

Ulla Vanhatalo, Heli Tissari & Anna Idström

Revisiting the Universality of Natural Semantic Metalanguage: A View through Finnish¹

Abstract

The Natural Semantic Metalanguage (NSM) is a method of semantic analysis, used for various tasks mainly in the field of linguistic research. A crucial part of the theory is the set of primes, minimal lexical units that are used to explicate words, cultural scripts and other concepts. Identifying the primes in a new language is an opportunity to reinforce and/or revisit the theory. The remarks presented in this paper resulted from the identification process of the Finnish-based NSM primes. The goal of this paper is to direct attention to some fundamental aspects in the Natural Semantic Metalanguage theory, especially to the relation between the universal language-independent NSM concepts and the English-based NSM. A number of remarks are made on the general system of the primes, as the paper points out issues related to e.g. the number, selection and mutual hierarchy of the primes. The economy and logic of certain prime constructions and the argumentation behind allolexy are discussed as well.

1. Introduction

The Natural Semantic Metalanguage (henceforth: **the NSM**) is an approach to linguistic meaning, originated by Anna Wierzbicka in the 1970's, and developed further mainly by Anna Wierzbicka and Cliff Goddard. The NSM is based on three fundamental assumptions: 1) *there exists a natural semantic metalanguage, through which all words in every language can be defined;*² 2) *this metalanguage is based on a set of indefinable atom-like words called primes, and a simple grammar;* 3) *these*

¹ The manuscript was submitted for the first time in March 2012. After three rounds of refereeing, the published version of this paper owes a lot to the most helpful and patient anonymous referees. We want to sincerely thank Prof. Cliff Goddard for very constructive comments on different versions of the manuscript.

² The result of an NSM analysis process is called *explication*.

primes,³ together with their associated grammar, can be identified in every natural language.

There are two main guidelines for identifying primes. Being a prime, a lexical unit must (a) be found in every natural language, and (b) be indefinable via other primes. To put it simply: All words (more precisely, lexical units) in all languages can be divided into two groups: the primes (64, plus some functional synonyms, *allolexes*) and the other words. The words in the latter group can, in principle, be explicated through the primes. As all NSM versions based on different languages are mutually fully translatable, any NSM version can be used for explicating any word in any language.

Although there is nowadays an increasing dialogue between the NSM and other semantic paradigms, there still remain disputed issues. Most importantly, the idea that primes are the core of a universal mental lexicon (along with the idea that most words other than primes are language-specific) has been challenged by many critics (e.g. Murray & Button 1988, see also Wierzbicka's reply 1988; van Driem 2004; Wawrzyniak 2010; Enfield 2002; Geeraerts 2010; McCawley 1983). Not enough attention has been paid to the unavoidable contradiction that the primes have firstly been identified in English, yet they are intended to be (and are considered by their proponents to be) language-independent. Indeed, the relationship between the primes as mental concepts and their manifestation in the world's languages deserves more extensive discussion than has been conducted so far. Moreover, a neutral analytic comparison between the NSM and other methods in the field of semantic analysis is also missing.

In the NSM approach, locating primes in a new language is seen as reinforcing the validity of the general theory of NSM. The primes are now considered to have satisfactory counterparts in more than 30 languages.⁴ Most of the findings in these papers seem to support the general theory of primes; however, some criticism has also been presented (e.g. Goddard & Karlsson 2008). An extensive literature already exists on careful

³ The counterparts of the primes in different languages are called *exponents*. If the exponent of certain prime in a certain language based NSM variant has multiple variants, they are called *allolexes*. For examples, see Table 1.

⁴ E.g. in Amharic (Amberber 2008), East Cree (Junker 2008), French (Peeters 1994), Japanese (Onishi 1994), Korean (Yoon 2008), Lao (Enfield 2002), Mandarin (Chappell 2002), Mangaaba-Mbula (Bugenhagen 2002), Malay (Goddard 2002), Polish (Wierzbicka 2002), Russian (Gladkova 2010, for the latest set, see the NSM home page), Spanish (Travis 2002), and Thai (Diller 1994).

translations of primes into many languages (e.g. Gladkova 2010; Yoon 2006). However the field test of these prime sets, i.e. their application into actual semantic explications remains less well examined. The vast majority of all NSM explications are made and published in the English-based NSM.⁵

In the Finnish project, the newly invented Finnish primes were tested with respect to their capability to explicate the meaning of certain words (Vanhatalo & Tissari, forthcoming).⁶ Many concerns reported in the present paper were exposed only in the translation process of explications (for a note on previous translation problems, see also Peeters 1994: 440). We needed to revisit our initial proposals and make adjustments (for example, in the case of *kind of* vs. *like*), and still some questions remained open.

The aim of the present paper is to revisit some fundamental aspects of the NSM theory based on observations made during the creation of a Finnish version of the NSM. As the remarks are ensuing from the identification work done with the Finnish-based primes, the Finnish primes are presented as a Table in section 2 (in more detail, Vanhatalo & Tissari, forthcoming). We have collected our comments in the following two sections under two closely related main headings: The relation between the universal NSM concepts and the English-based NSM (Section 3), and remarks on the general system of primes (Section 4). Section 5 will discuss the findings.

⁵ A lot of explications have been done in Polish and Russian (mainly Wierzbicka, also Gladkova), and some in French (Peeters), Arabic and Hebrew (Habib). Only part of the work has been discussed and published in English. The more NSM explications made in languages other than English will be analysed in English, the wider audience will be able to discuss, analyse, criticise and improve the work.

⁶ It is to be mentioned here, that the current paper is not the first endeavor to identify the NSM primes in Finnish. In her PhD thesis, Seija Tuovila (2005) published a selected set of NSM primes to be used in explications of Finnish emotion vocabulary. The thesis also included quite an extensive selection of compact NSM explications.

2. The Finnish⁷ based NSM primes

In the course of the identification process of the Finnish exponents, all the basic combinatorial possibilities of primes (from Goddard 2011c) were translated into the Finnish-based NSM (for examples, see Appendix 1.). After this, some NSM explications were translated into the Finnish-based NSM. These included the verb *promise* (Engl.), the discourse particle *well* (Engl.), the noun *a cup* (Engl.), the noun *God* (Engl.), and the interjections *Psst!* (Engl.), *Pst!* (Pol.), *Shh!* (Engl.) and *Sza!* (Pol.). Some examples are presented in Appendix 2. After the identification process, the NSM Finnish version has been successfully used in numerous explications of cultural scripts, cultural key words, states of emotions, social situations, etc. (Vanhatalo forthcoming).

The set of the Finnish primes is presented as currently⁸ seen in “state of the art” NSM research.

Table 1. The set of the Finnish primes, version 1.

English (NSM homepage 2013 ⁹)	Finnish (Vanhatalo & Tissari, forthcoming)
I	MINÄ
YOU	SINÄ
SOMEONE	JOKU~IHMINEN~HÄN
PEOPLE	IHMISET
SOMETHING~THING	JOKIN~ASIA
BODY	RUUMIS~KEHO
KIND	-LAINEN~-LÄINEN

⁷ Finnish, typologically located between fusional and agglutinative language types, is a Uralic language spoken by 5–6 million speakers mainly in Finland. In Finnish vocabulary, many words are created with derivational suffixes, with verbal suffixes in particular being extremely diverse. From the lexical semantic point of view, the Finnish language provides an interesting laboratory, as Finnish is mainly spoken by native speakers in a geographically delimited and linguistically relatively homogeneous area.

