

# On a Noun-Driven Syntactic Paradigm

Lauren M. Stuart, Julia M. Taylor, Victor Raskin

CERIAS, Purdue University, West Lafayette, IN, USA  
{lstuart, jtaylor1, vraskin}@purdue.edu

## Abstract

In addressing a verb bias in syntactic analysis we present the beginnings of a noun-driven analysis paradigm. Such a paradigm may complement and compensate for some weaknesses in the existing verb-driven paradigm (in which the noun is subordinated) as applied to information sources or contexts in which the data is structured in objects and more than in events.

## The Current Paradigm

In analysis of text and speech, the verb is privileged over the noun as the major determiner in syntactic and semantic structure. Noun and prepositional phrases that provide information on the participants, locations, and manners are analyzed as dependent upon the verb in some way. This is made more explicit in dependency grammars, where the verbs are typically the first nodes to be accessed in a traversal of the tree, and access to all of the other nodes proceeds through at least one verb. In constituency-grammar tree structures, the verb is “floated” nearer the root of the tree, as compared to nouns. For example, a noun may appear at the same height as the verb if it forms the subject noun phrase; the addition of adjectives, nouns, and other recursive structures continues to bury the “lowest” noun (which may be the most important, as it may be the “root” of the noun phrase: in “cat food tray”, the object referenced a tray and its preceding nouns further information about what [or which] tray it is). As well, since any prepositional phrases attached to the verb phrase, and all objects of the verb (direct and indirect), are attached under the verb phrase structure, we see the verb floated highest in the tree structure. For computational processing of this tree, the verb is encountered first, and determination of the meaning of the sentence must most efficiently make sense of the first information available; hence, even if the verb is not explicitly centered in semantic processing, its

primacy in the syntactic structures may enforce such centering.

In analyses on the syntax-semantics interface, noun and prepositional phrases (whose objects are also noun phrases) are analyzed as arguments to the major structure, which is the verb. One exception to this analysis is in the case of “to be”: in copula sentences, the verb is no longer as important: it is demoted and its syntactic complement is promoted to the status of predicate. This exception only proves the rule: verbs are predicates, unless they are not strong enough, in which case an adjective or nominal will be given special dispensation to act like a verb in assigning roles (Baker, 2005).

Further, and probably related, the semantic centering of verbs lines natural language reasoning up with first-order predicate logic (Raskin, 1983). In this case, the predicate can be thought of as a function, with its called arguments the syntactic arguments (there is probably no coincidence that they are named similarly) required in its surface form. Much work has been done in fitting natural language and predicate logic together, as in (Luuk, 2009).

## The Mismatch

The verb-centeredness of both syntactic and semantic processing is at odds with information spaces in which objects (represented by nouns) and their static, typical attributes or relationships (represented with adjectives, prepositions, and embedded phrases such as relative clauses), are of main concern.

In large databases, in which relationships are rigidly encoded via table schemas verb-centered analysis may underserve, or at least waste resources in processing, natural-language queries looking for key-value pairs, object-attribute pairs, etc. (Li et al, 2008). Computer programs, whether explicitly object-oriented or not, may be conceived of as interactions and transformations of data objects. Processing these instructions (in order to compare them to natural-language specifications or similar machine-

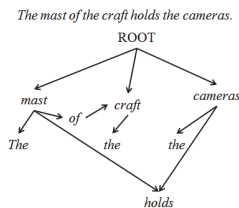
language programs) may, if the processing module is verb-driven, fail to account for object-based polymorphisms, or accommodate them at high cost.

Less-structured information repositories may also exhibit some bias toward the presentation of information in noun phrases, prepositional phrases, relative phrases, and weak verbs. Some of these sources we might assume have a semantic bias towards the definitions of objects, types, and typical relationships. For instance, an introductory article on geography may describe types and typical qualities and positions of geographical features in noun phrases and prepositional phrases, while a geographical history of a particular location may be more concerned with geographical events and exploration (action). Even then, the latter article will contain references to places (objects/nouns). Note through this article that events also manifest syntactically as nouns (e.g., “events”); the pairing of noun/object and verb/event is not so rigid, as many verb-concepts can be expressed as nouns.

## A Proposal

In an attempt to address this inequity – the dominance of verb-centered analysis – we introduce a beginning noun-driven syntactic analysis. The noun-driven syntax representation is only a part of the whole equation: the verb-centered paradigm has years of theory, application, and intuition, which each interact to re-entrench the paradigm. We present a first step in establishing a complementary paradigm for the noun, intending toward an object- rather than event-based semantic analysis.

The representation borrows from dependency grammars: the syntactic tree is mapped in terms of directional binary relationships between words (see Figure 1). In this case, the nouns are (with one exception) dominators of the (dependency-like) relationships. Links to adjectives, determiners, and other nouns in the noun phrase proceed outwards from the root noun, which we already expect from dependency grammar, as they depend upon the noun.



**Figure 1 A noun-driven syntax representation.**

However, the linkage of the nouns with the verbs has a reversed polarity: verbs depend upon the nouns. The number of noun phrase arguments to a verb is thus the number of incoming connections from nouns. Prepositional

phrases proceed outward from the word that they can be best said to modify, from nouns or verbs as appropriate. The noun object of a preposition is the only noun in this representation that does not dominate all of its relationships; the preposition dominates it so that we may establish a path from one root noun to the root noun of the prepositional phrase that further characterizes it.

Preliminary analysis of such a noun-centered dependency syntax representation suggests that it performs better (in terms of complexity of access of certain categories of words) than verb-driven representations derived from the same texts (Stuart et al, 2012).

The central tenet of the paradigm can undoubtedly be bent in many different directions. Future work will include investigation of the complexity and space tradeoffs between noun-driven dependency-like representations and constituency-like representations, after determination of how the latter might work from a noun-driven perspective. Another major question is whether a verb-driven or noun-driven syntax in particular is better suited to use with knowledge systems (in, say, processing large amounts of natural language data for summarization, reasoning, or learning) whose architectures may be more object- or event-centered. Hybrid spaces or systems may also benefit from using both paradigms in tandem.

Rectification of the mismatch – that is, restoring the noun to a place of importance in syntactic analysis – potentially improves natural language interaction between humans and non-human systems whose primary concerns are objects (physical or informational), their attributes, and their interactions.

## References

- Baker, Mark. 2005. *Lexical Categories: Verbs, Nouns, and Adjectives*. Cambridge: Cambridge University Press.
- Li, Zhanjun, Victor Raskin, and Karthik Ramani. "Developing Engineering Ontology for Information Retrieval." *Journal of Computing and Information Science in Engineering* 8(1):011003-13 (2008).
- Luuk, Erkki. 2009. The noun/verb and predicate/argument structures. *Lingua* 119 (2009): 1707–1727.
- Raskin, Victor. 1983. *A Concise History of Linguistic Semantics Pt 1*. Unpublished manuscript. Department of Linguistics, Purdue University, West Lafayette, IN.
- Stuart, Lauren M., Julia M. Taylor, and Victor Raskin. 2012. Towards Noun-Driven Parsing. In *Proceedings of 2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC) 1984-1989*. Seoul, South Korea.