ungrammaticality. (iii) forms a minimal pair with (iv), which is acceptable even if *leftward* movement takes place long distance (*a Gianni* is dislocated to the periphery of the matrix clause). (iii) is acceptable since the antecedent *a Gianni* precedes its trace and no backward localization is necessary.
(iii) A Gianni, mi sembra strano che abbia prestato un libro *t*
To Gianni (it) to-me seems strange that (he) has borrowed a book
(iv) *Che abbia prestato un libro *t*, mi sembra strano, a Gianni
That (he) has borrowed a book (it) to-me seems strange to Gianni

the sign of the verb. These features of sign languages may help to cope with the fact that, for visuo-spatial languages, a reduced number of items can be kept in the active buffer.

Whether or not this hypothesis is correct, these features, that LIS shares with many other sign languages, induce Cecchetto – Geraci – Zucchi (2007) to propose an explanation for why rightward movement of *wh* phrase is permissible in LIS. Due to the reduced level of embeddings, to the total lack of center embedding and to the fact that a LIS sentence is typically short (it contains a limited number of signs), backward localization of a trace in LIS is bound to be a local phenomenon: the syntactic parser has no need to reopen already closed units when it tries to insert a trace after the moved element has been identified. So, rightward *wh* movement is allowed, more or less as Heavy NP shift is allowed.

Cecchetto – Geraci – Zucchi's (2007) approach predicts that, while in spoken languages the fact that Spec,CP is linearized to the left allows extraction of *wh*-phrases out of subordinate clauses since this does not involve backward localization of the trace, in LIS, where Spec,CP is to the right, extraction of a *wh*-phrase out of a subordinate clause should be barred, because, when the parser reaches the right peripheral *wh*-phrase, the unit that contains its trace has already been closed off and the dependency should not be processed. This prediction seems to be correct for LIS, as they discuss.

In this section, we have seen that, if one assumes that the leftward placement of specifiers is a strategy to minimize the burden for the parser, there is a natural explanation for "exceptions", namely cases in which movement creates a configuration in which the moved category is linearized *to the right*. The existence of these "exceptions" suggests that rules of linearization are a relatively peripheral phenomenon and are not encoded as such in core syntax, in accordance with Chomsky's proposal that linearization is a PF phenomenon.

3. The linearization of Head and Complement

The Head Parameter is intended to account for harmonic languages in which the linear order between a head and its complement is consistent through the various categories. However, 'mixed' or 'disharmonic' word-order systems seem to be widespread, so, especially since Kayne's (1994) work which rejects the existence of the Head Parameter, the doubt has emerged that by assuming the Head Parameter one can explain the special cases (the harmonic systems) but neglects explaining the most common ones (disharmonic systems).

For this reason, I begin this section by briefly going over the typological results summarized in Dryer's (1992), which are directly relevant for the issue of the distribution of harmonic/disharmonic structures. Dryer, based on a large sample of 625 languages, tries to identify

those pairs of grammatical elements whose order correlates with the order of verb and object. More specifically, he proposes the following definition:

(1) If a pair of elements X and Y is such that X tends to precede Y significantly more often in VO languages than in OV languages, then <X,Y> is a Correlation Pair and X is a Verb Patterner and Y is an Object Patterner with respect to this pair.

In most cases, Dryer's investigation identifies Correlation Pairs that had already been recognized as such in the Greenbergian tradition (cf. Greenberg [1966]). For example, the following elements form Correlation Pairs (the list is not complete): adpositions and NPs (*to* + *John*), noun and genitive (*father* + *of John*), verb and (subcategorized) adpositional phrases (*slept* + *on the floor*), copula and predicate (*be*+ *a teacher*), tense/aspect auxiliary verb and VP (*has* + *eaten dinner*), complementizer (or adverbial subordinator) and embedded clause (*that* + *John is sick* or *because* + *John has left*). Interestingly, in all these cases it is plausible (or at least conceivable) to imagine that the Object Patterner in the Correlation Pair is a complement of the Verb Patterner. In at least two other cases a Correlation Pair is not obviously formed by head and complement, namely the pairs formed by noun and relative clause (*the man* + *who I saw*) and by adjective and standard (*taller* + *than John*) in comparative constructions. Probably the most substantial difference between Dryer's results and the former typological literature is introduced by the fact that Dryer can show that noun and adjective do *not* form a Correlation Pair.

