For special issue on "Lexical Typology", edited by Maria Koptjevskaja-Tamm

Semantic primes, semantic molecules, semantic templates: Key concepts in the NSM approach to lexical typology

Cliff Goddard, University of New England, Australia

Not surprisingly, in view of its long track record in cross-linguistic lexical semantics, the Natural Semantic Metalanguage (NSM) approach (Wierzbicka 1996, 1999; Goddard 1998, 2005, 2006, 2008; Harkins and Wierzbicka 2001; Goddard and Wierzbicka 2002; Peeters 2006; and other works) has a clear theoretical position on key issues in lexical typology and a well-developed set of analytical techniques. From a theoretical point of view, the overriding issue for lexical typology concerns the *tertium comparationis*. What are the optimal concepts and categories to support the systematic investigation of lexicons and lexicological phenomena across the world's languages? To this question, the NSM approach offers the following answer: the necessary concepts can – and must – be based on the shared lexical-conceptual core of all languages, which NSM researchers claim to have discovered over the course of a thirty-five year program of empirical cross-linguistic semantics. This shared lexical-conceptual core is the mini-language of semantic primes and their associated grammar. In addition, over the past 10 or so years, NSM researchers have developed certain original analytical constructs which promise to enhance the power and systematicity of the approach: in particular, the notions of semantic molecules and semantic templates. This paper sets out to explain and illustrate these notions, to report some key empirical findings, and to extrapolate their implications for the further development of lexical typology. I will also seek to highlight differences in assumptions and approach from some other prominent trends in lexical semantics.

1. General principles and approach

The NSM approach differs from most other work in cross-linguistic semantics in two fundamental ways. First and foremost, NSM semantics is based on reductive paraphrase, in a very strict and literal sense. An NSM explication of a sentence or sentence frame is a systematic reductive paraphrase, i.e. an attempt to "say the same thing" in a paraphrase composed of maximally simple, intelligible and translatable words (semantic primes), thereby laying bare the semantic content of the original sentence or sentence frame. NSM researchers do not attempt in the first instance to classify lexical meanings, but rather to paraphrase them without circularity. Classifications may emerge inductively, generalizations of other kinds

may emerge – but the first process is always paraphrase. A corollary to the reductive paraphrase technique is that no specialist or technical terms are allowed in formal NSM semantic explications, because to do so inevitably leads to unacceptable abstractness and obscurity and/or to circularity.

When semantic description is carried out in accordance with these principles, it can be viewed both as linguistic analysis and as conceptual analysis. In other words, the NSM claim is that a successful reductive paraphrase which satisfies native speaker intuitions and which predicts and/or explains natural usage (including entailments, implications, and so on) can be viewed as a conceptual model. Because it is carried out in terms that are known to speakers, that form part of their everyday linguistic competence, a paraphrase analysis can have a prima facie claim to conceptual authenticity, in the sense of representing what anthropologists call an "insider perspective". At the same time, the constraint that reductive paraphrases be carried out in the language concerned (or, equivalently, in words that have precise semantic equivalents in the language concerned) safeguards the analyses against terminological Anglocentrism, i.e. the imposition of Anglo conceptual categories onto the concepts of other languages.

The most fundamental NSM concept is the concept of semantic primes, i.e. meanings which cannot be paraphrased in simpler terms: the bedrock of linguistic meaning. To the extent that semantic primes can be identified and match up across languages, they provide a stable and language-neutral metalanguage for lexical typology, at least on its semantic side; for mapping out patterns of polysemy, patterns of structuring in the lexicon, the general architecture of semantic domains and fields, for investigating lexicon-grammar interactions, and so on (Lehrer 1992; Koch 2001; Koptjevskaja-Tamm 2008). Framing semantic analyses (explications) in semantic primes ensures that they are clear, translatable, and intuitively accessible, which of course make them more predictive and easier to test.

The current model of 63 primes is the result of an incremental program of empirical/analytical research that began with Wierzbicka (1972). Major benchmarks since then include Goddard and Wierzbicka (1994), Wierzbicka (1996), Goddard and Wierzbicka (2002), and Goddard (2008), along with numerous other publications. [Note 1] Needless to say, the claimed finding that these 63 meanings appear to be present as lexical meanings in all languages is itself a very substantial claim about lexical typology. The Table of semantic primes below (Table 1) is presented in its English version, but comparable tables have been drawn up for many languages, including Russian, French, Spanish, Chinese, Japanese,

Korean, Lao, Malay, Mbula (Mangaaba-Mbula), East Cree, and many others. (Russian, Japanese, and Spanish tables of primes are included in the Appendix.)

I, YOU, SOMEONE, SOMETHING~THING, PEOPLE, BODY	substantives
KIND, PART	relational substantives
THIS, THE SAME, OTHER~ELSE	determiners
ONE, TWO, SOME, ALL, MUCH~MANY	quantifiers
GOOD, BAD	evaluators
BIG, SMALL	descriptors
KNOW, THINK, WANT, FEEL, SEE, HEAR	mental predicates
SAY, WORDS, TRUE	speech
DO, HAPPEN, MOVE, TOUCH	actions, events, movement, contact
BE (SOMEWHERE), THERE IS, HAVE, BE (SOMEONE/SOMETHING)	location, existence, possession, specification
LIVE, DIE	life and death
WHEN~TIME, NOW, BEFORE, AFTER, A LONG TIME, A SHORT TIME, FOR SOME TIME, MOMENT	time
WHERE~PLACE, HERE, ABOVE, BELOW, FAR, NEAR, SIDE, INSIDE	space
NOT, MAYBE, CAN, BECAUSE, IF	logical concepts
VERY, MORE	intensifier, augmentor
LIKE~WAY	similarity

Notes: • Primes exist as the meanings of lexical units (not at the level of lexemes) • Exponents of primes may be words, bound morphemes, or phrasemes • They can be formally complex • They can have combinatorial variants or "allolexes" (indicated with ~) • Each prime has well-specified syntactic (combinatorial) properties.

Table 1: Semantic primes (English exponents), grouped into related categories

There is no space here to review or justify this inventory in detail, as has been done extensively in the publications previously mentioned. It perhaps bears repeating, however, that to be a plausible candidate as an NSM semantic prime, a word (strictly speaking, word-meaning) must be indefinable, i.e. ultimately simple, in addition to being well attested in a wide range of languages. A word like 'eat', for example, would be a non-starter on both counts, since it is clearly not undecomposable (it involves 'doing', the 'mouth', etc.) and it is known not to have equivalents in some languages (Wierzbicka 2009; cf. Newman 2009). The same applies to many other impressionistically basic items of English vocabulary, such as 'go', 'hot', and 'bird' (cf. Goddard 2001, 2002).

Polysemy is frequently a complication when trying to identify primes and match them up across languages. Often the range of use of exponents of the same prime do not coincide because aside from the identical shared meaning, the words in question also have additional meanings which differ from language to language, i.e. there is a match-up between the meanings of lexical units but not between whole lexemes. Though much remains to be done, NSM researchers have accumulated a lot of data about common patterns of polysemy over the

past 15 years. A selection of some widely attested patterns is summarized, with some inevitable over-simplification, in Table 2.

Semantic prime	Additional meaning(s)	Language and relevant lexical item
DO	'make'	Amharic (*adərrəgə*), Ewe (*wɔ*), Italian (*fare*), Kalam (*g-*), Malay (*buat*), Mbula (*-kam*), Russian (*delat'*), Spanish (*hacer*), Swedish (*göra*), Yankunytjatjara (*palyaṉi*)
FEEL	'taste and/or smell'	Ewe (*se le lãme*), Italian (*sentire*), Kalam (*nŋ*), Malay (*rasa*), Russian (*čuvstvovat'*), Spanish (*sentir*)
	'hear'	Amharic (*tə-səmma-*), Italian (*sentire*), Kalam (*nŋ*), Spanish (*sentir*)
	'feel by touch'	Acehnese (*rasa*), English (*feel*), Italian (*sentire*), Spanish (*sentir*)
BEFORE	'first'	Kalam (*nd*), Kayardild (*ngariija*), Lao (*kòòn*[1]), Mbula (*muŋgu*), Samoan (*muamua*)
	'ahead of and/or in front of'	Kalam (*nd*), Kayardild (*ngariija*), Russian (*do*), Samoan (*muamua*)
WORDS	'what is said and/or message'	Amharic (*k'al*), English (*words*), Malay (*perkataan*), Mbula (*sua*), Russian (*slova*)
	'talk and/or language'	Amharic (*k'al*), Kayardild (*kangka*), Korean (*mal*), Mandarin (*huà*), Mbula (*sua*)

Table 2. Selected common polysemies of exponents of semantic primes (data from studies in Goddard and Wierzbicka eds. 1994, 2002; Peeters ed. 2006; Goddard ed. 2008; and Gladkova 2010).

In NSM studies, language-specific evidence is always adduced to support claims for semantic primes which depend on a polysemy analysis. Of course, to establish polysemy requires a principled method of semantic analysis. The conventional wisdom (if one can call it that), according to which is often difficult or impossible to separate polysemy from semantic generality, or to separate lexically encoded information from contextual inference, is really just a symptom of the lack of an adequate systematic method of semantic description. If one does not have a method of stating even a single meaning, it is hardly surprising that one can make no headway when faced with multiple meanings. [Note 2]

The natural semantic metalanguage consists not just of a lexicon, but also of a syntax. Semantic primes are hypothesised to have certain universal combinatorial properties (a "conceptual syntax"), and available evidence indicates that these properties also manifest themselves in all or most languages. Space precludes an adequate treatment here, so the reader is referred to Goddard and Wierzbicka (2002) and Goddard (2008). To give a very brief indication of the kinds of properties involved, it can be mentioned that they include: (a)

basic combinatorics: e.g. that substantives can combine with specifiers – THIS THING, SOMEONE ELSE, ONE PLACE, TWO PARTS, MANY KINDS; (b) basic and extended valencies of predicates and quantifiers, e.g. that SAY has addressee and topic valencies (SAY SOMETHING TO SOMEONE, SAY SOMETHING ABOUT SOMETHING), that ONE allows a partitive option (ONE OF THESE THINGS); (c) the complement options of the "mental" primes, KNOW, THINK and WANT. After this thumbnail sketch of NSM assumptions, we can proceed to explore the applications to lexical typology.

2. Semantic molecules

An extensive body of published work shows that lexical meanings in many domains (including emotion terms, speech-acts, value terms, and discourse particles) can be explicated directly into semantic primes. Informally speaking, these domains can be characterised as "abstract" (non-concrete) areas of the lexicon, but there are also some items of concrete vocabulary that yield to this approach. I will illustrate with two English nouns from different semantic domains: *hands* (body-parts) and *children* (social categories). These examples have not been chosen at random. They will be relevant to subsequent argumentation.