⁸ With the prime SOMEONE, however, we have made an exception by using the earlier NSM version (Wierzbicka 1972). For more discussion, see 4.3.

⁹ The current table of NSM primes has been under development for decades. In the very first set (Wierzbicka 1972), as few as 13 primes were presented. Later on, there have been continuous updates (both withdrawals and additions) to the set of primes, the total number being currently (2014) 65. The very last change in the set of primes has been adding DON'T WANT as an individual prime (August 14th 2014 update on the NSM homepage). As the case of DON'T WANT would require deeper analysis than is allowed by the publishing process of this paper, we use the version 2013 with its 64 primes as our point of reference. Goddard describes the history of primes in terms of three generations (2011a).

PART	OSA
THIS	TÄMÄ~SE
THE SAME	SAMA
OTHER~ELSE	TOINEN~MUU
ONE	YKSI~ERÄS
TWO	KAKSI
SOME	JOKIN~JOKU~MUUTAMA
ALL	KAIKKI
MUCH~MANY	PALJON~MONI
LITTLE~FEW	VÄHÄN~HARVA
GOOD	HYVÄ
BAD	PAHA
BIG	ISO~SUURI
SMALL	PIENI
THINK	AJATELLA
KNOW	TIETÄÄ
WANT	TAHTOA
FEEL	TUNTEA~TUNTUA
SEE	NÄHDÄ
HEAR	KUULLA
SAY	SANOA
WORDS	SANAT
TRUE	TOSI~TOTTA
DO	TEHDÄ
HAPPEN	TAPAHTUA
MOVE	LIKKUA
TOUCH	KOSKEA
BE (SOMEWHERE)	OLLA (JOSSAIN)
THERE IS	OLLA (OLEMASSA)
HAVE ¹⁰	OLLA (OMISTAA)
BE (SOMEONE/SOMETHING)	OLLA (JOKU/JOKIN)
LIVE	ELÄÄ
DIE	KUOLLA
WHEN~TIME	MILLOIN~JOLLOIN~SILLOIN~ AIKA
NOW	NYT
BEFORE	ENNEN~AIKAA SITTEN~AIKAISEMMIN
AFTER	JÄLKEEN~AJAN KULUTTUA
A LONG TIME	KAUAN (AIKAA)~PITKÄN AJAN
A SHORT TIME	VÄHÄN AIKAA
FOR SOME TIME	JONKIN AIKAA
MOMENT	HETKI
WHERE~PLACE	MISSÄ~JOSSA~PAIKKA

¹⁰ In the last versions, the prime HAVE has been replaced with the prime BE (SOMEONE'S)

HERE	TÄSSÄ~TÄÄLLÄ
ABOVE	PÄÄLLÄ
BELOW	ALLA
FAR	KAUKANA
NEAR	LÄHELLÄ
SIDE	PUOLI
INSIDE	SISÄ-
NOT	EI
MAYBE	EHKÄ
CAN	VOIDA
BECAUSE	KOSKA~VUOKSI~TAKIA
IF	JOS
VERY	HYVIN~ERITTÄIN
MORE	ENEMMÄN~LISÄÄ (and ENÄÄ)
LIKE~AS~WAY	NÄIN~KUTEN

3. The relation between the universal NSM concepts and the English based NSM

One of the very first observations in the course of the Finnish project was that the exponents of the primes involve many kinds of polysemy. Polysemy may occur in two contexts at least: 1) The English-based NSM exponents may be polysemous, e.g. BAD or LIVE; 2) Any other language based NSM exponents may be polysemous, e.g. the Finnish TEHDÄ which is only allowed to function as ‘to do’ while ‘to make’ is prohibited although the verb has both meanings. From the point of view of a new NSM version, the crucial question is whether that polysemy is something that just happens to occur with the English version, or should the same explicative functions be found in other language based NSM variants, using some other lexical patterns. These questions led us to consider the role of the English language in the theory of NSM, to be more precise, the relation between the universal NSM concepts and the English-based NSM, the first one referring to the universal language-independent concept of primes, the second one referring to just another language variant of NSM.

It is not fully unambiguous to define the very core of NSM in a theory that claims itself language independent yet uses English-based primes as a metalanguage. This question warrants further attention. We will enlighten this concern in the following by analysing and discussing our findings.

The basic assumption in the NSM is that the primes are strictly identified by the proposed basic combinatorial possibilities (from Goddard 2011c, see Appendix 1). These sentences set the grammatical restrictions

for every prime, and the NSM explications can be made with these combinations only. The basic combinatorial possibilities are to determine that every prime is only used within the foreseen lexical relations. In other (and more conventional) words, only one sense of a certain word from a given language is selected to serve as an NSM prime. Being selected and confirmed as a prime, the lexical item is supposed to be an independent and indefinable (in the sense that it cannot be rephrased) meta-level unit. It is supposed to behave in the NSM explications cleanly without interference from the other lexical roles and relations it has in its normal use in the natural language from which it comes.

The situation is still not clear. Many problems caused by polysemy of the exponents of the NSM primes in particular languages are already well recognized. Within the English-based NSM, exponents of primes like FEEL and KNOW are indisputably polysemous (e.g. Goddard & Wierzbicka 1994: 31–32). The solution has been to describe the use of these words in detail in order to ensure both that the appropriate sense is used and that the appropriate lexical item is searched for in other languages. In some cases, cross-linguistic comparison has prompted a re-evaluation of certain primes. For example, the NSM understanding of THINK was challenged by the evidence from Swedish (Goddard & Karlsson 2008). The exercise showed that the semantic prime THINK has a more restricted grammar than the word *think* in ordinary English, but this had not been previously noted in the basic combinatorial possibilities. The situation with the English verb *think* and the prime THINK may be similar to the one with the English adjective *bad* and the prime BAD (see later in this section).

Based on the historical fact that the NSM primes were first used extensively in English, it is always possible that our understanding of them carries some unrecognized hidden features that are specific only to the English language. The only way to find out those hidden features is to carefully look at every set of primes identified in new languages. Even though already having the strict grammatical and contextual restrictions, the original exponents of the NSM primes are not yet free from ordinary language-dependent lexical relations, mainly polysemy. These relations, however, may only become visible through continuing the process of transposing explications from one NSM variant to another. The following examples from the Finnish language enlighten this question.

The case of ONE. The prime ONE (like TWO) is identified as purely a quantifier (Wierzbicka 1996: 44). In published explications, however, this

is not always the case, as can be seen in a part of the definition of *a cup* (Goddard 2011a: 229–230): *someone can hold_[m] one in one hand_[m]* (focus on the first use of *one*). The Finnish-based NSM does not support this kind of use of the prime ONE, as can be seen in the translation *ihminen voi pitää_[m] sellaista yhdellä kädellä_[m]*. In the same vein, expressions like *the other one* are being translated as *tuo toinen*, literally, ‘that other’. Or, in some cases, the translation could be even *tuo toinen tällainen*, literally, ‘that other this kind’, using the prime KIND. According to Cliff Goddard (personal discussion November 2011), this kind of secondary use of the quantifier ONE could in principle be replaced with expressions like *one something of this kind*, even though it could make the explication harder to understand.

The case of SOMETHING~THING. The Finnish language seems to be lacking the overall concept for the prime THING, which may well refer to concrete, as well as abstract, objects. The Finnish word *asia* is normally used about abstract subjects, while *esine* refers to concrete nouns. The latter has certain restrictions, as it is only used for relatively small and non-living objects. A house or a cat could not be referred as *esine*, while they usually do not belong to the abstract category of *asia* either. The solution in most of the cases is to use just the prime JOKIN (‘something’), to refer to the object in question.

