A Generative Grammar for Modal Syntagms

Francesco Galofaro

A Generative Grammar for Modal Syntagms

Francesco Galofaro

Università Politecnico di Milano, Italy

semioatp@gmail.com

Abstract. The contribution presents a generative grammar for Greimas' semiotic square, designed to represent modal syntax. In particular: (1) the grammar will generate both modal syntagms and their structural descriptions; (2) the articulation of the immanent narrative structure will be clarified; (3) structural features which are not manifested on the linear surface will become easier to examine. This model contributes to computational narratology theory, aiming at merging post-structuralist narratology with formal language theory.

Keywords: Computational Narratology · Generative Grammar · Modal Structures · Semiotic Square · Automaton

1 State-of-the-Art

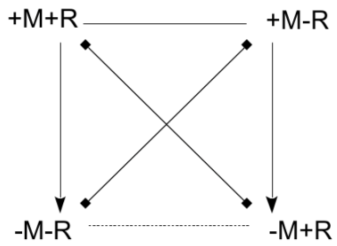
Greimas and Courtés [13] highlighted how formalization could be useful for proving logical coherence to the structural framing of semiotics, and for comparing different theories applied to the same cognitive object. Generative grammar and its success have a clear influence on Greimas' project. He developed his structural semantics by utilizing computational linguistics [11], formulating the semiotic square to describe the fundamental syntax of meaning [14], which should then generate the more complex structural levels. Following the same line of research, this paper aims to develop a generative grammar that formalizes fundamental syntax. In particular, we are interested in generating modal operators which feature narrative enunciates. In this manner we intend to study the immanent articulation of modality, and describe it in terms of standard algorithmic procedures which can be applied by researchers – cf. "automaton" and "algorithm" in [13].

Greimas' modal semiotic square[1] [12] is displayed in table 1:

[1] Petitot tried to demonstrate the inconsistency of the semiotic square by using Boolean logic (cf. 3.3.3 in [20]), yet committed a trivial error. He wrote that 1+1 = 0 whereas, given the idempotent laws, in Boolean algebra 1 + 1 = 1. For an outline of the semiotic debate on the consistency of the semiotic square, see [6], [9].

adfa, p. 1, 2011.

Table 1. The semiotic square.

Relations:	Square
+M+R – +M-R:= antinomy +M+R – -M+R := contradiction +M+R – -M-R := presupposition +M-R – -M+R := presupposition -M-R – -M+R := neutralization	

The semiotic square summarizes the respective relations between four kinds of simple semiotic oppositions: antonymy (a versus b); presupposition (if a then not b and/or vice-versa); contradiction (a versus not -a); and neutralization (not a nor b). This abstract, fundamental syntax articulates the modalities. The table presents the relationships between the various possible modal syntagms, composed by a modal operator (M) and an enunciate (R), which can express an action (to do) or a state (to be). Modal syntagms are combination such as “Wanting to do”; “Not being able to be”; “Not having to do” …

In the framework of a structural semantics, modalities are considered to be general values that are immanent to different semiotic manifestations, such as a literary text, a movie, as well as a technical object [7]. For example, the “Being able to do” modal syntagm can feature the hero of a movie or a real employee, when either of them find something that allows them to reach their goal (a magical potion; a new software).

Two distinct fields of research converge in Greimas’ model: modal logics [2], [21], and structural linguistics. Greimas and Kalinowski worked together on the link between them [17], until the second scholar decided that the two disciplinary souls could not inhabit the same body [18]. Seen through deontic logic, for example, if “I have to do something” then “I can do it”, “ought implies can” and therefore modal operators are inter-defined. In semiotics, on the contrary, the two modalities are largely independent, given the modal conflicts we find in texts. Kalinowski, instead, rejects this liberalization of the relationships between modalities, in spite of the fact that there are many logical systems in which the operators are not inter-defined[2].