Explications [A] and [B] below are taken from Wierzbicka (2007a) and Goddard and Wierzbicka (to appear), respectively. For present purposes, it is not necessary to argue for the details, but rather to draw out some general points about structure and nature of the representations. First, although the wording of the individual components may be relatively simple, an explication taken as a whole is a rather complex structure. Partly this is due to the fact that despite its small lexicon, the metalanguage of semantic primes allows a surprisingly rich flexibility of expression, and partly it is due to the way which explications must be crafted in terms of anaphoric and causal relations so that their various components cohere and make sense as a whole. It appears to be an empirical fact that many human concepts have this kind of intricate structure. It is also worth noting the range and diversity of semantic primes that typically occur in explications. Between them, explications [A] and [B] use nearly half the prime inventory – over 30 primes – drawn from all divisions of the prime lexicon. [Note 3] (In explication [A], component (e) differs from its counterpart in Wierzbicka (2007).)

[A] *hands (someone's hands)*

a. two parts of someone's body
b. they are on two sides of the body
c. these two parts of someone's body can move as this someone wants
d. these two parts of someone's body have many parts

e. if this someone wants it, these parts can move in many ways as this someone wants
f. because people's bodies have these two parts, people can do many things with many things as they want
g. because people's bodies have these two parts, people can touch many things as they want

[B] *children*

a. people of one kind
b. all people are people of this kind before they can be people not of this kind
c. when someone is someone of this kind, this someone has lived for a short time, not a long time
d. the bodies of people of this kind are small
e. when people are like this, they can do some things, they can't do many other things
f. because of this, if other people don't do some good things for them,
bad things can happen to them

While it is true that words from many domains (especially "abstract" domains) can be explicated directly into semantic primes, and that the same applies to some non-abstract words (as just shown), NSM researchers have long recognized (Wierzbicka 1991; Goddard 1998: Ch. 6) that for words from most domains of the concrete vocabulary, it is not possible to produce plausible explications directly in terms of semantic primes alone. Rather, such explications typically require a combination of semantic primes and complex lexical meanings known in NSM theory as semantic molecules. That is, semantic molecules are complex meanings which are decomposable into combinations of semantic primes but which function as units in the structure of other, more complex concepts. For example, explications for words like *sparrow* and *eagle* must include 'bird' as a semantic molecule; explications for *fork*, *spoon* and *plate* must include 'eat'; explications for *walk* and *run* must include 'feet' and 'ground'. The concept of semantic molecules is similar to that of intermediate-level concepts in the semantic practice of the Moscow School (Apresjan 1992, 2000; Mel'čuk 1989), but with the important additional constraint that NSM semantic molecules must be meanings of lexical units in the language. It appears that most of the concrete lexicon – nominal, verbal, adjectival – relies on semantic molecules. The exploration of semantic molecules promises to contribute much to a general theory of vocabulary structure, as well as to shed new light on conceptual structure. In particular, it can contribute to new ways of representing semantic complexity, new ways of depicting semantic dependencies and inter-relationships, and new ways of seeing the texture of semantic structure. In this section, I will illustrate these contentions with some concrete examples, starting with simple examples from the domains of body-parts and social categories.

Wierzbicka (2007a) is an extensive study of body-part semantics, including over 40 explications. For the most part, she found that body-part explications require specifications of

three different kinds: the "location" of the body-part, a partial characterisation of its shape, and an indication of its function. The aspect of interest at this moment is the need for a shape specification, because shape descriptors (such as 'long', 'round', and 'flat') are not semantic primes and if they are required in explications, this amounts to recognising them as semantic molecules. For example, explication [C] for *legs* utilises the molecule 'long'; explication [D] for *head* utilises the molecule 'round'. [Note 4] When semantic molecules appear in NSM explications, they are marked as such by the notation [m].

[C] *legs (someone's legs)*

a. two parts of someone's body
b. these two parts are below all the other parts of the body
c. these two parts are long [m]
d. these two parts of someone's body can move as this someone wants
e. because people's bodies have these parts, people can move in many places as they want

[D] *head (someone's head)*

a. one part of someone's body
b. this part is above all the other parts of the body
c. this part is like something round [m]
d. when a person thinks about something, something happens in this part of this someone's body

How then can shape descriptors be analysed? Can we be certain that they can be used safely in body-part explications without incurring circularity? Wierzbicka (2006a) provides a general treatment of shape descriptors, including explications for English *long*, *round*, *flat*, and *straight* (among others), and an account of the considerable polysemy of each of these words. The full details need not concern us. The key point can be drawn out from a single example, namely, the explication in [E] for 'long' (in the relevant sense, i.e. its shape descriptor sense).

[E] *something long (e.g., a tail, a stick, a cucumber)*

a. when someone sees this thing, this someone can think about it like this:
b. "two parts of this thing are not like any other parts
because one of these two parts is very far from the other"
c. if someone's hands [m] touch this thing everywhere on all sides,
this someone can think about it in the same way

It is immediately obvious that one body-part – 'hands' – plays a crucial role in this explication, and indeed, in all Wierzbicka's explications for shape descriptor concepts. This is because shape descriptors designate properties that are both visual and "tangible", and to spell

out the nature of the latter concept requires both semantic prime TOUCH (contact) and semantic molecule 'hands [m]'. As established earlier, however, 'hands' itself can be explicated directly into semantic primes, so there is no circularity here. [Note 5] Rather, there is a chain or hierarchy of semantic dependency that can be represented as follows:

{'legs', 'arms', 'head'} < {'long', 'round'} < {'hands'} < {semantic primes}

This diagram is intended to indicate that each word set enclosed in curly brackets depends semantically on all the word sets to the right of it. This dispels any assumption that the impressionistically basic words of a particular semantic domain (in this case, 'parts of the body') are more or less the same in their degree of semantic complexity.

Let us work through a second example from another domain, that of social categories, such as *men*, *women*, *boys*, *girls*, and *children* (Goddard and Wierzbicka to appear). We have already established that *children* can be explicated directly into semantic primes. Now consider explication [F], noting that in the final line the word 'child' appears as a semantic molecule. Essentially, the idea is that the concept of *women* depends on the idea that there are two kinds of people's bodies, *women* being people of the kind whose body type allows them to have children. In other words, the concept of 'women' depends semantically on the concept of 'child'. [Note 6]

[F] *women*

a. people of one kind
b. someone can be someone of this kind after this someone has lived for some time, not for a short time
c. there are two kinds of people's bodies, people of this kind have bodies of one of these two kinds
d. some parts of bodies of this kind are not like parts of bodies of the other kind
e. the bodies of people of this kind are like this:
   at some times there can be inside the body of someone of this kind a living body of a child [m]

Taking the analysis a step further, Goddard and Wierzbicka (to appear) go on to argue that the meaning of *men* incorporates 'women' as a semantic molecule. Subsequently, all three of these basic social categories, i.e. 'men', 'women' and 'children', are needed in the explications of numerous other words; for example, in the domain of kinship (Wierzbicka to appear). Some of these relationships can be depicted as follows:

{'father', 'mother', 'husband', 'wife'} < {'men'} <{'women'} < {'children'} < {semantic primes}

Lest this conclusion seem unremarkable, it is worth reminding ourselves that in the old structuralist canon 'men' and 'women' were analysed as [+MALE, +ADULT] and [–MALE, +ADULT], respectively. This depicted these two words as symmetrical in semantic structure and made no reference to 'children' whatever.

It will be evident by now that many complex concepts have multiple "nestings" of molecule within molecule. In explications for *cats* or *chairs*, for example, the most complex molecules are bodily action verbs like 'eat [m]' or 'sit [m]'. They contain body-part molecules such as 'mouth [m]' and 'legs [m]'. These in turn contain shape descriptors, such as 'long [m]', 'round [m]' and 'flat [m]', and they in turn harbour the molecule 'hands [m]', composed purely of semantic primes. A further nesting occurs when natural kind terms themselves function as semantic molecules at a shallow level of semantic structure. For example, words for unfamiliar species such as *tigers* and *zebras* contain a "likeness" reference to familiar natural kinds, such as 'cats' and 'horses', respectively; endonymic terms like *purr* and *saddle* also contain references to 'cats' and 'horses', respectively (Goddard 1998: 241–242).

As suggested by the last point, it is evident that semantic molecules can differ in their degree of productivity and in how widely they range across the lexicon. How many productive semantic molecules are there, it may be asked. At the current early stage of research, the answer is not very clear. For English, the number is probably between 150 and 250. It is known that for English productive semantic molecules can come from at least the categories listed in Table 3 (the examples given are non-exhaustive).

parts of the body:	'hands', 'mouth', 'eyes', 'ears', 'head', 'legs', 'feet', …
physical descriptors:	'long', 'round', 'flat', 'straight', 'hard', 'sharp', ...
physical activities:	'sit', 'stand', 'lie', 'eat', 'drink', 'hold', …
physical acts:	'kill', 'pick up', 'bite', …
expressive actions:	'laugh', 'make sounds', 'sing', 'write', 'read', …
topological:	'top', 'bottom', 'front', 'edges', 'ends', 'hole', …
life-forms:	'creature', 'animal', 'bird', 'fish', 'tree', …
environmental:	'ground', 'sky', 'sun', 'water', 'fire', …
times and places:	'year', 'day', 'country', 'home', 'school', …
materials:	'wood', 'stone', 'metal', 'glass', 'paper', …
mechanical parts:	'wheel', 'pipe', 'wire', 'engine', 'electricity', 'machine', …
social and family:	'men', 'women', 'children', 'mother', 'father', …
major cultural concepts:	'money', 'book', 'colour', 'number', …

Table 3: Selection of semantic molecules of English, grouped into categories (Goddard 2010)

To wind up this section, I will itemise some of the significant implications of the theory of semantic molecules for lexical typology. First, there may well be some universal or near-universal semantic molecules, particularly for concepts which are foundational for many other concepts and/or for large lexical classes. The molecule 'hands' is a prime candidate, and cross-linguistic surveys appear to support this position, once sufficient attention is focused on questions of language-specific polysemy. Other candidates are certain other body-parts such as 'eyes' and 'ears' (Wierzbicka 2007a), basic social categories like 'men', 'women' and 'children' (Goddard and Wierzbicka to appear), some basic kin concepts such as 'mother', 'father', 'husband' and 'wife' (Wierzbicka to appear), and some environmental molecules, such as 'sky', 'ground', 'fire' and 'water' (Goddard 2010). [Note 7]

Second, and just as interesting, it is equally clear that some semantic molecules are language-specific. This is only to be expected for high-level molecules such as taxonomic categories, since it is well established that there are languages which lack exact equivalents for words like 'animal', 'bird' and 'tree' (Goddard 2001), but the possibility that lower-level molecules such as shape descriptors and topological terms can also vary somewhat from language to language is more surprising. However, Wierzbicka (2006a) argues that English 'long [m]' does not exactly match the comparable Polish molecule 'podłużny [m]' 'elongated, oblong', and Brotherson (2008) argues that English 'ends [m]' differs subtly from its nearest counterpart 'tapu [m]' in Makasai (East Timor). The implications of these claims remain to be explored.