It seems fair to summarize Dryer's research by saying that it confirms the existence of a systematic correlation between the order of verb and object and the order of other pairs of elements. In this sense, the original motivation for the Head Parameter seems to persist. I take this to be a drawback of LCA-based theories. Still, disharmonic languages (or disharmonic configurations within harmonic languages) do exist and their existence cannot be simply neglected. An interesting perspective on disharmonic structures is investigating how disharmonic they can be, in order to see if there is some inner regularity in what seems a completely free word order system. For example, Holmberg (2000) by looking at some representative disharmonic languages observes that, although many alternative word orders are possible, it is *not* the case that any word order is allowed. In particular, there is a specific configuration that is not acceptable. This configuration is excluded by Final-over-Final Constraint in (2):

(2) *Final-over-Final Constraint*
If α is a head-initial phrase and β is a phrase immediately dominating α, then β must be head initial.

If α is a head-final phrase, and β is a phrase immediately dominating α then β can be head initial or head-final (from Holmberg [2000]; cf. also Julien [2000]).

The Final-over-Final Constraint is schematized in (3). Out of the four logically possible arrangements, only three are attested. The illicit arrangement is (3a), in which YP is the head-initial complement of the head-final projection XP (specifiers are omitted for simplicity):

(3) a. * $[_{XP} [_{YP}$ Y ZP] X]
b. √ $[_{XP} [_{YP}$ ZP Y] X]
c. √ $[_{XP}$ X $[_{YP}$ Y ZP]]
d. √ $[_{XP}$ X $[_{YP}$ ZP Y]]

Biberauer – Holmberg – Roberts (2007) mention, among other facts, the following evidence for Final-over-Final Constraint:

i) Old and modern Germanic varieties exhibit a mix of head-initial and head-final orders in VP and IP, with all permutations of Aux, V and Object attested except one (Den Besten – Edmondson [1983]; Hroarsdottir [2000]). The unattested order is the one that violates Final-over-Final Constraint.

ii) Sentence-final complementizers are not found in VO languages (Hawkins [1990]). This is so since both $[_{CP} [_{IP} [_{VP}$ V Obj] INFL] COMP] and $[_{CP} [_{IP}$ INFL $[_{VP}$ V Obj]] COMP] violate Final-over-Final Constraint.

iii) In the nominal domain, Finnish has mixed projections too: it has both pre- and postpositions and N-Complement as well as Complement-N order. All permutations are found except $[_{PP} [_{NP}$ N-Complement] P], the Final-over-Final Constraint violating order.

Biberauer – Holmberg – Roberts (2007) also mention some cases that do not conform to the Final-over-Final Constraint. These cases belong to the type illustrated by the German sentences in (4) and (5).

(4) Johann hat $[_{VP} [_{DP}$ den Mann] gesehen]
Johann has the man seen
(5) Johann ist $[_{VP} [_{PP}$ nach Berlin] gefahren]
Johann is to Berlin driven

Sentences in (4) and (5) violate Final-over-Final Constraint, since a head-final phrase (the VP) immediately dominates a head-initial phrase (the DP in 4 and the PP in 5). Biberauer – Holmberg – Roberts (2007) argue that (4) and (5) are representative of the known cases that do not

obey the Final-over-Final Constraint, since in all the Final-over-Final Constraint violations there is a categorical distinction between the head-final phrase β and the head-initial phrase α that is immediately dominated by β, with one of the two being part of the extended projection of the verb and the other being part of the extended projection of the noun (in the sense of Grimshaw [1991]). Biberauer – Holmberg – Roberts (2007) offer an explanation for the both the cases explained by Final-over-Final Constraint and the 'exceptions' like (4) and (5), but I will not try to summarize their accoun, since it assumes LCA, which is inherently incompatible with the Head Parameter. However, I will assume that their description is basically correct, including the observation that cases like (4) and (5) are different because α and β belong to different extended projections.