The case of LIVE. Even though the English verb *to live* is used in contexts like *they are living together* or *they live in the midst of a forest*, the prime LIVE does not include this type of adjunction, as pointed out by Wierzbicka (1996: 86–87). Making the difference in the English-based NSM seems still to be quite challenging, as seen even with the basic combinatorial possibilities: *many people live in this place, someone lives with someone else*. It is not always clear whether the sentences deal with Finnish *elää* ‘to live, to be alive’ or *asua* ‘to live, to stay, to reside’.

The point with the examples above is that the apparent polysemy of the primes under consideration may be specific to the English-based NSM exponents only, not necessarily to the universal NSM primes.

The case of MORE. From the Finnish point of view, the NSM prime MORE has, in addition to its main use (e.g. *joku haluaa enemmän / lisää* ‘someone wants more’) another specific use that we see as a separate meaning. In the list of the NSM primes, only MORE is listed, however, in practice *anymore* is used as well. Regardless of the fact that the word *anymore* in English contains the element *more*, the following two basic combinatorial possibilities *ei elä enää* ‘not living anymore’ and *ei enää*

kuten tämä ‘not like this anymore’ seem to deal with a different semantic concept. Namely, the English *more* refers to quantity, while *anymore* deals with (the length of) time. Quantity and time are different measures, and they cannot be combined at the level of primes without logical problems (see also the time expressions in Section 4). According to Cliff Goddard (personal discussion November 2011), one reason for linking *more* and *anymore* is the dynamic “heading forwards” feeling both of them have. The authors of this paper doubt whether this kind of consideration is sufficient, and moreover, whether the dynamicity is a property of the NSM prime as such or just a property that goes with the English language. As another “time-quantity” question arises in connection with the prime FOR SOME TIME, we suggest that both these primes deserve to be re-thought by the originators of the NSM. There is a certain mismatch in the identification of the primes.

The case of BAD. BAD is one of the very fundamental primes in the NSM, and it was among the most challenging translation tasks in the Finnish-based NSM project, hence it will be discussed in more detail.

The Finnish language has two main candidates for the exponents of BAD: PAHA and HUONO. The Finnish *paha* in many contexts comes close to English ‘evil or immoral’, while *huono* usually describes something as ‘low in quality’.¹¹ Both of them are opposites of *hyvä* ‘good’. In addition to *paha* and *huono*, there is also an adjective *tuhma* ‘incorrectly behaving, naughty’, the opposite of *kiltti* ‘correctly behaving’, both meanings expressed in English with *bad* and *good* (*good girl, bad boy*). As there are no obvious differentiated English counterparts for the Finnish *paha*, *huono* and *tuhma* (regardless of which of them is to be chosen as the official Finnish exponent of BAD), we have a good reason to believe that the semantic content of all these words are expressed by the NSM prime BAD.¹² As *tuhma* has remarkably different range of uses, we leave it aside from most of the further considerations (for a further discussion of ‘good girls, bad boys’ see Wierzbicka 2004).

The basic combinatorial possibilities can all be translated into the Finnish-NSM by using one or other of the candidates PAHA and HUONO,

¹¹ Explaining the Finnish *paha* with the English words *evil* or *immoral* does not actually help us to understand the very nature of this word, as both *evil* and *immoral* have their own complex meanings. Counterparts like these are to be taken with caution.

¹² An interesting aspect to BAD comes from e.g. the *Merriam-Webster Dictionary*: the first definition of English *bad* is quite close to the Finnish *huono*: “failing to reach an acceptable standard”.

and sometimes with both but with separate meanings (see Appendix 1; for details Vanhatalo & Tissari, forthcoming). One can also contrast between *paha* ‘bad, evil’ and *huono* ‘bad, low-quality’ in Finnish, saying, for example, about an alcoholic: *hän on huono ihminen, mutta ei hän paha ole* ‘s/he is a bad (not good) person, but s/he is not bad (evil)’. In the same vein, one can contrast ‘evil’ and ‘non-properly behaving’, saying, for example, of an poorly behaved student: *hän on tuhma, mutta ei paha* ‘s/he is bad (non-properly behaving), but not bad (evil)’. Furthermore, *ruoka ei välttämättä ole huonoa, vaikka se maistuisi pahalta* ‘food is not necessarily bad (in quality) even if it tastes bad’.

These examples led us to think about the true universality of the prime BAD. Is the very nature of BAD rather ‘paha’ or ‘huono’ – or something else, from some other language perspective? From the philosophical point of view the question is what the fundamental ideas of GOOD and BAD are (Goddard & Wierzbicka 1994: 47). Wierzbicka (1994: 496–497, 1996: 51–54) discloses her (and other researchers’ as well) earlier thoughts on these evaluative adjectives, mainly connecting the primes GOOD and BAD (which were not presented in the first set of primes Wierzbicka 1972) with the older prime WANT. She describes the problems with linking GOOD and WANT (the earlier explication of the English word *good* was ‘what someone wants’, and *bad* respectively ‘what someone doesn’t want’), and comes to the conclusion that WANT should not be linked with either of the evaluators. The main reason for breaking the older connection with the evaluators is the simple fact that sometimes one who wants can want something bad. Wishes and desires are subjective, while primes should be objective. Furthermore, even though people may not share their understandings of what is GOOD and what is BAD, they do agree that those concepts do exist (Wierzbicka 1996: 52).

Finnish is not the only language in which a problem arises with BAD. Russian, for example, has even more varieties to choose from: *ploxoj*, *durnoj*, *zloj* and *nexoros(hattu)ij* (Goddard & Wierzbicka 1994: 47, Gladkova 2007: 58). Studies on Mandarin Chinese (Chappell 1994: 142, referred by Wierzbicka 1996: 53), show how the primes GOOD and BAD are semantically asymmetrical, the exponent of BAD being narrower than the exponent of GOOD. According to Chappell (1994: 142), the Mandarin *huài* (exponent of BAD) seems to have some similarity with the Finnish *paha*, namely the aspects of immoral, nasty or evil. Wierzbicka explains the Mandarin variation as cultural rather than semantic (Wierzbicka 1994:

497), but for us the case is less clear. The use of Malay *buruk* (exponent of BAD) is also limited (Goddard 2002: 132).

Earlier in this section we referred to the consequences that research on the Swedish TÄNKÅ caused to the identification of the prime THINK. Yet there is a significant difference between the primes THINK and BAD: While THINK is a mental verb with highly specific syntactic features (complementation possibilities), BAD is used as an attribute or a predicate with less versatile syntax and much “meaningful” semantics. Nonetheless, we would encourage the NSM theory to sharpen the identification of BAD by distinguishing somehow between uses where English *bad* amounts to ‘evil, immoral, nasty, unpleasant’, on the one hand, and uses in which it conveys something like ‘low in quality’, on the other. Our suggestion is that the current prime BAD applies only to the ‘evil, immoral, nasty, unpleasant’ uses, while uses of *bad* that convey ‘low in quality’ are not semantically primitive but can be explicated – perhaps by using the negated version of the exponent for GOOD. The latter expression, furthermore, would include all other possible ‘not-good’ uses.