Major cultural concepts can also have profound implications for the vocabulary structure of particular languages, sometimes in non-obvious ways. For example, Wierzbicka (2006b, 2007b, 2008) argues that in English 'colour [m]' functions as a semantic molecule in words like *red*, *blue*, *green*, etc., but that many other languages lack "colour words" in the true sense, because their visual descriptor words do not involve any comparable molecule. The semantic molecule 'number' also has huge significance in English (and in many other languages), both in helping to constitute the productive lexical domain of number words (Goddard 2009a), and, less obviously, in contributing to diverse other concepts connected with quantification and measurement (such as, for example, categories like 'age', 'temperature', 'weight', units of measurement, words for measuring devices, arithmetical concepts, etc.).

3. Semantic templates: the example of physical activity verbs

In NSM theory, a semantic template is a structured set of component types shared by words of a particular semantic class – often applicable across many languages and hence with important applications for lexical typology. The concept was first employed in explications for artefact and natural kind terms (Wierzbicka 1985). It has since been elaborated and applied to adjectives of emotion, shape, colour, and physical qualities (Wierzbicka 1999, 2006a, 2006b, 2007b; Goddard and Wierzbicka 2007). In recent years, the semantic template concept has been extended to verbs. In this section I will describe these new developments, and in section 4 illustrate their application to lexical typology. Section 5 will consider affinities with work on lexical templates in several other frameworks.

NSM researchers have well-developed proposals for the template structure of several subclasses of physical activity verbs, including: (i) intransitive verbs of bodily motion (locomotion), such as *walk*, *run*, *jump* (Wong, Goddard and Wierzbicka to appear); (ii) routine bodily activities, like *eat* and *drink* (Wierzbicka 2009a, 2009b); and (iii) complex physical activity verbs typically involving instruments, such as *cut* and *chop* (Goddard and Wierzbicka 2009) [Note 8]. At the present time, these templates have been worked out and tested only for English and a small selection of other languages (Polish, Japanese, Kalam, Warlpiri), but NSM researchers believe that the same or similar template structures will be appropriate for many languages, hence their potential significance for lexical typology.

The proposed template structures for these three subclasses are very similar, as shown in Figure 1. In each case, the two top-most sections are termed, respectively: LEXICO-SYNTACTIC FRAME, and PROTOTYPICAL MOTIVATIONAL SCENARIO. The major differences lie in the final section(s). For bodily motion and routine physical activities, the MANNER section describes a coordinated set of body-part movements. For complex physical activities, the final sections of the template describe an INSTRUMENT and how it is used (USING THE INSTRUMENT), and the nature of the effect that the use of the instrument exercises on the object (WHAT IS HAPPENING TO THE OBJECT).

Locomotion verbs and routine physical activities, e.g. *walk, run, eat, drink*	Complex physical activities involving instruments, e.g. *cut, chop, grind*
LEXICO-SYNTACTIC FRAME	LEXICO-SYNTACTIC FRAME
PROTOTYPICAL MOTIVATIONAL SCENARIO	PROTOTYPICAL MOTIVATIONAL SCENARIO
MANNER	• INSTRUMENT • USING THE INSTRUMENT • WHAT IS HAPPENING TO THE OBJECT

Figure 1: Template structure for "physical activity verbs" of several subclasses.

Let us review these proposed template structures. LEXICO-SYNTACTIC FRAME refers to the top-most section, with different macro-classes having different frames. Figure 2 displays lexico-syntactic frames for three subclasses of physical activity verbs. The details in the frame determine the mapping from lexical semantics to morphosyntactic expression. The frames define core argument structure, inherent aspect, causal notions, and the controlled nature of the activities. Notice that no technical linguistic terms ('agent', 'patient', 'duration', 'control', or the like) are used in stating the frame.

Locomotion, e.g. *walk*, *run*	someone X is doing something somewhere for some time because of this, this someone's body is moving at this time in this place as this someone wanted
Routine physical activities, e.g. *eat*, *drink*	someone X is doing something to something Y for some time because of this, something is happening to this something at the same time
Complex physical activities, e.g. *cut*, *chop*	someone X is doing something to something Y with something else Z for some time because of this, something is happening at the same time to thing Y as this someone wanted

Figure 2: Lexico-syntactic frames for three verbal subclasses.

A notable feature of the lexico-syntactic frames displayed in Figure 2 is that they are phrased in the imperfective (note the durative component 'for some time'). Most treatments in other frameworks assume without discussion that perfective uses (*walked*, *ran*, *cut*, *chopped*, etc.) are basic, but NSM analysts agree with the tradition in Russian lexicology that, for physical activity verbs, the imperfective forms and uses are semantically simpler. This is because their perfective counterparts involve extra semantic components: e.g. 'at one time', and the specification that an outcome (related to that implied in the prototypical motivational scenario) has been achieved. Though we cannot go through the details here, the claim is that this analytical strategy enables a solution to the so-called imperfective paradox [Note 9] and to the problem of how to specify the semantic relationships between constructional variants (syntactic alternations) of a single verb (Goddard and Wierzbicka 2009).

The next section is PROTOTYPICAL MOTIVATIONAL SCENARIO. Against the externalist methodology of some work in lexical typology, e.g. Majid and Bowerman (2007), NSM researchers maintain that speakers conceptualize human activities by reference to their prototypical motivations. For example, the prototypical motivational scenario associated with English *walk* states that a person often does something like this (i.e. *walks*) when they want to

be somewhere after some time, not far from the place where they are. (This does not imply that people only ever *walk* with this motivation. Obviously, one can walk for exercise or pleasure, or for other reasons, but the claim is that the concept of *walking* makes reference to this particular motivation.) Likewise, the prototypical scenario associated with *eat* and with *drink* is that someone wants something to be inside their body. Complex physical activity verbs (*cut*, *chop*, *grind*, *knead*, etc.) have a richer cognitive structure than locomotion and other routine activities, because they prototypically involve something like conscious intention: an actor forming a "preparatory thought" directed towards changing the current state of some object. For example, for English *cut*, the prototypical motivational scenario involves wanting something not to be one thing anymore, but instead to be two things (and as well, wanting to control the separation process with some precision). Examples of prototypical scenarios for representative verbs of three subclasses are given in Figure 3.

walk (locomotion)	at many times someone does something like this when it is like this: this someone is somewhere at some time this someone wants to be somewhere else after some time this place is not far from the place where this someone is
drink (routine physical activity)	at many times someone does something like this when it is like this: this someone wants this something to be inside their body this something is something like water [m]
cut (complex physical activity)	at many times someone does something like this to something when it is like this: a short time before this someone thought like this about this something: "I don't want this something to be one thing anymore, I want it to be two things because of this, I want to do something to this something for some time when I do this, I want something to happen to this something all the time, as I want"

Figure 3: Prototypical motivation components for verbs of three different subclasses.

The next section of the template for physical activity verbs is MANNER: how the activity is carried out. Given the rational goal-directed nature of human action, it is not surprising that the details are closely linked to the prototypical motivation. People do things in a certain way in order to get the desired result. For English locomotion verbs like *walk* and *run*, the manner section includes a lot of detail about how the feet and legs move in relation to one another, and in relation to the ground (in other languages with more general motion verbs, much less detail is supplied). For routine physical activities like *eat* and *drink*, the MANNER section details how the parts of the mouth and (in some cases) the hands are used. For complex physical activities like *cut*, *chop*, and *grind*, the section that corresponds to "manner" is more complex, consisting of: first, a description of an INSTRUMENT; then of USING THE

INSTRUMENT, which can involve several interrelated sub-events; and then of how the object is affected by the action of the instrument (WHAT IS HAPPENING TO THE OBJECT). In the manner section, there is commonly provision for an iterative structure of repeated episodes ('the same thing happens many times; it happens like this ...').

4. Cross-linguistic comparison using semantic templates

The proposed templates for physical verbs were not preconceived notions, but emerged as a consistent organisational format during the painstaking and iterative process of semantic analysis (drafting and re-drafting multiple explications, testing them against range of use and native speaker intuitions, checking the coherence and well-formedness of the metalanguage, and so on). They seem to have a natural "internal logic" by which the causal and temporal interconnections between the various components can be ordered in a coherent fashion. On this account it seems likely that these templates will be similar across many languages. Albeit that the sample of languages is very small, this supposition seems to be borne out across the several non-English languages on which detailed work of this nature has already been done: Japanese, Polish, Kalam, and Warlpiri. Space permits only two cross-linguistic examples here: English *drink* compared with Kalam *ñb* 'eat/drink', and English *cut* compared with Japanese *kiru* "cut". The exposition will be abbreviated. Further justification and comparisons with other related verbs are given in Goddard and Wierzbicka (2009) and Wierzbicka (2009). (Note though, that the prototypical motivational scenario sections in the following four explications differ slightly from the previously published versions.)

English drink vs. Kalam ñb 'eat/drink'. Explication [G] below is for English *drink*. In terms of its overall structure, most of the relevant details have been introduced already. A couple of notable points are as follows: (i) the prototypical motivational scenario includes a characterisation of the object as 'something like water [m]', i.e. a liquid; (ii) the manner section depicts an iterative structure; more specifically, it involves doing something with the mouth that causes some of the "water-like" substance to be inside the mouth for a very short time, following which a further action of the mouth causes it to be somewhere else inside the person's body.

[G] *Someone X is drinking something Y:*

a. someone X is doing something to something Y for some time — LEXICO-SYNTACTIC FRAME
because of this, something is happening to this something at the same time

b. at many times someone does something like this to something when it is like this: — PROTOTYPICAL
this someone wants this something to be inside their body — MOTIVATIONAL SCENARIO

this something is something like water [m]

c. when someone does something like this to something for some time MANNER
the same thing happens many times
it happens like this:
this someone does something to this something with their mouth [m]
because of this, after this, part of this something is for a very short time inside this someone's mouth [m]
after this, this someone does something else to it with their mouth [m]
because of this, after this, it is not inside this someone's mouth [m] anymore,
it is somewhere else inside this someone's body for some time

If we were to compare this explication with that for English *eat*, we would see a slightly different prototypical scenario (involving 'something not like water [m]') and, consequently, a more elaborate manner section, with more detail about actions of parts of the mouth (related to chewing) and how these actions affect the substance in the mouth. As well, the manner section for *eat* involves some preliminary action with the hands (related to holding the food item and moving it to the mouth).

