

# Constraint on Merge: The Roots of the Lexical/Functional Divide

Ludovico Franco  
University of Venice, Italy  
Email: franco.ludovico@gmail.com

**Abstract**—This paper addresses the following questions: is (external) merge, the binary operation that combines two elements into a constituent in every variant of the Minimalist Program (Chomsky, 1993, 1995 and related works), an unconstrained operation? If so, what avoid generating ill-formed structures? I will argue here for a simple functional / lexical constraint on Merge, assuming a possible principled binary opposition for the items which enter the syntactic derivation. I will basically follow Kayne (2009), who assumes that the class of nouns (or *L-roots*) is the only open (lexical) class in grammar, updating the intuitions of Hale and Keyser (1993). This proposal leads to interesting structural and typological consequences.

**Index Terms**—merge, minimalist program, lexicon, biolinguistics, morpho-syntax

## I. INTRODUCTION

This work aims at investigating the properties of Merge, the operation that builds syntactic structures in the Minimalist program (Chomsky 1995, and related works). I will argue here that Merge, assumed to be the easiest, the first and, arguably, the only, step by which syntactic derivations take place is not a *free* step, once given a Numeration (see, for an alternative *constraining* hypothesis, Di Sciullo and Isac, 2008). I will hypothesize a simple functional / lexical constraint on (external) Merge, assuming a possible principled dichotomy / binary opposition (weakly relying on classic works in other subfields of grammar such as Jakobson and Halle, 1956) for all the items which enter the syntactic derivation.

The main inspirational works for the present proposal are (i) a recent paper by Richard Kayne (2009), which updates the intuitions of Hale and Keyser (1993) and assumes that the class of nouns is the only open (*lexical*) class in grammar; (ii) some recent Cartographic proposals (see Cinque, 2005; Cinque, 2010a; and for introductory purposes, Cinque and Rizzi, 2010); (iii) more broadly, those paradigms which assume that the elements within syntax and within morphology enter into the same kind of constituent structures (e.g. can be sketched via binary branching trees), such as Distributed Morphology (Halle and Marantz, 1993), the unifying paradigm of Manzini and Savoia (2006) or Nanosyntax (Starke, 2009; Caha, 2009).

Intuitively, approaching the architecture of the human faculty of language from its basis, a principle involved in a *directional/constrained* Merge must be necessarily simple and economic, and can be introduced as follow. Let's assume that our Lexicon stores only (underspecified, to some extent) lexical *roots* (let's call them *nouns* in an unorthodox fashion; see Barner and Bale, 2002 for a psycholinguistic anchorage): as for Merge, which combines items in syntax, a lexical root can target only a features' sets or functional item(s) and not *viceversa*: functional items can Merge to other functional items, leading to functional ordered sequences. Hence, *root Merge root* is banned (leading, at most, to exocentricity in compounds, see Progovac, 2009), while sequences of grammatical words, which crucially build syntax (see the *fseq* of Nanosyntactic paradigm), are allowed. Traditionally, functional items are those syntactic heads which are not defined in terms of [+Noun ; +Verb], marking grammatical or relational features, rather than picking out a class of objects (Abney, 1987).

Basically, if we assume that (only) functional items build syntax, we must say that lexical roots are *inert* in grammar: they do not project. That is the proposal in Kayne (2009), in which it is also argued that all verbs are functional *light verbs* (see also Franco *et al.* 2010, for clinical evidence from an anomic patient affected by Primary Progressive Aphasia, a degenerative syndrome marked by progressive deterioration of language functions and relative preservation of other cognitive domain). This is a basic fact, in order to implement a constrained Merge model. Thus, here, I will assume that only nouns, as lexical primitives, are *inert*. Since Jespersen (1965) the term "light verb" is a label used to refer to a class of verbs which is supposed to be semantically empty, thus lacking enough thematic strength to independently act as predicates.

