

Semantic Molecules

Cliff Goddard
School of Languages,
Cultures and Linguistics,
University of New England,
Armidale NSW 2351

This paper explains and explores the concept of “semantic molecules” in the NSM methodology of semantic analysis. A semantic molecule is a complex lexical meaning which functions as an intermediate unit in the structure of other, more complex concepts. The paper undertakes an overview of different kinds of semantic molecule, showing how they enter into more complex meanings and how they themselves can be explicated. It shows that four levels of “nesting” of molecules within molecules are attested, and it argues that while some molecules, such as ‘hands’ and ‘make’, may well be language-universal, many others are language-specific.

Keywords:

semantics, NSM, semantic molecule, semantic prime, lexical decomposition

1. Semantic primes: the “atoms” of meaning

As is well-known, the Natural Semantic Metalanguage (NSM) is a decompositional approach to semantics based on language-internal reductive paraphrase (Wierzbicka 1996; Goddard & Wierzbicka 2002; Goddard 2006). The key theoretical concept is that of semantic primes, i.e. elementary lexical meanings which cannot be further paraphrased in simpler terms. After a long program of empirical/analytical research—from Wierzbicka (1972), through Goddard and Wierzbicka eds (1994, 2002), and continuing—NSM researchers claim with some confidence to have identified 65 universal semantic primes. Semantic primes are the bedrock of linguistic meaning: the terminal elements of semantic analysis in any language.

A substantial and growing body of evidence indicates that these elementary meanings are lexicalised in all languages, i.e. that all languages possess concrete exponents (words, bound morphemes or phrasemes) which can be used to express the same inventory of primes. Semantic primes have an inherent syntax which also appears to be universal, in the sense that it manifests itself, albeit with formal variations, in all languages. One can therefore think of semantic primes and their associated grammar as the intersection of all languages.

The current inventory of primes (Wierzbicka 2006a, in press; Goddard 2007) is given in the Appendix, using English and Japanese exponents. Detailed “whole metalanguage” studies of Spanish, Chinese, Malay, Polish, Mbula, and Lao can be found in Goddard and Wierzbicka (eds, 2002). For similar studies of French, Italian, Portuguese, and Spanish, see Peeters (ed., 2006), and for Amharic, East Cree and Korean, see Goddard (ed., in press). A comprehensive bibliography of NSM publications can be accessed at: [www.une.edu.au/arts/LCL/nsm/index.htm].

2. Semantic molecules

It is well established that lexical items in many semantic domains (emotion terms, speech-act verbs, value terms, discourse particles, and “abstract” vocabulary in general) can be decomposed directly into semantic primes (cf. Wierzbicka 1992, 1996, and other works). It is equally apparent, however, that words of many kinds resist decomposition in one go, as it were, directly into semantic primes. For example, plausible explications for words like *cat*, *mouse* and *horse* must begin with the component ‘animals [M] of one kind’, and plausible explications for *oak*, *elm*, and *pine* must begin with the component ‘trees [M] of one kind’ (molecules are marked with [M] in explications). The concepts of ‘animals’ and ‘trees’ are themselves complex and further decomposable, but they function as units in the explications of many other concepts. By the term semantic molecule, then, we

understand a complex lexical meaning which functions as a semantic unit (or “chunk”) in the structure of other, more complex concepts.¹

To develop an idea of the role, range and nature of semantic molecules in the English lexicon, it is useful to consider several explications which have been developed independently of these questions. In this, we are following the normal NSM practice of trying to induce generalisations about semantic structure from empirical work. For reasons of space, we will have to confine ourselves to partial explications. Explication [A] below is a partial explication for a natural kind term—*cats* (Wierzbicka 1985; Goddard 1998). It follows the semantic template for natural kind terms: (a) category within the taxonomic hierarchy, (b) habitat, (c) size, (d) appearance, (e) behaviour, (f) relation with people.

