THE LITHUANIAN SEMANTIC METALANGUAGE
AND THE POLYSEMY OF KĀŽKAS

FRANCESCO ZAMBLERA
Kauno Vytauto Didžiojo universitetas
f.zamblera@uki.vdu.lt

1. Introduction

Cliff Goddard (2008b, p. 1) defines the Natural Semantic Metalanguage theory (NSM) as

a decompositional system of meaning representation based on empirically established universal semantic primes, that is, simple indefinable concepts which appear to be present as identifiable word-meanings in all languages.

This approach to natural language semantics, which originated in the early seventies with Anna Wierzbicka, is based on reductive paraphrasis and the postulation of a “semantic core” of universal primes which, by hypothesis, is common to all of the world’s languages.

Developed first with English, the theory has been applied to an ever growing number of languages, as typologically and genetically different as Polish (Wierzbicka 2002), Malay (Goddard 2002), Mangaaba-Mbula (Bugenhagen 2002), Amharic (Amberber 2008), Korean (Yoon 2008) and East Cree (Junker 2008). “Applying” the theory to a new language means essentially to find out how to express, in that language, the sixty-odd semantic primes and their basic combinatorial structures. Some semantic primes are, for example, SOMETHING, SOMEONE, GOOD, DO, OTHER (ELSE), while one of the basic syntactic structures of the metalanguage is SOMEONE DID SOMETHING GOOD TO SOMEONE ELSE. If the theory is viable, then every language must have a way to express such combinations unambiguously.

An important assumption of the NSM theory which is being tested by extensive analysis of natural languages is, therefore, that every language has, in some sense, the same subset of basic lexical items with the same combinatorial properties, and that texts which exploit only this subsystem can be transposed without any loss of meaning into any language.

This article presents some preliminary questions towards the definition of the Lithuanian Semantic Metalanguage. In an NSM context, Lithuanian has
been mentioned by Goddard (2001, p. 10), and there is also an interesting paper written at Aarhus University by Ugnė Krasauskaitė (2011).

For most primes, it is rather straightforward to find a Lithuanian equivalent, but there are some tricky issues, one of which this work will examine in detail; it represents an apparent challenge to two of the most important primes (SOMEONE and SOMETHING), namely, the fact that Lithuanian has only one word meaning both, the lexical item kažkas. The issue is discussed in sec. 4 and 5, where, following Goddard (2001), I will propose that kažkas is in fact polysemous and not ambiguous (so that Lithuanian has indeed two different concepts here, which happen to be expressed by the same word). After a short summary of the main points of the NSM theory (sec. 2), which is not intended to be exhaustive, sec. 3 presents the table of the proposed Lithuanian equivalents of the semantic primes, in the usual format. Finally, after a short list of possible further research (sec. 6), sec. 7 presents a Lithuanian Semantic Metalanguage text, the translation of the English NSM explanation of the Biblical concept of God, taken from Wierzbicka (2001, p. 21), a text which relies heavily on the contrast between someone and something.

As customary in NSM studies, semantic primes will be written in capital letters when quoted singularly, while texts written in NSM will be written without punctuation and indented to express semantic embedding.

2. The Natural Semantic Metalanguage

In this section I give a brief overview of the main points of the NSM theory, as it appears in such standard works as Goddard (2008, 2011), Wierzbicka (1996), Goddard and Wierzbicka (1994, 2002a, 2002b).

One of the most important claims made by the theory is that natural language is adequate as a semantic metalanguage: in the NSM theory, the technical language set up in order to talk about meaning is a restricted and properly defined subset of natural language. This subset is thought to be the “nucleus” of a natural language, which contains the undefinable terms, their logical-syntactical combinatorics and the rules for their morphological-allolexical variation. A grammar of this subset of a natural language should constitute a basic grammatical sketch of that language. The NSM theory aims to define the “ground subset” of every natural language, the one containing the undefinable lexical items and their basic combinatorial properties, their basic syntax.

According to Goddard and Wierzbicka (2002c, p. 41), “the fundamentals of language arise from the fundamentals of human thought, which are shared by all people and by all languages”. Moreover, we cannot escape from natural
language(s) to describe meaning: formalized systems used to describe meaning, as for example first-order logic, need to be resolved in terms of natural language in order to be understood and interpreted by human beings. That is, natural languages are their own ultimate semantic metalanguages. In this sense, Goddard (2008b, p. 3), speaks of the “meta-semantic adequacy” of natural languages.