---

[2] For example, in the intuitionistic framework, modal operators are usually independent [8]. An example of intuitionistic deontic logical framework has been proposed in relation to the Talmud [1]. Furthermore, since each system in modal logics represents a specific textual domain that assigns a specific non-ambiguous meaning: for example, Kripke’s structures for temporal logic have been converted into Moore's machines [23].

A typology of different possible modalities is presented in table 2:

Table 2. Modal typology. Each symbol represents a generative grammar category (upper case) or a terminal modality (lower case).

Modalities (S)	*Virtualising (V)*	*Actualising (A)*	*Realising (R)*
Exotactical *(X)*	Having to (h)	Being able to (a)	Doing (d)
Endotactical[3] *(N)*	Wanting to (w)	Knowing to (k)	Being (b)

The kind of modal syntagms we would like to produce are composed of all of the possibilities generated by the semiotic square in table (1), when M corresponds to virtualising or actualising modalities, and R to realising ones. For example: (+a+d); (-w+b); (+h-d) are well-formed syntagms because they can be generated by selecting a corner of the semiotic square and by substituting the values seen in table 2 for the categories M and R. We would also like to generate a structural description of these syntagms. There are numerous possibilities. If we want to accurately represent the typology in table 2, we require an unrestricted grammar. Otherwise, by making minor changes to the categories used to classify the modalities, we can design simpler grammars. This second option finds support in Hjelmslev's empirical principle, which privileges coherent, adequate, and simple theoretical choices [15].

2 An Unrestricted Grammar

In order to represent Greimas' typology (table 2), we will first consider an unrestricted grammar G1, that is able to generate a modal syntagm, and is represented by a combination of two lower case letters each of which is preceded by a positive or negative sign – see table 3. The grammar consists of two alphabets Z and T, variables (upper case letters) and terminals (lower case letters), and an initial variable $I \in Z$. It also contains a set of rules of substitution $\alpha \rightarrow \beta$ where α is a string of variables, and β is a string of both variables and terminal symbols.

A derivation starts from the initial symbol S and applies one of the rules of substitution. The language generated by the grammar is a string consisting only of terminals that can be derived from I with a chain of substitutions.

In order to provide an example of the way in which the grammar works, we derive the modal syntagm:

(a): -w+d ("not wanting to do")

[3] A modality is exotactical when it links enunciates with different subjects, endotactical when it attains a single subject. cf. "Modality" in [13].

Table 3. The unrestricted grammar G_1. Each symbol in Z represents a category within the grammar structure– cf. Tables 1 and 2. We have added the category C to represent the endo/exotactical category, and the category G (+/-) to represent the affirmative/negative opposition. The subset I of Z includes the starting symbol S, whereas T is the set of terminal symbols.

Symbolic repertoire	Rules	
Z:={S, M, R, C, G, V, A, X, N}	1 – S →	GCMGCR
I:={S}	2 – G →	+
T:= {h, w, a, k, d, b, +, -}	3 – G →	-
	4 – C →	X
	5 – C →	N
	6 – M →	V
	7 – M →	A
	8 – XV →	h
	9 – NV →	w
	10 – XA →	a
	11 – NA →	k
	12 – XR →	d
	13 – NR →	b

We can also design the tree of substitutions, which reveals the hierarchical relationships between the terminals (see Table 4). The tree shows how each symbol of the alphabet Z represents a category in the meta-language. For example, w is an endotactical virtualising modality, whereas d is an exotactical realising modality. This grammar is an unrestricted one, therefore some of the substitution rules are contextual. For example, according to rule 8, one can substitute h in V iff it is preceded by an X.

3 A Context-Free Grammar

We can reduce the complexity of this grammar if we observe how the endo/exotactical category is implicitly represented by the two separate rules which allows to substitute each V/A/R modality with two and only two terminal symbols, without changes in the generated language. The result is the context-free grammar G2, shown in table 5. Table 6 shows the derivation of (a).

Table 4. The derivation of (a). This figure was generated with JFLAP (www.jflap.org).