Running through the four sections presented here, we see that section (a) includes a taxonomic “life form” category, ‘animals [M]’, as a semantic molecule. Section (b) gets by without any additional semantic molecules: it is phrased almost exclusively in semantic primes.² Regarding section (c), the “size” section, Wierzbicka (1985) has argued that anthropocentrism pervades human construal of the physical world, and, in particular, that the human body furnishes a reference point for judgements of relative size. In the case of *cats*, this component depends on the potential to pick up a cat with two hands, with both ‘pick up [M]’ (a physical action verb) and ‘hands [M]’ (a body-part word) functioning as semantic molecules. The subsequent “appearance” section contains numerous molecules: yet more body-part terms (‘head’, ‘ears’, ‘eyes’, ‘mouth’, ‘tail’, ‘feet’),³ and terms of two additional kinds: shape descriptors (‘round’, ‘long’, ‘pointed’), and physical qualities (‘soft’, ‘sharp’).

[A] A partial explication for *cats*

- | | | |
|----|---|----------|
| a. | animals [M] of one kind | CATEGORY |
| b. | animals [M] of this kind live with people
sometimes they live in places where people live
sometimes they live near places where people live | HABITAT |
| c. | they are not big
a person can pick up [M] one with two hands [M] | SIZE |

¹ The Moscow School of semantics (Mel’cuk and colleagues) has long championed the need for intermediate-level decomposition (Apresjan 2000), but without the constraint that intermediate-level concepts exist as meanings of ordinary lexical items in the language concerned.

² *Cats* being “domestic” animals, the habitat component makes reference to them living with people or near where people live; and PEOPLE is a semantic prime. The comparable component for some other natural kinds may well require semantic molecules; for example, *whales* and *tuna* presumably ‘live in the sea [M]’.

³ The word ‘have’ which appears in the explications in association with body-parts, e.g. ‘they have soft [M] fur [M]’ is not the NSM prime HAVE, but a morphosyntactic realisation of the prime PART (HAVE PARTS).

- d. they have soft [M] fur [M] APPEARANCE
 they have a round [M] head [M]
 they have pointed [M] ears [M]
 their ears [M] are on two sides of the top [M] part of the head [M]
 their eyes [M] are not like people's eyes [M]
 they have some long [M] hairs [M] near the mouth [M], on two sides of the mouth [M]
 they have a long [M] tail [M]
 they have soft [M] feet [M]
 they have small sharp [M] claws [M]

Just as explication [A] can serve as an exemplar of natural kind terms, so explication [B] for *chairs* can stand as a (partial) exemplar of artefact terms (Goddard 2007). The top-level component contains a prolific semantic molecule: 'make [M]'. Any artefact term designates things which exist 'because people make [M] them'. The semantic template for artefact terms continues with a "purpose" section specifying people's motivation for making such things. In the case of *chairs*, this obviously involves 'sitting'; roughly speaking, chairs are things of the kind people make so as to be able to sit comfortably while doing things. To capture this notion, it is not necessary to employ 'comfortably' as a semantic molecule, because the required meaning can be rendered directly in semantic primes, as shown in the final line of [B]. What is necessary, however, is to employ 'sit [M]' as a semantic molecule, because the explication of 'sit' is much too complex (cf. Wierzbicka 2006b) to be substitutable directly into [B].

[B] A partial explication for *chairs*

- a. things of one kind CATEGORY
 things of this kind exist because people make [M] them
- b. people make [M] them because they want them to be in places where people live PURPOSE
 they want this because they want people to be able to sit [M] on them
 when they have to do something somewhere for some time
 people want people to be able to sit [M] on them at these times
 because they don't want them to feel something bad in their bodies

Other bodily activity verbs such as 'eat' and 'drink' are also known to be prolific semantic molecules. They are needed in words for artefacts such as *spoon*, *plate*, *cup*, and *bottle*, for food and drink words such as *meat*, *bread*, *tea* and *coffee*, and in many other contexts. It is not possible for reasons of space to discuss the structure of bodily activity verbs in any detail, but it is obvious that they call for the use of body-part terms as semantic molecules; for example, to explicate *eat* and *drink* one must include the molecule 'mouth'.