Taken together, the examples of polysemy¹³ presented above disclose the very fundamental question about the true universality of the current definition of the NSM primes. It is not always clear what we are actually dealing with: polysemy of the universal language-independent NSM primes or polysemy of the English exponents of the NSM primes. The question is crucial. According to Goddard (personal discussion November 2011), there is no overriding need to avoid inter-NSM (or inter any-language-based-NSM-version) polysemy. The burning question to us has been to make explicit the difference between ultimate inter-NSM polysemy and polysemy in the English-based NSM version, the first one having obviously more fundamental consequences for NSM based languages other than English than the latter one.

4. Remarks on the general system of primes

This section presents some remarks on the general system of primes. As the set of the NSM primes is very compact and internally closely related, also

¹³ Notably, polysemy is involved on the Finnish side, too (Vanhatalo & Tissari, forthcoming). As seen with English examples, sometimes the question about possible polysemy remains hidden until it pops up in a problematic translation process between NSM variants based on two different languages.

the questions raised up in this section and in the entire paper are inherently tied together.

4.1 The number, selection and mutual hierarchy of the primes

The question about the optimal number of NSM primes has been discussed since the early days of NSM. The number of primes has gradually increased over the past 40 years from 13 to 64, and some initially introduced primes have been removed (e.g. IMAGINE and WORLD, Wierzbicka 1972). Even though the primes are seen as the solid core of NSM, the general tendency seems to favor new primes slowly coming into the set.

The case of BE (SOMEWHERE), THERE IS, HAVE, and BE (SOMEONE/SOMETHING). The Finnish-based NSM challenges the existence of some of the current primes. In the Finnish set of primes, the verb OLLA ‘to be’ serves for four NSM primes, as it works for locational, existential, possessional or specificational verbal primes. The primes can still be distinguished by grammatical properties, as noted in the basic combinatorial possibilities. Some of these primes might be explicated with the other primes.

*The case of I and YOU.*¹⁴ Related to the expansion of the primes, a question arises about the internal relations between the primes. One can wonder whether all the primes are mutually equal or is it possible that some are e.g. more “primary” in some sense than the others. There seems to be no way to understand the prime YOU without first understanding the prime I, SINÄ without first understanding MINÄ. Arguably, YOU is always understood in relation to I, which is definitely something indefinable and could be considered to be a quintessentially primary prime.

The case of THIS and HERE. Another example of a possible hierarchy is provided by the Finnish TÄSSÄ~TÄÄLLÄ ‘here’. These forms are literally inner and outer locative cases of the pronoun *tämä* ‘this’, which raises a question about the relation between the primes THIS and HERE. As we cannot have TÄSSÄ~TÄÄLLÄ ‘here’ without TÄMÄ ‘this’, one of them seems to be more primary than another. The fact that in English HERE and THIS are different lexemes should not guarantee that they are different NSM primes. It should be carefully studied whether one of them

¹⁴ This remark is not specific to Finnish.

could be explicated by other primes. In that case, HERE might be just an allolex of THIS.

An indirect argument in support of the idea of recognising more and less primary primes comes from the history of the development of the prime set itself, as many of the early explications were made without certain currently proposed primes. NSM work over past decades has shown an unbelievable ability to produce sound explications with a number of different prime sets. Some of the current troubles with, for example, the apparent polysemy of some exponents of primes might result from an over-development of a compact and highly interrelated system. In other words, the number of primes in the current set is possibly too high. As all natural languages have their own issues with various lexical relations, we come again to the point that it is necessary to keep the various variants of NSM tightly tied to the universal NSM primes, and to avoid any unnecessary contact with, and contamination from, the English version (and lexical relations in the English language).

4.2 Economy and logic of primes

The case of time expressions. The primes expressing time raise questions about the logical consideration behind certain primes. The main question is whether the primes expressing time should be individual combinations as they are now or should they just be combined by separate primes as needed in use. Many languages have specific lexicalizations for expressing ‘for a long time’ and ‘for a short time’, with no morphological element corresponding to ‘time’ being included at all. This has been used in some NSM work as an argument in favour of the unitary nature of A LONG TIME and A SHORT TIME, i.e. against the idea that these meanings could be composites of TIME and a quantifier (such as MUCH and LITTLE). Actually the same feature is found in Finnish, where we have the alternative expression *kauan* ‘for a long time’. It is still not really valid to take such considerations into account when deliberating about the ultimate level of the NSM primes if these kinds of features are specific only to some languages.

Even though SOME + TIME wouldn’t work for the expression *for some time* (as the durational aspect would be missing without the word *for*), LITTLE + TIME might work for *a short time* or MUCH + TIME for *long time* in combinations like *a long time before, a short time after* (Goddard 2011c chart).

The case of SOME. According to the NSM policy, *some* in its semantically primary sense cannot be used with uncountable nouns (i.e. expressions like *some water* and *some time* are not allowed). The reason for this kind of restriction is that SOME in that sense is definable through other primes, *not much not little*. According to Cliff Goddard (personal discussion November 2011), the case of SOME is still currently not quite clear, and the identification of the prime may be changed in the future. But yet *some* still appears combined with the most uncountable expression in the independent prime FOR SOME TIME. One might then wonder whether it would be more useful to make this *some* as the actual prime SOME (allowing it to be used with any uncountable noun) and forget the narrow prime FOR SOME TIME.

The case of KIND and LIKE. The primes KIND and LIKE in Finnish draw attention to the surprisingly close relationship between these two primes and the very nature of each of them. The connection can be seen with English-based NSM expressions like *I have not seen someone like this* vs. *I have not seen someone of this kind*, which both could be translated to the Finnish-based NSM as *en ole nähnyt ketään tällaista*. Another example would be *someone says something like this* vs. *someone says something of this kind*, both could be translated into the Finnish NSM as *joku sanoo jotain tällaista / joku sanoo jotain näin*. Based on the similarities between these expressions, we could provocatively ask whether KIND and LIKE are truly different NSM primes or are they just English allolexes of one single mental prime. Even though cases like these are in minority in all of the uses of the prime KIND, the question of possible inter-NSM polysemy or synonymy itself is theoretically interesting, and worth deeper general discussion. Wierzbicka refers to the relationship between KIND and LIKE (1994: 494), but based on our evidence, we cannot fully agree with her arguments on the mutual independence of these concepts.

The case of KIND. Having a look at KIND independently, the English noun *kind* (and the English-based NSM prime KIND) is a much “stronger” and more wide-ranging concept compared with the Finnish suffix *-lainen~-läinen* (and thus also the prime -LAINEN~-LÄINEN). For example, it is quite hard to translate categorizing expressions like *natural kinds* into Finnish without a specific term *laji* (‘species’). One might even consider taking *laji* (‘species’) as an allolex of the prime -LAINEN~-LÄINEN.¹⁵

¹⁵ The Finnish *laji* ‘species’ is actually the etymological origin of the suffix *-lainen ~ -läinen*, as was kindly pointed out by one of the anonymous referees.

4.3 Grammatical number of the primes

The question about the grammatical number of primes is mostly involved with the much discussed prime PEOPLE, much less attention has been paid to the prime WORDS. These two primes are the only ones having no singular form. From the Finnish point of view, the grammatical number of these primes seems to be somewhat problematic. It is hard to see why we should use the plural forms IHMISET ‘people’ and SANAT ‘words’ instead of the simple singular forms IHMINEN ‘human being, person’ and SANA ‘word’. Logically, the singular form would be the primary one.