In the next section I will argue that one might make sense of Final-over-Final Constraint effects by extending to these cases the perspective that I have assumed in section 2 for the linearization of specifiers.

4. Final-over-Final Constraint Reduced to a Condition on Backward Localization

The gist of Ackema – Neeleman's (2002) explanation for why rightward movement is more constrained than leftward movement is that the parser has a preference for structures that do not require backward localization of a trace. Therefore, the linearization algorithm, which plausibly applies at the PF interface, favours leftward linearization of Specifiers. However, if this reasoning is correct, the condition on linearization is flexible enough, since cases of rightward movement are allowed under certain conditions, namely when rightward movement is local enough (clause bound).

Suppose that something like the Head Parameter exists and applies at the PF interface, when linearization of hierarchical structure takes place. A way to express a weak version of the Head Parameter might be saying that the linearization algorithm has a bias to favour a consistent linear order across categories. Of course, an important question is why the linearization algorithm should have such a bias. However, I cannot really deal with this question here[6]. What seems clear is that this bias can also be overcome, since disharmonic languages exist.

Crucially, the only configuration that does not seem to be attested is structurally similar to a case of backward localization, which we know to be problematic from the rightward movement case. Let me elaborate on this point. The Final-over-Final Constraint applies within the boundary of the clausal domain (the skeleton of which is the C-INFL-V heads) and within the boundary of the

[6] NESPOR – GUASTI – CHRISTOPHE (1996) claim that the Head Parameter is fixed based on phonological clues. It might be that the child develops a bias for harmonic structures due to the need to maintain a parallelism between the phonological structures and the syntactic one.

nominal domain[7]. It seems reasonable to assume that the heads inside these two domains stay in a special relation, with the c-commanding head selecting the c-commanded one. So, COMPL stays in a special relation with an inflectional node (it is standardly assumed that the tense/untensed character of the clause is reflected at the COMP level, for example languages have different complementizer for finite and non finite clauses) and INFL stays in special relation with the verb. Similarly, a preposition necessarily selects for an NP (or DP). The special character of the relationship between COMP, INFL and V is confirmed by the fact that in many languages these heads can be linked by the occurrence of head movement. Defining the exact character of this special relationship is another important issue that I cannot go into in this squib.

Be that as it may, the proposal that I would like to advance is that the parser might use this special relationship as a clue when the incoming sentence is processed. Then, suppose that the presence of COMP alerts the parsing of the occurrence of the INFL node in the immediate structure and the presence of the INFL node alerts the parsing of the occurrence of the verb (similarly in the nominal domain). This, together with the assumption that the parser closes off certain units of already parsed structure and removes them from the short-term memory, might be the basis to explain Final-over-Final Constraint effects. Let me explain this by making a representative example. Imagine that the parser has to link INFL to the Head that selects it (say COMP). If COMP precedes INFL, the parser "knows" that the incoming string will contain INFL. So, the link between the two Heads can apply at a longer distance. However, the situation is different if the Head that is selected (INFL) precedes the selecting Head (COMP). This configuration, which is the one excluded by Final-over-Final Constraint, is schematically illustrated in (6):

(6) * [$_{CP}$ [$_{IP}$ INFL VP] COMP]

I conjecture that what is wrong with (6) is that by the time the selecting Head COMP is met the portion of the structure containing INFL might have already been closed off. This is not unlikely if one considers that in principle the VP might contain a rich internal structure. In all the licit cases, either the VP does not intervene between COMP and INFL, as in (7) and (8), or the VP intervenes but COMP comes first, as in (9), so the parser is pre-alerted that INFL will arrive.