The Papuan language Kalam (Pawley and Bulmer in press) has no words equivalent in meaning to the English *eat* and *drink*. [Note 10]. Instead, both activities (as it seems from an English point of view) are designated by the verb *ñb-*, roughly, 'consume'. According to Pawley and Bulmer's dictionary, *ñb* is general in its semantics, rather than ambiguous; a sentence like *Tap etp nbsay*? 'What are they eating/drinking?' is genuinely vague.

Explication [H] shows how such an undifferentiated "eat/drink" meaning can be constructed (Wierzbicka 2009). It follows the same semantic template as for English *drink*, and many of the details also remain the same, while others differ. Notably (i) the prototypical motivation does not characterise the substance as either 'something like water [m]' (as with *drink*) or as 'something not like water [m]' (as with *eat*); and (ii) the period of time for which each mouthful of the substance remains in the mouth is described as 'a short time' (rather than 'a very short time', as with *drink*). Naturally, the elaborated manner details for *eat* are not appropriate.

[H] *Someone X is ñb-ing something Y:* [Kalam *ñb*]

[a] someone X is doing something to something Y for some time LEXICO-SYNTACTIC FRAME
because of this, something is happening to this something Y at the same time

[b] at many times someone does something like this to something when it is like this: PROTOTYPICAL
this someone wants this something to be inside their body MOTIVATIONAL SCENARIO

[c] when someone does something like this to something, MANNER
the same thing happens many times

it happens like this:
this someone does something to something with the mouth [m]
because of this, after this, part of this something is for a short time inside this someone's mouth [m]
after this, this someone does something else to it with their mouth [m]
because of this, after this, it is not inside this someone's mouth [m] anymore,
it is somewhere else inside this someone's body for some time

English cut vs. Japanese kiru. Explication [I] is a full explication for English *cut*, following the semantic template outlined in the previous section for complex physical activity verbs. Notice that the instrument section introduces an item that is 'not part of this someone's body' and that 'some parts of this something are sharp [m]'. In the lexico-syntactic frame and prototypical motivation sections, we see that there is an emphasis on the "controlled" nature of the intended separation effect on the object. How this is achieved is spelt out in the USING THE INSTRUMENT section, which describes how the actor's hand holds and guides the instrument so that it moves 'as this someone wants'. The next section shows how this results in the sharp edge of the instrument enacting a controlled effect on the object at the point of contact. [Note 11]

[I] *Someone X is cutting thing Y (e.g. some paper, a cake) with thing Z:*

a. someone X is doing something to thing Y with thing Z for some time — LEXICO-SYNTACTIC FRAME
because of this, something is happening at the same time to thing Y as this someone wanted

b. at many times someone does something like this to something when it is like this: — PROTOTYPICAL
a short time before, this someone thought like this about it: — MOTIVATIONAL SCENARIO
"I don't want this something to be one thing anymore, I want it to be two things
because of this, I want to do something to this something for some time
when I do this, I want something to happen to this something all the time as I want"

c. when someone does something like this, they do it with something — INSTRUMENT
this something is not part of this someone's body
some parts of this something are sharp [m]
when someone does something like this with something,
this someone holds [m] part of this thing with one hand [m] all the time

d. when someone does something like this with something, — USING THE INSTRUMENT
the sharp [m] parts of this thing touch this other thing for some time
during this time this someone's hand moves as someone wants
because of this, during this time the sharp [m] parts [m] of this thing touch
this other thing in some places as this someone wants
— WHAT IS HAPPENING TO THE OBJECT

e. because of this, something happens to this other thing in these places as this someone wants
because of this, after this, thing is not like it was before

The comparable scenario for English *chop* has a somewhat different prototypical scenario (it involves wanting 'something hard [m]' not to be one thing anymore, but instead to be many small things), with associated differences in the instrument-related sections (roughly, the sharp-edge instrument having a long handle, and being repeatedly raised above the object and brought down on it).

For an example of a similar-yet-different word in another language, we can turn to Japanese. The closest Japanese counterpart of *cut* is the verb *kiru.* Like *cut*, *kiru* usually refers to an activity performed with either a knife or scissors, and as in the case of *cut*, the prototypical intention appears to consist in dividing an object into two things in a controlled fashion. What is different is that in some situations Japanese *kiru* can refer to an action performed with one's fingers rather than with an instrument with a sharp edge. This applies in particular to paper, or objects made from paper, such as a sachet of powdered soup or sugar. Opening such a sachet with one's fingers can be described as a case of *kiru*, as in a sentence like: *Suupu-no fukuro-o te-de kitta* [soup-GEN sachet-ACC hand-INS kiru:PRT] '(He/she) opened the soup sachet with [his] hands.' Japanese dictionaries, e.g. Morita (1989), specify 'hand or cutlery' as the "instruments" associated with the verb *kiru*. When one folds a piece of paper, and then separates it along the crease with one's hand, the natural way to describe this in Japanese is with *kiru* – and, of course, this particular procedure is very commonly done in Japan on account of the widespread practice of origami (paper-folding); cf. Buisson (1992), Honda (1965). Origami involves the folding of paper and separating it into pieces of paper by hand and with precision.

Explication [J] below gives a unified interpretation of *kiru*. Despite the seemingly radical difference to *cut*, the explication follows the same template. The INSTRUMENT section presents the instrument of *kiru* as simply 'something', without specifying whether or not it is part of the agent's body. This unspecified 'something' can therefore stand equally well for a pair of scissors or for a person's fingers. Needless to say, this also means refraining from mentioning any sharp parts, though this an inherent part of the explication of *cut*.

The 'something' with which *kiru* is performed is also mentioned in the USING THE INSTRUMENT section. To interpret this part of the explication in a way which would allow reference both to a knife (or a pair of scissors) or to the agent's fingers, the thing in question needs to be conceptually dissociated from the hand. The relevant components read: 'when someone does something like this to something, one of this person's hands [m] moves as this person wants'; and then, 'because of this, this thing moves at the same time as this person wants'. This wording allows both interpretations: the references to the "instrument" as

'something' and as 'this thing' are capable of being interpreted as either 'something other than a part of this person's body' or 'a part of this person's body [the hand, or part of the hand]'.

[J] *Someone X-ga thing Y-o (Z-de)* ki-tte-i-ta. [Japanese *kiru*]

a. someone X was doing something to thing Y with thing Z for some time — LEXICO-SYNTACTIC FRAME
because of this, something was happening at the same time to thing Y as this someone wanted

b. at many times someone does something like this to something when it is like this: — PROTOTYPICAL
this something is not something very hard [m] — MOTIVATIONAL SCENARIO
a short time before, this someone thought like this about this something:
"I don't want this something to be one thing anymore, I want it to be two things
because of this, I want to do something to this something for some time
when I do this, I want something to happen to this something all the time, as I want"

c. when someone does something like this, they do it with something — INSTRUMENT

d. when someone does something like this with something, — USING THE INSTRUMENT
one of this someone's hands [m] moves as this someone wants
because of this, this thing moves at the same time as this someone wants
because of this, during this time parts of this thing touch this other thing in some places
as this someone wants

WHAT IS HAPPENING TO THE OBJECT

e. because of this, something happens to this other thing in these places as this someone wants
because of this, after this, thing is not like it was before

Again, the five-part template for complex physical activity verbs provides a guiding framework, despite the considerable differences in the content of individual sections.

Since we are focused in this section on physical activity verbs, it seems appropriate to draw out two specific conclusions pertaining to physical activity verbs, before turning to some more general reflections about the NSM contribution to the project of lexical typology. The first conclusion is that, unlike much of the nominal lexicon, the verbal lexicon generally has a non-hierarchical structure. [Note 12] Although informally it may be acceptable of speak of a set of verbs like *chop*, *slice*, *mince* as "verbs of cutting", it is not true that *chop, slice*, and *mince* are true semantic hyponyms of 'cut' – *chopping* is not a 'kind of cutting', *slicing* is not a 'kind of cutting', and so on. This follows because semantic analysis shows us that *chopping*, *slicing*, *mincing*, and so on, each have their own distinctive prototypical motivations and (related) manner specifications, and that these are not simply elaborated versions of the prototypical motivation and manner components for 'cut' (Goddard and Wierzbicka 2009).

Second, the most promising avenue for exploring the semantic typology of human activity verbs is to focus on the prototypical motivational scenario. This is because most of the other features of explications for verbs of this kind have their rationale in the nature of the prototypical motivation. For example, we can investigate whether all, or most, languages have a verb including in the following component: 'I want this something not to be one thing anymore'. If a verb with this component is found, we can ask if it also includes one of the following two components: (i) 'I want it (= this something) to be two things' (as with *cut*), (ii) 'I want it to be many small things' (as with *chop*). Concurrently, we can ask whether the prototypical 'something' involved in such scenarios is characterized in any way, and if so, in what way, e.g. as 'hard [m]' or 'very hard [m]', and whether the instrument is characterized in any way, e.g. as 'heavy [m]'. Proceeding in this way, it should be possible to build a lexico-semantic typology of verbs of physical activity based on universal semantic primes (such as WANT, SOMETHING, PART, ONE, TWO, MANY, SMALL, and so on) and on universal or widely attested semantic molecules, such as 'sharp [m]', 'hard [m]', and 'heavy [m]'.

5. Comparisons with other approaches

This section undertakes comparisons with three other programs that bear on systematic lexical typology. These examples have been chosen to illustrate a range of theoretical assumptions and a range of approaches to the metalanguage of semantic description.

FrameNet. Inspired and led by Charles Fillmore, the FrameNet program has been steadily documenting the English lexicon in line with the assumptions of frame semantics (Ruppenhofer 2006). In recent years, the program has extended to several other languages, including Spanish, German, and Japanese (Boas 2009c). The program is radically different to most, if not all, approaches to linguistic semantics because its architecture depends crucially on identifying "frames" of real-world knowledge. The argument is that speakers "can be said to know the meaning of the word only by first understanding the background frames that motivate the concept that the word encodes" (Fillmore and Atkins 1992: 76). The number of frames is very large, and they exist at different levels of generality. Any given word or set of words will belong to a rather specific frame, such as INGESTION, CUTTING, or TRAVEL, and to various higher-level frames, such as INTENTIONALLY_ACT. Lower-level frames are said to "inherit" higher-level frames. Frames are characterized by a set of frame-specific Frame Elements and a statement (a Frame Definition) about how they are interrelated.