Notice that many languages fails to incorporate the noun into a *light verb*, so that most 'verbal meanings' are expressed as V+N periphrases (see Amberber, Baker and Harley, 2010), probably demonstrating that most transitive and inergative verbs are not primitive but result from the incorporation of a noun into a limited class of light/general purpose verbs (e.g. 'do', 'give', 'take', 'put', 'hit'), and even the class of these primitive verbs may turn out to be closed and relatively small (Folli, Harley and Karimi, 2005; Cinque and Rizzi, 2010). One of such *light verb* languages is Persian, which is a crucial case also because it has been convincingly argued (Ghomeshi 1997) that Persian nouns

(remember: the only open class, according to our proposal) are non-projecting items. Notice that the not uncontroversial claim of a unique set of (nominal) roots, which go beyond traditional categories, finds many typological confirms also for other (alleged) open classes of items (Baker, 2003): in many languages it has been observed, for instance, that adjectives or adverbs can constitute a closed, often quite small class of elements (Dixon, 2004).

This brief paper is structured as follows. After an overall view on Merge, I will sketch my proposal from a structural viewpoint. Then, I will try to briefly show that a constrained Merge can easily explain in an economic way grammaticalization patterns (*clines*, see Heine and Kuteva, 2002; Von Stechow, 1995; Longobardi, 2004) and typological rarities. Notice that I am aware that, “despite the numerous attempts to uncover the principle(s) governing grammatical relations/orders, the concomitant demand of empirical accuracy with respect to actual languages, has reduced virtually all of the correlations proposed to mere statistical tendencies” (Cinque, 2010b, p.1) which are – however – interesting by themselves. Notice also that here, I will not address more technical details about Merge - given the simple constraint proposed - concerning e.g. the locality (topology) of the relations, and the interpretability of features (Collins, 2002; Bowers, 2010; Franco, 2008).

Finally, notice that in a related work under development (Franco 2011b), it will be addressed a problematic case study, *exocentric compounds*, melting up theoretical claims about their status and empirical evidence from clinical linguistics (see, for introductory purposes, Semenza and Mondini, 2006).

## II. MERGE

Syntactic structures in the Minimalist Program (Chomsky 1995) are built bottom-up by the operation Merge, which has two fundamental properties: (a) it is a *binary operation*, which combines two items into a constituent, and (b) it is *recursive*, so that the its output may subsequently be submitted to another Merge with other elements yielding a further syntactic unit. In the Minimalist program all the elements that are subjected to Merge are drawn from a set (namely, a list) called the *Numeration*. A Numeration is defined as a set of minimal pairs, a lexical item and an index, who signals the number of instances of the item along the derivation. Whenever items are selected from the Numeration in order to enter the syntactic derivation, their indices reduce by one. The derivation ends when every index scales down to zero.

Hence, syntax seems to be very simple, economic. This iterative operation of Merge is the sole responsible for building up syntactic structures (from bottom to top): the first input to the initial application of Merge consists of terminal items, and the last output of the final application of Merge expresses a hierarchical structure. The triggering step is illustrated below in (1):

$$(1) \text{ MERGE } (\alpha, \beta) \Rightarrow \{\alpha, \beta\} \Rightarrow \{\alpha, \{\alpha, \beta\}\}$$

Thus, Merge combines the input objects into a *set*. As an immediate consequence, it forms a hierarchy: the original input objects are *directly included* in the output object (de Vries, 2009).