The case of PEOPLE. The quite extensive discussion about PEOPLE is closely related to SOMEONE and the status of the word *person* in the NSM context (e.g. Goddard & Wierzbicka 2002: 44–45, 79; Goddard 2002: 20; Goddard & Wierzbicka 1994: 33). Goddard and Wierzbicka argue against having PEOPLE decomposed as the plural of SOMEONE or PERSON by saying that PEOPLE is restricted to humans while SOMEONE is not. Nonetheless, in the majority of the cases, isn’t the identification of SOMEONE exactly what we understand with human beings, even though the English language does not have an optimal word for it? Still we have to admit that in some cases, there can exist “someones” who cannot get the human label, namely, gods, aliens and other kind of beings (e.g. Habib 2011). The Finnish version of NSM ended up with adding IHMINEN ‘human being, person’ and HÄN ‘s/he’ as allolexes of SOMEONE.

The main reason for adding IHMINEN ‘human being, person’ and HÄN ‘s/he’ as allolexes of JOKU ‘someone’ is to ensure that Finnish NSM explications will stay clear and understandable. We are well aware of the slight but acceptable oddness of the formulation *this someone* in the English-based NSM, and regret the similar formulation *tämä joku* ‘this someone’ (with grammatical allolexes *tämän jonkun* (genitive), *tätä jotakuta* (partitive), *tänä jonakuna* (essive), etc.) would sound too odd in the Finnish-based NSM. Including the allolex IHMINEN ‘human being, person’ into the Finnish-based NSM does not go against the original thought of the NSM. As we still cannot see good enough universal reason to make the general distinction between singular and plural forms (even though this may be the case within the English language), we would suggest to consider removing the prime PEOPLE entirely from the NSM set of original primes, and adding allolexes PERSON and PEOPLE to the prime SOMEONE in the English version of NSM.

The case of WORDS. Another suggestion is related to the grammatical number of WORDS. According to Cliff Goddard (personal discussion November 2011), there is no fundamental reason to have the prime WORDS in the plural form. The prime is presented in the plural because it is most frequently used in plural (see Goddard 2011b). Thinking about the general philosophy of the NSM, however, we suggest considering having the prime in singular, as it is simpler than the plural. This agrees with Sinnemäki's (2011: 16) definition of complexity, one characteristic of which is that there are a number of elements in a structure.¹⁶

4.4 Argumentation behind the lexical allolexes

The lexical relations involved with the identification of primes are not restricted to polysemy. The cases of synonymy led us to think about both diachronic and synchronic lexical variation. The lexical meaning does change over time, space and context and it may have some effect on the identification of NSM primes.

The fact that the current written Finnish includes features from two main dialects, had some implications for the Finnish NSM project as well. Some of the synonyms that once served as counterparts in two different dialects, may have developed differentiated meanings after the formation of a common written Finnish language. In the study on Finnish primes (Vanhatalo & Tissari, forthcoming), it was not always clear what the best Finnish candidate for prime status is, and whether or not the erstwhile synonyms should both be taken into account as NSM allolexes (e.g. Western *iso* vs. Eastern *suuri* 'big', *haluta* vs. *tahtoa* 'want').

The term *allolex* is used in the NSM context to mean a variant of a prime. The allolexes have the same identification ('meaning') as the prime does, and they are used for grammatical or collocational reasons. Some of the allolexes are just morphologically different forms of the primes (*minä* 'I', *minun* 'my'), some can be different lexemes (*toinen*, *muu* 'other, else'). It is not always straightforward to decide whether some words should or

¹⁶ Both Goddard and an anonymous reviewer seem to be right, though, about the frequency of the plural *words* as compared to the singular form *word*. The 450 million-word *Corpus of Contemporary American English* (1990–2012) attests 78656 instances of the type *word* as against 98366 instances of the type *words*, while the 100 million-word *British National Corpus* (late 20th century) attests 18707 instances of the type *word* as against 23632 instances of the type *words*.

shouldn't be regarded as allolexes of certain prime, as seen in the following examples.

The case of BIG. The Finnish variant of NSM has allolexes ISO~SUURI, because they are claimed not to have remarkable meaning differences (this has been noticed in dialectological studies as early as Nirvi (1936: 30–32), although the situation may have changed since then). Further studies on these lexemes would be most welcome, as at least some specialized usage can be found, e.g. *hän on iso mies* 'he is a big man' vs. *hän on suuri mies* 'he is a great man', the first one referring to physical size while the latter one rather describes a mental or social feature. Interestingly, Wierzbicka does not seem to pay any attention to the English variation *big* ~ *large*, neither does she argue why it is just the English *big* that has been chosen to be the (only) exponent of BIG (1996: 54–55). According to Cliff Goddard (personal discussion November 2011), *large* is narrower in its range of use and can be defined through BIG.

The case of WANT. The search for the Finnish exponent for WANT gave us another pair of synonyms – HALUTA and TAHTOA, the meaning difference of these verbs possibly correlating with the difference between the nouns *desire* and *will* (respectively).¹⁷ The differences between these two candidates led us to an interesting question about the deepest aspects of wanting. According to Finnish dictionaries, *haluta* roughly means 'someone's action while trying to fill some need or to make some wish true', while *tahtoa* is 'determined and purposeful action when trying to reach some goal'. The verb *haluta* may have a more primitive and even sexual flavour (*haluan sinua/sinut* 'I want/desire you', the Finnish object can be either partitive or accusative), while *tahtoa* may be more cultivated and conscious and even controlling or controlled action.¹⁸ The verb *haluta* may have the component of owning and using, which *tahtoa* does not have at least in that degree. The nouns connected to the verbs are *tahto* and *halu*, the latter one is intuitively quite close to *himo* 'lust', which is the base for the verb *himoita* 'to desire, to lust (after)'. The differences come nicely visible through compound words like *tahdonvoima* 'willpower', *ruokahalu* 'appetite', or sayings like *luja tahto vie miehen vaikka läpi harmaan kiven* 'a strong will takes a man even through a grey stone'. The verbs *haluta* and *tahtoa* can be used as alternatives in many contexts, e.g. *tahdon/haluan*

¹⁷ We thank an anonymous referee for pointing out this correlation to us, although it may suffer somewhat from circularity of reference: *The Oxford English Dictionary* defines the noun *will* partly in terms of the noun *desire*.

¹⁸ Sound research results on the semantics of these synonyms are lacking.

muuttaa kaupunkiin ‘I want to move to a city’, *tahdon/haluan ruokaa* ‘I want to get food’. The verbs still do have many different uses; e.g. the official question and answer used when a couple is getting married: *Tahdotko – osoittaa – rakkautta –?* ‘Do you want to – express – love –?’ and *Tahdon* ‘I do’, to express the very conscious step being taken.

Without a deeper research (which would definitely be very much welcome), we propose that *tahtoa* might express more neutral wanting than *haluta*, and *haluta* could possibly be defined through *tahtoa*. Our current suggestion for the exponent of the prime WANT is thus TAHTOA.

The cases above should lead us to think about the lexical changes taking places around the border between a language and a dialect. The fact that this border is undoubtedly vague and in constant change, supports the idea of at least a certain level of vagueness of primes as well. Although the various cases with their roots in dialectology are not necessarily reported in this paper, this aspect might provide hints for a general discussion about the nature of the NSM primes – and the border between a language and a dialect. Furthermore, still related to the history of lexical changes, there are some cases where the current spoken language differs remarkably from the written one, e.g. the case of the Finnish pronoun *se* ‘it’ widely referring to humans.