At the lowest level, the Frame Elements are highly situation-specific (in effect, situation-specific semantic roles), but higher-level frames employ rather generic Frame Elements, such as Agent, Patient and Instrument. There are also numerous so-called extra-thematic frame elements, such as Manner, Time, Reason, Duration, Circumstances, and Reciprocation.

As mentioned, Frames are supposed to be extra-linguistic in nature. Petruck (1996:2) characterises the Frame as "a cognitive structuring device, parts of which are indexed by words associated with it". It seems to me, however, that in practice FrameNet analysts often proceed as if Frames and Frame Elements were linguistic meanings, especially at the higher-levels. Be that as it may, what is distinctive about the FrameNet approach is that: "[W]ords or word senses are not related to each other directly, word to word, but only by way of their links to common background frames and indications of the manner in which their meanings highlight particular elements of such frames" (Fillmore and Atkins 1992: 76–77).

As one might expect on these assumptions, FrameNet analysts are not much interested in individual word-meanings. Where definitions for individual words are given, they are often just taken from the Concise Oxford Dictionary and exhibit all the usual faults of conventional definitions (circularity, obscurity, inaccuracy, etc.). For example: *walk*: "move at a regular and fairly slow pace by lifting and setting down each foot in turn"; *drink*: "take (a liquid) into the mouth and swallow"; *chop*: "cut with repeated sharp, heavy blows of an axe or knife." Some words, such as *cut*, receive no separate definition because they are in effect lexical instantiations of a Frame (in the case of *cut*, the CUT Frame).

The character of low-level Frame Descriptions can be indicated with the INGESTION Frame, which includes *drink* and *eat,* and the CUT Frame, which includes *cut, chop, slice, mince,* and number of others. [Note 13] The capitalised words represent Frame Elements.

> INGESTION Frame: An Ingestor consumes food or drink (Ingestibles), which entails putting the Ingestibles in the mouth for delivery to the digestive system. This may include the use of an Instrument.
>
> CUT Frame: An Agent cuts an Item into Pieces using an Instrument (which may or may not be expressed).

The INGESTION Frame and the CUT Frame both inherit the higher Frames INTENTIONALLY_ACT, TRANSITIVE_ACTION, and INTENTIONALLY_AFFECT. In addition, the INGESTION frame inherits MANIPULATION.

The initial goal of the FrameNet project was to document the lexicogrammar of English, and especially to capture the complex interrelationships between lexical units (word-senses) and their valence patterns. However, since Frames are supposed to be conceptual structures ("an independently existing conceptual system that is not tied to any particular language" (Boas 2009a: 87)), and given the great amount of work that has already been devoted to identifying numerous frames and their interrelationships, they may seem to provide a valuable platform for cross-linguistic comparison. The Spanish, German, and Japanese FrameNet projects now underway have taken the existing (English-derived) Frames as a starting point: "[this] means that non-English FrameNets do not have to go through the entire process of frame creation" (Boas 2009a: 73). Of course, it is recognised that new frames may need to be invented where necessary, especially in highly culture-specific domains, but the general assumption is that the existing English-derived Frames will provide a solid foundation for cross-linguistic work.

Unfortunately, from the NSM vantage point this assumption seems quite unrealistic. Because the Frame elements and definitions are constructed in a technical English-based terminology (terms like 'agent', 'item', 'ingestor' and so on), we would expect there to be major problems ahead with implementing the cross-linguistic program. [Note 14] Even in its English work, the FrameNet "workflow" has shown little concern with metalanguage issues, apparently identifying and constructing frame descriptions on a rather ad hoc basis. The result is a maze of complex interrelated notions that will require a great deal of interrogation and re-interpretation before they reach a form in which they can be unproblematically mapped across to other languages. FrameNet practitioners are becoming aware of these problems. Boas (2009a: 92-3) states that the "applicability of semantic frames as a cross-linguistic metalanguage" remains to be tested, and that "to determine the feasibility of a truly independent metalanguage based on semantic frames for connecting multiple FrameNets [in different languages] … not an easy task."

For example, in the Japanese FrameNet, it was found necessary, in order to capture differences between Japanese verbs of motion (*wataru* 'go across' and *koeru* 'go beyond'), to divide the Frame Element PATH into two sub-categories: ROUTE and BOUNDARY (*wataru* can occur with both, but *koeru* only with BOUNDARY). Unfortunately, it seems inevitable that this kind of "splitting" procedure will lead to yet more subcategories with ill-defined relationships to each other and to the higher frames and frame elements. What is needed is not more poorly-defined technical terms, but a reductive paraphrase process that could constrain the proliferation of complex, non-translatable terms.

Despite the professional respect that NSM researchers accord to FrameNet for its commitment to thorough evidence-based semantic documentation, it should be evident that from a methodological point of view there are great divergences of practice between the two programs. The only rough alignment in organisational structure that I can see is between higher-level Frames (such as INTENTIONALLY_ACT and INTENTIONALLY_AFFECT) and the lexico-semantic frame of NSM templates. These two representational devices are designed to "get at" the same phenomenon, although FrameNet uses complex technical terms where NSM uses schematic paraphrases (as in Figure 2). Other than that, differences far outweigh similarities. The lower-level Frames, the basic organisational units of FrameNet, have no directly equivalent in NSM theory as it currently stands. One function of Frames, viz. coordinating and inter-relating sets of semantically similar words, is served by the prototypical motivational scenario section of the NSM verbal templates. Verbs with similar prototypical motivations (purposes) tend to be aligned, not only in terms of motivation, but also in terms of manner, instrument and use of instrument, and potential outcome.

Lexical Constructional Model (LCM). LCM (Mairal and Faber 2002, 2006; Ruiz de Mendoza and Mairal 2008) is a good example of how semantic analysis can be conducted in cognitive-functional approaches to language. It seeks to combine aspects of cognitive linguistics with aspects of Role and Reference Grammar (cf. Mairal and Van Valin 2001; Butler 2003: 231–2). Lexical templates are a key element in the model. They are similar to FrameNet frames in that they are regarded as schematic representations of world-knowledge and contain rich semantic detail, but the LCM researchers are more mindful of metalanguage considerations, seeking to borrow semantic primes from NSM and lexical functions from the Meaning-Text Model (Wanner 2007) in an effort to achieve more robust and cross-linguistically applicable analyses ("typological adequacy"). The LCM program has significant affiliations with the NSM program, inasmuch as it seeks both to identify the templates shared by semantically coherent classes, and also to decompose individual predicates into their full idiosyncratic detail.

Current LCM versions for physical activity verbs are not available to the present author, but we can consider the following lexical template for CUT verbs from Mairal and Faber (2002: 55).

[[**do´** (w, [**use.sharp-edged.tool**(α)**in**(β)**manner´** (w, x)])] & [BECOME **be-at´** (y, x)]] CAUSE [[**do´** (x, [**make.cut.on´** (x, y)])] CAUSE [BECOME **pred´** (y, (z))]], α = x.

Clearly the metalanguage leaves much to be desired. Most striking in this CUT template is the use of the self-evidently complex predicate **cut´** in the internal sub-event **make.cut´.on**. Other equally problematical elements appear in decompositions for individual verbs. For example, the iterative movement specification for *chop* is represented as **repeated´**, and the manner-effector specification for *hack* is represented as **[rough´ ∧ violent´]**; but the terms **repeated´**, **rough´** and **violent´** can be faulted as semantically complex and English-specific. [Note 15]

More subtle, but perhaps more important on this account, are differences between NSM and LCM on basic semantic components. Predicates like CAUSE and BECOME sound (to many linguists) as if they are plausible "atoms of meaning", but in the NSM literature it is argued that the supposed CAUSE predicate is much better represented by the semantic prime BECAUSE, and that the putative primitive BECOME can be decomposed in terms of semantic prime HAPPEN and other components. As for **do´**, this looks to the untutored eye to be equivalent to the NSM DO, as found in natural language, but it is not. Borrowed from Role and Reference Grammar, **do´** is an abstract "effector" predicate (Van Valin and Wilkins 1996), whose subject can be either animate or inanimate. In ordinary language, however, an inanimate object cannot be said to literally 'do' anything.

Another point of contrast with NSM is that LCM accepts a hierarchical structure for verb semantics. "Lexical hierarchies" is one of the operating principles of the model. [Note 16] Furthermore, like the FrameNet characterisations, LCM templates contain nothing analogous to the prototypical motivational scenario which plays a crucial role in the NSM semantic template for comparable verbs.

The Nijmegen School. Perhaps the major fault line in semantics concerns the nature of "meaning" itself. Are we primarily interested in conceptual meaning (sense, intension) or are we primarily interested in denotation (reference, extension)? In linguistics, the classical extensionalist or referentialist approach was Berlin and Kay's (1967) approach to color meanings, which relied on the Munsell color chips to map out the range of reference of color terms in different languages. The Nijmegen school of Stephen Levinson and associates position themselves squarely in the Berlin and Kay tradition. They have produced a series of influential publications, beginning with the spatial domain and extending to body-parts, physical activities, and landscape terms (Levinson et al. (2003); Majid, Enfield and van

Staden (2006); Majid and Bowerman (2007)). The Nijmegen program depends on grounding linguistic description in supposedly neutral "etic grids" (Levinson et al. 2003: 487), such as standardized kits of physical props and models, line drawings, video clips, and the like.

Briefly put, the NSM approach upholds a conceptualist approach to meaning and reliance on naturalistic language data. [Note 17] It would be foolish to deny that standardised stimuli can elicit useful data in some areas, but it is hardly possible to go directly from this kind of data to valid conceptual analysis without careful consideration of natural usage, without informed consultation with native speakers, and ultimately without a systematic metalanguage for representing meaning. One cannot bypass the problems of semantic representation and conceptual analysis by reference to a standardized set of stimuli. Aside from the fact that meaning does not equal denotation, there are multiple technical problems with devising supposedly neutral non-linguistic stimuli, such as the danger of unwittingly incorporating the categories of one's own language (as happened with the Munsell "color" chips (Lucy 1997; Wierzbicka 2008)). The extensionalist approach, furthermore, is inapplicable to abstract domains, such as cognitive, emotional and value terminology.