## III. CONSTRAINT

My proposal is the following: Merge is principled and it is sensitive to *categories* in a broad sense. Let's hypothesize that the only valuable distinction in grammar is between functional and lexical categories. Merge operates as a filter and bans all its applications that impair a syntactic derivation. Thus, if we label  $\checkmark$  lexical items and  $\alpha$  functional items, Merge works as follow:

- (2) a. Merge ( $\alpha, \checkmark$ )  $\Rightarrow \{\alpha, \checkmark\} \Rightarrow \{\alpha, \{\alpha, \checkmark\}\} \leftrightarrow$  output Ok
- b. Merge ( $\alpha, \alpha$ )  $\Rightarrow \{\alpha, \alpha\} \Rightarrow \{\alpha, \{\alpha, \alpha\}\}$  or  $\{\text{fseq}\} \leftrightarrow$  output Ok
- c. Merge ( $\alpha, \checkmark$ )  $\Rightarrow \{\alpha, \checkmark\} \Rightarrow \{\checkmark, \{\alpha, \checkmark\}\} \leftrightarrow$  bad output
- d. Merge ( $\checkmark, \checkmark$ )  $\Rightarrow \{\checkmark, \checkmark\} \Rightarrow \{\checkmark, \{\checkmark, \checkmark\}\} \leftrightarrow$  bad output

The combinations represented above get three crucial points: i) coherently with Kayne (2009) recent updates of Antisymmetry (for which Cinque, 2005; 2010 has given very strong typological evidence) lexical (*denotational*) items are not able to project; ii) merge between  $\checkmark$  items does not allow for syntactic derivations and probably if a combination of that kind is possible, it pertains to morphology, as with the example of the above-mentioned *exocentric* compounds; iii) ordered sequences of functional items (projecting *heads*) build grammar, which is coherent with promising paradigms within the contemporary theoretical linguistics, such as Cartography or Nanosyntax.

In order to work, my proposal has to make a not uncontroversial assumption: verbs do not exist, or in a less dramatic form, all verbs are *light verbs*. This belief has originated from the seminal works of Hale and Keyser (1993, 2002), it has been radically retrieved in the work of Kayne (2009) and it has been effectively interpreted in Cinque and Rizzi (2010) as a putative principle of cartographic researches.

Many structural questions can arise from the present hypothesis. However, further technical details will be omitted here because not relevant. For those who are interested, you can refer to Franco (2011a).

## IV. SOME NOTES ON GRAMMATICALIZATION

The proposal outlined above - leaving aside here structural questions involved within the generative framework (e.g. *asymmetry*, *dominance/hierarchy*, *selection*, etc.) and handled, as said, elsewhere - aims at providing possible unifying

answers related to various phenomena in subfields such as: i) language evolution and diachronical explanation; ii) statistical tendency in linguistic typology, with particular regards to implicational universals; iii) language acquisition and language loss; iv) language contact (e.g. how can a language absorb loan words into a *native* Lexicon?).

In this section I will briefly introduce the phenomenon of grammaticalization, trying to show that it is probably the most important factor for language evolution.

Grammaticalization is the historical development of function morphemes from lexical morphemes. One of the crucial properties of functional morphemes is that, in any natural language, their *inventory* is limited, as opposed to the virtually infinite lexicon of content items (Abney, 1987; Von Stechow 1995).

A list of some important kinds of functional morphemes, taken from Kay von Stechow (1995, p.176), may give an idea of what we are dealing with:

(3) *Noun Class - Gender - Number - Determiner - Quantifier - Case - Verb Class -Voice - Aspect - Tense - Modality - Negation - Complementizer - Conjunction - 'Wh'-Elements - Degree Words - Comparative - Superlative*

The notion of functional categories was introduced into the generative paradigm by the works of Fukui and Speas (1986) and Fukui (1986). Given the set in (3) it seems that “functional categories are what grammar is all about” (Von Stechow, 1995, p. 176). This intuition has been framed as a principle of natural languages: grammatically relevant cross-linguistic differences are confined to the properties of functional morphemes, but there must be an underlying regular pattern. A constrained version of Merge as given in (2) is, possibly, the more economical layout, if we consider Merge as the basic operation of language.