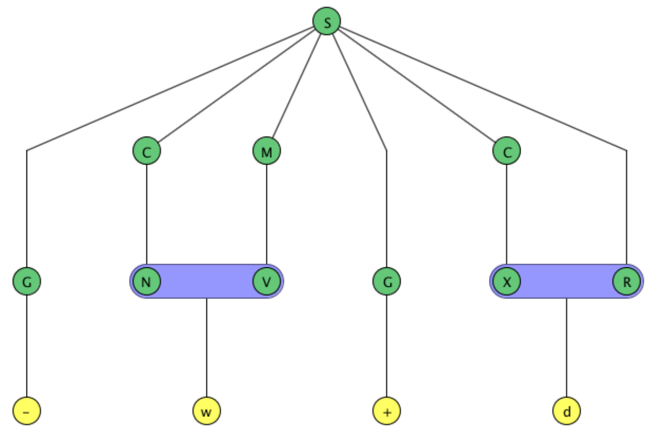
Rule	Steps of the Derivation	Tree
	S	
1	GCMGCR	
2	-CMGCR	
3	-CM+CR	
5	-NM+CR	
4	-NM+XR	
6	-NV+XR	
9	-w+XR	
12	-w+d	

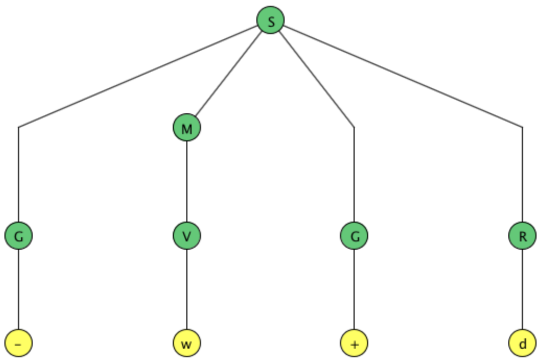
Table 5. The context-free grammar G_2.

Symbolic repertoire	Rules	
Z:={S, M, R, G, V, A}	1 – S →	GMGR
I:={S}	2 – G →	+
T:= {h, w, a, k, d, b, +, -}	3 – G →	-
	4 – M →	V
	5 – M →	A
	6 – V →	h
	7 – V →	w
	8 – A →	a
	9 – A →	k
	10 – R →	d
	11 – R →	b

G_2, is remarkably simpler than G_1 not only because it requires less rules and symbols, but also because it is context-free, it does not use contextual rules.

Table 6. The derivation of (a) in G_2.

Rule	Steps of the Tree Derivation
	S
1	GMGR
2	-MGR
3	-M+R
4	-V+R
7	-w+R
10	-w+d



4 A Right Linear Regular Grammar

If we drastically change the philosophy applied to the representation we can further reduce the complexity of the grammar (see G3 in table 7 and the corresponding derivation of (a) in table 8). This time we need only two categories (S, R) that respectively correspond to the modal enunciate and the doing/state enunciate. Both are preceded by a positive or negative symbol, and each category can be substituted with a terminal symbol.

Regular grammars generate regular languages: according to automata theory, a finite-state automaton can determine if a particular expression belongs to the grammar [16] – see table 9. The automaton is composed of three states. The initial one, q0, corresponds to the S category; the second one, q1, corresponds to the R category; and q2 is the final state. The symbols of the string (a) represent the transitions between states. As the reader can see: when the automaton reads the (+, -) symbols it remains in the same state; when it reads the (k, a, h, w) symbols it changes its internal states to q1; and finally, if it reads (d,b) it reaches the final state. Then, the automaton accepts the string as belonging to the grammar. If this does not happen, the string is rejected. Following Chomsky, the grammar can be considered a model of modal competence, whereas the automaton is a model of modal performance [5].

Table 7. The right-linear regular grammar G_3.

Symbolic repertoire	Rules	
Z:={S, R}	1 – S →	+S
I:={S}	2 – S →	-S
T:= {h, w, a, k, d, b, +, -}	3 – S →	wR
	4 – S →	hR
	5 – S →	aR
	6 – S →	kR
	7 – R →	+R
	8 – R →	-R
	9 – R →	b
	10 – R →	d

Table 8. The derivation of (a) in G_3.