Even with standardized extensionalist data sets, the problem of metalanguage does not go away. One still has to describe and analyse the data in terms of some language, and in practice researchers who use an extensionalist approach employ non-standardized English, including technical English, for this purpose. For example, in relation to "cutting and breaking events", Majid et al. (2007: 142-43) describe their results in terms of complex and language-specific parameters such as "relative predictability of the locus of separation", "use of a blade(-like) instrument" or "use of a sharp blow", and whether the separation is a "clean, cut-like fracture" or a "messy fracture". But if we do not use a descriptive metalanguage which is equally "implemental" in other languages, then in effect the language of the investigators becomes the default standard against which meanings in other languages are calibrated. Since every language (including English, and technical English no less than plain English) has its own local semantic peculiarities, this means giving up any prospect of precise and conceptually authentic semantic description. [Note 18]

6. Concluding remarks

In a recent overview, Koptjevskaja-Tamm (2008) identifies three major fronts for research in lexical typology: (i) patterns of structuring in the lexicon, (ii) the architecture of lexical fields (semantic domains), and (iii) lexicon-grammar interactions. I believe that the recently

developed NSM analytical concepts of semantic molecules and semantic templates have something to offer on each of these fronts.

In relation to patterns of structuring in the lexicon, the concept of semantic molecules leads to new ways of understanding lexical semantic complexity. Many semantic structures appear to have a kind of "gangly and lumpy" quality – lengthy strings of simple semantic primes interspersed with semantically dense molecules. Semantic molecules enable an incredible compression of semantic complexity, but at the same time this complexity is disguised by its being encapsulated and telescoped into lexical units embedded one in the other, like a set of Russian dolls. As far I can see, these qualities are unlike anything envisaged in structuralist, cognitivist or generativist approaches to lexical semantics.

Regarding the architecture of lexical fields, the relevance of the notions of semantic templates should be obvious. In establishing the semantic template for words of a particular semantic class in any language, one creates a framework for cross-linguistic comparison. On present indications it seems likely that many template structures will be the same or similar across languages. When comparing related words in different languages, it becomes a simpler and more tractable task to seek out the locus of differences between languages. To add an example to that of physical activity verbs, one can mention Goddard and Wierzbicka's (2007) study of physical quality adjectives, i.e. words like *sweet, hot, hard, heavy, rough, sharp*, in cross-linguistic perspective. The authors propose a semantic template that seems to work not only for English, but also for several other languages. Needless to say, whether and to what extent it would work for a much broader range of languages remains to be explored; but the point is that the groundwork has been done to make such a project possible. Presumably some semantic templates are widely attested across the languages of the world, while others are found only in languages of particular families, areas or grammatical types.

The idea that many semantic templates begin with a 'lexico-syntactic frame' relevant to macro-grammatical properties is obviously highly pertinent to lexicon-grammar interaction. We have seen examples of lexico-syntactic frames for just three classes of English verbs based on semantic prime DO in combination with other components; of course there are many others, corresponding to other subclasses of action and activity verbs. Classes and subclasses of verbs from other divisions of the lexicon have their own distinctive frames. For example, verbs and adjectives of emotion and sensation are based on FEEL, using components like (for emotions) 'someone X felt something (good/bad)' or (for sensations) like 'someone X felt something (good/bad) in part of his/her body'. Speech-act verbs are based on SAY, using components like 'someone X said something (good/bad/like this) to someone else Y'. Yet

other verbs are based on HAPPEN, using components like ‘something happened (to something X)’ or ‘something happened in place Z’, in combination with other elements. Although much remains to be done, the concept of lexico-syntactic frame would seem to offer a promising tool both for segmenting the verbal lexicon of individual languages and for comparing the organization of verbal lexicons across languages.

Overall, one can only agree with Koptjevskaja-Tamm (2008: 43) that the most urgent problems of lexical typology are methodological, including the need to “improve standards in cross-linguistic identification of studied phenomena and their (semantic) analysis”, and the need to “achieve a reasonable consensus on the meta-language used for semantic explications and on ways of representing meanings”. In my view, there is sufficient evidence in existing work to establish that the NSM metalanguage is an effective tool for analyzing and representing meanings across languages. [Note 19] Even if lexical typologists of other persuasions do not want to sign up to all the theoretical commitments of the NSM approach, I would urge them to consider the advantages of adding reductive paraphrase methodology to their descriptive toolkits.

Appendix: Exponents of semantic primes in three additional languages

Japanese	Spanish	Russian
WATASHI *I,* ANATA *you*, DAREKA *someone*, NANIKA~MONO~KOTO *something/thing*, HITO~HITOBITO *people*, KARADA *body*	YO *I*, TU *you*, ALGUIEN *someone*, ALGO~COSA *something~thing*, GENTE *people*, CUERPO *body*	JA *I*, TY *you*, KTO-TO *someone*, ČTO-TO~VEŠČ' *something/thing*, LJUDI *people*, TELO *body*
SHURUI *kind*, BUBUN *part*	TIPO *kind*, PARTE *part*	ROD~VID *kind*, ČAST' *part*
KORE *this*, ONAJI *the same*, HOKA *other*	ESTO *this*, LO MISMO *the same*, OTRO *other*	ÈTOT *this*, TOT ŽE *the same*, DRUGOJ *other*
HITO-~ICHI- *one*, FUTA-~NI- *two*, TAKUSAN *many~much*, IKUTSUKA *some,* MINNA *all*	UNO *one*, DOS *two*, MUCHO *much~many*, ALGUNOS *some*, TODO *all*	ODIN *one*, DVA *two*, NEKOTORYE *some*, VSE *all*, MNOGO *much~many*
II *good*, WARUI *bad*	BUENO *good*, MALO *bad*	XOROŠIJ~XOROŠO *good*, PLOXOJ~PLOXO *bad*
OOKII *big*, CHIISAI *small*	GRANDE *big*, PEQUEÑO *small*	BOL'ŠOJ *big*, MALEN'KIJ *small*
OMOU *think*, SHIRU *know*, HOSHII~—TAI~NOZOMU *want*, KANJIRU *feel*, MIRU *see*, KIKU *hear*	PENSAR *think*, SABER *know*, QUERER *want*, SENTIR *feel*, VER *see*, OÍR *hear*	DUMAT' *think*, ZNAT' *know*, XOTET' *want*, ČUVSTVOVAT' *feel*, VIDET' *see*, SLYŠAT' *hear*
IU *say*, KOTOBA *words*, HONTOO *true*	DECIR *say*, PALABRAS *words*, VERDAD *true*	GOVORIT'~SKAZAT' *say*, SLOVA *words*, PRAVDA *true*
SURU *do*, OKORU~OKIRU *happen*, UGOKU *move*, FURERU *touch*	HACER *do*, PASAR *happen*, MOVERSE *move*, TOCAR *touch*	DELAT' *do*, PROISXODIT'~SLUČAT'SJA *happen*, DVIGAT'SJA *move*, KASAT'SJA *touching*
(DOKOKA) IRU~ARU *be (somewhere)*, IRU~ARU *there is*, MOTSU *have*, (DAREKA/NANIKA) DEARU *be (someone/something)*	ESTAR *be (somewhere)*, HAY *there is*, TENER *have*, SER *be (someone/something)*	BYT' (GDE-TO) *be (somewhere)*, BYT'~EST' *there is*, BYT' U *have*, BYT' (KEM-TO/ČEM-TO) *be (someone/something)*
IKIRU *live*, SHINU *die*	VIVIR *live*, MORIR *die*	ŽIT' *live*, UMERET' *die*
ITSU~TOKI *when~time*, IMA *now*, MAE *before*, ATO *after*, NAGAI AIDA *a long time*, MIJIKAI AIDA *a short time*, SHIBARAKU NO AIDA *for some time*, SUGUNI *moment*	CUÁNDO~TIEMPO *when~time*, AHORA *now*, ANTES *before*, DESPUÉS *after*, MUCHO TIEMPO *a long time*, POCO TIEMPO *a short time*, POR UN TIEMPO *for some time*, MOMENTO *moment*	KOGDA~VREMJA *when~time*, SEJČAS *now*, DO *before*, OSLE *after*, DOLGO *a long time*, KOROTKOE VREMJA, *a short time*, NEKOTOROE VREMJA *for some time*, MOMENT *moment*
DOKO~TOKORO *where~place*, KOKO *here*, UE *above*, SHITA *below*, CHIKAI *near*, TOOI *far*, MEN *side*, NAKA *inside*	DÓNDE~SITIO *where~place*, AQUÍ *here*, ARRIBA *above*, DEBAJO *below*, CERCA *near*, LEJOS *far*, LADO *side*, DENTRO *inside*	GDE~MESTO *where~place*, ZDES' *here*, NAD *above*, POD *below*, DALEKO *far*, BLIZKO *near*, STORONA *side*, VNUTRI *inside*
-NAI *not*, TABUN *maybe*, DEKIRU *can*, -KARA *because*, MOSHI (BA) *if*	NO *not*, TAL VEZ *maybe,* PODER *can*, PORQUE *because*, SI *if*	NE *not*, MOŽET BYT' *maybe*, MOČ' *can*, POTOMU ČTO *because*, ESLI *if*
SUGOKU *very*, MOTTO *more*	MUY *very*, MÁS *more*	OČEN' *very*, BOL'ŠE~EŠČE *more*
YOO~DOO~YOONI *like/how/as*	COMO *like*	KAK~TAK *like*

• Primes exist as the meanings of lexical units (not at the level of lexemes) • Exponents of primes may be words, bound morphemes, or phrasemes. • They can be formally, i.e., morphologically, complex • They can have combinatorial variants or allolexes (indicated with ~) • Each prime has well-specified syntactic (combinatorial) properties.

Acknowledgements

I am most grateful to Anna Wierzbicka for valuable input to this paper. Earlier versions were presented at the ALT Workshop on Lexical Typology held in Paris, September 2007, and at departmental seminars at Stockholm University and Uppsala University in the same month. I received stimulating feedback from participants at all three gatherings. The present paper has benefited considerably from suggestions from several anonymous reviewers and from the editor of this special issue. This research was supported by the Australian Research Council.

References

Apresjan, Jurij D. 1992. *Lexical semantics: User's guide to contemporary Russian vocabulary*. Ann Arbor: Karoma. [Orig. published 1974 as *Leksiceskaja semantika – sinonimeceskie sredstva jazyka*, Moskva: Nauka.].

Apresjan, Jurij D. 2000. *Systematic lexicography*. Trans. Kevin Windle. Oxford: Oxford University Press.

Berlin, Brent & Paul Kay. 1969. *Basic color terms: Their universality and evolution*. Berkeley & Los Angeles: University of California Press.

Boas, Hans C. 2009a. Semantic frames as interlingual representations for multilingual lexical databases. In Hans C. Boas (ed.), *Multilingual framenets in computational lexicography*, 59–100. Berlin: Mouton de Gruyter. [First published *International Journal of Lexicography* 1995, 18(4). 445–478).