The case of ONE. The prime ONE can also be viewed through its allolexes. In the Finnish set of primes, there are two possible candidates for ONE, namely YKSI and ERÄS, the first one being a numeral and marking the number 1, while the latter one is more like the indefinite article *a/an* in English. According to the traditional Finnish grammar rules, *eräs* means a referent known to the speaker but not to the hearer while *yksi* refers to something more vague. The definitions of *eräs* and *yksi* have recently been softened to correspond more adequately to linguistic reality. Nowadays both are accepted as equivalents when used as determiners.

The case of KNOW. Certain kind of synonymy can also be recognised with the use of English KNOW with the basic combinatorial possibility *joku tietää jonkun toisen ihmisen (hyvin)* ‘someone knows someone else (well)’. When in Finnish *joku tietää* or *tuntee hänet*, the latter verb (which can be translated ‘feel’) is implicating deeper knowing. In Finnish, the verb *tietää* conveys knowing someone just a little (e.g. by name or face), while adding the particle *hyvin* ‘well’ means a little deeper degree of knowing. To express knowing someone truly well, one must use the verb *tuntea* (which can be translated ‘feel’), and one can emphasize the meaning with the particle *hyvin* ‘well’. For example, *tiedän naapurini kerrostalossa* ‘I know

my neighbours in the block of flats (by face)', *tiedän hyvin erään ihmisen joka pelkää lentämistä* 'I know someone well who is afraid of flying', *tunnen työtoverini* 'I know my work mate well', *tunnen hyvin itseni* 'I know myself very well'. The polysemy of the English exponent of the prime KNOW has been deeply discussed in Wierzbicka 1992, but studies around KNOW have been started again (in particular, "knowing someone" will possibly be removed from the set of potential combinations; Goddard personal discussion in March 2012).

4.5 Grammatical allolexy of primes

The complex Finnish nominal case system causes a very large number of grammatical allolexes¹⁹ (Goddard & Wierzbicka 2002: 20) for most of the substantive, verbal, adjective and numeral primes. In the context of NSM, not very much attention has been paid to the significance of extensive morphological phenomena of this kind.

The reason for this is that only a part of the English prepositions are allowed in the English based NSM – as some of them can be explicated with primes. Consequently, not all of the nominal cases (14–15 altogether) should be allowed in Finnish-based NSM explications, however, we find it notoriously challenging to define which ones. For example, can the Finnish elative and illative forms be used while *from* and *to* are not used in the English explications (instead, the formula "something was in one place/now it is in another place", is used)? The key question is: are these kinds of restrictions really driven by the universal language-independent nature of the NSM? It could also be logical to argue against the use of locative cases by noting that they are paraphrasable, i.e. not primitive (Cliff Goddard, personal communication in September 2014).

The case of the locative cases. The locative cases combined with spatial expression provide an enlightening example of this consideration. Let us have a look at BELOW (ALLA – *alle* 'to' + 'below', *alla* 'at' + 'below', and *alta* 'from' + 'below'), FAR (KAUKANA *kauas* 'to' + 'far', *kaukana* 'at' + 'far', and *kaukaa* 'from' + 'afar'), and NEAR (LÄHELLÄ *lähelle* 'to' + 'near', *lähellä* 'at' + 'near', and *läheltä* 'from' + 'near'). Use

¹⁹ We use the term *grammatical allolexy* to describe the different case forms of the primes (e.g. MINÄ (I) > *minä* (nominative), *minun* (genitive), *minua* (partitive), *minuna* (essive)), to mark the difference from the general NSM allolexy referring different lexemes (e.g. MUU~TOINEN (OTHER)).

of the cases add to the primes a temporal before-after perspective. So *kauas* ‘to somewhere far’, for example, implies that ‘after some time’ the moving person/thing will be ‘far from the place where it was before’. The English NSM would not allow directional ‘to’ and ‘from’ (noticed by Cliff Goddard, personal discussion in March 2012). We would suggest the NSM approach to consider the pros and cons of such practice.

The case of the abessive case. Another example comes from verbs. Problems occur with the prime CAN (VOIDA) in connection with abessive form (-*matta*/*-mättä*). The relatively widely used formulation ‘someone can’t not do something’ sounds quite awkward in the English-based NSM, and the strangeness of the Finnish *joku ei voi ei tehdä jotain* ‘someone can’t not do something’ is definitely at least as notable. In this case, we have decided to follow the NSM rule that exponents of the primes in different languages may look grammatically complex, even if expressing ‘simple’ meanings, and suggest the use of third infinitive abessive form: *joku ei voi olla tekemättä jotain* ‘someone can’t be without doing something’. One may ask why the English-based NSM does not use the expression *someone has to do something*, instead of *someone can’t not do something*. The English verb *have to* was in fact used in earlier NSM explications, but as pointed out by Goddard (2014), these two expressions (i.e. *can’t not do* and *have to do*) have different meanings. This can be seen, for example, in the context of someone who cannot prevent herself grinding her teeth at night: we could say that this person ‘can’t not grind her teeth at night’, but not that ‘she has to grind her teeth at night’. Carrying this argument over to the Finnish-based NSM, we can agree to recognise that there is a meaning difference between expressions like *ei voi olla purematta hampaitaan* ‘can’t not grind her teeth’ and *täytyy purra hampaitaan* ‘has to grind her teeth’.

The discussion about grammatical allomorphs of primes and the semantics of grammar raises a number of open questions about the general system of primes. For example, do the constructions like *I think* and *I’m thinking* have any difference that should be carried over to Finnish? The NSM grammar allows both expressions while generally trying to minimise the use of progressive. It seems that the current version of NSM may favor some languages by allowing or prohibiting certain type of complexity. With NSM variants other than English, it is not always clear what we are actually dealing with: legal allomorphs or illegal extended meanings or lexical elements other than primes. Without further studies on grammar, we

do not know whether e.g. a certain Finnish case has too much or too complex “meaning” to be freely used in NSM explications.

5. Discussion

The goal of this study was to revisit some basic principles of the NSM approach by using the exploration of the Finnish based NSM primes as the test bed. Sections 3 and 4 presented considerations on the relation between the universal NSM concepts and the English based NSM and, closely related, some remarks on the general system of primes.

The most important finding in our work with the Finnish prime system was that whenever the NSM primes are deployed in even the simplest explications in the English-based NSM (or in any other NSM version), various lexical relations specific to English (or to other languages) are immediately involved. Even the basic combinatorial possibilities referred to in this study (drawn from the Goddard 2011c chart) may be inexact or exclude some crucial elements. Most of the hidden lexical relations become visible only through a translation process to some other-language-based NSM (as with the case of THINK in Goddard & Karlsson 2008).

Our finding is important because attempts to find a convenient and flexible set of allolexes in any language based NSM version can always be hampered by claims like “this is not universal, this cannot be found in all languages”. It should be remembered in this context that the allolexes, lexical relations or grammatical aspects specific to the English-based NSM variant need not be found in other languages (even though some patterns may do so). Every language based NSM version has many allolexes or lexical relations specific to that particular language, which may not be found in any other language. The only requirement for these language-specific allolexes (with their possible lexical relations) is that they cannot be contradictory to the universal NSM primes – an emergent concept that is never 100% finished itself.

Partly due to the variation in allolexes, it is obvious that the NSM versions in different languages may differ in their flexibility and capability in analysing processes. One of our goals in the work with the Finnish NSM has been to make the Finnish-based NSM explications sound real, understandable and acceptable. We feel that at least some parts of the Finnish-based NSM explications were more natural than the originals, and the reason for this may be ignoring some English-specific lexical relations.