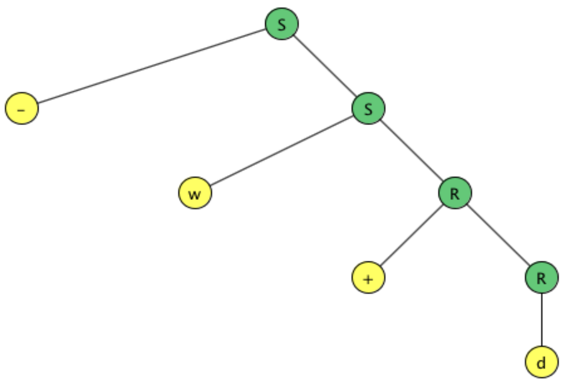
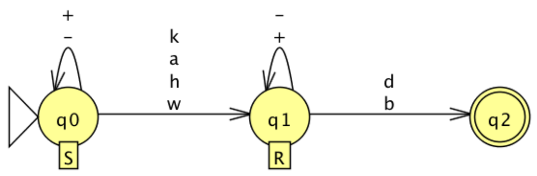
Rule	Steps of the Derivation	Tree
	S	
2	-S	
3	-wR	
7	-w+R	
10	-w+d	

Table 9. A finite-state automaton that can establish if a string of symbols belongs to G_2. The figure was generated with JFLAP (www.jflap.org).



5 Conclusions

What is the "true" grammar, among the three we proposed? There was a similar debate in the sixties concerning syntax grammar. Chomsky [24] proposed extra-linguistic criteria to facilitate the choice between different models, such as the "psychological plausibility". In my opinion, this only led to a circular reasoning basing cognitive linguistics on cognitive psychology and vice-versa [25]. A grammar is just one of the possible solutions to the problem of how to construct a formal demonstration in semiotic theory, showing an immanent structural articulation of meaning. As is shown by G2, some semantic features are incorporated in the syntax. This means that the more complex grammars explicit simpler semantic features of each modality, h = +V+X, "Having to do" is "Virtualising and Exotactical". In Greimas' structuralist perspective, "Semantic" means simply relational and functional (a given tract such as the "virtualising" tract cannot exist positively, i.e. independently from the presence of "actualising" and "realising" tracts). In G3, some semantic features of modality are lost: the automaton in table 9 is just a simple form of the semiotic square. As Marsciani wrote, even formal languages have categories that represent their semantics [9, pp. 24-27]. Thus, the identification of Hjelmslev's symbolic systems [15] with "asemantic" languages seems questionable.

A tree such as the one in table 8 at least reveals a structural feature which was unclear in the other grammars: the ipotactical relation between the (do/be) enunciate and the modal enunciate. Furthermore, given that modal syntagms are generated by a right-linear regular grammar, they can be considered regular languages, a well-known class of formal languages. This is not a trivial result due to the properties shared by regular languages: commutativity, associativity, distributivity, idempotency … [16]. Thanks to these properties it becomes possible to describe the complex chains of modal syntagms we find in semiotic processes as modal devices which "merge" smaller languages generated by simple grammars such as G3 [22, p. 70]. We observe another interesting structural property of regular languages, they are neither recursive nor self-embedded. Now this is puzzling because narrative programs are both recursive and self-embedded [10]. They share these properties with language syntax, as represented in classical generative linguistics [5]. Therefore, further important research questions concern the origins of recursiveness, given the fact that this property is not present in the simpler, deep fundamental structures of meaning described by Greimas, but appear only at a certain degree of structural complexity.