[C] and [D] below are partial explications for two complex physical activity verbs: English *cutting* and *chopping* (Goddard & Wierzbicka to appear a). Verbs

of this kind follow the semantic template: (a) lexico-syntactic frame, (b) prototypical motivation, (c) instrument, (d) manner, (e) what happens to the object, (f) potential outcome. The shared top-level component, the lexico-syntactic frame, can be phrased without recourse to semantic molecules. It characterises both verbs as activities by which a person produces some effect upon an object in a controlled fashion by means of an instrument.

Comparing the next two sets of components in their respective explications, certain key differences between *cutting* and *chopping* start to become apparent. The prototypical motivation in both cases includes the wish that ‘this something not be one thing anymore’ (i.e. the intent to create some kind of “separation” or “division”). Yet in the case of *chopping* it is specified that this motivation applies to a prototypical object which is ‘something hard [M]’, whereas with *cutting* there is no corresponding characterisation of the prototypical object. Continuing down the prototypical motivation component for *cutting*, it reads: ‘I want this something to be two things, I want these two things to have straight [M] edges [M]’. The word ‘straight’ is another example of a shape descriptor, but ‘edges’ represents a different kind of molecule from those we have previously observed. It can be termed an “ethnogeometrical concept” (Brotherson in press). The same molecule is found later in the “instrument” section of both explications: *cutting* and *chopping* require an instrument with a ‘sharp [M] edge [M]’.

[C] A partial explication for *cutting*:

Someone X was cutting something Y (e.g. some paper, a cake) with something Z

- a. someone X was doing something to thing Y with thing Z
 for some time LEXICO-SYNTACTIC FRAME
 because of this, at the same time something was happening to thing Y
 as this someone wanted
- b. this someone was doing it as people do when they do something
 to something PROTOTYPICAL MOTIVATION
 because a short time before they thought about this something like this:
 “I want this something not to be one thing anymore
 I want this something to be two things
 I want these two things to have straight [M] edges [M]”
- c. when someone does something like this because they think like this, INSTRUMENT
 they do it with something
 this something is not a part of this someone’s body
 this something has a sharp [M] edge [M]

[D] A partial explication for *chopping*:

Someone X was chopping something Y (e.g. some wood) with something Z

- a. someone X was doing something to thing Y with thing Z
 for some time LEXICO-SYNTACTIC FRAME

- because of this, at the same time something was happening to thing Y
as this someone wanted
- b. this someone was doing it as people do when they do PROTOTYPICAL MOTIVATION
something to something hard [M]
because a short time before they thought about this something like this:
“I want this something not to be one thing anymore
I want this one big thing to be many small things”
- c. when someone does something like this because they think like this, INSTRUMENT
they do it with something
this something is not a part of this someone’s body
it has two parts
one of these two parts has a sharp [M] edge [M]
the other part is a long [M] part

I have to emphasise that my focus is not on the validity or descriptive adequacy of these explications, but on their structure. No doubt numerous questions and objections have occurred to readers on matters of detail, which it would be impossible to anticipate and respond to here. The point of the exposition in this section has been to illustrate the role of semantic molecules in different kinds of explications, and to establish that certain particular kinds of molecules appear to be common: body-parts, shape descriptors, physical qualities, ethnogeometrical terms, and bodily actions. Before moving to other issues about semantic molecules, a matter of clarification should be attended to.

3. Recurrent “non-molecular” components

There are many recurrent components across the lexicon which are not semantic molecules, in the sense under discussion here, because they are not “encapsulated” as the meanings of surface lexical items. For example, many nouns begin with top-level categorical components such as: ‘one part of someone’s body’ (for body-part terms), ‘living things of one kind’ (for natural kind terms), ‘a place of one kind where people live’ (for words like *town*, *city*, *village*, etc). Likewise, many verbs contain high-level components related to semantic role or argument structure; for example: ‘someone did something’, ‘someone did something to something/someone else’, ‘someone did something to something with something’, ‘something happened to something else because of it’, and so. Many verbs also contain a component like ‘this someone did this because this someone wanted to do it’, corresponding to the technical notion of volitional action.

