Borut Pretnar

Nouns and verbs: a set-theoretical comparison

Abstract: Nouns and verbs are the key syntactic and semantic elements of human language. A set-theoretical comparison, based on their negation behaviour, is proposed and compared to existing essays on their different roles, mostly based on their syntactic and – to a lesser extent - semantic characteristics. A few elementary applications of set theory are demonstrated as an appropriate means to expose the distinguishing characteristics of nouns and verbs. Based on the preliminary set-theoretical investigation a hypothesis is proposed about the general capability of nouns to define members of sets. The hypothesis provides the base for a tentative definition of some lexical categories appended to nouns.

Keywords: Nouns, verbs, negation, set theory, lexical categories defined

1. Introductory remarks

The roles of nouns and verbs in human language have been investigated and compared since antiquity. Lyons (1968: 10-11, 14) observes that ancient Greeks defined nouns and verbs as constituents of propositions and that verbs and adjectives were put together in the same class. Panini is mentioned as the main grammarian of the Indian tradition where a distinction between verbs and nouns, similar to the Greek one, was made in Sanskrit.
In the past century, in pre-chomskian era, R. H. Robins published a thoughtful essay (Robins 1952). He recommended to extend the existing methodological approaches in order to investigate the (questioned!) universal presence of nouns and verbs across languages (Robins 1952: 297-298). The idea of universality across languages and the attempts to pin down the distinguishing characteristics of nouns and verbs remain challenged to the present time (Haspelmath: 2012). Nevertheless, the universality across languages is accepted in the mainstream linguistics. A contemporary generativist treatment of nouns and verbs is available in (Baker: 2003).

The present article makes use of an unconventional mathematical approach leaving the traditional syntax-based approaches aside. It is also an attempt to demonstrate that some linguistic issues can (and sometimes must) be handled in mathematical terms to be put in an appropriate explanatory model. Mathematics, as an exact science, may help to resolve disputes based on subjective judgement, frequently occurring in humanities. Of course, this is limited to the cases where the disputed problem can be “mapped” as a mathematical model.

2. Negation of nouns and verbs as a characteristic of their different nature

The starting point for the subsequent discussion is the plain observation: only a very limited class of nouns can be negated e. g, *non-smoker, nonconformity, non-truth* while it is completely nonsensical to negate others, e. g. **non-tree or *non-road.* In contrast, it is always, without exception, possible to negate a verb, particularly in its finite form as a predicate.
The mathematical theory of sets and closely related theory of logic can explain the reasons for this state of affairs. Among many textbooks on elementary level the books of I. Stewart (1995: 43-75) or Partee & al. (1993: 3-23; 87-99) may serve as an introduction.

Two key concepts are prerequisite for the subsequent discussion: the so-called *universal set*, i. e. the set containing all the elements relevant to the subject under discussion, and *negation***,** an operation defined in mathematical logic and simply (graphically) demonstrated e. g. by the so called Venn diagrams of sets. Based on these prerequisite notions, the operation of negation can be interpreted as *switching between two subsets whose union is the entire universal set*. (Mathematical term for “switching” is *taking complement or complementation of a subset*.) The following examples make these statements more understandable.

Thus in the first above example non-smokers can be represented as a subset of humans, and humans in turn can be represented as the union of the two subsets smokers and non-smokers**,** comprising all humans. This can be made clear by a concise equation, written in mathematical symbols for sets (curved brackets) and the operator of union (i. e. joining) of sets, the letter U:

[1]: {smokers} ∪ {non-smokers} = {all humans}

It is crucial to keep in mind that speakers are aware of the existence of the relation [1] between the three sets. This awareness makes it actually possible to use and understand this negation in the spoken communication. By analogy, the equations for the two additional examples may be written as

[2]: {conformities} ∪ {nonconformities} = {all results of comparing to the standard}

[3]: {truths} ∪ (non-truths} = {all statements}

Again, most speakers and listeners are aware of the relations [2] and [3].

The validity of the foregoing explanation can be further substantiated by turning the above "non-sensical" examples **non-tree* and **non-road* to sensible expressions by the following equations, bearing in mind that speakers are normally not aware of them:

[4]: {trees} ∪ {non-trees} = {plants}

[5]: {roads} ∪ {non-roads} = {traffic connections}

By creation of an appropriate universal set it is immediately clear that *non-trees* may include e. g. grass, flowers and bushes, and *non-roads* e. g. railroad connections and cartways.

The above reasoning can explain the additional fact that verbs, especially in their finite forms as predicates, can always be negated.
The predicate *by its very presence creates and defines two subsets* of the universal set defined by the noun (or noun phrase), normally the subject of the sentence: one by its affirmative form and the other (complementary) by its negated form. The same is true for adjectives (and actually attributes in general): they are always deniable. Note that some historical grammatical categorizations considered verbs and adjectives as members of a common category.

3. Further application of set formalism to the nature of word classes

The conclusions of the preceding discussion serve as a natural incentive for further questions:

- are nouns a necessary prerequisite for definition of sets?

- and, by broadened question, is *any* noun (or a noun phrase) capable (sufficient) to define the corresponding member(s) of a set?
- and, last not least, is mathematical set theory an appropriate means to establish a definition of noun or verb, standing the test across various languages?

Attempts to apply formal set definitions accompanied with appropriate semantic thought experiments may provide answers.