The technical device which NSM exploits to define the meaning of lexical items and expression in a language is that of reductive paraphrasis. In order to really define the meaning of a word or an expression, we must use words which are simpler than the one(s) we started from. To avoid both circularity and reductio ad infinitum, the NSM theory postulates the existence of a basic set of concepts, called semantic primes, whose meaning is undefinable (there are, by hypothesis, no simpler words with which to paraphrase their meaning).

These semantic primes are immediately understandable, without the need for any explanation. They are, moreover, universal, that is, expressible in every language. A particular language can express a prime with a word or an affix or a syntactic construction, but, according to this hypothesis, it must be always possible to express every prime in every language;

The lexical items which express the semantic primes are subject to each language-particular morphological and lexical rules. From a NSM point of view, we speak about allolexy and/or polysemy.

A word which expresses a semantic prime in a language can also have other meanings in that language (polysemy): for example, the Spanish word querer expresses the semantic prime want, but it also means love. Lithuanian has an interesting case of polysemy, in which both senses of the word are primes, namely, kažkas, which is the equivalent of both someone and something. This is not, however, the only case of polysemy among the primes: the Australian language Wambaya shows the same pattern of Lithuanian, with the stem gayini meaning both “someone” and “something” (different gender suffixes provide a clue to disambiguation, cf. Goddard 2001, p. 10). An extreme, and extremely interesting, example of multiple polysemy among the semantic primes is offered by the Australian language Bunuba (Knight 2008).

Morphological variation is treated as a case of allolexy in so far as it is automatically determined by the syntactic context. For example, in the English NSM, a “verbal” prime such as do is represented by various items in allolexical variation (do, does, did, done), and the selection of the correct form depends on the syntactic context (subject and time adverbial).

In a highly inflected language like Lithuanian, therefore, there will be extensive allolexy in the form of nominal cases and verbal tense-mode-persons. This is not a problem, as already shown by Wierzbicka 2002 for Polish, a typologically similar language.
3. The proposed semantic primes of Lithuanian

The following list of Lithuanian equivalents for the semantic primes includes also the newly proposed items FEW (Goddard, p. c., July 2010), “abstract THIS/IT” and BE (SOMETHING) (Goddard and Wierzbicka 2008), which in Lithuanian is expressed by the same lexical item as the “locational” BE (BE SOMEWHERE), a widespread pattern in the languages of the world (cf. Stassen 1997). The list of the proposed primes follows; I have been able to complete my original notes thanks to Krasauskaitė (2011, pp. 5-7).

Substantives
I aš; YOU tu; SOMEONE kažkas/asmuo; SOMETHING kažkas/daiktas/dalykas; PEOPLE žmon/uni0117s; IT tai; BODY kūnas.

Taxonomy and Partonomy
KIND OF rūšis; PART OF dalis.

Mental Predicates
THINK galvoti; KNOW žinoti; WANT nor/uni0117ti; FEEL jausti; SEE matyti; HEAR girdėti.

Speech
SAY sakyti; WORDS žodžiai; TRUE tiesa.

Actions, Events, Movements
DO daryti; HAPPEN nutikti ~ atsitikti; MOVE judėti; TOUCH liesti.

Existence, Possession
THERE IS (čia) yra; HAVE turėti; BE SOMETHING būti (kažkuo); BE SOMEWHERE būti (kažkur).

Life and Death
LIVE gyventi; DIE mirti.

Quantifiers
ONE vienas; TWO du; FEW mažai; MUCH/MANY daug; SOME keletas/keli; ALL visi.

Evaluators and Descriptors
GOOD geras; BAD blogas; BIG didelis/didysis; SMALL mažas.
**Determiners**
THIS (ši)tas, šis; THE SAME tas pats; OTHER kitas.

**Time**
WHEN/TIME kada/kai/laikas; NOW dabar; BEFORE prieš (tai); AFTER po (to); A LONG TIME ilgas laikas; A SHORT TIME trumpas laikas; FOR SOME TIME tam tikras laikas; MOMENT / (IN A) MOMENT akimirka/staiga.

**Space**
WHERE/PLACE kur/vieta; HERE čia; ABOVE virš(uje); BELOW apačioje ~ žemiau; INSIDE viduje; SIDE pusė; NEAR ariti; FAR toli.