To take one further example, from a smaller segment of the lexicon, emotion adjectives such as *sad*, *annoyed* and *homesick*, conform to a semantic template which begins: ‘someone feels something (good/bad), as people feel when they

think like this: – – ’ (followed by a prototypical cognitive scenario setting out certain characteristic thoughts and wants).

Such recurrent but “non-molecular” components can be extremely significant for the interface between lexical and grammatical semantics, and for the creation of lexical classes; but they are simple enough in their internal semantic structure to be spelt out in relatively short strings composed purely of semantic primes. They are not semantic molecules, because they are not the meanings of surface lexical items.

4. Numbers and kinds of semantic molecules

How many productive semantic molecules are there? At this formative stage of research, the answer is not very clear. Wierzbicka (in press) has hazarded an estimate (for English and Polish) of 100–200.

It is known that productive molecules are drawn from at least nine categories. The first six have already been briefly exemplified: (i) parts of the body, such as ‘hands’, ‘mouth’, ‘legs’, (ii) physical activities, such as ‘eat’, ‘drink’, ‘sit’, (iii) the verb ‘make’, (iv) shape descriptors, such as ‘long’, ‘round’, ‘flat’, (v) physical qualities, such as ‘hard’, ‘sharp’, and ‘straight’ (vi) ethnogeometrical terms, such as ‘edges’ and ‘ends’, and (vi) taxonomic concepts, such as ‘animal’, ‘bird’, and ‘tree’. To this list one can add at least three further categories: (vii) macro-terms from the natural environment, such as ‘ground’, ‘sky’, and ‘sun’, (viii) “elemental” concepts (for want of a better term), such as ‘water’ and ‘fire’, and (ix) basic social categories, i.e. kinds of people, such as ‘men’, ‘women’, and ‘children’.

Only the briefest justifications can be given here for the latter items. ‘Ground’ is needed (among other things) for explications of *walking* and *running*, for numerous geographical concepts, and for certain dimensional concepts, such as *tall* (Wong, Goddard & Wierzbicka to appear; Wierzbicka 2006a). ‘Sky’ is needed for words like *sun*, *moon*, *stars*, and *clouds*, and for the colour *blue*; and ‘sun’ for the explication of *day*, the colour *yellow*, and for *warm* (Goddard & Wierzbicka in press). ‘Water’ is needed in the explication of words like *rain*, *river*, *sea*, *swim*, *float*, *wash*, *pour*, *wet*, and *dry*; and ‘fire’ is needed for *hot*, *burn*, *smoke*, *ashes*, and for the colour *red*. ‘Men’ and ‘women’ are needed for kin and stage-of-life nouns like *father*, *mother*, *boy* and *girl*, for nouns designating gender-specific items of clothing, such as *dress* and *bra*, and body-parts such as *breast*, *penis* and *beard* (Goddard & Wierzbicka to appear b). ‘Children’ is needed for words like *baby*, *adult*, *mummy*, *daddy*, *toy*, *doll*, *school*, *play* (in one of its senses; Alexander 2006), and *innocent*.

5. Explications for some semantic molecules

This section presents a selection of explications for semantic molecules of several kinds. The relevance is two-fold: first, to show how some common semantic molecules can themselves be explicated; second, to establish some facts about the “nesting” of molecule within molecule.

A useful starting point is parts of the body (Wierzbicka 2007). Consider the word *head* (in the sense of a person’s *head*).⁴ As explicated in [D], it requires in its explication not only semantic primes establishing its parthood relationship with the body (‘one part of someone’s body’, etc.) and position (‘above all the other parts of this someone’s body’), but also an indication of its shape—in the form of the semantic molecule ‘round’. The meaning ‘head’ in turn functions as a molecule in certain other body-part meanings, such as *eyes* (shown in [E]), *ears*, and *hair*, among others.