Grammaticalization seems to be a unidirectional process and the counterexamples cited, for instance, in Norde (2009) are not unambiguous. Heine and Kuteva (2002) wrote that “grammaticalization is a unidirectional process, that is, it leads from less grammatical to more grammatical forms and constructions (p.4)”. This process, following Hopper and Traugott, 1993) may be interpreted as in (4):

(4) *content word > grammatical word > clitic > inflectional affix*

In our view this process (*cline*) may be the phylogenetic proof of a syntactic “Big Bang”, triggered by functional morphemes and it is essential for the study of language evolution. We think that, having in mind what is relevant for syntax and the hypothesis in (2), we may restate (4) as follows:

(5)  $\sqrt{ } > \alpha > \emptyset$

Guglielmo Cinque (2010b) has pointed out that the notion of ideal language(s) may be restated in terms of amount of movement of constituents (no movement vs. the most possible movement). From an asymmetric Merge perspective the ideal language would be the either (i) a radically *analytic* language, such as for example Riau Indonesia (described in Gil, 2004) or (ii) a radically *polysynthetic* language, such as Mohawk (described in Baker, 2001). In other words: (i) every functional morpheme presupposed vs. (ii) no presupposition at all (as concern for the inventory of functional categories in a given language, due to a grammaticalization process such as the one depicted in (5)). Since Pollock’s (1989) classical work, it has been postulated an abstract set of functional (un-spelled/presupposed) projections (e.g. by the means of the existence of certain systematic word order differences among languages).

Without entering into technical details, notice that our model fits the evidence of antisymmetric theory quite well, because it assumes that a lexical specifier and a lexical complement cannot be spelled out/linearized/parsed when adjacent. Hence, adjacency plays a role in grammar, contrary to common evidence, in the sense that it triggers a syntactic derivation, implying the necessity to avoid  $\{\sqrt{ }, \sqrt{ }\}$  (for a partially analogous proposal see the dynamic model of Moro 2000; 2008).

## V. TRIGGERING EMPIRICAL ISSUES

This idea, utterly speculative at first sight, can actually address interesting typological phenomena, otherwise quite unexplainable. One example could be provided by the morphosyntactic behavior, described in Heath (2007), of some languages of the Songhay family (a West-African language family of Mali), in which are present *bidirectional* case markers that specify both that the NP to the left is a subject and that the NP to the right is an object, without being bracketed uniquely with either. Another example, among many others, could be given by *coverbs*, as described for instance for the Australian language Jaminjung (Schultze-Berndt 2000; 2001). In Jaminjung coverbs form complex predicates with inflecting verbs, but can also act as main predicates in a clause subordinated by means of a case marker. However, coverbs constitute a distinct part of speech from nominals, which, unlike coverbs, take the full set of case markers and may occur in a noun phrase together with determiners or attributive adjectival nominals (Schultze-Berndt, 2000). Probably, they are, in our model, the best approximation of a pure  $\sqrt{ }$  (root). Furthermore, it is interesting to observe that also in Indo-European languages, such as Persian, are present underspecified “mismatching” words (Karimi-Doostan, 2011), that seem to function as roots, when isolated / unmerged with a functional complements.

A constrained version of Merge could also explain, from my viewpoint, universal tendencies in the morpho-syntax of language, from a typological viewpoint, investigating e.g. some Greenberg Universals (Greenberg, 1963), or make a basis for interesting recent investigation within the generative framework on *hierarchical universals* (see the *Final over Final Constraint*, assumed by Biberauer *et al.* 2010). Refer to Franco (2011a) if interested in the underlying argumentations.

## VI. CONCLUSION

In this brief paper I have proposed that Merge, the operation that builds syntax in the Minimalist program is principled and it is sensitive to *categories* in a broad sense. I have proposed that the only valuable distinction (*binary opposition*) in grammar is between functional and lexical categories. Merge operates as a filter and bans all its applications that *crash* a syntactic derivation. The only possible *array* is among functional categories, creating a functional sequence (*fseq*). This is coherent with challenging recent paradigms in theoretical syntax such as Cartography and Nanosyntax.