Boas, Hans C. 2009b. Recent trends in multilingual computational lexicography. In Hans C. Boas (ed.), *Multilingual framenets in computational lexicography*, 1–34. Berlin: Mouton de Gruyter.

Boas, Hans C. (ed.) 2009c. *Multilingual framenets in computational lexicography*. Berlin: Mouton de Gruyter.

Brotherson, Anna. 2008. The ethnogeometry of Makasai (East Timor). In Cliff Goddard (ed.), *Cross-linguistic semantics*, 259–276. Amsterdam: John Benjamins.

Buisson, D. 1992. *The art of Japanese paper*. Paris: Terrail.

Butler, Christopher S. 2003. *Structure and function: A guide to three major structural-functional theories. Part 1: Approaches to the simplex clause*. Amsterdam: John Benjamins.

Clark, Eve V. 2010. Learning a language the way it is: Conventionality and semantic domains. In Barbara C. Malt & Philip Wolff (eds.), *Words and the Mind: How words capture human experience*. Oxford: Oxford University Press.

Clark, Eve V. & Wong, Andrew D-W. 2002. Pragmatic directions about language use: Offers and words and relations. *Language in Society* 31. 181–212.

Enfield, N. J. 2003. *Linguistic epidemiology*. London: Routledge.

Fillmore, Charles J. & Atkins, B. T. S. 1992. Towards a frame-based lexicon: The semantics of RISK and its neighbors. In A. Lehrer & E. Kittay (eds.), *Frames, fields, and contrasts: New essays in semantic and lexical organization*, 75–102. Hillsdale, NJ: Erlbaum.

Fillmore, Charles J., Miriam R. L. Petruck, Josef Ruppenhofer & Abby Wright. 2003. 'Framenet in Action: The Case of Attaching'. *International Journal of Lexicography* 16(3). 297–332.

Gladkova, Anna. 2010. *Russkaja kul'turnaja semantika: ėmocii, cennosti, žiznennye ustanovki* [Russian cultural semantics: Emotions, values, attitudes.] Moscow: Languages of Slavic Cultures.

Goddard, Cliff. 1998. *Semantic analysis*. Oxford: Oxford University Press. [2nd ed'n in press]

Goddard, Cliff. 2001. Lexico-semantic universals: A critical overview. *Linguistic Typology* 5(1). 1–66.

Goddard, Cliff. 2006. Natural Semantic Metalanguage. *Encyclopedia of Language and Linguistics,* 2nd edn, Keith Brown (ed.), 544–551. Elsevier.

Goddard, Cliff. 2009a. The conceptual semantics of numbers and counting: An NSM analysis. *Functions of Language* 16(2). 193–224.

Goddard, Cliff. 2009b. *Vegetables*, *furniture*, *weapons*: Functional macro-Categories in the English lexicon. Paper presented at the *2009 Conference of the Australian Linguistic Society*. La Trobe University, Melbourne, 9–11 July 2009.

Goddard, Cliff. 2010. Semantic molecules and semantic complexity (with special reference to "environmental" molecules). *Review of Cognitive Linguistics* 8(1). 123-155.

Goddard, Cliff (ed.). 2006. *Ethnopragmatics: Understanding discourse in cultural context*. Berlin: Mouton de Gruyter.

Goddard, Cliff (ed.). 2008. *Cross-linguistic semantics*. Amsterdam: John Benjamins.

Goddard, Cliff & Anna Wierzbicka. 2007. NSM analyses of the semantics of physical qualities: *Sweet, hot, hard, heavy, rough, sharp* in cross-linguistic perspective. *Studies in Language* 31(1). 675–800.

Goddard, Cliff & Anna Wierzbicka. 2009. Contrastive semantics of physical activity verbs: 'Cutting' and 'chopping' in English, Polish, and Japanese. *Language Sciences* 31. 60–96.

Goddard, Cliff & Anna Wierzbicka. 2010. WANT is a linguistic and conceptual universal (response to Khanina). *Studies in Language* 34(1). 108-123.

Goddard, Cliff & Anna Wierzbicka. To appear. *Men*, *women* and *children*: The conceptual semantics of basic social categories. In Cliff Goddard and Anna Wierzbicka *Words and Meanings: Studies in Lexical and Cultural Semantics*. Oxford: Oxford University Press.

Goddard, Cliff & Anna Wierzbicka (eds.). 1994. *Semantic and lexical universals – Theory and empirical findings*. Amsterdam: John Benjamins.

Goddard, Cliff & Anna Wierzbicka (eds.). 2002. *Meaning and universal grammar – Theory and empirical findings.* Vols I and II. Amsterdam: John Benjamins.

Harkins, Jean & Anna Wierzbicka (eds.). 2001. *Emotions in cross-linguistic perspective*. Berlin: Mouton de Gruyter.

Honda, I. 1965. *The world of origami*. Tokyo & San Francisco: Japan Publications Trading Company & Elmsfor.

Jackendoff, Ray. 1990. *Semantic structures*. Cambridge, MA: MIT Press.

Khanina, Olesya. 2008. How universal is *wanting*? *Studies in Language* 32(4). 818–865.

Koch, Peter. 2001. Lexical typology from a cognitive and linguistic point of view. In Martin Haspelmath, Ekkehard König & Bernard Comrie (eds.), *Language typology and language universals*, Vol 2, 1142–1178. Berlin: Walter de Gruyter.

Koptjevskaja-Tamm, Maria. 2008. Approaching lexical typology. In Martine Vanhove (ed.), *From polysemy to semantic change: A typology of lexical semantic associations*, 3–54. Amsterdam: John Benjamins.

Lehrer, Adrienne. 1992. A theory of vocabulary structure: Retrospectives and prospectives. In Martin Pütz (ed.), *Thirty years of linguistic evolution*, 243–256. Amsterdam: John Benjamins.

Levin, Beth & Malka Rappaport Hovav. 1992. Wiping the slate clean: A lexical semantic exploration. In Beth Levin & Steven Pinker (eds.), *Lexical & conceptual semantics*, 123–151. Cambridge, MA: Blackwell. (Reprinted from *Cognition* 41. 1991, 123–151.)

Levinson, Stephen, Sérgio Meira and the Language and Cognition Group. 2003. 'Natural concepts' in the spatial topological domain – adpositional meanings in cross-linguistic perspective: an exercise in semantic typology. *Language* 79(3). 485–516.

Lucy, John A. 1997. The linguistics of 'color'. In C. L. Hardin & L. Maffi (eds.), *Color categories in language and thought*, 320–346. Cambridge: Cambridge University Press.

Mairal, Ricardo. 2002. Why the notion of lexical template? *Anglogermanica* Online No. 1. www.uv.es/anglogermanica/2002-1/mairal.htm. Accessed 6/4/06.

Mairal, Ricardo & Pamela Faber. 2002. Functional grammar and lexical templates. In R. Mairal & M. J. Pérez Quintero (eds.), *New perspectives on argument structure in Functional Grammar*, 39–94. Berlin: Mouton de Gruyter.

Mairal, Ricardo & Pamela Faber. 2007. Lexical templates within a functional cognitive theory of meaning. *Annual Review of Cognitive Linguistics* 5(1). 137–172.

Mairal, Ricardo & Robert D. Van Valin Jr. 2001. What Role and Reference grammar can do for functional grammar. In María Jesús Pérez Quintero (ed.), *Challenges and Developments in Functional Grammar*, 137–166. La Laguna, Tenerife: Universidad de La Laguna.

Majid, Asifa & Melissa Bowerman (eds.). 2007. *Cutting and breaking events across languages*. Special issue of *Cognitive Linguistics* 18(2).

Majid, Asifa, Melissa Bowerman, Miriam van Staden & James S. Boster. 2007. The semantic categories of cutting and breaking events: A crosslinguistic perspective. *Cognitive Linguistics* 18(2). 133-152.

Majid, Asifa, N. J. Enfield & Miriam van Staden (eds.). 2006. *Parts of the body: Cross-linguistic categorization*. Special issue of *Language Sciences* 28(2-3).

Mel'čuk, Igor A. 1989. Semantic primitives from the viewpoint of the Meaning-Text Linguistic Theory. *Quaderni Di Semantica* 10(1). 65–102.

Morita, Y. C. 1989. *Kiso Nihongo jito [Basic Japanese Dictionary]*. Tokyo: Kadokawa Shoten.

Newman, John (ed.). 2009. *The linguistics of eating and drinking*. Amsterdam: John Benjamins.

Parsons, Terence. 1990. *Events in the semantics of English: A study in subatomic semantics*. Cambridge, MA: MIT Press.

Pawley, Andrew. 1994. Kalam exponents of lexical and semantic primitives. In Cliff Goddard & Anna Wierzbicka (eds.), *Semantic and lexical universals – Theory and empirical findings*, 387–422. Amsterdam & Philadelphia: John Benjamins.

Pawley, Andrew & R. N. H. Bulmer. In press. *A dictionary of Kalam with ethnographic notes*. Canberra: Pacific Linguistics.

Peeters, Bert (ed.). 2006. *Semantic primes and Universal Grammar: Empirical evidence from the Romance languages*. Amsterdam: John Benjamins.

Petruck, M. R. L. 1996. Frame semantics. In J. Verschueren, J.-O. Ostman, J. Blommaert & C. Bulcaen (eds.), *Handbook of Pragmatics*, 1–13. Amsterdam & Philadelphia: John Benjamins.

Rappaport Hovav, Malka & Beth Levin. 1998. Building verb meanings. In Miriam Butt & Wilhelm Geuder (eds.), *The projection of arguments: Lexical and compositional factors,* 97–134. Stanford, CA: CSLI.

Ruiz de Mendoza Ibáñez, José Francisco & Ricardo Mairal. 2008. Levels of description and constraining factors in meaning construction: An introduction to the Lexical Constructional Model. *Folia Linguistica* 42(1). 355–400.

Ruppenhofer, Josef, Michael Ellsworth, Miriam R. L. Petruck, Christopher R. Johnson & Jan Scheffczyk. 2006. *Framenet II: Extended Theory and Practice*. http://framenet.icsi.berkeley.edu/index.php?option=com_wrapper&Itemid=126. Accessed 14/9/09.

Sibly, Ann. 2008. The semantics of verbs of human physical contact. BA Hons sub-thesis. The Australian National University.

Van Valin, Robert D. 2006. Some universals of verb semantics. In Mairal, Ricardo & Gil, Juana (eds.), *Linguistic universals*, 155–178. Cambridge: Cambridge University Press.