[D] *head*₁ (*someone’s head*)

- a. one part of someone’s body
- b. this part is above all the other parts of this someone’s body
- c. this part is round [M]

[E] *eyes* (*someone’s eyes*)

- a. two parts of someone’s body
- b. they are on one side of this someone’s head [M]
- c. because people’s bodies have these parts, people can see

Shape descriptors are needed in many other body-part concepts too. For example, the concept of *legs* requires the descriptor ‘long’, as shown in [F]. ‘Legs’ in turn is a molecule in verbs like *walk* and *run*, among others.

[F] *legs* (*someone’s legs*)

- a. two parts of someone’s body
- b. these two parts are below all the other parts of this someone’s 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

Already we can discern several levels of semantic nesting: shape descriptors (like ‘long’ and ‘round’) are molecules inside body-part concepts (like ‘legs’ and

⁴ Body-parts of other kinds of living things require distinct but related explications, i.e. words like *head*, *legs*, and so on are polysemous between a primary sense based on the human body, and a secondary sense applying to other species by analogy with the human body. For example, the *head*₂ of an animal or a snake can be explicated as follows: ‘one part of the body of a living thing of one kind (e.g. cow, snake); this part is like one part of people’s bodies; this part of people’s body’s is the head₁’.

‘head’); and body-part concepts in turn are molecules inside bodily action verbs (like ‘run’ and ‘eat’). But what about the semantics of shape descriptors?

Wierzbicka (2003, 2006a, 2007) has investigated these in depth, and has drawn a remarkable conclusion: that shape concepts depend in part on the concept of ‘hands’.⁵ The nub of the argument is that shape concepts designate properties which can be detected not only visually, but also by “touch”—i.e. by the touch of our hands. This aspect is crucial to capturing the tangibility and physicality of shape concepts (as compared with colour concepts, for example, which depend purely on seeing). A sample explication is given in [G].⁶

[G] *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 might seem at this point that we are facing a fatal circularity. How can shape descriptors depend on a body-part concept, namely ‘hands’, while at the same time body-part concepts depend on shape descriptors? In fact, there is no circularity on account of another remarkable result (Wierzbicka 2003, 2006a, 2007). Of all the body-part concepts, it seems that ‘hands’ alone can be explicated purely in terms of semantic primes, without recourse to any shape (or other) semantic molecules. The explication for ‘hands’ is shown in [H].

[H] *hands (someone’s hands)*

- a. two parts of someone’s body
- b. they are on two sides of this someone’s 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, all the parts on one side of one of these two parts can touch
 all the parts on one side of the other at the same time
- f. because people’s bodies have these two parts, people can do many things with many things
 as they want

⁵ Some physical quality concepts (Goddard and Wierzbicka in press) also depend on the concept of ‘hands’. For example, the concept of physical ‘hardness’ depends on the idea that if someone wants to “make an impact” on something hard (e.g., to break, deform, or scratch it), they cannot do this with the hands alone: they have to use some kind of instrument.

⁶ Wierzbicka (2006b) argues that the nearest Polish word *podłużny* ‘elongated, oblong’ differs slightly but discernibly from English *long* in component (b): *podłużny* requires only that two parts can be construed as ‘far’ apart (not necessarily as ‘very far’ apart). The difference accounts for differences in the range of use of the two words; a matchbox or a paperback book, for example, could be described as *podłużny*, but not as *long*.

- g. because people's bodies have these two parts, people can touch many things as they want

Wierzbicka (2003) argues that the foundational status of the concept of 'hands' makes sense from the point of view of an embodied concept of cognition. "[H]uman hands", she argues, "mediate, to a large extent, between the world and the human mind", because of the crucial role played by "handling" things and by touching things in an exploratory way with the hands. Equally, the hands are our principal "bodily instruments" for making things, for using things, and for doing things of many other kinds. It makes a lot of sense, therefore, that the concept of 'hands' is a foundational semantic molecule in so many human concepts.