**Logical Concepts**
BECAUSE dėl to (todėl, nes); IF jei (jeigu); NOT ne-; MAYBE galbūt; CAN galėti.

**Intensifier, Augmentor**
VERY labai; MORE daugiau;

**Similarity**
LIKE/WAY kaip/būdas.

4. **The polysemy of kažkas**

As has been seen, Lithuanian expresses both primes SOMEONE and SOMETHING by kažkas. Goddard (2001, p. 10), quoting Anna Wierzbicka, addresses the matter as follows:

> Very occasionally, it may appear that the same expression is used to cover both ‘someone’ (who) and ‘something’ (what). This is the case in Lithuanian where the relevant form is kas. But Tatanya Bulygina (Anna Wierzbicka, pc) argues that kas is polysemous: kas1 = ‘someone’, kas2 = ‘something’. Her arguments include the fact that kas1 has genitive1 kieno and genitive2 -ko, whereas kas2 has genitive1 -ko and genitive2 -ko. Kas1 and kas2 also have different agreement patterns; for example, Kas linksmas, tas mielas ‘who(ever) is cheerful is nice’ (adjectives with masculine agreement) does not equal Kas linksmus, tai mielius ‘what(ever) is cheerful is nice’ (adjectives with feminine/neuter agreement).

The main point for polysemy versus ambiguity lies thus in agreement phenomena: one of the basic combinations in NSM is the “attributive relation” (Goddard and Wierzbicka 2002c, p. 44), which allows the primes SOMEONE and
SOMETHING to combine with the evaluators GOOD and BAD and descriptors BIG and SMALL. These four items are adjectives in Lithuanian, agreeing in gender, number and case with the head noun. When they are used with kažkas, they show masculine or neuter gender, according to whether kažkas means SOMEBODY or SOMETHING, respectively. Another possibility is the use of the genitive singular of the adjective, again in the SOMETHING (but not in the SOMEONE) reading of the head kažkas, as in Italian qualcosa di buono or in Yiddish epes guts:

(1) kažkas geras/blogas/didelis/mažas (with the masculine singular nominative) “someone (*something) good/bad/big/small”

(2) kažkas gera/bloga/dideli/maža (with the neuter singular nominative) “something (*someone) good/bad/big/small”

(3) kažkas gero/blogo/didelio/mažo (with the singular genitive) “something (*someone) good/bad/big/small”

The same happens in the combination SOMETHING ELSE and SOMEONE ELSE:

(4) kažkas kitas “someone (*something) else”

(5) kažkas kito/kita “something (*someone) else”

However, agreement will help only when the head kažkas is in the nominative or accusative, which are the only cases for which adjectives show a distinct neuter form. A dative phrase like kažkam geram can mean both “to someone good” and “to something good”.

In some cases, the syntactic function of the noun phrase will determine the meaning of kažkas. For example, as the subject of HAPPEN or the complement of DO, kažkas can only mean SOMETHING:

(6) kažkas atsitiko “something (*someone) happened”

(7) kažkas bloga/*/blogas atsitiko “something (*someone) bad happened”

(8) padarei kažką blogo “you did something (*someone) bad”

As the subject of most other verbal primes, like DO, SAY, and the mental predicates, kažkas can only mean SOMEONE:
(9) kažkas padarė kažką gero "someone (*something) did something good"

In some contexts, however, we will have to resort to the nominal allolexes, that is, thing for something and person for someone. For example, the definition of the concept of God in Wierzbicka 2001 (p. 21, cf. sec. 6) begins with the NSM sentence:

God is someone (not something)

The most straightforward way to render this in Lithuanian is to recur to the nominal allolexes daiktas (thing) and asmuo (person)

(10) Dievas yra asmuo (ne daiktas)

Translating it as *Dievas yra kažkas (ne kažkas) would totally miss the point and sound unintelligible. As for the nominal allolexes, “thing” has two Lithuanian equivalents which can be considered in allolexical variation (see sec. 5), namely, daiktas (roughly, “concrete thing”) and dalykas (roughly, “abstract thing”). For person, the most probable candidate is asmuo. In many context žmogus would also fit; grammatically, it is the singular of žmonės, people. However, žmogus can only refer to human beings; it cannot be used, for example, to refer to God, while asmuo has no such restrictions.