(7) √ [$_{CP}$ [$_{IP}$ VP INFL] COMP]

(8) √ [$_{CP}$ COMP [$_{IP}$ INFL VP]]

(9) √ [$_{CP}$ COMP [$_{IP}$ VP INFL]]

[7] For simplicity, I will assume a simplified clausal structures with only three heads, V, INFL and COMP. A more careful examination should consider a finer structure in which these heads are decomposed in the *v*P shell, in the IP area (cf. CINQUE [1999]) and in the CP area (cf. RIZZI [1997]). I will leave this examination as a future work.

Note that, for this explanation to work, it must be the case that there is an asymmetry between COMP and INFL, namely the occurrence of COMP guarantees that an INFL node will be present somewhere in the incoming sentence, while the presence of INFL does not warrant the presence of COMP (in absence of this asymmetry, the configurations in 6 and 9 should have the same status, since VP intervenes between COMP and INFL in both of them). This hypothesis seems to be plausible, given the existence of clausal structures that lack the CP area (one obvious case is Exceptional Case Marking configurations like *I believe John to be smart*).

In the lowest section of the clause, a similar asymmetry should hold between the INFL node and the verb, with the former guaranteeing the presence of the latter, but not vice versa. Cases of VP not selected by INFL seem to exist as well. Candidate structures include absolute small clauses like (10) and reduced relative clauses like (11) in languages like Italian:

(10) Mangiata la mela, Gianni morì
Eaten the apple, Gianni died
"Having eaten apple, Gianni died"
(11) La mela mangiata da Gianni era avvelenata
The apple eaten by Gianni was poisoned

So, I would like to propose that the parser postpones closing off a structure only if it is sure that the structure is incomplete. The structural analogy with the rightward movement case should be apparent by now. In both configurations there is an asymmetric dependency in which one link of the dependency can alert the parser of the presence of the other link of the dependency. The configuration in which the "alerting link" comes first is not problematic. However, if the "alerting link" comes after, the configuration is acceptable only if the dependency is very local, because, otherwise, the relevant portion of the structure might have been closed-off to early for the dependency to be processed.

Note that 'exceptions' to Final-over-Final Constraint like sentences in (4) and (5) do not really come as a surprise adopting this perspective. It seems reasonable to assume that the relationship between COMP and INFL and between INFL and V is different from the relationship between V and its arguments. For example, there are verbs that do not take an internal argument, but there is no COMPL which does not select for INFL and no INFL that does not select for a verb. Furthermore, it seems intuitive that in (4) and (5) the V head selects for the entire category DP or PP and does so for semantic reasons, instead of entering into a special relationship with the head D or P. This is consistent with the fact that crosslinguistically there is no systematic head movement of P (or D) to V.

In this section I have proposed that the parser does not like structures that require backward localization of one link of the relevant dependency (the link that must be backward localized is the trace in case of a movement dependency and the selected Head in the cases just considered). Backward localization is allowed only if it a local process. This is the basis to explain both the restricted occurrence of rightward movement and the existence of Final-over-Final Constraint effects.

5. Conclusion

Although the issue of linearization of syntactic structures has been extensively investigated, we are still far from a comprehensive account. In this squib I have tentatively proposed that two limitations on word order (the restrictiveness of rightward movement and the Final-over-Final Constraint) can both have the same source, namely the fact that the parser tolerates configurations that require backward localization of a certain element only if the process of backward localization does not need to re-open fully analyzed categories which can count as atoms as far as on line syntactic processing is concerned.