6. Conclusions

Before summarising our conclusions, it may be worthwhile to reiterate that semantic molecules cannot be postulated *ad hoc*, but are subject to strict constraints. These include, first and foremost, a demonstration that any postulated molecule can be explicated without circularity down to the level of semantic primes. As we have just seen, this may involve working through several levels of semantic nesting, but it is crucial to be sure that no implicit circularity is introduced into the system by the addition of any given semantic molecule. Second, semantic molecules must exist as the meanings of lexical units in the language. Originally this constraint was motivated by cognitive considerations, i.e. by the notion that existing word-meanings have a *prima facie* claim to cognitive reality and are not simply inventions of the analyst. From another point of view, the lexicalisation constraint means that semantic molecules do not add any additional complexity to the overall semantic system, in the sense that the meanings which function as molecules are a sub-set of the lexical meanings already present in the language. Third, semantic molecules are only to be introduced when they are needed, i.e. when it emerges from the analytical process that the required semantic content cannot be represented directly in an intelligible fashion using semantic primes. In this connection, it is well to recall that many divisions of the lexicon can be explicated entirely without semantic molecules; for example, the lexical domains of emotion terms, speech-act verbs, and value terminology. The finding that semantic molecules are necessary in some domains is an empirical finding, not a matter of analytical convenience.

To return to the main theme of the present paper, we can conclude that there are as many as four levels of semantic nesting within highly complex concepts, such as those for natural kinds and artefacts. In the explication 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 level of 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).

It seems likely on current evidence that some semantic molecules are universal. This applies to 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 (Goddard 2001; Wierzbicka 2007). The same can be argued for 'make', given the universality of artefact concepts in the lexicons of all languages. Other candidates for universal semantic molecules are certain other body-parts such as 'eyes' and 'ears' (Wierzbicka 2007), basic social categories like 'men' and 'women' (Goddard & Wierzbicka to appear a), and perhaps the sociobiological concept 'mother', given the foundational status of the mother-child relationship in kinship systems (Wierzbicka 1992).

It is also clear that some semantic molecules are language-specific. This is only to be expected in the case of taxonomic categories: it is well established that there are languages which lack exact equivalents for words like 'animal', 'bird' and 'tree' (cf. Goddard 2001). Likewise, it is not too surprising that some languages employ particular concepts as semantic molecules while others do not. For example, in Polish the word *grzyb* 'mushroom' functions as a semantic molecule: there are many common Polish words for kinds of *grzyby* 'mushrooms', and various endonymic words which include 'grzyb [M]' in their meanings. Perhaps more surprising is the claim that lower-level molecules such as shape descriptors and ethno-geometrical terms can also vary somewhat from language to language, but as mentioned in Note 6, Wierzbicka (2006a, in press) argues that English 'long [M]' does not exactly match the comparable Polish molecule 'podłużny [M]' 'elongated, oblong'; and Brotherson (in press) argues that English 'ends [M]' differs from its nearest counterpart 'tapu [M]' in Makasai (East Timor). Wierzbicka (2006b) has argued that in English 'colour' functions as a semantic molecule in words like *red*, *blue*, *green*, etc., while many other languages lack "colour words" in the true sense, because their visual descriptor words do not involve any comparable molecule.

It should be clear that the concept of semantic molecules is an extremely fertile one, with multiple ramifications for our understanding of the overall structuring of the lexicon, for lexical typology, for language acquisition, and for language and cognition studies. In coming decades we can expect semantic molecules to be one of the most vibrant research fronts in semantics.