The nominal allolexes are needed independently, because kažkas cannot combine with a quantifier, so, as in English, we must say vienas daiktas/ dalykas “one thing” and vienas asmuo “one person” (with a plural quantifier, we can say both du asmenys “two persons” and du žmonės “two people”).

The combinations this someone, this something, extensively used in NSM explanations, are possible in Lithuanian (both are tas kažkas), though normally a personal pronoun would be used.

As a last observation, kažkas can be reduced to kas after the prime if (in Lithuanian jeigu), just as in Latin aliquis is reduced to quis in the same context: si quis veniat, “if someone should come”, Lithuanian, jeigu kas ateity.

5. The two items daiktas and dalykas

As has been seen, another interesting phenomenon in Lithuanian is the double nominal allolex of the prime something, that is, daiktas (roughly, “concrete thing”) and dalykas (“abstract thing”). From an NSM point of view, the
opposition daiktas/dalykas is similar to the Malay benda/perkara, discussed by Goddard (2002 p. 92-93): both substantives can be seen as manifestations of the prime something, the use of the one or the other depending on the syntactic contexts.

For example, we will find dalykas rather than daiktas as the subject of happen, or the complement of do, or the “topic” of know, think, say:

(11) noriu sužinoti apie vieną dalyką “I want to know about one thing”
(12) atsitiko vienas dalykas “something (one thing) happened”
(13) jis padarė du blogus dalykus “he did two bad things”
(14) jis kitų dalykų nesako “he didn’t say (other things) anything else”
(15) negali daryti tokio dalyko/*daikto “you cannot do something like this”

The concrete allolex will appear, for example, as the subject of move:

(16) šitas daiktas/*dalykas negali judėti “this thing cannot move”

The “daiktas and dalykas matter”, as Dagilytė (2013) shows, is, however, a little more complex. First of all, daiktas does not always refer to “concrete” objects: in the Online Corpus of Contemporary Lithuanian (Dabartinės lietuvių kalbos tekstynas), for example, there are clear cases of abstract meaning, for example, the widespread collocation galimas daiktas, kad “it is possible”, lit. “(it is) a possible thing that”. On the other hand, dalykas can have concrete referents, as in “skaičiuodami cholesterolio kiekį kiaušiniuose ar kituose labai materialiuose dalykose” (Dagilytė 2013, p. 22).

Furthermore, Dagilytė’s analysis of the concordances of the two terms in the Online Corpus of Contemporary Lithuanian and the Senųjų rašų tekstynas shows that the situation was different in earlier stages of Lithuanian. In the Older Literature Corpus, the word dalykas is almost absent, while daiktas is used with both concrete and abstract referents (Dagilytė 2013).

This interesting issue cannot be pursued further here; I will just point out that dalykas is polysemous, meaning also “(school or university) subject” (cf. again Dagilytė 2013, p. 23).
6. Further problems

To develop a full Lithuanian Semantic Metalanguage, there remain a couple of interesting issues. First of all, we are not yet done with *someone* and *something*, because in Lithuanian there are two other lexical items which can be translated as “someone/something”, namely *kai kas* and *kas nors*. The opposition *kažkas/kai kas/kas nors* is similar to the Russian *kto-to/koe-kto/ kto-nibud’* (someone), *cho-to/koe-cho/chto-nibud’* (something). The item *kai kas*, like its Russian counterpart *koe-kto/koe-cho*, implies that the speaker is talking about a concrete person or thing whose identity is known to him/her, but s/he does not want to reveal it. So *kai kas* can be considered as a semantically more complex form than *kažkas*; a possible NSM explanation could contain, among others, the following components:

- **kai kas**
  - someone/something
  - I know who/what this someone/something is
  - I don’t want to say who/what this someone/something is

- **kažkas**
  - žinai, kas yra
  - nenoriu pasakyti, kas yra

As for *kas nors*, the situation is a little more complex. Like its Russian counterparts *kto-nibud’/cho-nibud’*, it implies that the speaker has no actual person/thing in mind; s/he is talking about an hypothetical person/thing, who/which may not exist at all. It contains, in other words, an element of indefiniteness. Perhaps *kas nors* can be explained in terms of *kažkas*, though in some context implying indefiniteness *kas nors* might turn out to be an allolex of *kažkas*. The question deserves further investigation. Gladkova (2008, p. 303) gives *kto-to* and *cho-to*, which are the Russian equivalents of *kažkas*, as representing, respectively, the primes *someone* and *something*, in the Russian Semantic Metalanguage.