There are two equivalent ways of specifying a set. It is possible to list the members and enclose the list in curly brackets. The alternative way is the so called *predicative notation* based on the specification of a characteristic property of members (Stewart 1995: 45; Partee & al. 1993: 6). This can be illustrated by an example:

[6] List: {spoon, fork, knife} vs. predicative: {x│x is a hand-held device for eating}

Introductory textbooks of set theory usually start, for didactic reasons, with simple examples of sets. Their members are usually described by countable nouns denoting tangible objects or by proper names. In languages with inflection they are quoted in non-inflected nominative case. However, uncountable and abstract nouns (or both) can represent valid members of sets as well, e. g.:

[7] {sand, butter, straw} ; { x│x is whole number less than 7} ; {freedom, love, optimism}

However, now compare this:

[8] List: {to be, to do, to have} vs. predicative: { x│x is an English auxiliary verb)

The listed verbs somehow feel weird in their role as members of a set. They are quoted as infinitives, i. e. in a non-inflected form. Yet, it is questionable whether there is a valid parallel between non-inflected nominative case and non-inflected verb infinitive. As Widdows (2004: 87-88) points out, (undeclined) nouns and (unconjugated) verbs are not the same sort of things. In addition, the criteria for countability and abstract vs. tangible prove inapplicable. In contrast, the predicative version, defined by the phrase *“English auxiliary verb”* (where the term “verb” is a noun !) does not present the above concerns. Similar disparity between the listed and predicative definition would result with adjectives or prepositions instead of verbs. An in-depth explanation of this disparity exceeds the scope of this article.

Based on the preceding examples, tentative answers can be provided for the three initial questions above.

- The nouns are probably necessary (i. e. unavoidable) to define members of a set.
- The second question, whether any noun is capable (sufficient) could be denied by a counterexample, i. e. a noun (or noun phrase) not capable to define a set member. It appears unlikely that such counterexample could be found.
- If nouns are indeed necessary and sufficient for definition of members of sets this property could be used for their definition:
  Nouns are defined by their ability to define members of sets.

There is some supporting evidence for the above theses:
In cognitive psychology, (Medin et al. 2001: 197-201; 365-368; Solso & al. 2008: 254-259). nouns are hierarchically stored in brain by hypernym-hyponym links, closely resembling sets and subsets. [Of course, it is not assumed that human cognition (unconsciously) treats nouns as members of sets in the exact mathematical sense].
Additionally, the *titles* of books, stories, films, keywords of articles etc. are themselves kind of names of "sets", concisely subsuming their content and/or narrative. Nouns (including nouns modified by determiners, adjectives and prepositions) are far more frequent as *titles* than any other lexical category.
The assumption of sets defined by nouns creates a possibility to derive definitions of other lexical categories. Some examples of *tentative* definitions of adjectives and selected determiners may be:

Adjective is an operator on the set defined by a noun, it selects a *defined* subset.

The notion of *operator*, originally precisely defined in the mathematical context, is a very general concept. Speaking very approximately and loosely (but probably more understandably) an *operator* takes something and converts it to something else. A mathematical example are two numbers (*operands)* taken by the *operator* "*plus" for* *summation (or "times" for multiplication)* and converting them to sum (or product), respectively.

Determiner *some* is an operator on the set defined by a noun, it selects an *undefined* subset.

Definite article is an operator on the set defined by a noun, it selects the *defined or known* element of the set.

Indefinite article is an operator on the set defined by a noun, it selects an *arbitrary or unknown* element of the set.

The set theory is firmly founded on an axiomatic system and therefore provides a good logical base for definitions. However, the creation of an axiomatic system of linguistic definitions, similar to strict axiomatic systems in algebra and geometry, is almost certainly an unattainable goal.

References

Baker, Mark C. (2003): *Lexical Categories – Verbs, Nouns, and Adjectives Cambridge Studies in Linguistics Vol. 102*
Cambridge: Cambridge University Press

Haspelmath, Martin (2012): *How to compare major word-classes across the world's languages.* In Thomas Graf & Denis Paperno, Anna Szabolcsi & Jos Tellings (eds.), *Theories of everything: in honor of Edward Keenan*, 109-130 (UCLA Working Papers in Linguistics 17)
Los Angeles: UCLA

Lyons, John (1968): *Introduction to Theoretical Linguistics*
Cambridge: Cambridge University Press

Medin, Douglas L., Brian H. Ross, Arthur B. Markman (2001): *Cognitive Psychology.* Fort Worth TX: Harcourt College Publishers, 3rd edn.

Partee, Barbara, Alice ter Meulen, Robert E. Wall (1993): *Mathematical Methods in Linguistics – Studies in Linguistics and Philosophy Vol. 30*
Dordrecht: Kluwer Academic Publishers

Robins, Robert Henry (1952): *Noun and Verb in Universal Grammar*
Language 28 289-298

Solso, Robert L., Otto H. MacLin, M. Kimberley MacLin (2008): *Cognitive Psychology*
Boston: Pearson Education Inc. 8th edn.

Stewart, Ian (1995): *Concepts of Modern Mathematics.* Mineola NY: Dover Publications.

Widdows, Dominic (2004): *Geometry and Meaning*
Stanford: Center for the Study of Language and Information

About the author:

Dr.-Ing. Borut Pretnar

ORCID: 0000-0001-6004-8522

Independent (no formal affiliation), long-standing member of Linguistic Circle at Philosophical Faculty of University of Ljubljana, with main interest in mathematical linguistics.
Linguistic Circle is an academic institution with decades long tradition and with »open access« membership policy. Until his recent death it was run under the professional auspices of academician prof, Janez Oresnik.

Author′s educational background is metallurgical engineering and physical chemistry.