Appendix

Table 1: Semantic primes—English exponents

Substantives:	I, YOU, SOMEONE, SOMETHING/THING, PEOPLE, BODY
Relational substantives:	KIND, PART
Determiners:	THIS, THE SAME, OTHER/ELSE
Quantifiers:	ONE, TWO, MUCH/MANY, SOME, ALL
Evaluators:	GOOD, BAD
Descriptors:	BIG, SMALL
Mental predicates:	THINK, KNOW, WANT, FEEL, SEE, HEAR
Speech:	SAY, WORDS, TRUE
Action, events, movement, contact:	DO, HAPPEN, MOVE, TOUCH
Location, existence, possession, specification:	BE (SOMEWHERE), THERE IS/EXIST, HAVE, BE (SOMEONE/ SOMETHING)
Life and death:	LIVE, DIE
Time:	WHEN/TIME, NOW, BEFORE, AFTER, A LONG TIME, A SHORT TIME, FOR SOME TIME, MOMENT
Space:	WHERE/PLACE, HERE, ABOVE, BELOW, FAR, NEAR, SIDE, INSIDE
Logical concepts:	NOT, MAYBE, CAN, BECAUSE, IF
Augmentor, intensifier:	VERY, MORE
Similarity:	LIKE/AS

Table 2: Semantic primes—Japanese exponents

WATASHI <i>I</i> , ANATA <i>you</i> , DAREKA <i>someone</i> , NANIKA/MONO/KOTO <i>something/thing</i> , HITO/HITOBITO <i>people</i> , KARADA <i>body</i>
SHURUI <i>kind</i> , BUBUN <i>part</i>
KORE <i>this</i> , ONAJI <i>the same</i> , HOKA <i>other</i>
HITO-/ICHI- <i>one</i> , FUTA-/NI- <i>two</i> , TAKUSAN <i>much/many</i> , IKUTSUKA <i>some</i> , MINNA <i>all</i>
II <i>good</i> , WARUI <i>bad</i>
OOKII <i>big</i> , CHIISAI <i>small</i>
OMOU <i>think</i> , SHIRU <i>know</i> , HOSHII/-TAI/NOZOMU <i>want</i> , KANJIRU <i>feel</i> , MIRU <i>see</i> , KIKU <i>hear</i>
IU <i>say</i> , KOTOBA <i>words</i> , HONTOO <i>true</i>
SURU <i>do</i> , OKORU/OKIRU <i>happen</i> , UGOKU <i>move</i> , FURERU <i>touch</i>
(DOKOKA) IRU/ARU <i>be (somewhere)</i> , IRU/ARU <i>there is</i> , MOTSU <i>have</i> , (DAREKA/NANIKA) DEARU <i>be (someone/something)</i>
IKIRU <i>live</i> , SHINU <i>die</i>
ITSU/TOKI <i>when/time</i> , IMA <i>now</i> , MAE <i>before</i> , ATO <i>after</i> , NAGAI AIDA <i>a long time</i> , MIJIKAI AIDA <i>a short time</i> , SHIBARAKU NO AIDA <i>for some time</i> , SUGUNI <i>moment/in one moment</i>
DOKO/TOKORO <i>where/place</i> , KOKO <i>here</i> , UE <i>above</i> , SHITA <i>below</i> , CHIKAI <i>near</i> , TOOI <i>far</i> , MEN <i>side</i> , NAKA <i>inside</i>
-NAI <i>not</i> , TABUN <i>maybe</i> , DEKIRU <i>can</i> , -KARA <i>because</i> , MOSHI (BA) <i>if</i>
SUGOKU <i>very</i> , MOTTO <i>more</i>
YOO/YOONI <i>like/as</i>

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 different morphosyntactic properties, including word-class, in different languages • They can have combinatorial variants (allolexes) • Each prime has well-specified syntactic (combinatorial) properties.

Acknowledgements

I would like to thank Anna Wierzbicka for many valuable suggestions and discussions about matters dealt with in this study, and two anonymous reviewers for their helpful comments. I am also indebted to Akiko Torri for valuable assistance in identifying the Japanese exponents of primes given in Table 2.