Carlo Cecchetto
Università degli Studi di Milano-Bicocca
Dipartimento di Psicologia
Piazza dell'Ateneo Nuovo
I – 20126 Milano
carlo.cecchetto@unimib.it

References

Ackema, P., Neeleman, A. (2002) Effects of Short-Term Storage in Processing Rightward Movement. In Nooteboom, S. et al. (eds.) *Storage and Computation in the Language Faculty*. Dordrecht. Kluwer. 219-56.

Den Besten, H., Edmondson, J. (1983) The Verbal Complex in Continental West Germanic. In Abraham, W. (ed.) *On the Formal Syntax of the Westgermania*. Amsterdam. Benjamins. 155-216.

Biberauer, T., Holmberg, A., Roberts, I. (2007) *Structure and linearization in disharmonic word orders*. Talk given at WCCFL 26.

Cecchetto, C. (1999) A Comparative Analysis of Left and Right Dislocation in Romance. In *Studia Linguistica*. 53/1. 40-67.

Cecchetto, C., Geraci, C., Zucchi, S. (2006) Strategies of Relativization in Italian Sign Language. In *Natural Language and Linguist Theory*. 24. 945-75.

Cecchetto C., Geraci, C., Zucchi, S. (2007) *Another way to mark syntactic dependencies. The case for right peripheral specifiers in sign languages*. Manuscript. University of Milan and University of Milan-Bicocca.

Chomsky, N. (1995) *The Minimalist Program*. Cambridge, Mass. MIT Press.

Cinque, G. (1999) *Adverbs and Functional Heads. A Cross-linguistic Perspective*. New York-Oxford. Oxford University Press.

Dryer, M. (1992) The Greenbergian Word Order Correlations. In *Language*. 68. 81-138.

Geraci, C., Gozzi, M., Papagno, C., Cecchetto, C. (2007) How grammar can cope with limited short-term memory. In *Cognition*. doi:10.1016/j.cognition.2007.04.014.

Greenberg, J.H. (ed.) (1966) *Universals of Language*. Cambridge, Mass. MIT Press.

Grimshaw, J. (1991) *Extended projection*. Unpublished paper. Brandeis, University.

Haider, H. (1997) Typological Implications of a Directionality Constraint on Projections. In Alexiadou, A., Hall, T.A. (eds.) *Studies on Universal Grammar and Typological Variation*. Amsterdam. Benjamins. 17-33 (= *Linguistics Today*. 13).

Hawkins, J. (1990) A parsing theory of word-order universals. In *Linguistic Inquiry*. 21. 223-61.

Holmberg, A. (2000) Deriving OV order in Finnish. In Svenonius, P. (ed.) *The derivation of VO and OV*. Amsterdam. Benjamins. 123-52.

Hroarsdottir, Th. (2000) *Word order change in Icelandic: from OV to VO*. Amsterdam. Benjamins.

Julien, M. (2000) *Syntactic heads and word formation: a study of verbal inflection*. Oxford. OUP.

Kayne, R. (1994) *The Antisymmetry of the Syntax*. Cambridge, Mass. MIT Press.

Nespor, M., Guasti, M.-T., Christophe, A. (1996) Selecting word order: the Rhythmic Activation Principle. In Kleinhenz, U. (ed.) *Interfaces in Phonology*. Berlin. Akademie Verlag. 1-26.

Petronio, K., Lillo-Martin, D. (1997) *Wh*-Movement and the Position of Spec-CP: Evidence from American Sign Language. In *Language*. 73/1. 18-57.

Rizzi, L. (1997) The Fine Structure of the Left Periphery. In Haegeman, L. (ed.) *Elements of Grammar*. Dordrecht. Kluwer. 281-337.

Rochemont, M., Culicover, P. (1990). *English Focus Constructions and the Theory of Grammar*. Cambridge. Cambridge University Press.

Ross, J.R. (1967) *Constraints on Variables in Syntax*. MIT. Ph. Doctoral dissertation.

Zeshan, U. (2004) Interrogative Constructions in Sign Languages – Cross-linguistic perspectives. In *Language*. 80/1. 7-39.