## REFERENCES

- [1] Abney, S. (1987). The English noun phrase in its sentential aspect, Ph.D. dissertation, MIT.
- [2] Amberber, M., Baker, B. & M. Harvey. (2010). Complex predicates: cross-linguistic perspectives on event structure. Cambridge: Cambridge University Press.
- [3] Baker, M. (2001). The atoms of language. New York: Basic Books.
- [4] Barner D. & A. Bale. (2002). No nouns, no verbs: Psycholinguistic arguments in favor of lexical underspecification. *Lingua*, 112, 771-791.
- [5] Biberauer, T., A. Holmberg & I. Roberts. (2010). A syntactic universal and its consequences. Ms, Universities of Cambridge and Newcastle.
- [6] Bowers J. (2010). Arguments as relations. Cambridge, MA: MIT Press.
- [7] Caha, P. (2009). The Nanosyntax of Case. Ph.D. dissertation, University of Tromsø.
- [8] Chomsky, N. (1986). Barriers. Cambridge, MA: MIT Press.
- [9] Chomsky, N. (1995). The Minimalist Program. Cambridge, MA: MIT Press.
- [10] Cinque, G. (2005). Deriving Greenberg's universal 20 and its exceptions, *Linguistic Inquiry*, 36, 315-332.
- [11] Cinque, G. (2010a). The syntax of adjectives. A comparative study. Cambridge, MA: MIT Press.
- [12] Cinque, G. (2010b). Word Order Typology. A Change of Perspective, ms. Università di Venezia.
- [13] Cinque, G. & L. Rizzi. (2010). The cartography of syntactic structures. In B. Heine & H. Narrog (eds.), *Oxford Handbook of linguistic analysis* (pp. 51-65), Oxford: Oxford University press, 51-65.
- [14] Collins, C. 2002. Eliminating labels, in S.D. Epstein & T.D. Seely (eds.) *Derivation and explanation in the minimalist program*, Oxford: Blackwell, 42-64.
- [15] Di Sciullo, A. M. & D. Isac. (2008). The asymmetry of merge. *Biolinguistics* 2, 260-290.
- [16] Dixon, R. M. W. (2004). Adjective classes in typological perspective. In R. M. W. Dixon & A. Y. Aikhenvald (eds.), *Adjective classes*. Oxford: Oxford University Press, 1-45.
- [17] Fintel, K. von. (1995). The formal semantics of grammaticalization. In *Proceedings of NELS 25: Workshop on Language Change*, 175-189.
- [18] Folli, R., Harley H. & S. Karimi. (2005). Determinants of event structure in Persian complex predicates. *Lingua* 115,1365-1401.
- [19] Franco, L. (2008). Graph Theory and Universal Grammar, Ph.D. dissertation, Università di Firenze.
- [20] Franco, L. (2011a). The strict asymmetry of Merge. ms. Università Ca' Foscari, Venezia.
- [21] Franco, L. (2011b). Exocentric compounds, syntax and asymmetric Merge ms. Università Ca' Foscari, Venezia.
- [22] Franco L., Zampieri, E., Garzon, M., Meneghello, F., Cardinaletti, A. & C. Semenza. (2010). Noun-Verb Distinction as a Consequence of Antisymmetry: Evidence from Primary Progressive Aphasia. *Procedia - Social and Behavioral Sciences*, 6, 45-46.
- [23] Fukui, N. (1986). A theory of category projection and its application. Ph.D dissertation, MIT.
- [24] Fukui, N. & Speas, P. (1986). A theory of category projection and its applications, *MIT Working Papers in Linguistics*, 8, 128-172.
- [25] Ghomeshi, J. (1997). Non-projecting nouns and the ezafe construction in Persian. *Natural Language and Linguistic Theory* 15, 729-788.
- [26] Gil, D. (2004). Riau Indonesian sama: Explorations in macrofunctionality. In M. Haspelmath, (ed.) *Coordinating Constructions*. Amsterdam: John Benjamins, 371-424.
- [27] Greenberg, J. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In J. Greenberg (ed.), *Universals of language*. Cambridge, MA: MIT Press, 73-113.
- [28] Jespersen, O. (1965). A Modern English Grammar on Historical Principle. London: George & Unwin.
- [29] Hale K. & S. J. Keyser. (1993). On argument structure and the lexical expression of grammatical relations. In K. Hale & S. J. Keyser (eds.), *The view from Building 20*, Cambridge, MA: MIT Press, 53-109.
- [30] Hale K. & S. J. Keyser. (2002). Prolegomenon to a theory of argument structure. Cambridge, MA: MIT Press.
- [31] Halle, M., & Marantz, A. (1993). Distributed morphology and the pieces of inflection In K. Hale & S. J. Keyser (eds.), *The view from Building 20*, Cambridge, MA: MIT Press, 111-176.
- [32] Heat, J. (2007). Bidirectional Case-marking and Linear Adjacency. *Natural Language and Linguistic Theory* 25, 83-101.
- [33] Heine, B. & T. Kuteva. (2002). World Lexicon of Grammaticalization Cambridge: Cambridge University Press
- [34] Hopper, P. J. & E. Traugott. (2003). Grammaticalization. Cambridge: Cambridge University Press.
- [35] Jakobson, R. & M. Halle. (1956). Fundamentals of language. The Hague: Mouton.
- [36] Karimi-Doostan, Gh. (2011). Lexical categories in Persian, *Lingua* 121, 207-220.
- [37] Kayne, R. (2009). Antisymmetry and the Lexicon, *Linguistic Variation Yearbook 2008*, 1-32.
- [38] Longobardi, G. (2004). Formal Syntax, Diachronic Minimalism, and Etymology: The History of French Chez. *Linguistic Inquiry* 32, 275-302.
- [39] Manzini, M.R. & L.M. Savoia. (2007). A unification of morphology and syntax. Studies in Romance and Albanian dialects. London: Routledge.