Furthermore, there is the question of verbal aspect (the prime *do*, for example, can be expressed both by *daryti* and *padaryti*). Wierzbicka (2002, pp. 105-124) exhaustively discusses the question with reference to Polish, and her solution will probably be applicable to Lithuanian with little or no change, although this is still to be done.
7. The “God” script

Here is the definition of the (Biblical) concept of God as presented in Wierzbicka (2001, p. 21), in the original English NSM and in the corresponding Lithuanian.

I acknowledge that the Lithuanian of the translation does not always sound idiomatic, just as its English counterpart does not, for that matter. However, we must keep in mind that the English NSM is not, properly speaking, naturally spoken English, but a sort of technical language based on English, and the same is true for the Lithuanian one. An interesting analogue to the NSM would be the technical Latin of medieval philosophers and semanticists, such as Buridan and Ockham (cf. Spade 2002, Klima 2009). The NSM can also be compared to a controlled language in the computational linguistics sense. NSM texts do not always sound idiomatic (although they must be comprehensible and grammatically acceptable), because in naturally spoken language, we hardly stick to simple and undefinable terms, but use the full expressive power of the language at our disposal.

In the translation, I have slightly departed from “pure” NSM usage by substituting Jis (lit. He) for the NSM expression THIS SOMEONE, because repeating tas kažkas in almost every sentence would sound particularly cumbersome in Lithuanian, though it would not be ungrammatical. Jis is used with no intended reference to the “masculinity” of God, but only because the Lithuanian word Dievas happens to have masculine gender; in an NSM context, jis can simply be seen as a portmanteau morph for the combination THIS + SOMEONE. As is customary in referring to God, I have capitalized the initial letter of Jis.

Sentence (b) would sound much better without kažkas, that is, “this someone is good” (Jis yra geras); I have finally resolved to maintain the original form in the translation. So far as sentences (f) and (g) are concerned, the most straightforward equivalent of the prime BECAUSE would be dėl to, kad (as in sentence h); I have used the (almost-)synonym nes in order to avoid the “double kad” (viskas yra dėl to, kad Jis nori, kad būtų), which would sound odd in Lithuanian and hinder comprehension. Finally, sentence (h) would sound more natural by substituting priežastis “cause” for dalykas. The item “cause”, however, is semantically complex, and cannot be used in NSM explanations.

I am indebted to Jolanta Kovalevkaitė for many useful suggestions in completing this translation. Any shortcomings are of course my entire responsibility.

I give Wierzbickas’s definition followed by the Lithuanian NSM equivalent.

(a) God is someone (not something)
(b) this someone is someone good
(c) this someone is not someone like people
(d) there isn’t anyone else like this someone
(e) this someone exists always
(f) everything exists because this someone wants it to exist
(g) people exist because this someone wants them to exist
(h) this someone exists because this someone exists, not because of anything else
(i) this someone lives

(a) Dievas yra asmuo (ne daiktas)
(b) Jis yra kažkas geras
(c) Jis nėra tokš kaip žmonės
(d) nėra nieko kaip Jis
(e) Jis yra visada
(f) viskas yra, nes Jis nori, kad būtų
(g) žmonės yra, nes Jis nori, kad būtų
(h) Jis yra dėl to, kad yra, o ne dėl bet kokio kito dalyko
(i) Jis gyvena

References


Lietuvių semantinė meta-kalba ir žodžio „kažkas“ daugiareikšmingumas
F. Zamblera (Kauno Vytauto Didžiojo universitetas)

Šitame straipsnyje pristatomas preliminarus „natūralios semantinės meta-kalbos“ ( angl. Natural Semantic Metalanguage, NSM) teorijos pritaikymas lietuvių kalbai. Pateikiamas pirmųjų semantinių elementų sąrašas ( angl. semantic primes) bei vienas tekstas, išverstas iš anglų natūralios semantinės meta-kalbos. Be to, nagrinėjamas viena iš
pagrindinių problemų NSM teorijos pritaikymui lietuvių kalbai – žodžio kažkas reikšmė. Straipsnyje nurodoma, kad žodis kažkas yra iš tikrųjų daugiareikšmis (angl. polysemous: 
kažkas1 = SOMEONE; kažkas2 = SOMETHING), o ne dviprasmiškas (angl. ambiguous) (SOMEONE ir SOMETHING).