References

- Alexander D 2006 *Literal, figurative, metaphorical: A semantic inquiry into the semantic field of game and play* in English PhD thesis University of New England.
- Apresjan J 2000 *Systematic Lexicography* [Trans. Kevin Windle] Oxford Oxford University Press.
- Brotherson A In press 'Ethnogeometrical concepts in Makasai' in Cliff Goddard (ed.) *Cross-Linguistic Semantics* Amsterdam John Benjamins.
- Goddard C 1998 *Semantic Analysis: A Practical Introduction* Oxford Oxford University Press.
- Goddard C 2001 'Lexico-semantic universals: A critical overview' *Linguistic Typology* 5(1): 1–66.
- Goddard C 2006 'Natural Semantic Metalanguage' *Encyclopaedia of Language and Linguistics* 2nd edition Keith Brown (ed.) Oxford Elsevier: 544–551.
- Goddard C 2007 'Semantic primes and conceptual ontology' in *Ontolinguistics: How Ontological Status Shapes the Linguistic Coding of Concepts* AC Schalley & D Zaefferer (eds) Berlin Mouton de Gruyter: 145–174.
- Goddard C (ed.) In press *Cross-Linguistic Semantics* Amsterdam John Benjamins.
- Goddard C & Wierzbicka A 2002 'Semantic primes and universal grammar' in C Goddard & A Wierzbicka (eds) *Meaning and Universal Grammar – Theory and Empirical Findings* Vol I Amsterdam John Benjamins: 41–85.
- Goddard C & Wierzbicka A In press 'NSM analyses of the semantics of physical qualities: *Sweet, hot, hard, heavy, rough, sharp* in cross-linguistic perspective' *Studies in Language* 31(4).
- Goddard C & Wierzbicka A To appear a 'Towards a semantic typology of physical activity verbs: 'cutting' and 'chopping' in English, Polish, and Japanese'.
- Goddard C & Wierzbicka A To appear b '*Men, women and children: The semantics of basic social categories*'.
- Goddard C & Wierzbicka A (eds) 1994 *Semantic and Lexical Universals – Theory and Empirical Findings*. Amsterdam: John Benjamins.
- Goddard C & Wierzbicka A (eds) 2002 *Meaning and Universal Grammar – Theory and Empirical Findings* Vols I and II Amsterdam John Benjamins.
- Peeters B (ed.) 2006 *Semantic Primes and Universal Grammar: Empirical Findings from the Romance Languages* Amsterdam John Benjamins.
- Wierzbicka A 1972 *Semantic Primitives* [trans. A Wierzbicka & J Besmeres] Frankfurt am Main Athenäum.
- Wierzbicka A 1985 *Lexicography and Conceptual Analysis* Ann Arbor Karoma.
- Wierzbicka A 1992 *Semantics: Culture and Cognition* New York Oxford University Press.
- Wierzbicka A 1996 *Semantics: Primes and Universals* New York Oxford University Press.
- Wierzbicka A 1999 *Emotions Across Languages and Cultures: Diversity and Universals* Cambridge Cambridge University Press.
- Wierzbicka A 2003 Empirically established universals of language and their importance for cognitive science: The human conceptualisation of shape Keynote paper at the 4th International Conference on Cognitive Science University of New South Wales Sydney July 2003.

- Wierzbicka A 2006a 'Shape in grammar revisited' *Studies in Language* 30(1): 115–177.
- Wierzbicka A 2006b 'There are no "color" universals, but there are universals of visual semantics' *Anthropological Linguistics* 47(2): 217–244.
- Wierzbicka A 2007 'Bodies and their parts: An NSM approach to semantic typology' *Language Sciences* 29(1): 14–65.
- Wierzbicka A In press 'The theory of the mental lexicon' *HSK (Handbücher zur Sprach- und Kommunikationswissenschaft) Slavic Languages* Berlin/New York Walter de Gruyter.
- Wong JO, Goddard C & Wierzbicka A To appear 'Walk, run and jump: The semantics of verbs of bodily motion'.