Van Valin, Robert D. & David P. Wilkins. 1996. The case for 'effector': Case roles, agents, and agency revisited. In Masayoshi Shibatani & Sandra A. Thompson (eds.), *Grammatical constructions: Their form and meaning*, 289–322. Oxford: Clarendon Press.

Vendler, Zeno. 1967. *Linguistics in philosophy*. Ithaca, NY: Cornell University Press.

Viberg, Åke. 2002. Polysemy and disambiguation cues across languages. The case of Swedish *få* and English *get*. In Bengt Altenberg & Sylviane Granger (eds.), *Lexis in Contrast*, 119–150. Amsterdam: John Benjamins.

Viberg, Åke. 2006. Towards a lexical profile of the Swedish verb lexicon. *STUF* (Sprachtypol. Univ. Forsch.), 59 (1). 103–129.

Wanner, Leo (ed.). 2007. *Selected Lexical and Grammatical Issues in the Meaning-Text Theory: In honour of Igor Melčuk*. Amsterdam: Benjamins.

Wierzbicka, Anna. 1972. *Semantic primitives*. Frankfurt: Athenäum.

Wierzbicka, Anna. 1985. *Lexicography and conceptual analysis*. Ann Arbor: Karoma.

Wierzbicka, Anna. 1991. Semantic complexity: Conceptual primitives and the principle of substitutability. *Theoretical Linguistics* 17. 75–97.

Wierzbicka, Anna. 1992. *Semantics, culture, and cognition: Universal human concepts in culture-specific configurations*. Oxford: Oxford University Press.

Wierzbicka, Anna. 1996. *Semantics: Primes and universals*. Oxford: Oxford University Press.

Wierzbicka, Anna. 1999. *Emotions across languages and cultures: Diversity and universals*. Cambridge: Cambridge University Press.

Wierzbicka, Anna. 2006a. Shape in grammar revisited. *Studies in Language* 30(1). 115–177.

Wierzbicka, Anna. 2006b. There are no 'color universals', but there are universals of visual semantics. *Anthropological Linguistics* 47(2). 217–244.

Wierzbicka, Anna. 2007a. Bodies and their parts: An NSM approach to semantic typology. *Language Sciences* 29. 14–65.

Wierzbicka, Anna. 2007b. Shape and colour in language and thought. In Andrea C. Schalley & Drew Khlentzos (eds.), *Mental states: Vol 2: Language and cognitive structure*, 37–60. Amsterdam: John Benjamins.

Wierzbicka, Anna. 2008. Why there are no 'colour universals' in language and thought. *Journal of the Royal Anthropological Institute* 14. 407-425.

Wierzbicka, Anna. 2009. All people eat and drink: Does this mean that 'eat' and 'drink' are universal human concepts? In John Newman (ed.), *The linguistics of eating and drinking*, 65–89. Amsterdam: John Benjamins.

Wierzbicka, Anna. In press. 'Eating' and 'drinking' in Kalam. In John Bowden & Nikolaus Himmelmann (eds.), *A Festschrift for Andrew Pawley*. Canberra: Pacific Linguistics.

Wierzbicka, Anna. To appear. Universal of kinship: Overcoming Eurocentrism in kinship semantics using seven universal semantic molecules.

Wong, Jock, Cliff Goddard & Anna Wierzbicka. To appear. *Walking, running, jumping*: The semantics of human locomotion.

Notes

1: A bibliography and downloads can be found at the NSM Homepage: www.une.edu.au/bcss/linguistics/nsm

2: Khanina (2008) is a case in point. The author concludes after a survey of 73 languages that WANT cannot be a semantic prime in the NSM sense, because in many languages the exponent of WANT expresses other meanings as well, especially 'like', 'love', and 'seek' or, less commonly, modal-like meanings, such as 'can'. The conclusion does not follow, however, because Khanina adopts a "no polysemy" assumption. Declining to engage in language-internal semantic analysis, she simply accepts the range of dictionary glosses for a given item in any language as representing a single general meaning or 'macrofunction'. For a defence of WANT as a conceptual and linguistic universal, see Goddard and Wierzbicka (2010).

3: They are: SOMEONE, PEOPLE, BODY, SOMETHING/THING; PART, KIND; THIS, THE SAME, OTHER; ONE, TWO, MANY, SOME, ALL; GOOD, BAD; SMALL; WANT, DO, HAPPEN; MOVE, TOUCH; LIVE; TIME, FOR SOME TIME, A LONG TIME, A SHORT TIME, BEFORE; SIDES; LIKE/AS; CAN, IF, NOT, BECAUSE. There is a lesson here: a fairly well developed metalanguage is needed to address even a small selection of examples.

4: Since it is framed as 'one part of someone's body', explication [D] obviously does not apply to the 'heads' of various kinds of animals; furthermore, an animal's head is typically not positioned above the other parts of an animal's body. The NSM view is that body-part words when applied to animals involve a polysemic extension from an anthropocentric prototype. The pattern governing the extension is of some interest in itself, because it depends on an analogical relation; e.g. *head$_2$ (e.g. a cat's head* or *a snake's head)*: 'one part of the body of a living thing of one kind; this part is like one part of people's bodies; this part is of people's bodies is the head [m]'.

5: From a cognitive point of view the discovery that 'hands' is a foundational semantic molecule provides a striking point of convergence between empirical semantics and "embodiment theory".

6: The explication for the singular ‘child [m]’ is the same as for ‘children [m]’, except that the initial component is ‘someone of one kind’, rather than ‘people of one kind’.

7: Goddard (2001) expressed some doubt about the universality of ‘father’, especially in comparison with ‘mother’. See Wierzbicka (to appear) for arguments that appear to vindicate the universal status of the meaning ‘father’. Goddard (2001) also claimed on the basis of Japanese, which appears to make a categorical distinction between *misu* (roughly) ‘water generally’ and *yu* ‘hot water’, that ‘water’ was not a universally lexicalised concept. See Goddard (to appear) for arguments supporting a retraction of this claim.

8: Sibly (2008) is a study of English verbs of physical contact, such as *hit*, *punch*, *slap*, *kiss*, and *kick*.

9: The paradox (cf. Levin and Rappaport Hovav 1992; Jackendoff 1990; Parsons 1990) turns on the fact that for a certain class of verbs (Vendler’s (1967) accomplishment verbs) the simple past tense entails a result-state, yet this entailment fails when the verb appears in the imperfective (progressive). For example, to say that *John cleaned the table* entails that the table ended up clean, whereas to say that *John was cleaning the table* does not carry this entailment.

10: For discussions of Kalam from an NSM perspective, see Pawley (1994), Goddard (2001); see also Wierzbicka (1996: 200-202), Goddard (2001: 27-28).

11 : The component ‘I want this something to be two things’ might meet with the objection that (obviously) one can *cut* an object into more than two pieces. This objection does not seem to be valid, however. For one thing, the reference to ‘two things’ is embedded in a prototypical frame (‘at many times someone does something like this when it is like this: ...’), so there is no claim that *cutting* is always intended to produce two pieces. Just as importantly, the first “cut” would normally separate two pieces from one another, and the same can be said about every subsequent “cut”. On the other hand, there are various lexicalized English expressions with generic mass noun objects (*cutting grass*, *cutting hair*), which do not sufficiently match the prototypical motivational scenario. Such expressions do not necessarily have literal counterparts in other languages, even when those languages have verbs which

approximate English *cut* in meaning. For example, in Polish one cannot say *ciąć trawę* (for 'to cut grass') or *ciąć włosy* (for 'to cut hair'); one says *kosić trawę* and *obcinać włosy*, respectively. This suggests that the English expressions are lexicalised items, and do not reflect the semantics of *cut* as an independent verb.

12: Notwithstanding the preponderance of taxonomic ('kind of') and partonomic ('part of') relations in the nominal lexicon, it is well to remember that there are many non-taxonomic macro-categories as well (Wierzbicka 1985; Goddard 2009b).

13: For the sake of completeness, the SELF_MOTION Frame (which includes *walk*, *run*, and a great many others) is defined as follows: The Self-Mover, a living being, moves under its own power in a directed fashion, i.e. along what could be described as Path, with no separate vehicle.

14: Fillmore et al (2003) describe the steps in the FrameNet workflow as follows: "(1) characterizing schematically the kind of entity or situation represented by the frame (2) choosing mnemonics for labelling the entities or components of the frame, and (3) constructing a working list of words that appear to belong to the frame, where membership of the same frame will mean that phrases that contain the LUs [lexical units] will all permit comparable semantic analyses". There is no reference to any requirement for standardisation of descriptors, let alone any notion of a rigorous and controlled metalanguage.

15: Van Valin (2006: 164-170) proposes a system of cross-linguistic semantic representation using English-derived terms such **consumed´**, **exist´**, **afraid´**, CAUSE, PROC (process), INGR (ingressive), while at the same time asserting – inexplicably, as it seems to me – that "the vocabulary of the semantic metalanguage used in the decomposition … are not words from any particular human language, despite their obvious similarity to English words" (p167).

16: FrameNet's hierarchical structure is implicit in its mechanism of frame inheritance. In effect, FrameNet regards *chopping*, *slicing*, *mincing*, etc. as "kinds of cutting", because the verbs in question are all assigned to the CUT frame.

17: Despite my objections to extensionalism, I would like to record my appreciation of the studies in the Majid and Bowerman (2007) collection, especially in their efforts to go beyond the elicited data to propose conceptual interpretations, drawing on their own expert knowledge of the languages concerned.

18: Majid et al (2007: 145-6) reject "formal *intensional* representations" on the grounds that they typically consist of primitive predicates like CAUSE and BECOME, together with unanalysed constants such as BROKEN. Since such supposed constants carry most of the burden of distinguishing one meaning from another, they carry a tremendous burden and yet, according to Majid et al (2007), they are "essentially black boxes, since, so far, there has been no account of what it actually *means* to "break" something or to "smash" something". Needless to say, however, this criticism does not apply to the NSM research tradition; see, for example, Goddard (1998: 281) and Goddard (2006) for analyses of English *break* fully into semantic primes.

19: From a practical point of view, wider use of the NSM vocabulary of semantic description will help to minimize the proliferation of hard-to-reconcile individual terminologies. As an example of how divergent descriptive metalanguages "create obstacles for evaluating cross-linguistic connections even between studies of high semantic and lexicographic quality", Koptjevskaja-Tamm (2008) adduces Enfield's (2003) and Viberg's (2002, 2006) studies of Swedish *få* and Lao *daj4*, respectively. Though the two verbs would appear to be similar in their basic meaning (something like 'get' or 'acquire') and in their patterns of grammaticalization, it is impossible to directly compare the studies on account of their divergent descriptive terminologies.