This remark should be very welcome, as it proves the vitality of NSM as a universal concept.

The theoretical risk of ending up with an NSM which may not be fully coherent but which has a slightly different toolbox for every language should not obstruct practising and developing the method. There are more pros of having more accurate semantic analysis than cons of possibly compromising with some aspects of the original ideas of NSM. While a critical approach to every method is important, one should be able to compare the shortcomings of the NSM with the ones of other analysis methods. Semantic analysis is among the most challenging tasks in the field of linguistic research, and this far there have not been too many perfect methods available.

Interestingly, even though the official policy regarding the allolexes in the English based variant of NSM is quite strict, the practice has more variation. This can be seen in the original English based explications where some non-primes are used without the molecule²⁰ mark and with no commentary or other indication that they are intended as allolexes. E.g. the explication of *a cup* (Goddard 2011a: 229–230): *stuff* in the phrase *these other things are made of_[m] the same hard_[m], smooth_[m] stuff*; *during* in the phrase *during this time, this someone's fingers_[m] move as this someone wants*; *bit* in the phrase *because of this, a little bit of something like hot_[m] water_[m] moves*; *way* in the phrase *sometimes when someone is drinking_[m] something in this way*; *as* in the phrase *part of the edge_[m] at the top_[m] of this thing touches one of this someone's lips_[m] for a short time, as this someone wants*. According to Cliff Goddard, most of these cases are allolexes of different primes. The word *stuff* should be replaced with the expression *the same something*, while *during* could be seen as an allolex of the expression *at this time*, used in relation to time periods (personal discussion November 2011). The word *bit* was used in some earlier explications, instead, it should have been replaced with the relatively recently-proposed prime LITTLE. The word *way* is an allolex of the prime THIS. Similar questions arise from the explication of *God* (Wierzbicka 2001: 21; see Appendix 2), as words other than primes or molecules occur there as well (*exist, always, they (them)*). Solutions like these are undoubtedly understandable and even desirable, but they should be clearly marked, reasoned – and opened to wider use. Various small and

²⁰ Molecules are mini explications that function as units in the semantic structure of other, yet more complex words.

unimportant adjustments may be needed even in the explications made by the real experts. The same kind of flexible attitude could be shown with more visible and frequent things, as suggested above in relation to the English allolexes of the prime SOMEONE. It is understandable, though, that any changes in the current prime inventory would require major consideration and extensive exploration.

At the present point in time, one may be justified in asking whether it is even possible to think about any ultimate pure NSM explications if all the NSM versions are making them with slightly different toolboxes. We would not bother too much about this question. What we do think, however, is that the current set of the NSM primes is, after the decades of careful description, identifiable in a relatively language independent way, and that this set can be regarded as the pure core of the NSM.

As soon as the English-based NSM version gets any kind priority among NSM versions, significant risks arise at two directions: First, favoring underlying lexico-grammatical features typical only to the English language would lead to Anglocentrism. On the other side, restricting the English version of NSM from creating sufficient allolexy would lead to too complicated explications in the most popular version of NSM.

From one important aspect, the English version of NSM is very special compared to the others. Namely, it is through the English-based NSM that most linguists (and other potentially interested people) form their attitude to the whole approach. Adding appropriate allolexes and grammatical features to the English variant would simply make the English-based NSM explications easier to comprehend by a wider audience. The English-based NSM should be treated as any other version of the NSM, it is not the universal NSM concept itself. To make it concrete, we could add another column to Table 1 in section 2, and separate the universal NSM primes and the English based NSM variants into different columns, the latter one having more and flexible allolexes.

Taken together, the main theoretical message of this paper is that the difference between the underlying ultimate universal NSM concepts and the English-based NSM version should be kept clear. This requirement may sound theoretical, unachievable, and even unnecessary, but this is the final point to which all of the practical question marks and concerns raised in the present study have led us. The findings in this paper pointed out some quite vague underlying constructions within the general system of primes. Regardless of the reported shortcomings, the NSM has shown indisputable

ability to deal with an extremely challenging task, namely semantic analysis.

References

- Amberber, Mengistu (2008) Semantic primes in Amharic. In Cliff Goddard (ed.), pp. 83–120. Amsterdam: John Benjamins.
- Bugenhagen, Robert D. (2002) The syntax of semantic primitives in Mangaaba-Mbula. In Cliff Goddard & Anna Wierzbicka (eds.), Vol. 2. pp. 1–64. Amsterdam/Philadelphia: John Benjamins.
- Chappell, Hilary (1994) The universal syntax of semantic primes in Mandarin Chinese. In Cliff Goddard & Anna Wierzbicka (eds.), Vol. 1, pp. 243–322. Amsterdam/Philadelphia: John Benjamins.
- Diller, Anthony (1994) Thai. In Cliff Goddard & Anna Wierzbicka (eds.), pp. 149–170. Amsterdam: John Benjamins.
- van Driem, George (2004) Review: Cliff Goddard & Anna Wierzbicka (eds.), *Meaning and universal grammar: Theory and empirical findings*. 2 vols. *Language* 80 (1): 163–165.
- Enfield, Nick J. (2002) Combinatoric properties of Natural Semantic Metalanguage expressions in Lao. In Cliff Goddard & Anna Wierzbicka (eds.), Vol. 2. pp. 145–256. Amsterdam/Philadelphia: John Benjamins.
- Geeraerts, Dirk (2010) *Theories of Lexical Semantics*. Oxford University Press.
- Gladkova, Anna (2007) *Russian Emotions, Attitudes and Values: Selected Topics in Cultural Semantics*. PhD Thesis. Australian National University.
- (2010) *Russian cultural semantics: Emotions, values, attitudes*. Moscow: Languages of Slavonic Cultures. (in Russian)
- Goddard, Cliff (2002) Semantic primes and universal grammar in Malay (Bahasa Melayu). In Cliff Goddard & Anna Wierzbicka (eds.), Vol. 1, pp. 87–172. Amsterdam/Philadelphia: John Benjamins.
- (ed.) (2008), *Cross-Linguistic Semantics*. Studies in Language Companion Series 102. Amsterdam: John Benjamins.
- (2011a) *Semantic Analysis: A Practical Introduction*, second, revised edition. Oxford: Oxford University Press.
- (2011b) The lexical semantics of *language* (with special reference to *words*). *Language Sciences* 33: 40–57.
- (2011c) Chart of NSM semantic primes <http://www.griffith.edu.au/__data/assets/pdf_file/0003/347565/nsm-chart-01-2011.pdf> (4 February 2013).
- (2014) Have to, Have got to, and Must: NSM analyses of English modal verbs of “necessity”. In Maite Taboada & Radoslava Trnavac (eds.), *Nonveridicality and Evaluation: Theoretical, computational and corpus approaches*, pp. 50–75. Leiden: Brill.
- Goddard, Cliff & Karlsson, Susanna (2008) Re-thinking THINK in contrastive perspective: Swedish vs. English. In Cliff Goddard (ed.), pp. 225–240. Amsterdam: John Benjamins.