- [40] Moro, A. (2000). *Dynamic Antisymmetry*. Cambridge, MA: MIT Press.
- [41] Moro, A. (2008). *The Boundaries of Babel. The Brain and the Enigma of Impossible Languages*. Cambridge, MA: MIT Press.
- [42] Norde, M. (2009). *Degrammaticalization*. Oxford: Oxford University press.
- [43] Progovac, L. (2009). Layering of grammar: Vestiges of proto-syntax in present-day languages. In G. Sampson, D. Gil & P. Trudgill (eds.), *Language Complexity as an Evolving Variable*, Oxford: Oxford University Press, 203–212.
- [44] Schultze-Berndt, E. (2000). Simple and complex verbs in Jaminjung: A study of event categorization in an Australian language. Ph.D. dissertation, Radboud University, MPI Series in Psycholinguistics.
- [45] Schultze-Berndt, E. (2001). Ideophone-like characteristics of uninflecting predicates in Jaminjung (Australia). In F. K. E Voeltz & C. Kilian-Hatz (eds.) *Ideophones*. Amsterdam: John Benjamins, 355-373.
- [46] Semenza, C & S. Mondini, (2006). Neuropsychology of compound words. In G. Libben & G. Jarema (eds.), *The representation and processing of compound words*. Oxford: Oxford University press, 71-95
- [47] Starke, M. (2009). Nanosyntax: A short primer to a new approach to language, *Nordlyd*, 36, 2-6.
- [48] Vries, M. de. (2009). On Multidominance and Linearization. *Biolinguistics* 3, 344–403.

**Ludovico Franco** received his MA degree in Linguistics from the University of Siena, Italy in 2005. He obtained a PhD in Theoretical Linguistics from the University of Florence in 2008. He is currently an ESF PhD fellow at the University Ca' Foscari of Venice, performing researches in the field of neurolinguistics, with particular regard to experimental and theoretical morpho-syntax.