References

1. Abraham, M., Gabbay, DM, Schild, U. "Obligations and prohibitions in Talmudic deontic logic", Artificial Intelligence and Law, Springer, 2011
2. Béziau, J.Y. New light on the Square of Oppositions and its Nameless Corner, in "Logical Studies", vol.11., 2003
3. Chomsky, N. "Formal Properties of Grammars", in Duncan Luce, R., Bush, R., and Galanter, R. (eds.) Handbook of Mathematical Psychology, II, pp. 323-418, John Wiley and sons, New York, 1963
4. Compagno, D. "Due quadrati semiotici per due tipi di relazioni semantiche: contrarietà e indipendenza", in Proceedings of the XVI national congress of the SFL, CUEC, Cagliari, pp. 75-85, 2010
5. Deni, M. Oggetti in azione, Franco Angeli, Milano, 2002.
6. Fischer-Servi, G. "Axiomatizations for some Intuitionistic Modal Logics", Rend. Sem. Mat Polit. de Torino, 42, 179-194, 1984
7. Galofaro, F. "Dal quadrato semiotico alle controversie bioetiche. Una discussione formale?" in Proceedings of the XVI national congress of the SFL, CUEC, Cagliari, pp. 86-99, 2010
8. Galofaro, F. "Formalizing narrative structures. Glossematics, Generativity, and transformational rules", Signata, 4, Liège, 2013
9. Greimas, A.J. "Esquisse d'une morphologie du français en vue de sa description mécanographique", in Linguistics, Volume 4, Issue 22, pp. 34–59, 1966
10. Greimas, A.J. "Pour une théorie des modalités" in Langages, 43; vol.10, pp- 90-107, 1976
11. Greimas, A.J. and Courtés, J. Sémiotique. Dictionnaire raisonné de la théorie du langage, Paris, Seuil, english trans. Semiotics and language: an analytical dictionary, Indiana University Press, Bloomington, 1983
12. Greimas, A.J. and Rastier, F. "The Interaction of the semiotic constraints", in Greimas, A.J., Du sens, Seuil, Paris, 1970, english trans. On meaning, Frank Collins and Paul Perron (eds), Minneapolis: University of Minnesota Press, 1987
13. Hjelmslev, L. Omkring sprogteoriens grundlasggelse, Munskgaard, København, 1943; english trans. Prolegomena to a Theory of Language, F.J. Whitfield (ed.), University of Winsconsin Press, Madison (Wis.),1961
14. Hopcroft J.E., Motwani, R., Ullman, J.D. Introduction to Automata Theory, Languages, and Computation, Pearson Paravia Bruno Mondadori, Milan, 2009
15. Kalinowski, G. "Du mètalangage en logique. Réflexions sur la logique déontique et son rapport avec la logique des normes", Documents de Travail et prépubblications n.48 serie A, Centro Internazionale di Semiotica e di Linguistica, Università di Urbino, Urbino, 1975.
16. Kalinowski, G. Carré sémiotique et carré logique. Les remarques d'un logicien sur le carré sémiotique, in Landowski E. (ed.), Le carré sémiotique, Actes Sémiotiques – Bulletin, 17, EHESS – CNRS, 1981
17. Marsciani, F., Ricerche semiotiche II, Esculapio, Bologna, 2012
18. Petitot, J. Morphogenèse du sens, Presses Universitaires de France, Paris, 1985, english trans. Morphogenesis of Meaning, Peter Lang, Bern, 2003

19. Ricoeur, P. Lectures 2, Seuil, Paris, 1999
20. Greimas, A.J., Fontanille, J. Sémiotique des passions. Des états des choses aux états d'âme, Paris, Seuil, 1991
21. Baumgarten, J.; Tripp, A.; Aziz, A.; Singhal, V.; Andersen, F. "An Abstraction Algorithm for the Verification of Generalized C-Slow Designs", in LNCS 1855, Berlin: Springer, pp. 5-19, 2000
22. Chomsky, N. "Some Empirical Issues in the Theory of Transformational Grammar." In Goals of Linguistic Theory, edited by S. Peters, Prentice-Hall Inc., Englewood Cliffs, NJ, 1972
23. Galofaro, F. "We have to change mind. Neural Plausibility and the Crisis of Cognitive Explanation", in Erica Cosentino (ed.) Un conto aperto. Il neoculturalismo dopo la svolta cognitiva, RIFL, Special number, pp. 101 – 115, 2012