- Goddard, Cliff & Wierzbicka, Anna (eds.) (1994) *Semantic and Lexical Universals*. Studies in Language Companion Series 25. Amsterdam: John Benjamins.
- (eds.) (2002) *Meaning and Universal Grammar: Theory and Empirical Findings*, Vol. 1–2. Studies in Language Companion Series. Amsterdam/Philadelphia: John Benjamins.
- Habib, Sandy (2011) Ghosts, fairies, elves, and nymphs: Towards a semantic template for non-human being concepts. *Australian Journal of Linguistics* 31: 411–443.
- Junker, Marie-Odile (2008) Semantic primes and their grammar in a polysynthetic language: East Cree. In Cliff Goddard (ed.), pp. 163–204. Amsterdam: John Benjamins.
- McCawley, James D. (1983) Review of Anna Wierzbicka's *Lingua Mentalis: The semantics of natural language*. *Language* 59 (3): 654–659.
- Murray, D.W & Button, Gregory (1988) Human emotions: Some problems of Wierzbicka's 'Simples'. *American Anthropologist* 90 (3): 684–686.
- Nirvi, Ruben E. (1936) Kansankielen synonyymeistä ja synonyymisluonteisista sanoista. *Virittäjä* 40: 24–88.
- Onishi, Masayuki (1994) Semantic primitives in Japanese. In Cliff Goddard & Anna Wierzbicka (eds.), pp. 361–386. Amsterdam/Philadelphia: John Benjamins.
- Oxford English Dictionary <<http://www.oed.com>> (29 September 2014)
- Peeters, Bert (1994) Semantic and lexical universals in French. In Cliff Goddard & Anna Wierzbicka (eds.), pp. 423–442. Amsterdam: John Benjamins.
- Sinnemäki, Kaius (2011) *Language Universals and Linguistic Complexity*. PhD thesis. General Linguistics, Department of Modern Languages, University of Helsinki.
- The NSM home page: <<http://www.griffith.edu.au/humanities-languages/school-languages-linguistics/research/natural-semantic-metalanguage-homepage>>
- Travis, Catherine (2002) *La Megalengua Semántica Natural: The Natural Semantic Metalanguage of Spanish*. In Cliff Goddard & Anna Wierzbicka (eds.), Vol. 1, pp. 173–242. Amsterdam/Philadelphia: John Benjamins.
- Tuovila, Seija (2005) *Kun on tunteet: Suomen kielen tunnesanojen semantiikkaa* [Such emotions]. The semantics of emotion words in the Finnish language] Doctoral dissertation. Oulu: Oulun yliopisto.
- Yoon, Kyung-Joo (2008) The Natural Semantic Metalanguage of Korean. In Cliff Goddard (ed.), pp. 121–162. Amsterdam: John Benjamins.
- (2006) *Constructing a Korean Natural Semantic Metalanguage*. Seoul: Hankook Publishing Co.
- Vanhatalo, Ulla (forthcoming): *Things you can do with the Natural Semantic Metalanguage*.
- Vanhatalo, Ulla & Tissari, Heli (forthcoming): *Natural Semantic Metalanguagen suomenkielinen versio 1.0* [Natural Semantic Metalanguage. The Finnish version 1.0].
- Wawrzyniak, Jan K. (2010) Native speakers, mother tongues and natural semantic metalanguages. *Language Sciences* 32: 648–670.
- Wierzbicka, Anna (1972) *Semantic Primitives*. Frankfurt: Athenäum.
- (1988) Semantic Primitives: A Rejoinder to Murray and Button. *American Anthropologist, New Series* 90 (3): 686–689.

- (1992) *Semantics, Culture and Cognition: Universal Human Concepts in Culture-Specific Configurations*. New York: Oxford University Press.
- (1994) Semantic primitives across languages: A Critical review. In Cliff Goddard & Anna Wierzbicka (eds.), pp. 445–500. Amsterdam: John Benjamins.
- (1996) *Semantics: Primes and Universals*. Oxford: Oxford University Press.
- (2001) *What Did Jesus Mean? Explaining the Sermon on the Mount and the Parables in Simple and Universal Human Concepts*. New York: Oxford University Press.
- (2002) Semantic primes and universal grammar in Polish. In Cliff Goddard & Anna Wierzbicka (eds.), pp. 65–144. Amsterdam/Philadelphia: John Benjamins.
- (2004) The English expressions *good boy* and *good girl* and cultural models of child rearing. *Culture & Psychology* 10(3): 251–278.

Contact information:

Ulla Vanhatalo
 School of Language, Translation and Literary Studies
 33014 University of Tampere

The Department of Modern Languages
 PO Box 24
 00014 University of Helsinki
 Finland
 ulla(dot)vanhatalo(at)helsinki(dot)fi

Heli Tissari
 University of Eastern Finland
 Philosophical Faculty
 Joensuu Campus
 School of Humanities
 PO Box 111
 80101 Joensuu
 Finland
 heli(dot)tissari(at)uef(dot)fi

Anna Idström
 University of Helsinki
 anna(dot)idstrom(at)welho(dot)com

Appendix 1. Some examples of the basic combinatorial possibilities of NSM primes (from Goddard 2011c). For the full set in Finnish, see Vanhatalo & Tissari, forthcoming.

MINÄ (I), SINÄ (YOU)

minä en tiedä ‘I don’t know’

minä haluan sinun tekevän/tietävän jotakin ‘I want you to do/know something’

jotain paha voi tapahtua minulle/sinulle ‘something bad can happen to me/you’

joku kuten minä/minunlaiseni ‘someone like me’

PALJON~MONI (MUCH~MANY)

paljon ihmisiä / monet ihmiset ‘many people’

paljon asioita / monet asiat ‘many things’

paljon osia / monet osat ‘many parts’

monenlaisia ‘many kinds’

monina aikoina ‘at many times’

monissa paikoissa ‘in many places’

paljon jotain tällaista (esim. vettä) ‘much something of this kind (e.g. water)’

paljon enemmän ‘much/many more’

PAHA (BAD)

jotain paha ‘something bad’

pahat ihmiset ‘bad people’

jotain paha tapahtuu ‘something bad happens’

tehdä jotain paha (jollekulle) ‘do something bad (to someone)’

tuntea jotain paha ‘feel something bad’

tämä on paha(a) ‘this is bad’

on paha jos... ‘it is bad if...’

TAHTOA (WANT)

joku tahtoo jotakin ‘someone wants something’

joku tahtoo tehdä/tietää/sanoa jotakin ‘someone wants to do/know/say something’

joku tahtoo jonkun muun tekevän/tietävän jotakin ‘someone wants someone else to do/know something’

joku tahtoo jotakin tapahtuvan ‘someone wants something to happen’

Appendix 2. Some examples of the Finnish based NSM in use.

Discourse particle *well* (Engl.)

Well, —

(Goddard 2011a:173)

<p>you said something a short time before because of this I want to say something after a very short time I'm thinking about it now because I want to say it well [I say: —]</p>	<p>sinä sanoit jotain vähän aikaa sitten tämän takia minä tahdon sanoa jotain hyvin lyhyen ajan kuluttua minä ajattelen sitä nyt koska minä tahdon sanoa sen hyvin [minä sanon: —]</p>
---	--

Noun *God* (Engl.)

God

(Wierzbicka 2001: 21)

<p>(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</p>	<p>(a) Jumala on joku (ei jokin) (b) hän on joku hyvä (c) hän ei ole sellainen kuin ihmiset (d) ei ole ketään muuta sellaista kuin hän (e) hän on olemassa aina (f) kaikki on olemassa koska hän tahtoo sen olevan olemassa (g) ihmiset ovat olemassa koska hän tahtoo heidän olevan olemassa (h) hän on olemassa koska hän on olemassa, ei minkään muun takia (i) hän elää</p>
---	---